Setting Up Open Access Repositories

Challenges and Lessons from Palestine

Rawia Awadallah
IUG
Palestine rradi@iugaza.edu.ps
0000-0001-6317-9747

Iyad ALAgha
IUG
Palestine ialagha@iugaza.edu.ps
0000-0002-7516-4902

Tomasz Miksa, Andreas Rauber
TU Wien Austria
miksa@ifs.tuwien.ac.at
0000-0002-4929-7875

Joy Davidson
UoG UK
Joy.Davidson@glasgow.ac.uk
0000-0003-3484-7675

Adnan Yahya
BZU
*Palestine yahya@@*birzeit.edu
0000-0002-8769-1961

Anna Maria Tammaro, Stefano Caselli
University of Parma Italy
annamaria.tammaro@unipr.it
0000-0002-9205-2435

Janet Anderson, David Anderson
UoB UK
J.Delve@brighton.ac.uk
0000-0003-2673-4830

Nael Salman
PTUK
Palestine N.Salman@ptuk.edu.ps
0000-0002-5190-3889

Yousef AbuZir
QOU
Palestine yabuzir@qou.edu
0000-0002-1220-1411

Abstract – Research outputs produced by developing countries lack visibility. Common reasons are high costs of publishing, reluctance to share, as well as lack of journals focusing on research challenges specific to developing countries. The ongoing change towards open access and the rising popularity of institutional repositories allows for bridging the gap to developed countries. However, this still requires cultural, organizational and technical changes. In this paper, we describe a holistic approach for deploying open access repositories and building research data management services and increasing data curation skills. We describe how we identified users' needs and necessary supporting systems and services. We also explain the rationale and challenges faced when implementing popular repository systems and share experiences in developing institutional data management policies. Finally, we provide common goals for a national roadmap. All these actions are the first step towards the preservation of both research outputs and cultural assets. The paper is based on our experiences from ROMOR project that helped in establishing policies, systems, and organizational workflows at four Palestinian universities.

Keywords – Open Access, Institutional Repositories, RDM, Policy, Developing Countries.

Conference topics – 1. Collaboration: a Necessity, an Opportunity or a Luxury? 4. Building Capacity, Capability and Community

I. Introduction

In the Digital Age, more and more people are accessing and contributing digital content which has the potential to improve people's lives by making information better available, increasing opportunities for political and economic engagement, and making government more transparent

16th International Conference on Digital Preservation
iPRES 2019, Amsterdam, The Netherlands.

and responsive. However, global participation in the Digital Age still remains uneven for many and thus can lead to a growing knowledge divide among research communities or whole countries.

Research outputs of Palestine researchers are often essential for the development of appropriate programs to solve local problems, e.g. estimating underground water reserves in Gaza strip. Their research is often essential for policymakers and development practitioners to identify the exact needs of the local community and to optimize their policies and investments according to these needs.

However, developing countries such as Palestine are at low rank in terms of research output and its visibility (same applies for digital preservation of cultural heritage,

e.g. digitization of rare manuscripts). This not only can be explained by the low budget allocated to research and the fluctuating political climate but also by lack of proper practices, policies, infrastructure and culture of sharing research outputs.

To close the gap in sharing and access to scientific data and publications, developing countries need to participate in developing knowledge exchange anisms and systems that benefit them by making their outputs Findable, Accessible, Interoperable and Reusable (FAIR)[1]; universities should encourage publishing through cooperative, peer-reviewed open access platforms; and the governments and development agencies should invest in developing more effective knowledge sharing systems and digital repositories.

mech-Table 1: OAIRs Developed within ROMOR

IUGSpace:	iugspace.iugaza.edu.ps
FADA:	fada.birzeit.edu
OSOL:	dspace.qou.edu
PTUK Repository:	scholar.ptuk.edu.ps

This paper aims to share the experience of ROMOR project in building and improving Open Access Institutional Repositories (OAIRs) of four universities in Palestine (see Table 1. In particular, the paper describes the development of OAIRs, and their planning, implementation/improvement, and evaluation. These include the repository, scope and coverage, metadata standards and interoperability, policies, and business models. These repositories include not only staff publications and theses, but also hundreds of scanned rare books that are freely available to the public[1], and will soon include environmental research datasets including geological, hydrological, and meteorological datasets. The paper also identifies next steps needed for introducing a national roadmap to include further universities and provide common services.

Despite the focus on Palestine, this paper can be seen as a blueprint for other developing countries that share similar challenges and consider establishing OAIRs.

The paper is organized as follows. Section 2 describes related work on open access and research visibility in Palestine. Section 3 presents requirements identification and training material development that laid the foundation for OAIRs development. Section 4 deals with business modeling and preparation of OAIR deployment plans. Section 5 presents the process of installing OAIRs, challenges, and lessons learned. Section 6 focuses on data management policy development. Section 7 presents goals for a national roadmap. Conclusion and future work are provided in Section 8.

II. Related work

A. *The Visibility of Research in Palestine*

Developing countries such as Middle Eastern Arab countries are at low rank in terms of research output. However, these countries have witnessed a considerable growth in research size and impact in the last decade due to the growing investments in research ([2]; [3]).

