

Investigating the Potential for Open Government Data (OGD) in Qatar

Ali Albinali

Department of Computer Science
Loughborough University
Loughborough, UK
e-mail: A.Albinali@lboro.ac.uk

Russell Lock

Department of Computer Science
Loughborough University
Loughborough, UK
e-mail: R.Lock@lboro.ac.uk

Iain Phillips

Department of Computer Science
Loughborough University
Loughborough, UK
e-mail: I.W.Phillips@lboro.ac.uk

Abstract—Data are growing rapidly, and technological breakthroughs are adding new methods and strategies to tackle big data. Thus, research challenges exist to explore the value of big data in supporting improved decision-making and public service delivery across various industries. In recent years, many countries have started to utilize Open Government Data (OGD) in developing open environments and platforms to improve their economies, enabling small and medium enterprises (who may lack the resources to undertake their independent market analysis) to use them. Qatar is no exception; it has begun to promote and make it available to everyone, including small and medium-sized companies. That is due to the potential economic returns and enhanced transparency. This research looks at aspects of OGD and the situation in Qatar. In addition, we collected data from various organizations and stakeholders using two questionnaires. The key goals of the two independent questionnaires were to investigate the amount of knowledge of OGD as a concept, how Small and Medium Enterprises (SMEs) in Qatar perceive present OGD platforms, and how OGD may be improved to fulfill the requirements of SMEs. Further, regardless of gender or educational level, most Qatari youth and adults reacted positively to using OGD.

Keywords- Big data; Open Government Data (OGD); Small and Medium Enterprises (SMEs); Data Analytics; Data Analytics Framework (DAF).

I. INTRODUCTION

SMEs adopt data analysis and management tools, enabling them to visualize information by synthesizing data from multiple value chain processes [1]. In addition, introducing and utilizing Big Data in any SME requires establishing a model to gather and analyze the data. A good data analysis model will help analyze raw data sets to understand and realize the information they contain and discover patterns in the data to derive valuable insights. Furthermore, adopting a data analysis model aims to overcome data handling difficulty and exploit the volume, variety, velocity, and veracity elements that necessitate a significant data flow in a constantly evolving system [2]. It generates a convenient analysis to solve the challenges caused by a growing amount of data. Further, with the appropriate data analysis model, SMEs can maximize their performance effectively, cutting costs by developing more efficient ways to conduct business and maintaining enormous volumes of data. SMEs using readily available data to provide high-quality services will advance in a competitive climate. Therefore, utilizing Big Data is an opportunity for SMEs to innovate and offer value-added services to their customers.

Furthermore, governments are seeking new ways and methods to support diversification and promote SMEs in the

economy. Using data analysis and management platforms allows governments to understand the needs of their citizens better, eliminate systemic flaws, and improve operations, cutting costs and boosting the services provided by any government entity [3]. OGD can help maximize the benefits of using big data. In addition, OGD is a subset of open data and is government-related data open to the public [4]. It refers to publicly accessible information about the government [2]. Multiple datasets, including those related to finance, population, geography, the public, transportation, traffic, education, etc., may be included in government data.

The economy of Qatar is among the most robust in the region and one of the most promising in the world. Qatar maintained balanced growth rates in the face of global concerns, and its GDP increased by more than 12% in the third quarter of 2021 compared with the first. According to the policies put in place by the Qatar National Vision 2030, which aims to lay the groundwork for a competitive and diversified knowledge-based economy, Qatar has succeeded in bolstering its economic standing on the global stage over the past few years [5]. Moreover, the Qatari government promotes using Open Data to encourage openness, public participation, better governance, inclusive economic growth, and innovation in Qatar [5].

SMEs utilize OGD for different reasons. However, due to limited resources, SMEs, particularly small start-ups, need help to use big data analytics effectively [4]. SMEs throughout the Gulf Cooperation Council (GCC) and Qatar are still in the early stages of using OGD [6]. They encounter a number of difficulties when implementing OGD, including the fact that the datasets are not updated frequently and are only available in unprocessed formats, there is no interpretation or clear description for these datasets, the majority of the datasets are only available in Arabic, some GCC countries lack a clear OGD policy and classify OGD under other topics, users are discouraged from adding to the datasets, and the formats and interactive maps and usability are limited.

The current OGD platforms, generally and in GCC countries mainly, only employ data from government agencies; they do not mix data from other sources, such as third-party data and social media, with analytics. The OGD lifecycle classification was one of the most significant categories lacking. In other words, these difficulties should be categorized according to the various stages/phases of the OGD lifecycle. Furthermore, they might be classified based on the actions performed at each step of the OGD lifecycle. As a result, the decision maker or end user may effectively track, monitor, and address the difficulties.

As a result, the Qatari government needs to provide SMEs with an OGD platform to use big data analytics to achieve

more innovation and growth. The required OGD platform must give more than the available platforms and state-of-the-art currently allows. The government, companies, and the general public via social media are the most known data sources. With the help of these various data sources, SMEs will be able to capture the information needed that will aid them in analyzing trends and seeking new business opportunities [7]. Information acquired from social networks and other sources should be merged with public data.

This paper provides research into the use of OGD and the deployment of OGD in Qatari SMEs to investigate the Qatari government's and SMEs' readiness for employing OGD. Two separate surveys were undertaken to measure the amount of understanding of OGD as a concept, as well as how SMEs in Qatar perceive the present OGD platforms and how they could be improved to fit their needs.

To do this, we developed the following research questions, which will be addressed in this paper:

- What is the level of awareness in the public and private sectors about OGD?
- Are there policies and practices to encourage the use and publication of GD in Qatar?
- Is data mainly supplied directly, indirectly, or both on Qatar's Open Data Portal?

In this research, firstly, we selected different types of organizations to participate in the data collection. Organizations may be governmental, semi-governmental, or private. We assume that the selection of varying organization types may affect the data collection method and the procedure for requesting accessibility for both organization's stakeholders and data; secondly, the choice of stakeholders that will participate from these organizations. The organization's type and the stakeholder's background and experience may impact their responses, recommendations, decisions, and selections.

The rest of the paper is structured as follows. Section 2 presents the related background for OGD, its terms, OGD pros and cons. Section 3 introduces the methodology used in our study and the surveys that were undertaken. Section 4 presents the evaluation process and the analysis of the results. Finally, we give our conclusions in Section 5.

II. BACKGROUND AND RELATED WORK

A. OGD Terminology

Many concepts are related to OGD, such as open data, public data, e-government, linked data, and data portal. Firstly, there is a need to differentiate between open data, public data, and OGD.

