

## Data archiving for dissemination within a Gulf nation

Brian W. Mandikiana<sup>1</sup>, Lois Timms-Ferrara<sup>2</sup>, and Marc Maynard<sup>3</sup>

### Abstract

Since 2008, Qatar University's Social and Economic Survey Research Institute (SESRI), has been collecting nationally representative survey data on social and economic issues. In 2017, SESRI leadership established an Archiving Unit tasked with data preservation and dissemination both for internal purposes and with the intent of disseminating select data to the public for secondary analysis.

This paper reviews the lessons learned from creating a data archive in an emerging economy where both cultural and political sensitivities exist amid diverse groups of stakeholders. Challenges have included recruiting trained personnel, developing policies for data selection and workflow objectives, processing restricted and non-restricted datasets and metadata, data security issues, and promoting usage. Additionally, there is hope that the presence of the Archiving Unit adds value for other SESRI research staff involved in the design, collection, documentation, and processing of studies. After successfully addressing these challenges over the past year, the Archive met its objective to launch a data center at the Institute's website (<http://sesri.qu.edu.qa>) and to make multiple datasets available for public download from it. Also, to be discussed are the tools, processes and leveraging of resources that are being implemented as the archiving process continues to evolve.

### Keywords

Data archive, Qatar, Middle East, open data, digital preservation workflow, repositories

### 1. Introduction

The Middle East and North Africa (MENA) region presents a unique environment for survey research throughout the data lifecycle. Research institutes in the region face the challenges of navigating cultural and political sensitivities and the availability of research capabilities. These research institutes also have to figure out how these aspects among other factors affect planning and design, data collection, data processing, data dissemination, and data re-use. In Qatar, considerable progress has been made in promoting survey research (Gengler, Le, and Howell, 2018), and has resulted in the collection of several studies. Similarly, over the past decade, many research organizations have emerged within the Gulf region and the MENA region as a whole, generating dozens of datasets. Given the insufficient efforts in data sharing initiatives across the Gulf region, access to data for replication of published research output to confirm findings is limited. Data sharing throughout this region would improve the credibility of statistics generated from the region, improve relevance and reliability of such data, lessen data collection costs incurred due to fielding similar surveys, and build a body of knowledge that policy-makers could use in addressing various issues facing the region.

Government and in particular, national statistics offices throughout the MENA region have been the driving force for survey data collection, data processing and sharing of emerging statistics. As more government offices and data collection agencies within the region are adopting open data and FAIR (findable, accessible, interoperable, and re-usable) data principles, the sharing of data and related metadata is gaining momentum (Saxena, 2018). Organizations beyond government (private foundations, companies, research centers) are beginning to sponsor data collection, and these institutions require a review of existing data to inform and supplement their efforts on behalf of their stakeholders (Elsayed and Saleh, 2018). Although findings from a recent study by Shaon, Straube, and Chowdhury (2017) highlight deficiencies in capabilities, particularly skills required to execute data sharing activities, efforts to address the lack of data sharing are underway, beginning with the establishment of the first open data archives for survey research in the region. Below we give a brief description of the Social and Economic Survey Research Institute (SESRI), an independent survey research organization that recently established a data archive and how it has approached promoting data sharing given the contextual challenges that the organization faces.

### **1.1 The Social and Economic Survey Research Institute a brief history**

The Social and Economic Survey Research Institute (SESRI) is an independent academic research organization at Qatar University. Since its inception in 2008, with the assistance of the Institute for Social Research (ISR) based at the University of Michigan, it has developed a robust survey-based infrastructure in order to provide high-quality survey data for planning and research in the social and economic sectors. The data are intended to inform planners and decision makers, as well as the academic research community.

Since 2008 SESRI has conducted more than 50 household surveys on such topics as entrepreneurship, education, food security, marriage delay, social capital, labor, migration, tourism, health, among others. Also, the on-going omnibus surveys that cover many timely subjects.

With the development of a fully operational and sustainable survey research program at Qatar University, SESRI has accomplished much in line with its purpose and mission. The mission of the Social and Economic Survey Research Institute (SESRI) is;

*“to contribute to the development of Qatari society by providing high-quality survey data to guide policy formulation, priority setting, and evidence-based planning and research in the social and economic sectors.”*

The above mission statement highlights the importance of data access. It is from such a background that the need for a microdata data archive emerged.