Palestine is a developing country located at the south- eastern side of the Mediterranean coast. It is only since the 1970s that universities came into existence in Palestine. According to the latest publications and statistics of the Palestinian Ministry of Higher Education (MOHE), there are 16 universities. Most of the universities started as 2-year College institutions and then were developed to provide university-level

[1] fada.birzeit.edu/handle/20.500.11889/2836

16th International Conference on Digital Preservation
iPRES 2019, Amsterdam, The Netherlands.

education. Currently, many of these universities are involved in research and offer graduate programs in various scientific disciplines including engineering, health, basic science, economics, humanities, and other fields [4].

Despite the affordance of research activities in Palestine, access to the results of implemented research has been severely limited due to publishing and access inabilities ([5]; [6]; [7]). Because of the common trend among scholars to publish their research findings in highly reputed journals, most of the Palestinian scholarship is either unpublished or delayed. In addition, the traditional system of scholarly publishing and the high costs of publications have contributed to the limited growth of the Palestinian scholarship ([8]). Apart from the research, the visibility of research outcomes is also kept to the minimum, due to financial restrictions limiting the publication and distribution of national research outcomes. In addition, information important for the resolution of problems specific to developing countries is not often published in journals from the developed world.

As a result, the research outcomes, which most of is imperative to address local and regional developmental issues, die at the institutional level as they remain invisible to those who may need them. Although some results may eventually get published in local journals, the outcomes may not be widely disseminated due to the poor distribution and recognition of these journals. This disappointment comes despite the so much commitment of efforts and resources that may be devoted to undertaking research in Palestine.

B. Open Access in Palestine

Little attention has been paid to the development of Open Access Institutional Repositories in the Arab academic institutions. OAIRs can be defined as digital archives of intellectual outputs created by faculty, research staff and students of an institution, and accessible to end-users both within and outside the institution, with few if any barriers to access ([9]; [10]). One of the main benefits of OAIRs is to maximize the availability, accessibility, discoverability, and functionality of scholarly outputs at no cost to the user ([11]; [12]). In addition, OAIRs often provide technologies and methods to support the preservation of research outputs and ensure that they remain accessible and reusable over time ([13]; [14]).

A number of academic institutions in Arab countries have taken the initiative to build OAIRs to foster the dissemination of their research outputs ([15]; [16]). Few studies have tried to track and assess the progress of institutional repositories in the Arab world ('[17]; [16]; [18]). All these studies agreed that OAIR developments in the Arab region are still at early stages. Carlson [17] assessed a sample of repositories from the Arab world in terms of accessibility and transparency and reported that they were notably lacking in explicit policies, metadata, and preservation. Studies also reported that repositories have not been growing consistently due to the lack of mandating policies or the lack of awareness of OAIR benefits.

There is a positive response toward OA in Palestine. This is indispensable in order to raise the profile of research performed in the Palestinian higher education institution and integrate it into the international knowledge pool. As of January 2019, there are 23 OA journals indexed in the Directory of Open Access Scholarly Resources. The Open Access initiative has further strengthened in Palestine through embarking digital projects funded by international entities. Until recently, only two OAIRs from Palestine were oZcially registered with the Registry of Open Access Repositories[2]) or the Open Directory of Open Access Repositories (OpenDOAR[3]).

C. ROMOR Project

One recent, ongoing and important initiative for OA in Palestine is through a project titled "Research Output Management through Open Access Institutional Repositories in Palestinian Higher Education" ROMOR. ROMOR was kicked off on January 2017 by four Palestinian universities and four EU universities with EU funding from Erasmus+. ROMOR aims to improve the management, visibility, and accessibility of scientific research outputs in Palestinian HEIs by establishing new or enhancing existing OAIRs, improving institutional capacity for

[2] http://roar.eprints.org

[3] http://v2.sherpa.ac.uk/opendoar

16th International Conference on Digital Preservation
iPRES 2019, Amsterdam, The Netherlands.

the management and sharing of research outputs, and developing and/or refining curricula to ensure that emerging researchers are better able to manage their work across the entire research lifecycle.

ROMOR is the first project that seeks to build capacity for developing OAIRs at Palestinian universities. Without the input of knowledge from research that tackles national challenges and priorities, the development stakeholders and policymakers in Palestine and beyond may not be able to act effectively, and development initiatives may suffer from inappropriate programs. Considering the mobility restrictions between the West Bank and Gaza, the OAIRs is providing a unique opportunity for cross-border coordination and exchange of research findings between Palestinian universities, thus leading to more complementary and less duplicated research. In addition, the project is innovative in that it embraces a variety of capacity building activities that include: conducting needs assessments, holding workshops, building OAIRs for both publications and research data, providing vocational and academic teaching, implementing policies and exploring sustainable solutions for digital curation and preservation. The overall activities aim to bridge theory with practice. The Project seeks to build OAIRs not only by using existing solutions but also by exploring innovative solutions that are tailored to the local needs such as the need to handle the Arabic content.

III. Identifying Needs and Developing Training Material

To help scope the OAIR service development and/or refinement, user requirements were gathered at the start of the ROMOR project through two surveys targeted to both researchers and support staff. The first [19] aimed to assess researchers' current practices, and the second explored institutional support staff capacity. The four participating institutions include: Islamic University of Gaza, Al-Quds Open University, Birzeit University, and Palestine Technical University-Khadoori.