1) Open data: The open definition states the principles and guidance that open data should conform to regarding data and content [8]. Open indicates how anyone can freely access, use, reuse, and redistribute data for any purpose regarding the requirements that preserve provenance and openness. Therefore, data is published in open data format, machine-readable, platform-independent, and open to the public without restrictions or under an open license [8].

Therefore, open data refers to data that is free of charge to the public without limitations [9]. Open data is considered a key enabler of open government [2].

2) Public data: public data is made freely available to the public but only sometimes open. An example of public data is the archive of legal documents, which are accessible freely. On the other hand, if these public data are organized in a digital format, sorted and indexed, and made available online in a standard format. This public data will be open also. Open data contains heterogeneous data from several sources. So, there is a need for a body to host these data centrally, and the government is a clear choice.

3) OGD: OGD is a subset of open data and is government-related data open to the public [2]. Government data might contain multiple datasets such as finance, population, geographical, public, transportation, traffic, education, etc. [10]. [11] defined another definition for OGD as follows: "Open data is data that anyone can access, use, or share. Simple as that. When big companies or governments release non-personal data, it enables small businesses, citizens, and researchers to develop resources which make crucial improvements to their communities."

Secondly, other terms such as e-government, linked data, and data portal are defined, which also relate to OGD.

1) E-government: there are many definitions of e-government existing in the literature; the one related to the government's use of technology is used to improve its offered services to other entities, including citizens, employees, partners, suppliers, and other government agencies [12]. Therefore, by supporting the connection between citizens and their government, e-government can develop better relationships and deliver information and services more efficiently. While initially, e-government just referred to the presence of government on the Internet as an informative website; the concept has since evolved. With the introduction of the 'open government' concept, open government data initiatives are considered a subset or an extension of e-government [13].

2) Linked data: it is the process of following a set of best practices for publishing and connecting structured data on the web [14]. Linked data refers to data that is published on the web, and it is also connected to other external datasets.

3) Data portal: the open data movement targets making data open for government and public sector information to boost its reuse. A typical implementation is to gather and publish datasets into central data portals or data catalogs to provide a "one-stop-shop" for data consumers [15]. While a data catalog would act as a registry of data sources, providing links, a data portal is more commonly a single entry point hosting the actual data, where end users can, search and access the published data and interact with it suitably [16]. One of the main functions of a data portal is the administration of metadata for the datasets, potentially including metadata harmonization. Different tools are

enabled on government data portals, for example, data format conversion, visualization, query endpoints, etc. Therefore, Open Data Portals (ODP) are essential, and the solution will provide an ODP with a data analytics framework for SMEs.

III. USE SMES MOTIVATIONS FOR THE UTILIZATION OF OGD

A. *Generated Economic Value Through Open Data*

Open data is already contributing to the economic growth of countries worldwide [23][24]. They also support creating and strengthening new markets, companies, and jobs [19]. Government plays a vital role in creating value from open data, not only in its publication stage. Organizations can create value with open data in various companies and industries in three ways [20]: for traditional companies or new non-technological startups to make decisions, in the same way as their governments use open data to improve decision-making, the same can happen for the vast private sector; to generate new products or services that create value for the clients of the companies; and to be accountable in a market where consumers require more information and reward transparent companies: by releasing data, companies can guarantee that their actions are transparent [13].

B. *Promote Greater Openness of Public Data*

One of the mechanisms available to the government to encourage the use of open data by the private sector is to strengthen the supply of these data in quantitative and qualitative terms. In South Korea, the government has promoted a series of measures to promote open data, allowing the development of many digital applications from public open data. One of the most active Open Data sites is the Seoul Open Data Plaza (data.seoul.go.kr), managed by the metropolitan government of Seoul [21]. In 2012, Seoul initiated an open data initiative sharing public information to create diverse business opportunities for the private sector and develop IT industries. This portal is an online channel to share and provide citizens with all public data of Seoul, such as real-time bus schedules, subway schedules, Wi-Fi public service places, and facilities for disabled people, among others [22], [23].

C. *Promote or Regulate the Opening of Data in Other Sectors*

Governments, international organizations, and civil associations have been at the forefront of open data proliferation and openness [18]. As governments have adopted the open data agenda, citizens and consumers demand transparency in other sectors, such as business, academia, and government organizations [24]. As noted in this document, data has become the currency of modern economics. A recent study published by the "Future of Privacy Forum" 30 projects that the global data volume will grow from approximately 0.8 zettabytes (ZB) in 2009 to more than 35 ZB in 2020 [25]. Likewise, the government can play a role in encouraging companies to share their data safely and respectfully regarding the privacy of consumers and citizens [26]. Universities and academia (in their various institutions such as science and

technology agencies) also have to take a step forward to publicize and give access to their data in different formats to other societal actors [27].

D. *Promote Data Entrepreneurship*

The recent success story of ODINE (Open Data Incubator Europe) [32] and the well-known emergence of data ventures in the United Kingdom and the United States have demonstrated the opportunity to generate value, scalability, and profits with open data ventures. In Latin America, there are also cases of successful data ventures that have received foreign investment and have grown in the last 5 years, for instance, OPI, Data4, and Atlantia [25]. Moreover, in the United Kingdom, the government has offered open government data of the highest quality through data.gov.uk. The Open Data Institute (ODI), in its Open Data Means Business research, has analyzed 270 companies in the United Kingdom that use, produce, or invest in open data as part of their business strategy. These companies (also called "open data companies") invoice more than 110 billion dollars a year and employ more than 500 thousand people [26].

IV. CHALLENGES AFFECTING SMES FULL UTILIZATION OF OGD

A set of classifications and categories for the challenges that prevent SMEs in the GCC region from utilizing the OGD effectively are discussed. Saxena [6] applied the models which have been developed to the status of OGD in the GCC countries [28]. Sieber and Johnson have introduced four models of open data that define the relationship between citizens and government (also called the Citizen engagement model of OGD) [29]; for more details about the benefits and costs of these four models (See [29]).

1) *Data over the wall* - Government publishing of open data: the government publishes open data via an online open data portal that acts as a unidirectional conduit from the data owner/collector (government, community, organization) to the end user (citizen, community organization, or private sector).

2) *Code exchange* - Government as open data activist: the government supports the reuse of open data to directly extract or create value from its offering, e.g., through app development contests.

3) *Civic issue tracker* - Data from citizen to government: the government accepts direct feedback from citizens on a limited range of issues in a crowdsourcing paradigm. Data may or may not also come from the government.