### **1.2 SESRI Data Archive: A (Very) brief history**

Founded in 2017, the SESRI Data Archive, housed at Qatar University is one of the few microdata archives in the Persian Gulf region providing access to high-quality socio-economic household survey data. The archive was created due to the recognition of the need to maintain and re-use SESRI's own created resources, and the desire to share data to advance evaluation of social and economic policies in Qatar, the Gulf region, and beyond. Furthermore, Angel-Urdinola, Hilger, and Ivins (2011)

highlighted the need to enhance access to microdata in the Middle East and North Africa region. Since its establishment, the archive has supported SESRI survey data management, preservation, internal use, and external dissemination. The SESRI Data Archive's mission is:

*"To advance globally accepted standards of economic and social survey research data for; storage, management, preservation, and dissemination."*

Few studies (Angel-Urdinola, Hilger, and Ivins, 2011; Saxena, 2016; Saxena, 2017; Thompson, 2009) report microdata archiving initiatives in the Middle East and North Africa or the Gulf region. In this paper, we describe a case study of the establishment and development of the SESRI Data Archive in the State of Qatar and explore the approach and challenges faced during the first year of operation. After briefly reviewing the state of survey data sharing in the Gulf region, we preview the challenges confronted and, in some cases, anticipated, followed by initial development of the archive in terms of staffing and working policy development. Next, we review the development of the archive data processing workflows. We conclude with some brief comments on the partners' plans for strengthening the infrastructure of the archive and promoting its use and growth.

## 2. Data sharing culture in the Gulf region

Availability of microdata from countries in the Persian Gulf region has been limited both in terms of comparative analysis across countries, but also within individual countries. Several outside efforts to investigate views of societal structures and policy perspectives have taken place over the years. Some of the well-established include; the Arab Barometer (2005-ongoing), the World Values Survey (1981-ongoing), and more recently the Pew Global Attitudes project (2001- on-going). Survey research efforts to collect and publicly share microdata within individual countries have been less ambitious and met with moderate success (Angel-Urdinola, Hilger, and Ivins, 2011; Saxena, 2017). Although data collection through household survey has been a standard way of supporting governments' decision-making processes in the region, data sharing has not been a critical feature in the Persian Gulf. Other institutions involved in social science survey research include the Emirates Center for Strategic Studies and Research (ECSSR) and the Bahrain Centre for Studies and Research, Bahrain (BCSR). However, when looking at the MENA region, notable successes include the Economic Research Forum (ERF), focusing on the Middle East and North Africa (Thompson, 2009).

In Qatar, recent noteworthy data sharing developments have emerged. First, the Ministry of Information Communication and Transport in Qatar published the Open Data Policy (ICTQATAR, 2014). This policy stresses the importance of making data available, while at the same time using the linked open data model as a guide to further the degree of openness. It is argued that a data sharing culture will facilitate not only efficient public service delivery but building a less extensively hydrocarbon-based economy that is knowledge-based and shares recent data with the public in an accessible manner (ICTQATAR, 2014). Second, in addition to the open data policy, the data management policy was published to inform issues of data handling in the State of Qatar. Among other issues highlighted is the need for professionals to support data archival related processes. These policies demonstrate the leadership's willingness to support a culture of open data sharing in the State of Qatar.

The Social and Economic Survey Research Institute (SESRI) was founded on a principle of data sharing for research and policy planning purposes. Furthermore, the research institute can provide a model for how microdata and associated documentation may be shared from within a unique and sometimes challenging environment. SESRI is in a unique position in year ten of its research program. It has conducted dozens of high quality, timely and substantively relevant survey research projects that have already seen primary use in public policy planning in Qatar, as well as, in peer-reviewed policy-oriented academic journals (Al-Emadi et al., 2017; Diop et al. 2016; Gengler and Mitchell, 2018). SESRI management sees opportunities presented by providing materials for reproducible research and the growth in open-access publishing. The underlying infrastructure for data sharing is still in its infancy. Policies to support data use, data management and access have not been fully developed. Internal research networks are still nascent and need further development and encouragement.