We used the DCC[4] research data management service model [21] which helps break down RDM infrastructure into specific activities to assist with planning and implementation. 'Soft' infrastructure aspects, including policies, business planning, and training, underpin the more technical infrastructure requirements in the center which is based around the data lifecycle. It is important to note here that this is an idealized view of an RDM service model. In reality, most institutions will have some services in place as well as many gaps.

The gaps were identified through analysis of the ROMOR academic staff and managerial staff surveys. The findings are mapped to the DCC's RDM Service model and results are compared to the RDM landscape in the UK. Overall, results revealed the lack or complete absence of RDM policies at the institution's level. In addition, there is a lack of business and sustainability plans for data repositories and RDM services. As a result, adopted data archiving procedures were mostly immature and not consistent with known good RDM practices. Most researchers still have to store and manage their research outputs on their own, with limited support provided by the institution. Results also showed a lack of awareness of the potential of OA publishing and OAIRs to promote the visibility of research outputs.

These results were used to make decisions on the training structure and the educational material to be produced. The areas of competencies related to Data Management and Data Engineering, developed by the EDISON Project [22], were matched to the objectives/needs of the ROMOR Project. The EDISON competencies of these two areas have been compared with the educational and training needs of the Palestinian Universities. ROMOR partners have agreed to employ the DCC's RDM service model to structure the training including an emphasis on softer infrastructure aspects including policies, business planning and training; and also on more technical infrastructure requirements which are based around the data lifecycle.

The training included each of the model components, considering aspects of EU partners' good practices along with Palestinian universities specific needs and expectations. Following the DCC's RDM service model, three areas of training and education are highlighted in the ROMOR RDM process in Palestinian universities:

[4] http://www.dcc.ac.uk

1. *Data Governance:* a set of activities including strategic planning, supervision, and enforcement that governs the process and methodologies that are carried out to ensure and improve the quality of RDM, including organizational structure and business planning, policy, and advocacy awareness to data stakeholders.
2. *Data Management:* consists of data selection and handover, data storage in repositories, data catalogues, data sharing, data reuse, and data appreciation.
3. *Data literacy:* embedded in R&D data flow and related to research data life-cycle. R&D management and data utilization are main perspectives in data literacy; practical skills such as data analysis, data description (metadata and methods) and tools are the main focus in data literacy.

IV. Developing Business Models

Once a general picture of user needs was identified, the Palestinian partners began to scope out the OAIR services that would be delivered at each institution. A series of practical exercises were run as part of dedicated training workshops and continued at each Palestinian institution. Each training workshop aimed to build upon the work carried out during the previous session.

In the first workshop participants were introduced to the Business Model Canvas (BMC)[5]. The BMC helps an organization to create a business model by considering among others: what value the service will deliver to customers, how to maintain a relationship with customers, what are the key activities and operational units involved in delivering the service, etc. The key value of the BMC approach is that it allows a group of stakeholders to work collaboratively to develop a high-level view of the business model for a given service. The graphical presentation of the BMC makes it easier to identify any gaps or disconnects that may hinder the value proposition(s) being realized.

Some Palestinian partners were just starting out with scoping RDM services while others are further down the line and are currently rolling out pilot services or are delivering more established RDM services. Through the short exercises, each group developed a BMC for their RDM service. The BMCs were completed collaboratively within working groups involving main stakeholders to allow different viewpoints to be captured. It was agreed on that in the Palestinian context, the structural funding of the institution may be the best option for sustainability. Resource sharing or even shared operation and shifting the data entry work burden to material producers and integrating the effort into the management process may work well to reduce the costs to sustainable levels. What makes such an approach reasonable is that the archiving volumes are quite low and are most likely to remain so for several more years. Potential funding bodies (e.g. MoHE) is supportive of structuring the costs of archiving into the research costs that can be supported. We also believe that the pioneering institutions, in terms of OAIR development and adoption, can offer services to latecomers that can help offset some of the costs. This includes: (1) Consultancy services both technical, policy drafting and needs assessment. Such services can be offered by Library, IT and maybe legal staff. (2) Development services by offering a running system that suits the needs of the client institution against a fee, possibly with maintenance and training provisions. (3) Hosting services for the OAIR of institutions who find it too costly to run their independent OAIR facilities.

The Palestinian partners started their BMC during the first workshop and then further refined it with input from relevant stakeholders back at their own institutions.Below we outline high-level recommendations for developing a BMC, highlighting the main recommendations provided by EU partners on each of the Palestinian partners' BMCs:

- Create an inventory of your current infrastructure (hardware and technical solutions but also to the range of policies, guidance, training, and support that are provided) before you start.
- Aim to keep the number of potential value propositions included in the BMC to a minimum.
- With OAIRs, end-users do not usually provide a resource to support activities, so try to make sure that those who will provide the resource for carrying out activities (i.e., who pays for these) can see a value in doing this such as making internal

[5] https://drive.google.com/open?id=122M11PC-CRW2YXq-ArKiQ6ap-46dyWUbO

reporting easier or more eZcient.
- Bear in mind that value can be realized indirectly as well directly.
- Do not underestimate the importance of good communication channels and establishing relationships with end-users (customers). Remember that communication does not always mean person-to- person, it can also be realized through automated systems.
- Consider how best to leverage key partnerships – both within your own institution but also with other universities to deliver value.

Building upon the BMC and policy development work, representatives of the Palestinian universities developed an Institutional Implementation Roadmap.