4) *Participatory open data* - Open data as open government: the government-citizen co-production of data where open data becomes a direct conduit between citizen and government, where citizen contributions are dynamic, and the government becomes responsive to demand-side requests for data.

According to [6], most GCC countries fall into the first model outlined above, "Data over the wall". Therefore, all GCC is still at an early stage or phase for utilizing and implementing OGD [6]. They face a number of challenges in

the OGD implementation, such as the datasets are not regularly updated, the available datasets are in unprocessed format, there is no interpretation or clear description for these datasets, and most of the datasets are available in the Arabic language only, some GCC countries have no clear OGD policy and classify OGD under other topics, discouraged users from contributing to the datasets, and limited formats and interactive maps and user-friendly formats.

Saxena discussed drivers and barriers to reusing OGD in Oman as one of the GCC countries [30]. A qualitative approach has been applied to the national OGD portal of Oman (<https://data.gov.om>). The national data portal of Oman is a free data-sharing portal where anyone can access data relating to the Sultanate of Oman. The data portal provides datasets from different entities for everyone - citizen, investor, researcher, or developer [31]. The national OGD portal of Oman has published over 56 data sets across 12 sectors. Moreover, there are 17 data providers or entities and three mobile apps which may be used by different users [30], [32]. Saxena concluded that Oman's OGD initiative could be classified as a hybrid of the three models [29]; Data over the wall, Code exchange, and Participatory open data [30].

Having discussed the challenges and barriers of SMEs so far, especially for GCC, various challenges/issues and classifications still need to be mentioned. For example, the OGD platforms use exclusive data from governmental entities; they do not incorporate data from different sources, such as third-party data and social media combined with analytics. Therefore, one of the essential classifications that needed to be included in the OGD lifecycle classification. Moreover, it could be categorized by the different activities in each stage of the OGD lifecycle. Therefore, the end user could track, monitor, and tackle the challenges effectively.

V. METHODOLOGY

Data collection facilitates and improves the decision-making process and the quality of those decisions. Thus, to answer our research questions related to the role of SMEs in Qatar in utilizing and spreading the use of OGD, we used a mixture of quantitative and qualitative data collection methods, such as surveys [33]. The primary objectives of the two independent surveys were to examine the level of awareness of OGD as a concept, how SMEs in Qatar view the current OGD platforms, and how OGD may be enhanced to meet the needs of SMEs.

We targeted two categories of stakeholders for the survey. The first category is the public, i.e., citizens and residents, and the second category is SMEs and Investors. So, we have designed a survey for each category as follows.

The first survey was called the "OGD Awareness Survey", which aims to evaluate the awareness of citizens and residents in Qatar of the Open Government Data. Further, the second survey was called "OGD – SMEs and Investors Survey", which aims to evaluate the awareness and utilization of OGD by SMEs and Investors in Qatar. The surveys were divided into several sections.

VI. RESULTS AND EVALUATION

A. Selection of Organizations and Stakeholders

We have selected various organizations and stakeholders participating in the data collection process. Examples of organizations are the Ministry of Interior (MOI), Ministry of Justice (MOJ), Qatar Development Bank (QDB), Qatar International Court and Dispute Resolution Centre (QICDRC), Ministry of Commerce and Industry and Hukoomi.

B. Survey Data Analysis Procedure and Used Tools

The surveys were introduced in English to manage the different terms included in the survey. Two channels for the targeted stakeholders of the first survey (OGD Awareness Survey) were used. The first channel used was a request for an official email list of consumers or end users of the Ministry of Interior (MoI) and State of Qatar services, either citizens or residents, through an SMS application. After receiving the list of four hundred emails from MoI, these emails were added to an email group and sent an email containing the survey's purpose and URL. The second channel was conducted through a series of visits to MoI service locations, connecting us with participants willing to participate in our study. Then, we met with another one hundred participants to complete the survey using an iPad. Therefore, this survey was distributed to 500 participants. The survey was designed to be online using a tool called Microsoft Forms, which we have authorized access to through Loughborough University within Microsoft Office 365. As a result, 422 responses received a return rate of 84%. The output of Microsoft Forms is a Microsoft Excel file containing participant records.

In the distribution plan of the second survey (OGD – SMEs and Investors Survey), we requested an authorized email list of SMEs from the MoI and State of Qatar services. After receiving the list of 125 emails from MoI, we added them to an email group and sent them an email containing the survey's purpose and URL. The survey was designed to be completed online using Microsoft Forms. Finally, we received 101 responses, a return rate of 81%.

Moreover, we performed an initial analysis in Microsoft Excel files to check which responses should be included in the statistical analysis. For the first survey, we found that 94 records of the total records answered "No" in the consent section, which means they did not participate in our survey. As a result, we found that 328 were valid for analysis. For the second survey, only one record of the total records answered "No" in the consent section, which means they did not participate in our survey. Thus, 100 SMEs or investors were considered valid for analysis. Finally, cleaning and transformation steps were necessary to prepare an Excel file for statistical analysis. Firstly, we changed or renamed the names of columns in the first row or header to meaningful names. Secondly, we carried out two main transformation steps in Microsoft Excel using a Power Query Editor such as the following:

- Remove a set of columns related to Microsoft Forms, such as Start time, completion time, and email.

- Replace the null values in columns with the Not Available (NA) value.

Furthermore, the transformed Microsoft Excel files for the two surveys were imported into IBM Statistical Package for the Social Sciences (SPSS) to provide further analysis of the survey results. SPSS is one of the best-known statistical techniques researchers use to deliver advanced statistical analyses, including the Chi-Square Test, which provides a reliable estimation of research results and can define the relationships between research variables [40][41].

C. OGD Awareness Survey - Data Analysis Findings

We selected all questions from Section A: Demographic Information and questions Q8, Q9, and Q10 from Section B: OGD Awareness. First, general analysis shows the number of responses that express the relationship between two variables using cross-tabulation analysis. Then, we performed an in-depth study that shows the dependency or independency between two questions or variables using the Chi-Square as statistical analysis.