### 3. Challenges

As in many cases, the great promise can be accompanied by significant and diverse challenges as the SESRI Data Archive set down its development path. Challenges faced by the archive can be grouped into three areas: technical, organizational, and contractual or legal. Technical issues tend to be common to many digital preservation efforts and focus mainly on the status of data and metadata resources and the ability to harness those resources to provide valuable materials for future researchers. Organizational and contractual challenges can be unique to a particular context and may require distinctive approaches and innovative solutions to address. The solutions to these challenges can be found by looking at exemplary institutions and proven standards and best practices.

#### 3.1 Technical Challenges

Survey metadata is compiled from a variety of sources including published summary results, survey reports, and methodological reports. Recovery of complete and updated survey data and documentation files and supporting metadata provided the first primary challenge to the archive. Since the surveys were conducted many years ago and are subsequently used for not only primary analysis but also reused in succeeding analyses, identifying complete and authoritative version files required additional effort working with principal investigators (PIs), analysts and information technology (IT) staff. Unfortunately, standard file-naming conventions and version control were not implemented across survey projects and, further, multiplicities<sup>4</sup> or multiple versions of the same file were found and needed to be rectified. While files were stored on a shared network drive, external contextual metadata was not captured. The lack of unique and permanent study identifiers also hindered the development of an authoritative survey inventory.

For each survey, several relevant files had to be reviewed and evaluated for completeness based on consultation with appropriate research unit staff, technical evaluation of file formats and comparisons of datasets and related documentation (i.e., questionnaires, Blaise export, etc.). This tended to be an iterative process due to the loss of context because sufficient metadata was not captured originally at its creation. Blaise is the computer-assisted telephonic interviewing (CATI)/computer assisted personal interviewing (CAPI) program utilized at SESRI and while Blaise exports both Arabic and English questionnaires, if available, these files needed to be cross-referenced to typically incomplete English-only variable labels as stored in the exported Stata data file (.dta).

Concerning dissemination efforts, SESRI had no external data dissemination channel other than the availability of publications via its main website.

### **3.2 Organizational Challenges**

Throughout its research, SESRI had developed a generally accepted set of procedures concerning data management throughout the lifecycle of the project, but did not have formal data management policies to guide data use beyond the completion of the project and publication of the final report(s). Research and Policy Units worked in tandem, but data preservation and archival work were not built into the process. Lead Principal Investigators and research assistants worked with the resulting data primarily for analytical purposes and developing scholarly research outputs and did not see the value of maintaining a type of persistent, reusable version of the resulting data files. This situation is not unique; many survey research organizations complete a project and move on immediately to the next one and do not have time to describe and document the just-completed project.

The challenge for the Archive Unit was to re-define the division of labor and re-configure the requirements around the survey files hand-off from the Research Unit.

### **3.3 Legal and contractual challenges**

SESRI is committed to preserving the confidentiality and privacy of individual survey respondents. As with all its survey projects, protecting respondent confidentiality and privacy is critically important. The Archive Unit extends this protection beyond the initial collection and use of the data from individual interviews to data re-use and exposure beyond SESRI staff. Various approaches are employed to enforce this commitment including the creation of a public release version of datasets that suppress appropriate variables or uses other mechanisms to mitigate the risk of violating response confidentiality.

Additionally, contractual arrangements with funding partners can impose restrictions on the preservation, accessibility, and use of SESRI datasets. Each survey must undergo a thorough review of underlying funding commitments and business considerations as a critical component when determining if and how the data can be documented, stored and released.

## **4. Recruitment and staff**

Skilled human capital is one of the essential factors in building a data archive. It is recommended that a minimum of four dedicated staff be recruited for a well-functioning data archive. Currently, the SESRI Data Archive is under-resourced. The lead archivist has the primary role in managing the day-to-day operations of the data archive. Although recruiting more staff was approved, securing local hires has been a challenge for several reasons. First, the concept of microdata archiving is commonly unknown or misunderstood by most people within the region. Second, few institutes equip students with a set of skills that are essential for data archiving. Finally, academic qualifications such as information technology, information systems, statistics, are not as prominent as engineering studies in the Gulf region. On the other hand, recruiting non-local staff to take up the posts is challenging for many reasons. Due to changes in policy, there has been more focus on hiring locally. However, only when that fails are international candidates considered. In as much as that provision is available, getting

multilingual candidates with working knowledge of Arabic is yet another challenge. This adds to the length of the recruitment process, and as a result potential good candidates end up getting hired by other companies while the recruitment process is still in progress.