Its goal was to move from broad aims to a defined program of activities that would help to realize the value proposition(s) of the OAIR and associated services. Participants were asked in particular to consider time frames for delivering each activity and to define mechanisms for assessing progress towards completing each activity (milestones).

The process of developing implementation roadmap was continued locally. The implementation plans were reviewed by the EU partners and feedback was shared with individual partners.

General recommendations for preparing an implementation plan based on the individual feedback reports:

- Scope a pilot implementation before rolling out wider implementation across the entire institution to see if the approach is viable and to identify potential problem areas are likely to scale.
- Provide a specific deadline for each activity to be completed and include relevant milestones to help monitor progress.
- Be realistic about time frames associated with each activity.
- Be clear about who is involved in carrying out each activity (e.g., IT, Library, Research OZce) and where possible, include named staff.
- Keeping track of time, effort and resource required to carry out each activity in the roadmap is a good idea. This will help with assessing the feasibility of scaling up the pilot with the resources available.
- Define metrics relating to each of the roadmap activities (i.e., provide targets for content uploads; number of users).

V. Setting Up OAIRs

In this section we discuss and justify our choice of the repository framework, and present our detailed implementation plan. The main challenges faced during the implementation process are also discussed, focusing on how we managed to approach and resolve these problems by automating services where possible. We believe that our experience and adopted solutions can be informative to other institutions who plan to establish OAIRs.

A. Setting up the repository software (DSpace)

There is plenty of off-the-shield repository software frameworks that are widely used to develop OAIRs around the globe. These frameworks can be either commercial or free to use, and often have different capabilities and limitations. Examples of these frameworks include DSpace[6] and EPrints[7], which are of the most popular repository frameworks. Other frameworks provide only back-end solutions for archiving and data preservation, but the front-end component needs to be developed based on the institution's needs. Examples of these frameworks include Fedora[8] and Islandora[9].

The partner Palestinian universities have decided to go for open source and free-to-use repository frameworks due to the limited financial resources. We believe that it may be risky to start the transition towards OA publishing by investing large budgets on commercial repository frameworks, considering the limited budgets of Palestinian universities.

DSpace was chosen to be used by the fours Palestinian universities due to its popularity, well-documentation, and the huge technical support available online. In addition, the lack of skilled software

[6] https://duraspace.org/dspace

[7] https://www.eprints.org/uk

[8] https://duraspace.org/fedora

[9] https://islandora.ca

16th International Conference on Digital Preservation
iPRES 2019, Amsterdam, The Netherlands.

developers who have experience with developing repository services has made us compelled to use a "turn-the-handle" solution such as DSpace.

Having decided to use DSpace, it important to report the limitations that DSpace has, and how these limitations conflict with the needs of Palestinian institutions in particular. First, the customization of the user interface and back-end services of DSpace is diZcult. In addition, DSpace, like most open-access frameworks, has limited support to the Arabic language. DSpace only supports the localization of control buttons, menus, and toolbars, but it does not support the adaptation of metadata presentation and repository structure based on the user's language of interest.

The setup of the repository software was assigned to the IT teams in all universities. The teams were guided to online resources about the implementation steps. To coordinate activities between teams at different universities, a special online forum was launched so that IT staff can discuss implementation details, share experiences and troubleshoot technical problems.

B. Define and set up metadata catalogues

When planning for metadata catalogues, we had two objectives in mind: First, standard metadata catalogues should be adopted so that the OAIRs contents can be easily discovered and indexed by web crawlers and harvesters. Dublin Core metadata was chosen for building metadata catalogues, but was slightly extended to cope with the specific needs of Palestinian research. In fact, large volumes of Palestinian research outputs, including staff publications and theses, are published in Arabic, especially from Art and humanities departments. Therefore, it is important to have metadata for capturing both Arabic and English bibliographic data. We decided that each Palestinian research output, which can be of any type, should have metadata values in both Arabic and English. For example, a journal article to be deposited into the OAIR should have titles and abstracts in both Arabic and English. This is important to maximize the reach to OAIR contents by both the Arab and wider communities.

Second, one of the objectives of ROMOR is not only to build standalone OAIRs, but also to establish the infrastructure needed to link these repositories in the future. The ultimate goal is to build a national metadata repository that aggregates and enables access to the collective Palestinian research volumes through a single user interface. This goal cannot be achieved without unifying the metadata catalogues to achieve the desired interoperability, and assure that these catalogues are compliant with the FAIR principles. Several meetings have been conducted at both institution's level and between institutions to identify and standardize, where possible, the required metadata catalogues.

C. Customization of user interface to address Arabic language needs

The DSpace front-end theme was modified so that metadata values relevant to the language selected by the user are retrieved and presented. Metadata values of the same item can be rendered in either English or Arabic by simply switching the preferred language from the top bar of the DSpace interface.

Another common problem with repository systems is the inability to distinguish alternative names used by the same author. Authors may write their names in Arabic or English in their publications or may have multiple names due to variations of spelling across different papers. As a result, the same author registered with DSpace may end up with two or more collections of publications, each of which is associated with a different name.

Solutions have been proposed to approach this problem by adopting unique identifiers for authors, such as ORCID, and then grouping repository items by using these identifiers instead of author names. However, Palestinian partners decided not to adopt ORCIDs at this stage because the use of ORCID is still not common in Palestine, and there is a lack of awareness about it among the research community. In addition, many Palestinian researchers publish in local and regional journals not indexed by popular digital libraries, and thus their publications cannot be retrieved by using ORCIDs. We decided at this stage to use the oZcial university email addresses of authors as identifiers in DSpace. Variant names of the same author will be mapped to one oZcial email address, and publications will be grouped by the email address rather than by the author's name.

