ISSN: 2253-0347 مجلـة علوم الإنسان و المجتمع EISSN: 2602-781X
Journal of Human and Society Sciences
volume: 14 /N°: 01 /Year:2025 /pp: 1310-1344

Research Data Management in the Arab Repositories Indexed in Re3data Registry: Analytical Study

Received date:09/01/2024 Accepted date:07/04/2025 Published date: 25/04/2025

Badia Boukhobza [1], Kamel Boukerzaza [2] , Meriem Boulahlib[3]

[1,2,3] Laboratory : New technologies of information and their role in national development, Instute of library science & documentation,

University constantine2- Algeria, ***Email:***

[1]Badia.boukhobza@univ-constantine2.dz

[2]Kamel.Boukerzaza@univ-constantine2.dz

[3]Meriem.Boulahlib@univ-constantine2.dz

Abstract:

This study aims to analyse different aspects of research data management available via (06) Arab repositories indexed in Re3data. We employed a descriptive method to examine data sourced from Re3data.

The study concluded that research data management in Arab repositories encounter challenges primarily related to infrastructure and quality management standards. Nevertheless, Arab repositories respond positively to the open access movement for academic content and open science practices in academic communication.

Keywords: Research data sets, Research data management,digital repositories, Arab research data repositories, Re3data registry.

Corresponding Author: Badia Boukhobza , Badia.boukhobza@univ-onstantine2.dz

1 .INTRODUCTION:

Research data management has become a vital and modern subject, attracting considerable interest from researchers and academic institutions because of its role in raising the value of scientific research and improving openness and trustworthiness. This renders research results valuable, capable of producing original scientific knowledge, and potentially reusable and shareable for developing novel research programs. (Nwabugwu & Godwin, 2020, p. 7) to achieve socio-economic development for societies.

The term "research data" encompasses all digital data sets created during the scientific research process, which may include textual or statistical data, questionnaires, images, audiovisual data such as films and audio recordings, laboratory observations, experimental results, workflow systems, survey measurements, and algorithms (Elsayed & Saleh, 2018, p. 281). Thus, research data encompasses all types and forms of raw scientific content without analyzing or interpreting phenomena.

Research data management is conducted within the framework of FAIR principles, which are guidelines and recommendations ensuring efficiency in processing and analyzing data and effectiveness in making it available to achieve quality and wide visibility (Azeroual & others, 2022, p. 4). These principles form a practical basis for embodying open science through the Findability, Accessibility, Interoperability, and Reusability of research data sets via a technological infrastructure called Research Data Repositories (RDR).

Research Data Repositories "are associated with all aspects of the dataset lifecycle, including data collection, evaluation, organization, archiving for long-term preservation, and reuse" (Mohamed, Abdulbasset, & Taha, 2021, p. 446). The management of research data collections in these repositories is linked to the availability of metadata and its standardized storage, as well as the creation of management plans (Nabti, Achouri, & Atrouz, 2017, p. 15) and persistent identifiers.

The Registry of Research Data Repositories (Re3data) initiative, which contains descriptive information about 3,299 research data repositories worldwide across all knowledge disciplines, aims to enable the academic community to discover and access research data, choose the appropriate repository when depositing data, and govern research data repositories by research institutions and funding agencies (Re3data, 2024).

In this endeavour, some Arab countries have worked on building and implementing research data repository projects and indexing them in Re3data to manage research data and make it easily accessible to researchers globally and in the Arab world. This is especially important as it is a key part of the open science movement that many global scientific organizations, such as UNESCO, seek to embody and promote the culture of its use.

Based on the above, we aim in this research paper to present a landscape of Arab research data repositories listed in Re3data guided by the following objectives:

To delineate the features of research data collections accessible through Arab repositories focusing on content and file formats.

Highlight access levels capabilities in Arab research data repositories listed in Re3data and the licenses used.

Discover the technical infrastructure utilized in managing research data available through Arab repositories.

Discover the extent to which Arab repositories listed in Re3data adhere to quality management standards .

2. Literature Review:

Several studies related to research data were reviewed, which helped us define the study's dimensions and various indicators for comprehensive analysis and Evaluation of Arab repositories. We observed that academic literature in the Arab world lacks research addressing the analytical and evaluative dimensions of research data repositories, as this is a recent topic in research practices.

In 2022, Cintia De Azevedo Lorenço, Marcello Mundim Rodrigues, and Guilherme Ataide studied the nature of research data in South American repositories. They aimed to determine the characteristics of research data repositories created and funded by higher education institutions in South America and describe the formats and extensions of files containing deposited datasets.

The study relied on eight (08) repositories in Re3data, including a random sample of 1,115 research datasets. The datasets were analyzed using tables and charts to obtain results.