For the past one year since the establishment of the SESRI Data Archive, the existing lead data archivist together with external collaborators from other organizations have worked together on many items essential for the normal operation of the data archive. Also, financial support, IT services support, and human resources services have been provided within SESRI and by other departments within the housing institution, that is, Qatar University.

## 5. Approach and architecture

Development of the SESRI Archive proceeded on two tracks: general policy development and practical data processing workflow implementation. On the policy track, general information gathering and best practice reviews were conducted with a particular focus on generally accepted frameworks to learn and evaluate broad organizational structures and policies. This included a review of data management and preservation policies from a variety of data archives and preservation institutions. Data management policies for primary data collection organizations were reviewed and discussed. From these, a draft data management policy was created and discussed with SESRI administration and presented in a training setting to SESRI staff (Research Unit, Policy Unit and IT).

In summary, the data management policy sets out to achieve the following objectives:

*To support research collaboration through efficient access to survey data;  
To warrant that SESRI confirms with requirements of funding bodies, especially Qatar  
National Research Foundation.*

The National Digital Stewardship Alliance (NDSA) levels of preservation (Ashenfelder, 2016) were subsequently reviewed to provide a departure point for further discussions around minimal and aspirational archive goals. The NDSA levels lend themselves to this type of evaluation. In all discussions it was recognized that the Archive Unit should adhere to and implement methods to meet the minimal requirements (NDSA level one) and develop plans and actions to reach aspirational goals along the continuum on various dimensions, including storage location, data integrity, metadata development, standardized file formats, as well as, data access.

A reasonably aggressive schedule to release up to four studies in the first nine months of the archive was set early in the process. These first four studies allowed for discussion around general issues of public release both on the administrative level as well as within the archive team. Exploring multiple dimensions of concerns for releasing microdata for the first time required a thoughtful and deliberate process on the one hand, but also a flexible and iterative process on the other. Clear communications with stakeholders (including management, unit heads, and PIs) were critical. Documentation of decision-making processes was also crucial.

Once the studies were selected and agreed upon, an efficient data processing workflow needed to be identified and developed. The team divided this process into three processing phases: 1) pre-ingest evaluation and initial preservation; 2) data processing for preservation; and 3) data processing for public release. It was understood that the public release files would not include confidential variables, contractually restricted variables, as well as variables that are restricted for other reasons. Internally, though, the Archive Unit's charge is to preserve data and documentation from all SESRI projects regardless of whether a public release version would be made available.

Archive staff used the Open Archival Information Systems (OAIS) Reference Model (Consultative Committee for Space Data Systems, 2012) for thinking about and planning the processing workflow. While work had already begun on several data files, process tasks were identified and reviewed for inputs, outputs, and dependencies. These were mapped to the OAIS model for further discussion and modification.

## 6. Data processing workflow

Data processing, both in terms of data cleaning and metadata development, are critically important to the overall success and sustainability of any data archive. Consistent, reliable and reproducible actions are threaded together to provide standard pathways that can be managed efficiently. The ability to successfully clean and preserve one specific data resource must be, itself, preserved through testing and documentation. The short-term effort involved in thinking through required steps, identifying dependencies, and documenting their effectiveness, pays dividends in the long-term as the data collection and the Archive Unit grows.

For the SESRI Archive workflow, data processing was designed around capabilities and features of Stata<sup>5</sup>, and metadata development efforts rely on Colectica<sup>6</sup> Designer File naming conventions and network file storage structures were established and documented taking into consideration the anticipated data collection scope and versioning. Data cleaning and modularized, standardized labeling and recoding schemes were designed and documented by Archive Unit staff to provide consistent and reliable data handling routines. Additionally, guidelines were established concerning code documentation and recording of decision-making rationale for future reference.