16th International Conference on Digital Preservation
iPRES 2019, Amsterdam, The Netherlands.

D. Populating repository through batch import

Batch import is a service offered by most repository systems to allow a fast population of repositories. Multiple items with associated metadata can be imported and published in one process rather than submitting items one by one through an input form.

We decided to use batch import only in the first stage in order not to launch empty repositories. In later stages when repositories stabilize and users familiarize with OA practices, relevant policies will be applied to mandate researchers to deposit their research papers. We believe that using batch import is necessary at the beginning of the repository life because it will be frustrating to ask individual researchers to submit dozens of publications, which perhaps were published over decades, one by one. In contrast, if the researcher can find most of his/her publications already deposited into the OAIR, he/she will become more motivated to keep his/her research records up to date.

Above all, the main obstacle for performing batch import into Palestinian OAIRs is to find collections of research outputs to be imported. Only academic theses and staff publications published in university journals are archived and published through the library and journal websites. In addition, it is very diZcult to contact the whole research community at Palestinian universities to collect their research belongings.

Therefore, we decided to collect staff publications and university theses by means of automatic data scraping. The project team developed tools to automatically download theses and staff publications along with associated metadata from different university sites, and deposit them into the OAIR. As the library and university journals websites are not compatible with DSpace in terms of metadata schemes and database structures, the developed tools can restructure collected data into the format required by the DSpace. As a result, thousands of publications could be imported into DSpace from different university's legacy systems. It is worth mentioning that the developed tools included solutions to avoid duplicate submissions of items, to exclude items not published by university staff, and to resolve the problem of variant, multi-lingual and misspelled author names by using text similarity algorithms.

E. Fix and Complete Missing Metadata

Publications retrieved from the web or from the university's legacy systems may have many missing or incorrect metadata. Also, they are unlikely to have metadata values in both Arabic and English. Most importantly, keywords were of the most missing metadata fields because they often exist inside documents but are not embedded in webpages accessible to web crawlers. Keywords are very important for OAIRs to enable subjective classification and browsing of contents. Additional effort is still needed to fix or complete missing metadata of deposited items. As we were dealing with thousands of repository items, we decided to automate this process as possible. To do so, we tried to extract missing metadata from documents by using a machine learning library called Grobid[10]. Grobid is trained to extract bibliographic data from PDF documents automatically. It takes a well-structured publication as input and gives bibliographic data, such as the title, abstract, author names, keywords, journal or conference details, as output. We developed a pipeline to which we inputted thousands of publications into Grobid, and then repository items were directly updated with the extracted metadata. Overall, we found that our metadata correction steps updated or corrected about 23% of OAIR items. However, Grobid is not trained on Arabic text, thus were not applied to Arabic publications.

F. Integrate OAIR with RDM process and other university systems

After being deployed and populated, it is very important for the OAIRs to be part of the research workflow currently implemented at Palestinian universities. In other words, the research oZce, library, and different faculties should start using and contributing to the repository. At the same time, responsibilities should be clearly defined to avoid conflict between university entities and assure a seamless flow of deposited items from the submission step until they get published. After planning workflows on paper, the repository software of OAIRs has been configured to apply the planned workflows. This involves representing the workflow in a machine-understandable format, creating users groups associated with workflow tasks, and assigning responsibilities.

[10] https://grobid.readthedocs.io/en/latest/Introduction

Another important service is to link repository content with the institutional and personal pages of academic and research staff. It may be advantageous for researchers if they can retrieve the list of their publications from the OAIRs and present them on their pages. Linking the repository content with staff pages presents the following benefits: 1) it will release the staff from the burden of creating and editing publication lists manually. 2) It will keep the staff motivated to deposit content regularly to keep their publication lists. 3) Most importantly, it will increase access to the repository content since the visitors of staff pages will eventually land on the repository pages. Therefore, we developed a service that shows on each researcher's site the list of publications he/she deposited into the IR. All that the researcher needs to do is to insert a short script inside his/her site, and the list of publications will be rendered automatically (see Example [11]).

One challenge we identified from the needs assessment studies was the inability of the research oZces and faculties to monitor the research impact and progress. OAIRs provide valuable metrics that can be used to track accessibility and reach to deposited items such as the number of views, the number of downloads and the countries from which visits originate. DSpace provides item-based metrics, but there are no similar metrics for faculties or communities. In addition, there is currently no support for detailed analysis such as tracking the progress of faculties or individuals over a period of time or classifying deposited items based on types or publishers. Therefore, we extended OAIRs by building a reporting service that enables to submit queries similar to the aforementioned scenarios and to present results in both tabular and visual representations. An example of the reporting service can be found here . This reporting service can be easily extended in the future to handle different queries upon request.

VI. Policies

Another aspect of data management being addressed by ROMOR is the development of an institutional policy. The success of ROMOR initiative at Palestinian universities and at the national level in the future will depend strongly on the existing policies and regulations related to the deposition of the different research outputs types, especially on the OA mandate policy at a national and an institutional level. No policy for research data management has been in place before ROMOR. In this section, we describe our efforts in the development of policies in ROMOR Palestinian institutions striving for OA.