The authors found that the content formats of files containing research data available through the studied repositories primarily focus on textual data. This data can be found in a single file with

a single extension or heterogeneously across diverse files with multiple extensions.

The researchers emphasized the increasing need to describe files better and specify their formats and extensions so that users of research data repositories can retrieve and access them, ensuring best practices in their management (Mundim Rodrigues, De Azevedo Lourenço, & Ataide Dias, 2022).

In a 2019 study, Raj Kumar Bhardwaj from the University of Delhi tracked the development of research data repositories in India and analyzed their content to study subject coverage, content formats, author identifier systems, persistent digital entity identifiers, and software and application programming interfaces.

The author used content analysis to study forty-five research data repositories indexed in Re3data. The collected data was exported to a Microsoft Excel file for appropriate statistical analysis.

Of 45 repositories, 30 were identified as open, 12 as restricted, and 03 as closed. The majority of Indian repositories—28 in total — cover the field of life sciences. Content formats focus on statistical data in 31 repositories, accounting for 68.9%. The study also revealed that 80% of Indian research data repositories do not implement metadata standards (Bhardwaj, 2019).

2017 Maxi Kindling and others reviewed and described research data repositories indexed in Re3data in 2015. Their study's objective was to provide statistical insights regarding metadata standards, content types, subject areas covered by these repositories, and access conditions and services offered.

The authors adopted a statistical analytical approach for the descriptive data of 1,381 repositories across all research disciplines, accessed by searching the Re3data database.

The study revealed variations in the services provided by repositories, the languages they accommodate, the available content formats, and the standards they adhere to. Thus, data repositories are heterogeneous in terms of their content nature. The researchers also found that standards about information and data management infrastructure need to be implemented, as evidenced by the limited number of repositories applying author and entity identification systems and inadequate compatibility and utilization of application programming interfaces (Kindling & Others, 2017).

Upon reviewing the previous literature, we found that the studies exhibited commonalities in examining repository content concerning language, structure, file formats, extensions, and subject specialization. Nevertheless, only one study addressed the repository's technical and operational environment.

The current study consolidates various indicators to assess and analyze Araanalyzeitories thoroughly. Therefore, It not only provides theoretical insights within the realm of library and information science regarding research data repositories but also attempts to provide practical information for researchers, research institutions, and information institutions to ensure best practices in managing research data in digital repositories. This is particularly important as managing research data repositories is an emerging and nascent topic that has not received adequate attention from Arab research institutions, with many Arab countries still lacking a research data repository. Consequently,

this study endeavours to illuminate the current state of research data management in Arab repositories and identify strengths to enhance and weaknesses to address.

3- MITHODOLOGY AND SCOPE OF THE STUDY:

In this section, we will elucidate the methodology and procedures of the study to achieve the desired objectives.

We utilized the descriptive method, which was deemed the most suitable approach for examining the characteristics of the research data repositories under investigation. This entailed collecting data from Re3data concerning repositories created by institutions in the Arab countries and indexed as of December 1, 2024. Subsequently, we examined the collected data and tabulated it to facilitate discussion of the results according to five dimensions that align with the study objectives and cover the most important aspects of research data management in digital repositories. These dimensions are as follows:

• Repository content: Through indicators of the subject specialization of research data, its languages, and forms.

• Access to research data collections: Through indicators of research data access patterns and licenses.

• Technical infrastructure of the repository: Identifying the Software, metadata standards, persistent identifiers, and application programming interface (API) that support repositories' operational environment and ensure research data management within them.

• Research data quality management: Discovering the extent to which the repository applies quality management standards improves the publication.

After directly browsing the Re3data registry website and using the filtering feature to search for repositories by country, we discovered that there are 07 Arab research data repositories geographically distributed across (05) countries, as shown in the following table:

Table 1: Arab Research Data Repositories Indexed in Re3data

Repository name	Country	Repository type	URL
ACSS Data verse	Libanon	Diciplinary	https://dataverse.theacss.org/
AFDB Statistical data portal	Tunisia	Diciplinary	https://dataportal.opendataforafrica.org/
AUBScolar works	Libanon	Institutional	https://scholarworks.aub.edu.lb/
Cloud Flame	Saudi arabia	Diciplinary	https://cloudflame.kaust.edu.sa/
Egypt's Information Portal	Egypt	Other	http://www.eip.gov.eg/Default.aspx
NHRSP	Saoudi arabia	Diciplinary	https://nhrsp.shc.gov.sa/
Wadi AbuDom investigations	Sudan	Diciplinary	http://en.wadi-abu-dom.de/webgis/

Source: Authors' preparation based on Re3data registry

Table 1 shows that Lebanon has two research data repositories:

1. ACSS Data verse: A specialized repository in social science research data produced in the Arab world. It is one of the projects of the Arab Council for Social Sciences, helping researchers and research institutions to preserve and facilitate access to research datasets in social sciences (ACSS, 2024).