TABLE I. IN-DEPTH ANALYSIS OF RELATIONSHIPS BETWEEN OGD AWARENESS SURVEY QUESTIONS

| Question | Expected Related Questions | Chi-Square | P-value | Relationship Status |
|-------------------------------|--|------------|---------|---------------------|
| Q2 - Age Group | Q8. OGD- Reaction | 39.681 | 21.026 | Dependent |
| | Q9. OGD Qatar Gov Usability | 29.023 | | Dependent |
| | Q10. OGD Qatar Third Parties Usability | 22.790 | | Dependent |
| Q3 - Gender | Q8. OGD- Reaction | 9.878 | 9.488 | Dependent |
| | Q9. OGD Qatar Gov Usability | 5.589 | | Independent |
| | Q10. OGD Qatar Third Parties Usability | 2.158 | | Independent |
| Q4 - Nationality | Q8. OGD- Reaction | 4.062 | 9.488 | Independent |
| | Q9. OGD Qatar Gov Usability | 9.595 | | Dependent |
| | Q10. OGD Qatar Third Parties Usability | 3.979 | | Independent |
| Q5 - Highest Qualification | Q8. OGD- Reaction | 38.548 | 21.026 | Dependent |
| | Q9. OGD Qatar Gov Usability | 22.812 | | Dependent |
| | Q10. OGD Qatar Third Parties Usability | 13.270 | | Independent |
| Q6 - Computer Knowledge Level | Q8. OGD- Reaction | 30.359 | 21.026 | Dependent |
| | Q9. OGD Qatar Gov Usability | 18.237 | | Independent |
| | Q10. OGD Qatar Third Parties Usability | 19.323 | | Independent |
| Q7 - Employment Situation | Q8. OGD- Reaction | 45.495 | 26.296 | Dependent |
| | Q9. OGD Qatar Gov Usability | 47.567 | | Dependent |
| | Q10. OGD Qatar Third Parties Usability | 55.700 | | Dependent |

1) We examined the relationship between Q2- Age Group and the three questions: Q8- OGD Reaction, Q9- OGD

of QatarGov Usability and Q10- OD Qatar Third Parties Usability. Figure 1 shows the responses of the Q2- Age Group to question Q9 (OGD of QatarGov Usability). Both responses from age groups range 18-30 and 30-44 indicate that they are mostly Extremely likely to use OGD. These numbers indicate that OGD from the Qatar government is trusted and could be used or utilized by youth and adults. Moreover, we performed the Chi-Square analysis using IBM SPSS between questions Q2, Q8, Q9 and Q10. For example, the Chi-Square value is 29.023 for the relation between Q2 and Q9. Moreover, the degree of freedom "df" is 12, mapped to a p-value of 21.026 with a confidence of 0.95 according to the Chi-Square distribution. Therefore, there is a dependency between the two questions because the Chi-Square value is greater than the p-value. We found that there is an indication that a high percentage of youth and adults show their interest in OGD as a positive topic in responses to Q8, and they will utilize both governmental and third-party's open data. Moreover, these findings are confirmed by performing the Chi-Square analysis between Q2 and the three questions as in Table I.

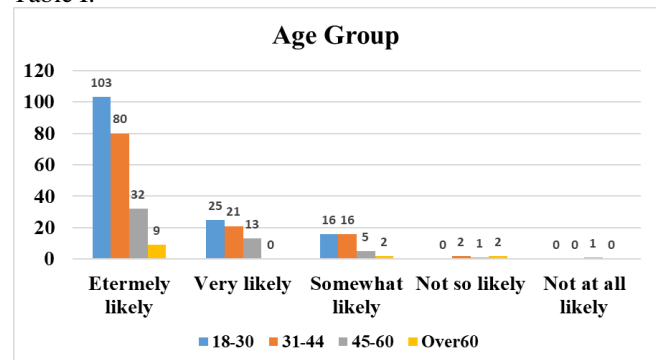


Figure 1. OGD Usability per Age Group.

2) We checked the relationship between Q3- Gender and the same three questions Q8, Q9 and Q10. Figure 2 shows the responses to question Q8. Both responses from gender, either male or female, mostly range between Very positive 117 for males and 73 for females, and somewhat positive 52 for males and 60 for females. These numbers represent a good indication that OGD is something required by both genders. Moreover, the Chi-Square value between the questions Q3- Gender and Q8, Q9 is 9.878, and the degree of freedom "df" is 4, which is mapped to a p-value of 9.488 with confidence 0.95 according to the Chi-Square distribution. Therefore, there is a dependency between the two questions because the Chi-Square value is greater than the p-value. Truly, there is an indication that a high percentage of males and females show their interest in OGD, and they may utilize both governmental and third-parties open data as in responses to OGD of QatarGov Usability and OGD QatarThirdParties Usability, respectively. Moreover, these findings are

confirmed by performing the Chi-Square analysis between Q3- Gender and the three questions, as shown in Table I.

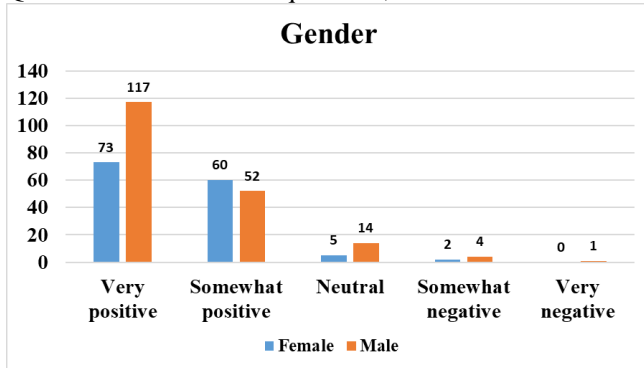


Figure 2. OGD Reaction per Gender.

3) We examined the relationship between Q4- Nationality and the three questions Q8, Q9 and Q10. Figure 3 shows the responses of Q4- Nationality to the variable. Both responses from Qatari or Non-Qatari are mostly Extremely likely to use OGD (140 Qatari and 84 Non-Qatari), and the other values have a slight difference for citizens. These numbers indicate that OGD from the Qatar government is trusted and could be used or utilized by citizens and residents. Furthermore, there is evidence of significant interest in using OGD from citizens and residents. These findings are presented in Table I.

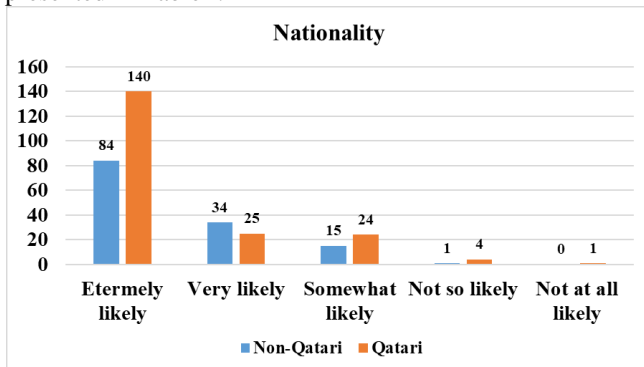


Figure 3. OGD Usability per Nationality.