[12]

A. *Development Phase*

During the training workshops, participants were introduced to the range of policies that may need to be in place to support their OAIR and associated services. They have been introduced to RCUK Common Principles on Data Policy[13]. The DCC and the Jisc MRD projects guidance and support materials[14] have also been introduced to the participants. The participants have been also trained using the guidance for developing a research data management Policy provided by LEARN Toolkit of Best Practice for Research Data Management.[15]

The process of developing a data management policy in Palestinian universities required extensive consultation with stakeholders and the approval of the relevant committees. In Palestinian universities, the management of research outputs is usually led by deans of scientific research and in collaboration with the libraries, and the IT centers. Natural they took part in developing the data management policy with ROMOR team in each university. The experiences of the local universities showed that such policies need to be in Arabic language and need to be accompanied both by activities to raise awareness of what is being advised, and institutional support services to enable researchers to actually implement the recommendations.

Participants have been working on the development and refinement of their policies iteratively over the life of the ROMOR project. A policy document has been produced and is currently is either pending

[11] http://site.iugaza.edu.ps/ialagha/dspace_pub

[12] https://iugspace.iugaza.edu.ps/DSpaceDBProject/testreport.jsp

[13] http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx

[14] http://www.dcc.ac.uk/resources

[15] http://learn-rdm.eu/wp-content/uploads/RDMToolkit.pdf

approval or approved in each partner Palestinian university. In addition, a policy model in Arabic language has been approved by Palestinian partners. It should serve as a template that can be used by other universities who would like to start the process of developing a data management policy[16].

General recommendations for developing policies have been distilled through aggregation from the individual feedback reports received form EU partners on each of the Palestinian partners' policy:

- It is wise to get senior management buy-in from the outset when developing or revising policies in relation to the OAIR.
- The scope of the policy should be very clearly stated (e.g., the policy covers the deposit of journal articles, research data, and/or theses).
- Be very clear about who is expected to comply with the policy (staff, students).
- Avoid drafting very long or complex policies with legal sounding language. Aim to produce concise policies written in plain language.
- Make a clear delineation between the content of the policy itself and specific details that should be shared as separate guidance.
- It is advisable to avoid referring to named tools or products in the policy itself as these can change over time. Again, this sort of specific information should be provided as guidance.
- Test the policy out before seeking ratification. This will help to ensure that the scope of the policy is realistically achievable with the resources available. Be sure to involve researchers as well as operational units.
- Be sure to consider workflows and institutional infrastructures (hard and soft) that will be required and/or impacted by the policy and be sure to consult with all relevant stakeholders (e.g., operational units that may need to support services to enable compliance with the policy).
- Be sure to spell out any terms that may lack clarity (i.e., if the policy covers 'active' researchers, specify what this term includes).
- Consider who will be responsible for monitoring compliance with the policy and how this will be carried out.

[16] https://drive.google.com/open?id=1a08etl1MTabC2VP3vb-N7QNyKSR5QuaQg

B. Challenges

There are some factors leading to diZculties in effectively implementing the policies:

Compliance with the policy: If OA is only recommended in the developed policies, some researchers will deposit their publications and research data, and some will not. It is not only about personal preferences, but there are also significant differences in OA practices among disciplines too. From the Palestinian universities experience and point of views, if the institution wants to store and preserve its intellectual contribution to human knowledge, the best way is to declare deposition manory. It would be much easier for an institution if the framework for such a mandate exists on a national level, in-laws, codes, regulations, and rules. In the next section of this paper, we will describe ROMOR activities in order to realize all these efforts at the national level. However individual Palestinian universities, for example, mandated that the repository is the main and the only means for offcially recording the researchers' publications activities. These universities generate automatic reports about the research activities of its researchers and consider these reports for evaluations and academic promotions. Researchers are also encouraged to publish their publications in OA journals, however, the preprint, the final version of the accepted or published paper must be deposited, and OA provided in the accordance with the right holders' provisions. If the rights are partially or totally transferred to the publishers, the author is obliged to deposit an eligible version of the paper and provide OA to it as soon as possible. On the other hand, the mandate for depositing the master's and the doctoral thesis is stated in the developed policies.

The capacity of the university to implement the policy: for policy implementation to be successful, it is important for institutions to have the necessary administrative and other abilities required to carry out the implementation process. These necessary skills and resources include the availability of suZcient numbers of qualified staff within the different responsible units. However, the lack of suZcient staff was found to be of particular concern during the implementation process and was seen as a burning issue and a major stumbling block. To overcome this problem, some partner Palestinian

16th International Conference on Digital Preservation
iPRES 2019, Amsterdam, The Netherlands.

universities are decentralizing the RDM efforts. They are including graduate students, researchers, research representatives, and support staff from the different faculties and other research units in the RDM process by involving them first in a series of RDM training workshops. Second, this decentralization of is RDM is reflected in the workflow of the repositories. Through the repository, the students or the researchers first deposit their publications, which are then received by the research representatives or support staff in their faculties, who in return check the deposited item and forward them to the responsible person in the library or in the dean of scientific research, before actually being deposited to the repository.