2. AUBScholar Works is an institutional repository that provides digital services to help collect, preserve, and distribute digital materials, especially research data. The repository also includes scientific articles in the disciplines it covers to increase the dynamics of scientific communication (AUB-SW, 2024).

The Tunisian Republic has a specialized research data repository called AFDB Statistical Data Portal. It supports the development of statistics in Africa as a basis for data management, considering it a tool that helps development in the continent and achieves community participation as the data is transparent and reliable. The African Development Bank is the institution that oversees the funding of economic research and related research data, making it the institution that supports the management of economic statistical data. This repository is considered a shared digital space that helps economic researchers discover, examine, and download research datasets (AFDB, 2024).

The Kingdom of Saudi Arabia has two specialized digital repositories:

1. Cloud Flame repository: Based on cloud infrastructure, it manages, processes, and digitally preserves research to design an

open, collaborative academic environment between researchers (CPC, 2024).

2. National health research and studies portal: An initiative to build a digital infrastructure for national studies and research in the Kingdom of Saudi Arabia and facing future health challenges. It also aims to make health studies more sustainable and effective by promoting a culture of sharing research data within scientific research ethics and standards to generate new scientific knowledge (NHRSP, 2024).

Re3data lists a research data repository in the Arab Republic of Egypt with its descriptive data. However, after examining the repository's website, it was found that it is no longer available due to the update of the repository from the Egypt information portal to the Egypt data portal, available at https://opengov.unes7cwa.org/node/1211. However, the registry still needs to update the changes that occurred to the repository. Therefore, it was excluded from the study due to the need for more recent and correct information about it.

Sudan oversees a research data repository specializing in geographical and spatial systems as part of a collaborative research project to study the "Masarwadi Abu Dom" area to develop scientific research about the region (W.A.D.I, 2020).

4. Results and discussion:

In this section, we will address the applied aspect of the study, which revolves around analyzing Arab research data repositories and evaluating their management according to the indicators specified in the methodological procedures of the study as follows:

4.1. Repository content:

Re3data provides descriptive information about the content of repositories (research data), focusing mainly on subject specialization, language, and file format.

The registry classifies subjects into four main sections, each section divided into sub-topics hierarchically. These subject sections are Humanities and Social Sciences, Life Sciences, Natural Sciences, and Engineering.

The registry also distinguishes between ten content formats: scientific and statistical data formats, Standard office documents, Images, Raw data, Plain text, Structured text, Structured graphics, Audiovisual data, Source code, and others.

Regarding language, it distinguishes between 62 official and local languages worldwide. Concerning the content features of Arab research data repositories, we present them in the following table:

Table 2: Content features of Arab research data repositories listed in Re3data

Repository name	Countent Subject	Countent language	Countent type
ACSS Data verse	Humanities and social sciences	English	Raw data Sandard office documents Structured text
AFDB	Humanities and	English	Data base

Statistical data portal	social sciences: Social sciences Economics Natural sciences: Geography		plain text scientific and Statistical data formats Standard office documents Structured graphics
AUBScolar works	Humanities and social sciences Life sciences Natural sciences Engineering sciences	English	Audio visial Images Scientific and statistical data formats Standard office data formats
Cloud Flame	Natural sciences: Chemistry - Geochemistry Mineralogy- Cristallography	English	Raw data Csientific and Statistical data formats Standard office documents Structured graphics
National health research and studies portal	Life sciences: Medecine	English	Standard office documents
W-A-D-I Wadi AbuDom	Humanities and social sciences:	German	Data base Images- Plain text-

EISSN: 2602-781X——————————ISSN: 2253-0347

investigations	scientific and statistical data formats Structured text

Source: Authors' preparation based on Re3data registry

Table 02 illustrates the characteristics of research data in terms of content language, type, and subject. It is observed that the majority of Arab repositories listed in Re3data make their data available in English, with only the Sudanese repository making it available in German. The table also shows that the most shared content formats are:

Scientific and statistical data formats, Standard office documents Raw data, Structured graphics, Structured text, Images, and Plain text are distributed across only two out of 06 repositories, and the Lebanese repository AUBScholar Works is the only one that provides data in Audiovisual format.

Regarding subject coverage, Arab repositories mostly share the field of Humanities and Social Sciences, except for the Cloud Flame repository, which focuses only on Natural Sciences. The latter, along with the AFDB Statistical data portal and AUBScholar Works repositories, also shares research data in Life Sciences with the Saudi NHRSP repository. The field of Economics is specific to the AFDB Statistical data portal repository.