4) We checked the relationship between Q5- Highest Qualification and the three Q8, Q9 and Q10. Figure 4 shows the Q5- Highest Qualification responses to the Q9- OGD of QatarGov Usability. Both responses for a Bachelor's degree or a Postgraduate degree mostly range between Extremely likely 99 for a Bachelor's degree and 87 for a Postgraduate degree, and Very likely 26 for a Bachelor's degree and 32 for a Postgraduate degree. These numbers indicate that OGD from the Qatar government is trusted and could be used or utilized by highly educated people in Qatar. Moreover, the Chi-Square value is 22.812, and the degree of freedom "df" is 12, which is mapped to a p-value of 21.026 with a confidence of 0.95 according to the Chi-Square distribution. Therefore, there is a dependency between the two variables

because the Chi-Square value is greater than the p-value, as in Table I.

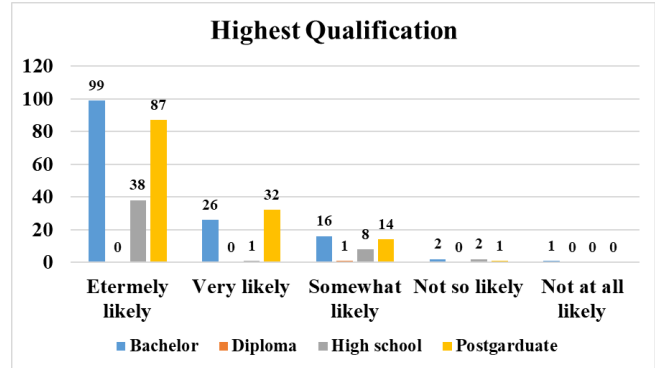


Figure 4. OGD Usability per Highest Qualification.

5) We examined the relationship between Q6- Computer Knowledge Level and questions Q8, Q9 and Q10. Both responses from Computer Knowledge Level Ranges Expert or Intermediate mostly range between Very positive 91 for Expert and 52 for Intermediate, and Somewhat positive 62 for Expert and 39 for Intermediate, as illustrated in Figure 5. These numbers represent a good indication that OGD is something required by people who have expert or intermediate computer knowledge levels. Furthermore, facilitating data access for persons with limited computer abilities is a barrier and should be addressed in the available OGD platforms. The Chi-Square value is 30.359, and the degree of freedom "df" is 12, mapped to a p-value of 21.026 (see Table I) with a confidence of 0.95 according to the Chi-Square distribution. Therefore, there is a dependency between the two variables because the Chi-Square value is greater than the p-value of expert or intermediate computer knowledge levels interested in utilizing advanced features from government and third-parties open data.

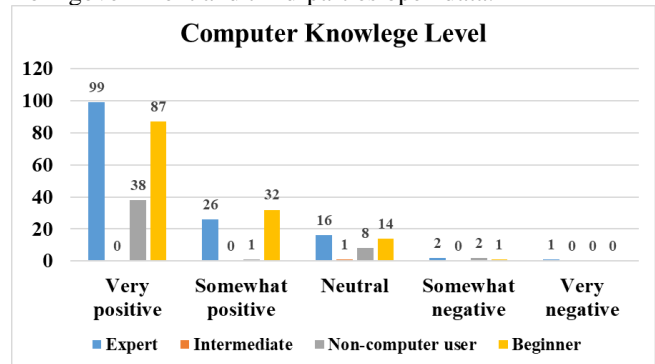


Figure 5. OGD Reaction per Computer Knowledge Level.

6) We checked the relationship between Q7- Employment Situation and the three questions Q8, Q9 and Q10. Figure 6 shows the responses of Q7- Employment Situation to question Q10 - OD QatarThirdParties Usability. Both responses from qualification ranges Expert or Intermediate mostly range between Very likely as 127 for

Working-Full-Time and 41 for in Education. These numbers indicate that open data from Qatar Third-parties are trusted and could be used or utilized by people working Full-Time or in Education. Moreover, the Chi-Square analysis between the question Q7- Employment Situation and the three questions (Q8, Q9 and Q10). For instance, the Chi-Square value for Q10 is 55.700, as shown in Table I, and the degree of freedom "df" is 16, which is mapped to a p-value of 26.296 with confidence 0.95 according to the Chi-Square distribution. Therefore, there is a dependency between the two variables because the Chi-Square value is greater than the p-value.

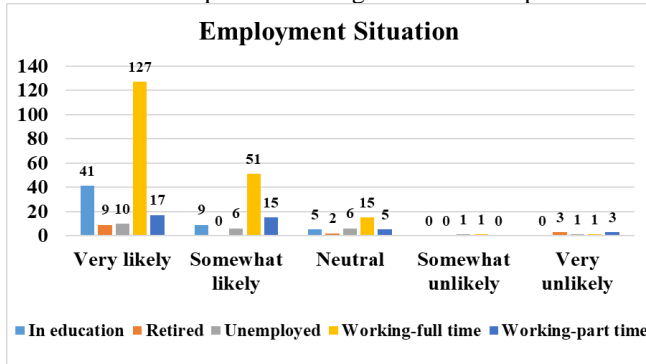


Figure 6. OGD Usability per Employment Situation.

D. OGD – SMEs and Investors Survey - Data Analysis Findings

This section discussed how we performed the statistical analysis for the second survey, " OGD – SMEs and Investors Survey," General Analysis, and In-Depth Analysis. We selected Q3, Q6, and Q7 from Section A (Demographic Information), question Q9 from Section B (OGD Awareness), Q12 and Q13 from Section C (OGD Organizational Information), Q18, Q19, Q21, and Q23 from Section D (OGD Technical Information) as in Table V. Through our analysis both general and In-depth, we used keywords of each question to express its meaning for simplicity.