Copyright issues: there is a lack or almost complete absence of awareness about copyright literacy among most Palestinian researchers. There seems to be a general misunderstanding amongst academics about copyright and the rights they have to re-use their publications after they have signed copyright transfer agreements. Explaining these policies and talking through these issues is often the most diZcult and frustrating part of encouraging people to deposit material. Offering training on copyright literacy and the use of tools for copyright checking such as SHERPA/RoMEO[17] can help indeed. Also during the submission process, the support staff can do any copyright checking on behalf of authors which edge, the best way is to declare deposition mandatory. It would be much easier for an institution if the framework can be done easily using the SHERPA/RoMEO or by contacting publishers directly with a standard e-mail. However, the problem becomes more severe when it comes to publications in Arabic journals. Most of these journals have no clear OA statement and are not included in SHERPA/RoMEO. These diZculties have highlighted the importance of having an institutional statement of copyright, which supports and encourages authors to retain control over their work.

Quality control: Peer-review, is outside the scope of the repository itself. In order to ensure a certain level of quality control, Palestinian universities recommend peer-reviewed submissions to their repositories. However, many Arabic language publications are not peer- reviewed. On the other hand, these publications focus on local problems on most of the time, and therefore there is an essential need to have them visible. Restricting depositing to peer-reviewed publications will leave many important publications invisible.

Cultural issues: some researchers tend to be conservative in their publications habits. They have the impression that OA publishing reflects low-quality research. Some researchers also, have technophobia problem especially those from humanitarian fields. We approach such researchers most of the time personally and using the word-of mouth method. We also offer to deposit in the repository on their behalf until they feel confident they can do it themselves.

VII. National Roadmap

While institutions and their libraries are moving ahead to establish OAIRs to maximize the visibility of their academic output and make it as widely available as possible, there are still considerable obstacles that inhibit academic community from participating in this initiative. Among challenges are low bandwidth, technophobia, technical support and security, finance and legal aspects, lack of skilled personnel to manage the repository, and lack of adequate power supply (as the electric power sector of Gaza is poor and it really affects the servers that house the repository).

A need for a national road map for research output management and a national repository in Palestine was clearly articulated by stakeholders of workshops organized by ROMOR. The participants represented different stakeholders such as the institutions' libraries, research centers, Ministry of Higher Education (MoHE), the Accreditation and Quality Assurance Commission (AQAC), and numerous higher education institutions.

The institutions that already have OAIRs stress the need to identify ways to automatically ingest objects into the repository to minimize the workload on researchers and supporting services. They also mentioned the need to have better external visibility (through integration with hubs like OpenAIRE, etc.), and to support research data that include types beyond theses and publications (such as Databases,

[17] http://www.sherpa.ac.uk/romeo.php

datasets, source code, etc.). However, research data requires special policies and procedures that go beyond the capabilities of current OAIRs. Research data comes in different formats that require special metadata for interpretation and demand distinct formats and description as required by different research communities. In addition, special solutions for digital preservation are required as research data are often stored in formats that rapidly become technologically obsolete. The sensitivity of research data may also influence the security needs that require special consideration in policies. The workshops participants have also discussed the fragmented environment of OAIRs, in which effort and costs are duplicated, numerous software platforms and versions are managed individually by each institution, metadata are applied inconsistently, users are served poorly, and the ministry and the universities are unable to take advantage of collective data about content and users. Besides, digital works come with preservation and storage issues that are only just being realized in Palestine. Institutions need to begin tackling the preservation of digitally produced and recorded material. But the enormity of the task is only just beyond the capacity of many institutions.

A committee was established that comprises representatives from ROMOR team, the Palestinian MoHE, and the AQAC. The committee was responsible for shaping the national roadmap with inputs from the discussions within ROMOR working groups and workshops. The committee hoped by this to foster future efforts towards the organizational, informational and technical development of the national infrastructure.

The national roadmap defines the following objectives:

- identify the needs to integrate scholarly publishing and dissemination of the universities
- identify the requirements to build a shared platform for research outputs management
- build capacity needed for building, managing and sustaining the shared platform
- adopt a common metadata schemes that would be essential for interoperability and for collecting and comparing data across institutions
- establish, maintain and manage a national network of repositories to allow discovery and reuse
- establish, maintain and manage a central infrastructures to enable reliable and interoperable OAIRs
- establish a national center for research data management that will provide infrastructure and services for archiving and sharing research data of different types
- develop research output management and OA policies at the national level
- deploy preservation technology framework
- promote collaboration and partnerships among a large number of institutions and individuals from the academic community in Palestine and in the world
- improve the criteria used for assessment and accreditation of HEIs and academic programs to be research-driven and transparent. AQAC/MoHE can benefit from the metrics gained from OAIRs to assess the research performance and activity of HEIs.

VIII. Conclusion

In this paper, we discussed methodology, solutions, and challenges when developing Open Access Institutional Repositories and research data management services at four universities in Palestine. We have highlighted specific technical challenges, e.g. customization of repository software to support Arabic metadata, as well as institutional challenges relating to research data management policy definition, ratification and adoption. For each of the thematic areas, we have provided specific lessons learned that can be of high value to other developing countries which aim at increasing visibility of their research outputs and strengthening competences of their researchers in data management and preservation.

The ongoing and future work focuses on implementing a national roadmap for integrated services that allow not only to promote publications but also to share and preserve data. This will enable to reduce costs and facilitate the exchange of expertise and highly skilled human resources. We are also working towards better integration with initiatives such as the European Open Science Cloud or PlanS.