From the above, we conclude that Arab research data repositories do not aim to valorize the Arabic language but seek global visibility by making data available in English. Arab

research data repositories also lack diversity in content in terms of form and nature of topics addressed, which makes the research data ineffective in building innovative research projects, especially in applied fields that are rapidly evolving. Thus, Arab repositories are similar and homogeneous in content form, languages, and topics, contrary to the global landscape of data repositories indexed in Re3data, according to the study conducted by kindling,M and others.

In parallel, Arab research data repositories seek to valorize specialized content in Humanities and Social Sciences due to the cultural diversity in the Arab region rooted in history. This has led to an interest in this type of research and the archiving of its data in repositories that ensure its preservation and availability in studying Arab societal behaviours.

4.2. Access levels in Arab research data repositories and licenses:

Re3data categorizes research data repositories according to indicators related to the level of access to research data, the adopted licenses consistent with the openness practices described in the repository policy, and the permissions embodied by authors in those licenses.

The level may be open without any legal or financial restrictions preventing researchers from accessing, deriving from, or republishing research data as they see fit. The level may be embargoed, meaning data is made available but with certain restrictions for a specific period before becoming open. Legal or technical controls may also restrict the level through the use of licenses that require the researcher to be affiliated with the

research institution overseeing the repository or the need to register in the repository or pay a fee. The level may be closed entirely, meaning the content is unavailable, and only descriptive references are provided.

4.2.1.Research Data access:

The registry distinguishes between four access levels to research data: Open, Embargoed, Restricted, and Closed. This table illustrates access to research data available in Arab repositories. we illustrate it in this table.

Table 03: Levels of access to research data available through Arab repositories listed in Re3data.

Repository name	Open	Embargoad	Restricted	Closed
ACSS Data verse	✓	✗	✗	✗
AFDB Statistical data portal	✓	✗	Registration	✗
AUBScolar works	✓	✗	Institutional member ship	✗
Cloud Flame	✓	✗	Registration	✗
National health research and studies portal	✗	✗	Unknown	✗
W-A-D-I	✓	✗	✗	✗

Wadi
AbuDom
investigations

Source: Prepared by the authors based on Re3data registry

Table 03 demonstrates that the Lebanese repository ACSS Dataverse provides open access to its data while other access modes are unavailable. This indicates that this repository supports the concept of research data sharing and enhances its visibility among researchers. The Tunisian repository, AFDB Statistical Data Portal, offers its data openly or with restricted access based on registration requirements. This approach allows for tracking researchers' behaviours and needs in utilizing research data and monitoring the usage of research data provided by the repository. This repository does not support the closed or prohibited access mode.

The Lebanese repository AUBScholar Works provides its data openly or with restrictions requiring researchers to be affiliated with the research institution. Upon careful examination of the repository, we observed that most datasets, especially recent ones containing detailed analyses and precise information, are restricted. Conversely, outdated research data is openly available, indicating a limited level of openness for this repository.

The Saudi repository Cloud Flame offers research data either openly or with restricted access. Registration is required to identify researchers using the data.

The National Health Research and Studies Portal provides restricted access to its research data. However, the restriction

mechanism must be specified, posing a barrier to researchers' and users' access to the repository's datasets. Open, prohibited, or closed data access modes are not indicated.

The Sudanese repository W.A.D.A Abudom Investigations offers its research datasets exclusively through open access, with other access modes not supported.

Consequently, Arab research data repositories predominantly support open access or restricted access mechanisms, similar to Indian repositories, as evidenced by Bhardwaj,R.K's study results. Restrictions are implemented through registration or institutional affiliation requirements, as observed in the Lebanese AUBScholar Works and Saudi health research repositories. These two repositories are the most restrictive regarding access among Arab repositories and, thus, the least visible.

The Arab repositories' support for open access to research data can be attributed to the overseeing institutions' inclination to integrate into the open access movement and implement open science practices in the Arab world. Based on our observations of most studied Arab repositories, the nature of subject coverage, which tends towards humanities and social studies, is not highly current.

4.2.2. Research Data and Database Licenses:

Re3data categorizes research data repositories into twelve classes based on the licenses governing the research data and the repository database. This section highlights the nature of the licenses employed in Arab repositories, as illustrated in the following table.

Table 04: Research Data Licenses Available in Arab Repositories Listed in Re3data

Repository name	Data licenses	Data base licences
ACSS Data verse	CC Copyrights	✗
AFDB Statistical data portal	Copyrights	✗
AUBScolar works	CC Copyrights	✗
Cloud Flame	✗	✗
National health research and studies portal	✗	✗
W-A-D-I Wadi AbuDom investigations	Copyrights	✗

Source: Authors' Preparation based on Re3data registry

Table 04 refers to the nature of the licenses employed in providing research data and the licensing framework of the repository databases. It is noted that the Lebanese repositories, ACSS DataVerse and AUB Scholar Works, provide their data either under a Creative Commons (CC) license, which aligns with the open access environment for research data, thus permitting retention, derivation, modification, or the possibility of republishing and distributing the data under the same original research data license. This also includes the potential for commercial exploitation by the permissions established by the

authors, as represented by various types of Creative Commons licenses. Alternatively, datasets may be made available under Copyright licenses, whereby the data becomes restricted, allowing researchers only to review the data without the right to manipulate it, as the researchers or publishers retain all rights. This is the licensing model adopted by the AFDB Statistical Data Portal and the W.A.D.I. Wadi Abudom repository.