TABLE II. IN-DEPTH ANALYSIS OF RELATIONSHIPS BETWEEN OGD SMEs SURVEY QUESTIONS

| Question | Expected Related Questions | Chi-Square | P-value | Relationship Status |
|----------------|--|------------|---------|---------------------|
| Q3 - Age Group | Q9. SMEs Utilization Published OGD | 2.539 | 12.592 | Independent |
| | Q12. Verify Qatar Gov OGD Policy | 4.397 | 12.592 | Independent |
| | Q13. Qatar Gov Verify OGD Access | 20.747 | 12.592 | Dependent |
| | Q18. Qatar ODP Usability Purpose | 4.392 | 16.919 | Independent |
| | Q19. Qatar ODP Registration Provided Data Method | 6.756 | 21.026 | Independent |
| | Q21. Qatar ODP Registration Dissuade Status | 4.614 | 12.592 | Independent |
| | Q23. Qatar ODP Verify Data Analytics Usability | 3.493 | 12.592 | Independent |

| | | | | |
|-------------------------------|--|--------|--------|-------------|
| Q6 – Highest Qualification | Q9. SMEs Utilization Published OGD | 10.743 | 9.488 | Dependent |
| | Q12. Verify Qatar Gov OGD Policy | 3.658 | 9.488 | Independent |
| | Q13. Qatar Gov Verify OGD Access | 10.379 | 9.488 | Dependent |
| | Q18. Qatar ODP Usability Purpose | 13.608 | 12.592 | Dependent |
| | Q19. Qatar ODP Registration Provided Data Method | 38.066 | 15.507 | Dependent |
| | Q21. Qatar ODP Registration Dissuade Status | 7.608 | 9.488 | Independent |
| | Q23. Qatar ODP Verify Data Analytics Usability | 6.925 | 9.488 | Independent |
| Q7 – Computer Knowledge Level | Q9. SMEs Utilization Published OGD | 13.132 | 9.488 | Dependent |
| | Q12. Verify Qatar Gov OGD Policy | 16.909 | 9.488 | Dependent |
| | Q13. Qatar Gov Verify OGD Access | 22.256 | 9.488 | Dependent |
| | Q18. Qatar ODP Usability Purpose | 17.619 | 12.592 | Dependent |
| | Q19. Qatar ODP Registration Provided Data Method | 31.504 | 15.507 | Dependent |
| | Q21. Qatar ODP Registration Dissuade Status | 8.729 | 9.488 | Independent |
| | Q23. Qatar ODP Verify Data Analytics Usability | 10.199 | 9.488 | Dependent |

1) We examined the relationship between Q3- Age Group and the seven questions: Q9- SMEs Utilization Published OGD, Q12- Verify QatarGov OGD Policy, Q13- QatarGov Verify OGD Access, Q18- Qatar ODP Usability Purpose, Q19- Qatar ODP Registration Provided Data Method, Q21- Qatar ODP Registration Dissuade Status and Q23- Qatar ODP Verify Data Analytics Usability. Figure 7 shows the responses of Q3- Age Group to question Q13- QatarGov Verify OGD Access. Both responses from age groups range 18-30 and 31-44 are mostly free of charge, as 37 for (18-30) and 35 for (31-44). These numbers indicate that responses from these age groups need free-of-charge access to Qatar government OGD. Moreover, the Chi-Square analysis between the question Q3- Age Group and the seven questions mentioned above are presented in Table II. For example, the Chi-Square analysis between Q3 and Q13- QatarGov Verify OGD Access is 20.747. According to the Chi-Square distribution, the degree of freedom "df" is 6, mapped to a p-value of 12.592 with confidence 0.95. Therefore, there is a dependency between the two questions as the Chi-Square value is greater than the p-value. After we performed the analyses between Q3 and these questions, we found no difference between age groups. Q9, Q12, Q13, Q18, Q19, Q21 and Q23. Furthermore, there is an indication that a high percentage of youth and adults need free-of-charge

access to Qatar government OGD as in responses to QatarGov Verify OGD Access. Thus, access to Qatar government OGD should be free of charge according to youth and adults' feedback.

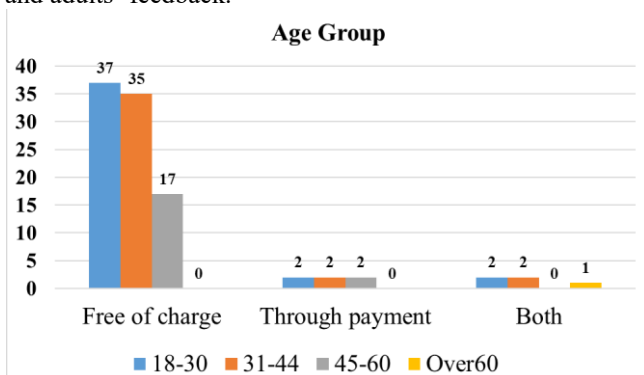


Figure 7. OGD Access per Age Group.

2) We examined the relationship between Q7- Computer Knowledge Level and the seven questions (Q9, Q12, Q13, Q18, Q19, Q21 and Q23). Figure 8 shows the responses of Q7 - Computer Knowledge Level to question Q23- Qatar ODP Verify Data Analytics Usability. Both responses from computer knowledge levels range between Expert and Intermediate or mostly between Yes 34 for Expert and 10 for Intermediate, and 33 No for Expert and 16 for Intermediate. These indicate that responses from these computer knowledge levels had no difference, which may have happened because they did not use an analytics platform. Moreover, the Chi-Square analysis using SPSS between the variables Q7 and the seven questions is illustrated in Table II. For example, the Chi-Square analysis between Q7 and Q23- Qatar ODP Verify Data Analytics Usability is 10.199, and the degree of freedom "df" is four which is mapped to a p-value of 9.488 with confidence 0.95 according to the Chi-Square distribution. Therefore, there is a dependency between the two variables because the Chi-Square value is greater than the p-value. After we performed both general and In-Depth analysis between Q7 and these questions, we found no difference between computer knowledge levels in their responses to the registration process ODP will dissuade the SMEs from utilizing the OGD.