References

[1] M. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. O. Bonino da Silva Santos, P. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. Evelo, R. Finkers, and B. Mons, "he FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data*, vol. 3, 2016. DOI: 10.1038/sdata.2016.18.

[2] S. Zyoud, S. Al-Jabi, and P. W. Sweileh, "Scientific publications from Arab world in leading journals of Integrative and Complementary Medicine: A bibliometric analysis," *BMC Complementary and Alternative Medicine*, vol. 15, 2015. DOI: 10.1186/s12906- 015-0840-z.

[3] S. H. Zyoud, "Dengue research: a bibliometric analysis of worldwide and Arab publications during 1872–2015," *VirologyJournal*, vol. 13, pp. 1–10, 2016.

[4] M. of Higher Education, *Annual Statistical Directory of Palestinian Higher Education Institutions* 2017-2018, 2018. [Online]. Available: https://www.mohe.pna.ps/services/statistics.

[5] IAP, "Scientific Research in Palestine," *Palestine Academy for Science & Technology*, 2002.

[6] W. M. Sweileh, S. H. Zyoud, S. W. Al-Jabi, A. F. Sawalha, and S. A. Khalil, "Research Output from Palestine (1995–2012): A Bibliometric Study," *International Information & Library Review*, vol. 46, no. 3- 4, pp. 99–112, 2014. DOI: 10.1080/10572317.2014. 943070.

[7] W. M. Sweileh, S. H. Zyoud, S. Al-Khalil, S. W. AlJabi, and A. F. Sawalha, "Assessing the Scientific Research Productivity of the Palestinian Higher Education Institutions: A Case Study at An-Najah National University, Palestine," SAGE Open, vol. 4, no. 3, 2014. DOI: 10.1177/2158244014544287.

[8] M. Qumsiyeh and J. Isaac, "RESEARCH AND DEVELOPMENT IN THE OCCUPIED PALESTINIAN TERRITORIES: CHALLENGES AND OPPORTUNITIES," *Arab Studies Quarterly*, vol. 34, no. 3, pp. 158–172, 2012. [Online]. Available: http://www.jstor.org/stable/41858700.

[9] X. Y. C. W. Yongchao ZHAO, "Academic institutional repositories in China: A survey of CALIS member libraries," *Journal of Data and Information Science*, vol. 5, no. 2, 2012.

[10] S. Ali, S. Jan, and I. Amin, "Status of Open Access Repositories: a Global Perspective," *International Journal Of Knowledge Management And Practices*, vol. 1, no. 1, pp. 35–42, 2013.

[11] D. C. Prosser, "Institutional repositories and Open Access: The future of scholarly communication," *Inf Serv Use*, vol. 23, 2003. DOI: 10.3233/ISU- 2003- 232-327.

[12] R. Cullen and B. Chawner, "Institutional Repositories, Open Access, and Scholarly Communication: A Study of Conflicting Paradigms," *The Journal of Academic Librarianship*, vol. 37, no. 6, pp. 460 –470, 2011. DOI: https://doi.org/10.1016/j.acalib.2011.07.002.

[13] H. Hockx-Yu, "Digital preservation in the con text of institutional repositories," *Program*, vol. 40, pp. 232–243, 2006.

[14] O. Saini, "Understanding the Role of Institutional Repository in Digital Preservation in Academic Libraries: A Review of Literature," *Library Philosophy and Practice,* vol. 2018, 2018.

[15] J. Mavodza, "A review of the open access concept in the UAE," *New Library World*, vol. 114, 2013. DOI: 10.1108/03074801311326885.

[16] S. S. Ahmed and S. Al-Baridi, "An overview of institutional repository developments in the Arabian Gulf Region," *OCLC Systems & Services: International digital library perspectives*, vol. 28, pp. 79–89, 2012. DOI: 10.1108/10650751211236613.

[17] S. Carlson, "An Assessment of Institutional Repositories in the Arab World," *D-Lib Magazine*, vol. 21, no. 5/6, 2015. DOI: 10.1045/may2015-carlson.

[18] A. aalYateem and N. Bn-Hameed, "Digital Repositories in the Arab Universities: A Comparative Analytical Study," *Procedia Computer Science*, vol. 65, pp. 768 –777, 2015. DOI: https :// doi . org / 10 . 1016/j.procs.2015.09.022.

[19] I. AlAgha and R. Awadallah, *Assessing Current Open Access and Research Data Management Practices and Services in Palestinian HEIs,* Project number: 573700- EPP-1-2016-1-PS-EPPKA2 -CBHE-JP. Co-funded by the Erasmus+ programme of the European Union., 2017. DOI: 10.5281/zenodo.801735.

[20] R. Awadallah and I. AlAgha, *Assessing Research Data Output Management at the Managerial Levels at Partner PS HEIs*, Project number:573700-EPP-1-2016-1- PS-EPPKA2-CBHE-JP. Co-funded by the Erasmus+ programme of the European Union, 2017. DOI: 10. 5281/zenodo.801773.

[21] J. Rans and A. Whyte, *Using RISE, the Research Infrastructure Self-Evaluation Framework*, Edinburgh: Digital Curation Centre, 2017. [Online]. Available: www.dcc.ac.uk/resources/how-guides.

[22] W. Los, "EDISON Project: Building Data Science Profession for European Research and Industry," 2016.