The Saudi repositories, Cloud Flam and the National Health Research Repository do not provide information on the licenses utilized, reflecting a lack of awareness regarding the importance of licenses for disseminating research data content within the Kingdom. Furthermore, the table indicates that the databases of the Arab repositories do not rely on any specific licenses, which results in the interoperability and migration of data or exchange among digital systems being unregulated by legal or ethical standards.

Consequently, we conclude that the majority of the studied Arab repositories rely on licenses that safeguard authors' rights and protect the intellectual property of research data. These licenses also regulate research data use and enforce legal and ethical obligations in its handling while encouraging its sharing.

4.3. Technical Infrastructure of Repositories:

The technical infrastructure of research data repositories consists of three main components: Software, Metadata, and APA. Collectively, these components form an integrated operational environment that facilitates the secure and reliable management of research data based on standardized criteria that

ease its flow and retrieval. In this section, we will explore the supportive infrastructure of the Arab repositories listed in Re3data according to the mentioned indicators.

4.3.1.Repository Software:

The Software utilized in research data management systems serves as a fundamental indicator for assessing the effectiveness and efficiency of a repository in managing, organizing, and processing research data in a manner that allows for the provision of high-quality services to users, particularly in the search process. Below is an overview of the Software employed by the Arab repositories listed in Re3data.

Table 05: Software Used in the Management of Arab Repositories Listed in Re3data

Repository name	Software
ACSS Data verse	Data verse
AFDB Statistical data portal	Unknown
AUBScolar works	Dspace
Cloud Flame	Unknown
National health research and studies portal	Unknown
WadiAbuDom investigations	Unknown

Source: Authors' Preparation based on Re3data registry

As indicated in Table 05, the Lebanese repository ACSS DataVerse was constructed and designed using the Dataverse software, a proprietary system developed by Microsoft specifically for managing datasets. It is characterized by its

ability to process, integrate, and preserve data for reuse, effectively responding to all aspects of the data lifecycle through advanced cloud services (dataverse, 2024).

In contrast, the AUB Scholar Works repository utilizes DSpace software, an open-source system employed by numerous global institutions for research data management. DSpace is recognized for its capability to employ metadata and the Handle tool for defining digital entities with persistent hyperlinks, as well as for supporting long-term preservation and efficient retrieval. This is mainly due to its reliance on the OAI-PMH protocol, which allows it to harvest metadata across open archives (Zang & Chen, 2015).

However, the other Arab research data repositories have not specified the Software utilized for their management, as evidenced by the 1,182 repositories listed in the registry.

From the above, it is evident that Arab repositories urgently need to adopt well-known and widely used Software. Such adoption would help mitigate challenges related to software management, particularly with open-source solutions, given the high development potential provided by a broad community of developers, such as those for DSpace. Additionally, it would enable the management of their research data in standardized ways, enhancing global visibility and utility.

4.3.2 Repository Metadata Standards:

Metadata is a crucial indicator in evaluating research data repositories. It provides guidelines that describe and define the structure of dataset elements (Jian, 2013). This includes aspects related to intellectual property, context, formats, and management according to standardized plans.

The repositories indexed in the registry utilize a variety of metadata standards, with the most notable being Dublin Core (596 repositories), Cite Metadata Schema (437 repositories), Data Documentation Initiative (279 repositories), Repository-Developed Metadata Schemas (220 repositories), ISO 19115 (208 repositories), Open Archives Initiative Object Reuse and Exchange (145 repositories), and For Digital Geospatial Metadata (102 repositories). The following illustrates the application of metadata standards in the Arab repositories listed in Re3data..

Table 06: Metadata Standards in Arab Repositories Listed in Re3data

Repository name	Repository meta data standard
ACSS Data verse	DC-DDI-DM
AFDB Statistical data portal	✗
AUBScolar works	DC
Cloud Flame	✗
National health research and studies portal	✗
WadiAbuDom investigations	✗

Source: Authors' Preparation based on Re3data registry

Table 06 illustrates the nature of metadata standards employed in describing research datasets. The ACSS DataVerse repository relies on clear standards capable of accommodating various

details related to datasets, thereby ensuring their suitability for bibliographic citation. These standards include DM, DDI, and DC. Conversely, the other Lebanese repository, AUB Scholar Works, utilizes the Dublin Core (DC) standard for its metadata description.