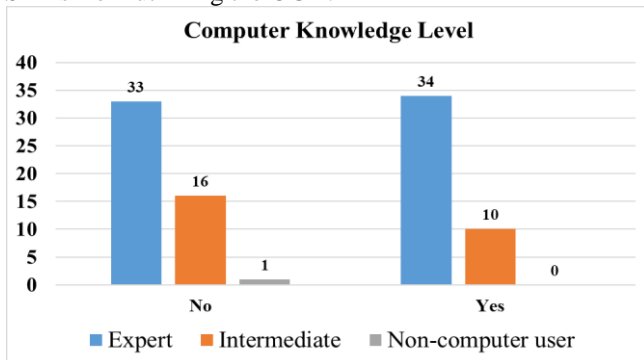


Figure 8. Data Analytics Usability per Computer Knowledge Level

3) We checked the relationship between Highest Qualification and the seven questions. Figure 9 shows the responses of Q6- Highest Qualification to question Q9- SMEs Utilization Published OGD. Both responses from the highest qualification range from a Postgraduate degree and a Bachelor's degree are between 35 Yes for Postgraduate degrees and 16 for Bachelor degrees, and 20 No for a Postgraduate degree and 21 for Bachelor's degree. These numbers indicate that SMEs' highest qualifications (Postgraduate and Bachelor degrees) utilize the published OGD. Moreover, the Chi-Square analysis using SPSS between question Q6 and the seven questions Q9, Q12, Q13, Q18, Q19, Q21, and Q23 are shown in Table II. For example, a Chi-Square analysis between Q6 and Q9 - SMEs Utilization Published OGD is 10.743, and the degree of freedom "df" is four which is mapped to a p-value of 9.488 with confidence 0.95 according to Chi-Square distribution. Therefore, there is a dependency between the two variables because the Chi-Square value is greater than the p-value. After we performed the analysis between Q6 and these questions, we found that there is an indication that a high percentage of highly educated people will utilize Published OGD, need free-of-charge access for Qatar government OGD, believe in the full or semi-full utilization of the open data portal and believe in that Qatar government OGD is provided from both direct (i.e., through ODP) and indirect (i.e., through the website of the ministry or the OGD source).

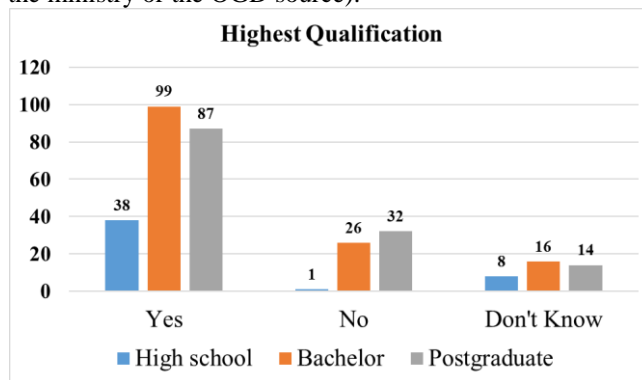


Figure 9. SMEs' Utilisation of Published OGD per Highest Qualification.

VII. CONCLUSION AND FUTURE WORK

This paper explored the existing literature regarding the definition of OGD, its history, benefits that could be brought by using it, and any drawbacks. As a result, this paper paved the way for future researchers to develop prominent data analytics theories that would benefit the Qatari economy and company. Furthermore, we have conducted intensive surveys to gather information regarding the readiness to use and enjoy the fruits of OGD. The results showed how this open data could impact society and SMEs. The analysis is concluded with a set of key findings as follows:

1) Most youth and adults in Qatar reacted positively to utilizing OGD, as seen in their responses to the OGD Reaction and how they will use governmental and third-party open data. Therefore, there is a need for a plan to increase awareness of OGD between different age groups.

2) Gender appears not to be a factor in the willingness to utilize OGD in Qatar.

3) It does not matter what nationality is (i.e., Qatari or Non-Qatari) in Qatar towards the OGD, i.e., residents and citizens show interest in OGD as a positive topic with no distinction.

4) Regardless of education, gender, and employment status, people in Qatar have a positive reaction and are keen on OGD. For this reason, the OGD portal must be simple and easy to use by all people, regardless of their education level.

5) According to the replies to Qatar Gov Verify OGD Access, many youths and adults who are SME investors or owners want free access to OGD. As a result, OGD should be made available to SMEs at an accessible cost to help the Qatari private sector, as they may not utilize it if they have to pay.

6) Owners and investors of SMEs of different ages did not differ regarding the following questions: SMEs Use Published OGD, Verify QatarGov OGD Policy, Qatar ODP Usability Purpose, Qatar ODP Registration Provided Data Method, Qatar ODP Registration Dissuade and Qatar ODP Verify Data Analytics Usability.

7) Highly educated investors and owners of SMEs will utilize Published OGD, need to have free access to Qatar government OGD, and believe in full utilization (i.e., accessed, downloaded and used) or semi-full utilization (i.e., accessed and downloaded) of the ODP, and believe in that Qatar government OGD is provided from both direct (i.e., through ODP) and indirect (i.e., through the website of the ministry or the OGD source).

8) All qualifications are the same regarding their belief that the policy of OGD and the registration process of ODP will dissuade SMEs from utilizing the OGD.

9) Many expert and intermediate computer knowledge levels who are SME investors or owners will utilize OGD. They believe that the Qatar government OGD policy should exist and need free-of-charge access for Qatar government OGD. Additionally, they believe in the open data portal's full utilization (i.e., accessed, downloaded, and used) or semi-full utilization (i.e., accessed and downloaded). Moreover, they think that the Qatar government OGD is provided directly (i.e., through ODP) and indirectly (i.e., through the ministry's website or the OGD source).

10) There is no difference between computer knowledge levels in their responses to the registration process of the ODP will dissuade the SMEs from utilizing OGD.

The Qatari government should create a national-level, centralized service where SMEs utilize big data analytics tools and examine open data supplied by the government and

others to improve business decision-making and discover new chances for expansion and innovation. We are now working completion completing the development of an OGD platform that encourages the use of OGD and overcomes the issues raised in this article. With big data analytics, the planned OGD platform would assist SMEs in achieving more innovation and growth. Additionally, it establishes a centralized, national service where SMEs may use big data analytics tools and techniques and evaluate open data provided by the government and others to enhance corporate decision-making and find new opportunities for growth and innovation. Additionally, data from social networks and outside sources will be combined with open data to give additional information to SMEs, reflecting the whole economic picture and assisting decision-makers in delivering better conclusions.