The current SESRI Archive Unit processing workflow (see Table 1) can be thought of as an iterative process with three major phases: 1) pre-ingest review; 2) data processing for preservation; and 3) data processing for public release. Each of these phases had several quality control reviews built-in, and each ultimately ends with a newly preserved data resource and appropriate documentation.

Table 1. SESRI Workflow Summary

1	Pre-Ingest Review	Submission Information Package (SIP) evaluation; Review for confidentiality and potential public release
	Ingest	Study level cataloging; File level technical review
	Preservation	Files to archive storage; Metadata to repository
2	Data Processing	Initialize administrative checkpoints; Build study metadata; Dataset checks and cleaning; Build variable metadata; Record actions
	Data Quality Review	Review dataset and processing log
	Codebook Creation	Generate and review English and Arabic codebooks
	General Quality Review	Review Codebook and Dataset(s) based on processing log
	Preservation	Files to archive storage; Metadata to repository
3	Public Release Screening	Policy review; Contractual review; Confidentiality review; Variable level screening; Final approval
	Prepare Public Use Files	Remove restricted cases, variables, administrative variables, etc.; Prepare usage notes; Citation requirements; Review Terms and Conditions
	Preservation	Files to archive storage; Metadata to repository

Concerning public data release screening, an advisory committee provides oversight and guidance to the Archive Unit to balance respondent confidentiality concerns and contractual restrictions with the overarching goal to provide open access to SESRI research output. The Archive Unit employs a variety of methods to protect confidential data including de-identifying or anonymize sensitive information and aggregating indicators at appropriate geographic levels.

## 7. Data availability

In order for the archive to advance appropriately SESRI's goal of sharing data with scholars and thereby contributing to scientific research on important issues, a data portal was created offering public release versions of the metadata and datasets to select surveys, with access provided directly from the website. This portal would become the Data Centre.

### 7.1 Data center portal for dissemination

The SESRI website had been well established and designed to introduce the Institute to a variety of stakeholders including the Qatar University community, policy analysts around the country, media outlets, and the broader research community interested in Qatari public opinion. In an organized format, the site is easily navigated with links provided to key analytical reports, SESRI sponsored events, training workshops. The Data Center, however, was a different endeavor for the website and communications unit.

As a primary research facility within Qatar University, known as the national institute of higher education in Qatar, there are inherent dependencies and relationships with stakeholders that must be cultivated across departments and takes time to achieve. One unanticipated mandate was that Qatar University Web Services had to build the Data Center website and upload the prepared studies

for dissemination to assure adherence to standards and branding protocols. This involved significant involvement of the SESRI communications and archive staff with the University Web Services team to realize a fully functional Data Center that in January 2018 became the official public face of the SESRI Data Archive.

Archive staff is actively adding materials to the Data Center website with new studies becoming available every two to three months. There are extensive plans for expanding the website to include efficient data discovery tools, including a catalog of publicly released studies once there is a substantial collection uploaded.

## 7.2 User and access support

Staffing requirements for providing researcher support are being assessed and is anticipated to grow with the size of the public offerings. A list of Frequently Asked Questions (FAQs) has been, and a variety of additional support documentation is being considered as more information is gathered on user needs. Plans are in place to secure DOI space to assign unique identifiers and manage version support.

## 7.3 Promotion marketing the archive

A multistage approach for promoting the use of the archive has been developed, but not yet tested. The plan involves a combination of traditional outreach with public and targeted announcements, press releases, and email communications, along with components of social media. The plan further invites partnerships with organizations of mutual interests and a set of planned paid promotions.

## 7.4 Measuring success

Given institutional objectives to share data among an intellectual community of research scholars in order to inform discussions that advance social science, the metrics for measuring success include both short term and long term outcomes. In the short run, success will be gauged by Data Center traffic reports, detailed records of downloaded materials, and the amount of interest in the Qatari opinion as measured by an optional form to be added to the mailing list. As the Data Center becomes more complete, the team plan to build a bibliography of scholarly articles, conference papers, and teaching materials that utilize SESRI data and develop a portal for submitting completed works and citation links.

## 8. Discussion

The discussion above has shown the processes involved in building a microdata archive. In summary, to increase the success of the data archive establishment, many factors need to be considered.