The DDI (Data Documentation Initiative) standard is notable for its ability to provide a detailed description of statistical and social data concerning their collection methods and analytical methodologies (Verdigan, 2013). As for the DM (Dublin Core Metadata), it does not significantly differ from the DC standard, which offers general information about research data, such as package title, author, publication date, abstract, keywords, hyperlink, and usage rights while lacking further details related to the nature of the research sample, the tools used for data collection, and the analysis methods employed. The table further indicates that the remaining repositories must adhere to metadata standards.

From the preceding observations, it can be inferred that Arab repositories do not assign substantial importance to metadata standards, leading to diminished efficiency in retrieving research datasets during search processes due to a lack of descriptors and reliance on free indexing, Reliance on metadata standards ensures the standardization of dataset descriptions and the identification of all descriptive and thematic attributes and characteristics on one hand. On the other hand, metadata standardization helps achieve compatibility and integration in data exchange between repositories or even enables sharing with other digital platforms through protocols and metadata harvesting that allow for interoperability. Simultaneously, it

improves retrieval efficiency through search engines that heavily depend on standardized descriptive data through precise content indexing. Additionally, there is an increased risk of violating intellectual property rights due to random and non-standardized descriptions. This finding aligns with the conclusions drawn by Marcello Mundim Rodrigues and his colleagues in their study of research data repositories in South America, as well as Bhardwaj Raj Kumar in his examination of Indian research repositories, where they collectively noted the necessity for these repositories to adopt metadata standards for file description, similar to the situation in Arab repositories.

4.3.3 Application Programming Interfaces (API) Tool:

Application Programming Interfaces (APIs) allow Software and applications to communicate with one another using standardized protocols and languages by providing instructions to databases for execution (Kichinin, 2022, p. 2). Thus, APIs act as intermediaries facilitating the exchange of research datasets between repositories and digital spaces without data migration from one system to another. The following section examines the current use of APIs among Arab repositories.

Table 07: Use of APIs in Arab Repositories Listed in Re3data

Repository name	API
ACSS Data verse	✗
AFDB Statistical data portal	✗
AUBScolar works	✗
Cloud Flame	✗
National health research and studies portal	✗

W-A-D-I Wadi AbuDom investigations		✗

Source: Authors' Preparation based on Re3data registry

Table 07 indicates that the studied Arab research data repositories do not support Application Programming Interfaces (APIs). This limitation significantly reduces the opportunities for interoperability of datasets, thereby weakening the prospects for collaboration among research institutions in data exchange due to the absence of protocols and technical standards that would facilitate such interactions. Despite all repository databases being open, the provision of API protocols in Arab research data repositories is not afforded the necessary importance. However, this issue is not confined solely to Arab repositories; it is also prevalent among research data repositories globally, as noted by Kindling, Maxi in findings up to 2015.

4.4. Research Data Management Quality:

Re3data emphasizes the importance of quality in research data management by providing details related to quality management standards . In this section, we aim to explore the application of quality management standards in Arab research data repositories.

Research data management according to quality standards is a significant focus for Re3data. Adherence to this standard reflects the transparency and reliability of research data. The registry differentiates between research data repositories based on three levels: 1,727 repositories apply quality standards, 1,439 repositories have not specified their adherence, and 58

repositories do not comply with quality management standards. The following presents the application of quality management standards in Arab repositories:

Table 08: Quality Management Standards in Arab Repositories Listed in Re3data

Repository name	Quality management
ACSS Data verse	✗
AFDB Statistical data portal	✓
AUBScolar works	Unknown
Cloud Flame	Unknown
National health research and studies portal	Unknown
W-A-D-I Wadi AbuDom investigations	Unknown

Source: Author Preparation based on Re3data registry

Table 08 indicates that the Tunisian repository, AFDB Statistical Data Portal, is the only repository that supports a quality management standard. In contrast, the repositories AUBScolar Works, Cloud Flame, W.A.D.I. Wadi AbuDom Investigations and the National Health Research and Studies Portal have yet to specify a quality management indicator. At the same time, the ACSS Data Verse repository does not comply with the quality standard. Consequently, more attention must be paid to quality management in these repositories.

The limited focus on quality management in these repositories can be attributed to a deficiency in the skills and competencies of specialists in the field of research data management, as well as a lack of awareness regarding the culture and significance of

quality and the necessity of adherence to it in digital repositories. Additionally, the non-standard operational environment of most Arab repositories exacerbates this issue. Therefore, there is a pressing need to enhance efforts to align with most repositories listed in the registry that implement a quality management standard.

Re3data indicates that in the United States alone, 718 repositories out of 1,193 apply quality standards, making it the country with the highest support for data availability according to this metric. Germany follows with 310 repositories out of 537. This situation makes research data in Arab repositories collectively weak in terms of quality compared to leading countries such as the United States and Germany.

5. CONCLUSION:

In conclusion, this study generally finds that the management of research data available through the Arab repositories listed in Re3data requires greater attention in order to provide added value to Arab research and studies across various scientific disciplines. This conclusion is based on the analysis of the examined repositories.

Arab repositories, particularly the Lebanese repository ACSS Data Verse and the Sudanese repository, do not prioritize diversifying their research data collections in terms of language and specialized topics.

The lack of content diversification reduces the opportunities to benefit from research data repositories in feeding intelligent research systems based on intersection with big data concepts

and intelligent analysis techniques, thus limiting the opportunities for scientific innovation, technical development, and investment in emerging technologies. These challenges place Arab repositories in front of future stakes that are difficult to face.

Arab repositories significantly support open access policies for research data and repository databases. In contrast, data from these repositories is often accessed under restricted conditions, either through registration or requiring researchers to be affiliated with supporting institutions.

Regarding access licenses, Arab repositories primarily rely on Creative Commons and Copyright licenses; however, they do not implement any licenses for accessing the repository database.

The two Lebanese repositories are the most standardized among Arab repositories, as their technical infrastructure relies on clear repository management systems. They also support standardized descriptions of data sets by adopting the Dublin Core metadata standard and using persistent identifiers, particularly DOIs. Conversely, Arab repositories are far from utilizing application programming interfaces (APIs), which reduces interoperability opportunities.

Due to the general weakness of technological infrastructure in managing research data in Arab repositories, it is difficult to integrate artificial intelligence-based technologies due to the lack of standardization that would help machine learning or deep learning algorithms mine and retrieve data according to desired

needs, as well as reducing opportunities for intelligent data analysis.

Arab repositories receive support from various scientific and technological institutions, which enables them to provide technological and financial resources and develop policies related to research data management.

Despite the availability of conducive logistical conditions for repository management, Arab repositories generally do not adhere to quality standards in managing research data sets, implement tools that promote their content, or enhance improved publication practices.

Arab research data repositories need to exert more effort according to a collaborative vision from scientific and research institutions in order to reflect on the economic and social level of Arab societies, especially in light of the trend towards economies based on scientific knowledge that rely on research data as a fundamental and vital resource for them.

The primary aim of this study is to leverage the strengths and weaknesses of Arab repositories to build and develop research data repositories in the Arab world, particularly in Algeria, which currently lacks such a repository. Planning and implementing these projects requires studying models similar to academic environments and socio-economic climates, where strategic goals are accurately defined, and optimal investment in necessary functional requirements is ensured.

6. STUDY PROPOSAL:

-In light of the results obtained from our study of the Arab repositories indexed in Re3data, we propose several measures that could enhance the efficiency and effectiveness of research data management:

-Promote awareness of the importance of research data within the research community and encourage its deposit, publication, and sharing through investments in digital repositories.

-Ensure the training and development of competencies capable of managing research data repositories, including information specialists and IT experts, through continuous training programs that integrate standardized practices and various technologies supporting the infrastructure of research data repositories.

-Organize scientific events and forums to discuss various policies and requirements necessary for building research data repositories and managing their content, as well as to review leading international models in an evaluative, analytical context to leverage strengths and address weaknesses based on their experiences.

-Work to attract a more significant number of researchers and research teams to enrich repositories with research data representing the primary and raw outputs of research while also ensuring the registration of repositories in global directories and registries, particularly Re3data, which specializes in cataloguing research data repositories to enhance visibility and scientific communication.

-Diversify research data in scientific topics, languages, and content formats to meet the needs of researchers and academic

institutions, particularly in specialized fields requiring detailed collections on various phenomena.

-Encourage researchers to assign their research data with open licenses that facilitate access, reuse, and interoperability within the framework of FAIR principles.

-Adopt strategic planning when building data repositories to define objectives and inventory resources, mainly financial and technical, to provide a standardized operational environment. This includes repository management systems (Software) to manage data and provide related services according to quality management standards. It also supports using metadata standards to describe all aspects of data sets, such as DDI and DM, and employing digital identifiers related to digital entities or authors to facilitate data retrieval and citation.

- Coordination between Arab governments and finding mechanisms that allow the establishment of multi-sectoral institutions with a shared Arab vision for supervision and cooperation in the field of research data management and building data repositories.

-Encourage using application programming interfaces (APIs) and the OAI-PMH protocol, considering repositories as open archives to enhance interoperability and build partnerships with research institutions in research data accessibility.

-Support initiatives for open access to research data as part of efforts to realize the practices of the open science movement, given that open research data is one of its fundamental pillars.

-Continuously update the descriptive and informational content of Arab repositories in Re3data to provide a comprehensive and accurate descriptive framework for the repositories such as research data plan schema.