### *Financial investment*

Research institutes and organizations need to plan for the financial investment required for data dissemination. Financial investments may range from the cost of data documentation, special handling of restricted data, expert consultation fees to address particular issues, and appropriate data processing of data awaiting dissemination. In situations where the research institution is involved in all parts of the data lifecycle, there are more cost-cutting opportunities.

### *Workflow documentation*

Data processing workflows must guide data processing staff in a manner that supports consistency and reproducible research. Furthermore, similar to data files, it is important to implement good preservation of syntax used for data processing. Annual reviews of workflows are equally important. Updating these processes can significantly contribute to efficiency.

### *Confidentiality*

A common concern amongst many research Institutions is meeting the requirements for maintaining confidentiality. Data centers planning to share data will need to invest time in this process. Notably, besides removing data such as respondent residential areas, information on the day, month, and year of birth, names, other exploratory statistical techniques and aggregation can be applied to reduce the risk of re-identifying survey participants.

### *Data publications*

Many researchers have proposed data publications as a way of promoting data sharing. This strategy could help unlock data enclaves in the Gulf region where for many years data sharing has not been part of the norm. Hsu et al. (2015) highlights the rapid rise in the data publications as a concept and incentivizing strategy for data sharing. However, for this strategy to work Hsu et al. (2015) proposes making data papers part of the mainstream academic output, such as journal articles.

### *Training*

A feature of most statistical related training programmes is that students usually work with clean data. In order to build capacity, students should receive training data management as well as sound exploratory data analysis techniques. Similar to Shaon, Straube and Chowdhury (2017), in Qatar and other countries in the region, research institutions need to provide training opportunities to students. Besides, data centers should encourage researchers to participate in data management related programmes as a way of retooling as well as keeping up to date with the evolving statistical programming software.

## **9. Conclusion**

At SESRI the need for an institutional social sciences data archive has been recognized. Qatar's open data policy and data management policy suggest that there is much needed support at the national level to promote data sharing. Given the limited microdata archives in the Gulf, the OAIS reference model is critical in promoting a better understanding of archival processes in the region. Data Archiving is beneficial to the growth of quantitative research in the Gulf. Equally important, conducting this case study has allowed us to share insights on setting up a data archive, the challenges involved, and lessons learned in the process. Ways to improve prospects for data sharing could include data management plans with understandable terms of use for gathered data files. Equally important, particularly in the social science field, sharing data anonymizing techniques can contribute towards reducing the risks of data sharing. Moreover, when Data Archive staff offer training on best practices and standards for efficient hand-offs to depositors, it dramatically improves ingest processes.

The potential for growth in data sharing exists in Gulf countries. Re-use of nationally representative household survey data needs to be encouraged in Gulf states to further data quality, smoother access and metadata documentation that supports analysis. Research institutions in the Gulf need to build on the sparse data dissemination programmes that have been established thus far and making data related products interoperable. These are foundations upon which they may build data archives or at least deposit survey data with established archives in the Gulf region. Thus, we need to start implementing archival processes, tap into innovative infrastructures that are emerging to preserve survey data for generations to come.

## References

Al-Emadi, A., Kaplanidou, K., Diop, A., Sagas, M., Le, K.T. and Al-Ali Mustafa, S. (2017) '2022 Qatar World Cup: Impact Perceptions among Qatar Residents', *Journal of Travel Research*, 56(5), pp. 678-694.

Angel-Urdinola, D.F., Hilger, A. and Ivins, I.B. (2011) 'Enhancing access to micro-data in the Middle East and North Africa' *Arab World Brief*, 4. Washington, DC: World Bank. Online: <https://openknowledge.worldbank.org/handle/10986/9451>

Ashenfelder, M. (2016) *Expanding NDSA Levels of Preservation*. Available at: <https://blogs.loc.gov/thesignal/2016/04/expanding-nds-a-levels-of-preservation/> (Accessed: 14 March 2018).