- Improving research data management by working to find technological and legal mechanisms to integrate emerging technologies, especially big data analysis and artificial intelligence.

Bibliography list :

1. ACSS. (2024). *ACSS data verse*. Retrieved 12 01, 2024, from Arab council of the social scieces: https://dataverse.theacss.org/

2. AFDB. (2024). *AFDB data finder repository*. Retrieved 12 01, 2024, from Africa information highway: https://dataportal.opendataforafrica.org/

3. AUB-SW. (2024). Retrieved 12 01, 2024, from American university of Beirut - Scolar works: https://scholarworks.aub.edu.lb/

4. Azeroual, o., & others. (2022). Puting FAIR Principles in the context of research information: FARness for CRIS and Cris for FAIRness. *Knowledge Management and information systems.* Valletta - Malta. Retrieved 11 20, 2024, from https://hal.science/hal-03836525/file/KMIS2022_Azeroual-Schopfel-Polonen-Nikiforova_FAIR-principles.pdf

5. Bhardwaj, R. (2019). Content analysis of indian research data repositories: Prospects and possibilities. *DESIDOC journal of library and information technology, v.29*(n.6). Retrieved 11 23, 2024, from https://publications.drdo.gov.in/ojs/index.php/djlit/article/view/15137/7174

6. CPC. (2024). *Combustion cyberinfrastructure*. Retrieved 12 01, 2024, from Combustion pyrolysis chemistry group: https://cpc.kaust.edu.sa/research/detail/combustion-cyberinfrastructure

7. dataverse, M. (2024). *Microsoft dataverse*. Retrieved 12 09, 2024, from Qui sommes nous?: https://www.prodwaregroup.com/fr-fr/solutions/microsoft-power-platform/microsoft-dataverse/#

8. Elsayed, A., & Saleh, e. (2018, 12). Research data manegement and sharing among researchers in arab universities: an exploratory study. *IFLA journal, V.44*(N.4). Retrieved 11 20, 2024, from https://repository.ifla.org/handle/123456789/346

9. Jian, Q. (2013). Infrastructure,standards, and policies for research data management. *Sharing of scientific and technical resources in the Era of big data.* Retrieved 12 10, 2024, from https://oceanexpert.org/downloadFile/35260

10. Kichinin, L. (2022). Validation des modeles avec PyRevit et Power BI. Retrieved 12 10, 2024, from https://static.au-uw2prd.autodesk.com/Class_Handout_CS500222_ClassHandoutCS500222.pdf

11. Kindling, M., & Others. (2017, 03). The landscape of research data repositories in 2015: A Re3data analysis. *D-lib magazine, v.23*(n. 3/4). Retrieved 11 23, 2024, from https://mirror.dlib.org/dlib/march17/kindling/03kindling.html

12. Mohamed, A., Abdulbasset, S., & Taha, A. (2021, 10). Research Data and Its Repositories in Universities: A Theoretical Study. *The Egyptian Journal of Information Siences, V.8*(N.2). Retrieved 11 20, 2024, from https://jesi.journals.ekb.eg/article_221530.html

13. Mundim Rodrigues, M., De Azevedo Lourenço, C., & Ataide Dias, G. (2022). The nature of scientific datasets in south american repositories: A survey of formats and extensions. *Encontros bibli Revue, v.27*. doi:https://doi.org/10.5007/1518-2924.2022.e85148

14. Nabti, M., Achouri, N., & Atrouz, B. (2017). Management and Preservation of Digital Collections in Open Research Data Repositories: An Analytical Study of the Status of African Repositories in Re3data. Retrieved 11 20, 2024, from https://search.mandumah.com/Record/870150

15. NHRSP. (2024). Retrieved 12 01, 2024, from National health research studies portal: https://nhrsp.shc.gov.sa/

16. Nwabugwu, M., & Godwin, L. (2020, 05 21). Research data management (RDM) services in libraries: lessons for academic libraries in Nigeria. *Library phelosophy and practice*. Retrieved 11 20, 2024, from https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=7915&context=libphilprac

17. Re3data. (2024). *About Re3data*. Retrieved 11 20, 2024, from Re3data.org: https://www.re3data.org/about

18. Verdigan, M. (2013). Stratigic priorities of the data documentation initiative (DDI) alliance. Retrieved 12 10, 2024, from https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.40/2013/WP6.pdf

19. W.A.D.I. (2020). *WEBGis*. Retrieved 12 01, 2024, from Wadi Abudom investigations: http://en.wadi-abu-dom.de/webgis/

20. Zang, Y., & Chen, H. (2015). Data management and curation practices: the case of using dspace and implication. *The 78th annual meeting of the association for information science and technology*. Retrieved 12 09, 2024, from https://core.ac.uk/download/pdf/229305526.pdf