Consultative Committee for Space Data Systems (CCSDS) (2012) *Reference Model for an Open Archival Information System (OAIS)*. Magenta Book. Issue 2. June 2012. <https://public.ccsds.org/Pubs/650x0m2.pdf>

Diop, A., Le, K.T., Johnston, T. and Ewers, M. (2016) 'Citizens' attitudes towards migrant workers in Qatar', *Migration and Development*, 6(1), pp. 144-160. DOI: 10.1080/21632324.2015.1112558

Diop, A. Al Ansari, M., Le, K.T., Elmaghraby, E., Al Bloshi, A., Al Qassas, H., Al Khulaifi, B. and Mustafa, S. (2017) 'From the "Fareej" to Metropolis: Qatar Social Capital Survey II.' SESRI, Qatar University. Available at: <http://hdl.handle.net/10576/6386> (Accessed: 20 January 2018)

Elsayed, A. M. and Saleh, E. I. (2018) 'Research data management and sharing among researchers in Arab universities: An exploratory study.' *IFLA Journal*, 44(4) pp. 281–299. DOI: 10.1177/0340035218785196.

Gengler, J., Le, K. T., and Howell, D. (2018) 'Survey challenges and strategies in the Middle East and Arab Gulf regions.' In T. P. Johnson, B.-E. Pennell, I. A. L. Stoop, and B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional and multicultural Contexts (3MC)* (pp. 555-568). New York, NY: John Wiley.

Gengler, J. and Mitchell, J.S. (2018) 'A Hard Test of Individual Heterogeneity in Response Scale Usage: Evidence From Qatar', *International Journal of Public Opinion Research*, 30(1), pp. 102–124. Available at: <https://doi.org/10.1093/ijpor/edw025> (Accessed: 20th January 2018)

Houghton, B. (2016) 'Preservation challenges in the digital age', *D-Lib Magazine*, 22(7/8), pp. 1-6. DOI: 10.1045/july2016-houghton

Hsu, L., Martin, R.L., McElroy, B., Litwin-Miller, K. and Kim, W. (2015) 'Data management, sharing, and reuse in experimental geomorphology: Challenges, strategies, and scientific opportunities', *Geomorphology*, 244, pp. 180-189.

ICTQATAR (2014) Open Data Policy. Ministry of Information Communication and Transport, Qatar. Available at: <http://www.ictqatar.qa/en/documents/document/open-data-policy> (Accessed: 20 January 2018)

ICTQATAR (2015) Data Management Policy. Ministry of Information Communication and Transport, Qatar. Available at: <http://www.ictqatar.qa/en/documents/document/data-management-policy> (Accessed: 20 January 2018) Saxena, S. (2018) 'Drivers and barriers towards re-using open government data (OGD): a case study of open data initiative in Oman', *Foresight*, 20(2), pp.206-218.

Saxena, S. (2017) 'Open public data (OPD) and the Gulf Cooperation Council (GCC): Challenges and prospects', *Contemporary Arab Affairs*, 10(2), pp. 228-240.

Saxena, S. (2016) 'Integrating Open and Big Data via 'e-Oman': prospects and issues', *Contemporary Arab Affairs*, 9(4), pp. 607-621.

Shaon, A., Straube, A. and Chowdhury, K. R. (2017) 'Setting up a National Research Data Curation Service for Qatar: Challenges and Opportunities', *International Journal of Digital Curation*, 12(2), pp. 146–156. DOI: <https://doi.org/10.2218/ijdc.v12i2.515>

Thompson, K. A. (2009) 'Data in development: An overview of microdata on developing countries', *IASSIST Quarterly*, 33(4), pp. 25-30.

---

## End-notes

<sup>1</sup> Brian Mandikiana is the Senior Research Data Analyst / Lead Archivist at The Social and Economic Survey Research Institute (SESRI), Qatar University and can be reached by email:

[bmandikiana@qu.edu.qa](mailto:bmandikiana@qu.edu.qa).

<sup>2</sup> Lois Tims-Ferrara is the Chief Executive Officer at Data Independence, Ellington, Connecticut, USA

<sup>3</sup> is the Director of Technology at Data Independence, Ellington, Connecticut, USA

<sup>4</sup> Preservation Challenges in the Digital Age,

<http://www.dlib.org/dlib/july16/houghton/07houghton.html>

<sup>5</sup> Stata – [www.stata.com](http://www.stata.com) (viewed 14 March 2018)

<sup>6</sup> Colectica – [www.colectica.com](http://www.colectica.com) (viewed 14 March 2018)