REFERENCES

- [1] M. Van Rijmenam, T. Erekhinskaya, J. Schweitzer, and M.-A. Williams, "Avoid being the Turkey: How big data analytics changes the game of strategy in times of ambiguity and uncertainty," in *Long Range Planning*, 2019, vol. 52, no. 5, p. 101841, doi: <https://doi.org/10.1016/j.lrp.2018.05.007>.
- [2] J. Kučera, D. Chlapek, and M. Nečeský, "Open Government Data Catalogs: Current Approaches and Quality Perspective," in *Technology-Enabled Innovation for Democracy, Government and Governance*, 2013, pp. 152–166.
- [3] R. Fernandez and S. Ali, "SME contributions for diversification and stability in emerging economies - an empirical study of the SME segment in the Qatar economy," *J. Contemp. Issues Bus. Gov.*, vol. 21, no. 1, pp. 23–45, Jul. 2015, [Online]. Available: <https://search.informit.org/doi/10.3316/informit.103635030363705>.
- [4] S. Akter, S. F. Wamba, A. Gunasekaran, R. Dubey, and S. J. Childe, "How to improve firm performance using big data analytics capability and business strategy alignment?," *Int. J. Prod. Econ.*, vol. 182, pp. 113–131, 2016, doi: <https://doi.org/10.1016/j.ijpe.2016.08.018>.
- [5] T. Mohammed, "Economic Policy," *Government Communication Office*, 2022. <https://www.gco.gov.qa/en/focus/economic-policy/>.
- [6] S. Saxena, "Significance of open government data in the GCC countries," *Digit. Policy, Regul. Gov.*, vol. 19, no. 3, pp. 251–263, Jan. 2017, doi: 10.1108/DPRG-02-2017-0005.
- [7] C. A. Ardagna, P. Ceravolo, and E. Damiani, "Big data analytics as-a-service: Issues and challenges," in *2016 IEEE International Conference on Big Data (Big Data)*, 2016, pp. 3638–3644, doi: 10.1109/BigData.2016.7841029.
- [8] "Open Definition," 2019. <https://opendefinition.org/> (accessed May 23, 2022).
- [9] K. J. Reiche and E. Höfig, "Implementation of Metadata Quality Metrics and Application on Public Government Data," in *IEEE 37th Annual Computer Software and*

- Applications Conference Workshops*, 2013, pp. 236–241, doi: 10.1109/COMPSACW.2013.32.
- [10] White house, “Open Government Partnership,” *Open Government Initiative*, 2011. <https://obamawhitehouse.archives.gov/open/partnership#:~:text=President Obama launched the Open,hundreds of civil-society organizations.>
- [11] ODI, “ODI,” *Open Data Institute*, 2020. <https://www.theodi.org/> (accessed Mar. 01, 2020).
- [12] K. Layne and J. Lee, “Developing fully functional E-government: A four stage model,” *Gov. Inf. Q.*, vol. 18, no. 2, pp. 122–136, 2001, doi: [https://doi.org/10.1016/S0740-624X\(01\)00066-1](https://doi.org/10.1016/S0740-624X(01)00066-1).
- [13] T. Jetzek, M. Avital, and N. Bjørn-Andersen, “Generating Sustainable Value from Open Data in a Sharing Society BT - Creating Value for All Through IT,” in *International Working Conference on Transfer and Diffusion of IT*, 2014, pp. 62–82.
- [14] C. Bizer, T. Heath, and T. Berners-Lee, “Linked data - The story so far,” *Int. J. Semant. Web Inf. Syst.*, vol. 5, no. 3, pp. 1–22, 2009, doi: 10.4018/jswis.2009081901.
- [15] M. Osorio-Sanabria, J. Brito-Carvajal, H. Astudillo, F. Amaya-Fernández, and M. González-Zabala, “Evaluating Open Government Data Programs: A Systematic Mapping Study,” in *2020 Seventh International Conference on eDemocracy & eGovernment (ICEDEG)*, 2020, pp. 157–164, doi: 10.1109/ICEDEG48599.2020.9096755.
- [16] C. Alexopoulos, L. Spiliotopoulou, and Y. Charalabidis, “Open Data Movement in Greece: A Case Study on Open Government Data Sources,” in *Proceedings of the 17th Panhellenic Conference on Informatics*, 2013, pp. 279–286, doi: 10.1145/2491845.2491876.
- [17] A. Stott, “Open data for economic growth,” 2014. [Online]. Available: <https://www.worldbank.org/content/dam/Worldbank/document/Open-Data-for-Economic-Growth.pdf>.
- [18] M. S. Altayar, “Motivations for open data adoption: An institutional theory perspective,” *Gov. Inf. Q.*, vol. 35, no. 4, pp. 633–643, 2018, doi: 10.1016/j.giq.2018.09.006.
- [19] G. Michener and O. Ritter, “COMPARING RESISTANCE TO OPEN DATA PERFORMANCE MEASUREMENT: PUBLIC EDUCATION IN BRAZIL AND THE UK,” *Public Adm.*, vol. 95, no. 1, pp. 4–21, 2017, doi: <https://doi.org/10.1111/padm.12293>.
- [20] G. V. Pereira, M. A. Macadar, E. M. Luciano, and M. G. Testa, “Delivering public value through open government data initiatives in a Smart City context,” *Inf. Syst. Front.*, vol. 19, no. 2, pp. 213–229, 2017, doi: 10.1007/s10796-016-9673-7.
- [21] I. Safarov, A. Meijer, and S. Grimmelikhuijsen, “Utilization of Open Government Data: A Systematic Literature Review of Types, Conditions, Effects and Users,” *Info. Pol.*, vol. 22, no. 1, pp. 1–24, Jan. 2017, doi: 10.3233/IP-160012.
- [22] F. Welle Donker and B. van Loenen, “How to assess the success of the open data ecosystem?,” *Int. J. Digit. Earth*, vol. 10, no. 3, pp. 284–306, Mar. 2017, doi: 10.1080/17538947.2016.1224938.
- [23] A. Zuiderwijk and C. C. Hinnant, “Open data policy-making: A review of the state-of-the-art and an emerging research agenda,” *Inf. Polity*, vol. 24, pp. 117–129, 2019, doi: 10.3233/IP-190160.
- [24] J. P. Tennant, F. Waldner, D. C. Jacques, P. Masuzzo, L. B. Collister, and C. H. J. Hartgerink, “The academic, economic and societal impacts of Open Access: an evidence-based review.,” *F1000Research*, vol. 5, pp. 1–39, 2016, doi: 10.12688/f1000research.8460.
- [25] S. S. Dawes, L. Vidasova, and O. Parkhimovich, “Planning and designing open government data programs: An ecosystem approach,” *Gov. Inf. Q.*, vol. 33, no. 1, pp. 15–27, 2016, doi: <https://doi.org/10.1016/j.giq.2016.01.003>.
- [26] J. Bates, “The strategic importance of information policy for the contemporary neoliberal state: The case of Open Government Data in the United Kingdom,” *Gov. Inf. Q.*, vol. 31, no. 3, pp. 388–395, 2014, doi: <https://doi.org/10.1016/j.giq.2014.02.009>.
- [27] A. Meijer and S. Potjer, “Citizen-generated open data: An explorative analysis of 25 cases,” *Gov. Inf. Q.*, vol. 35, no. 4, pp. 613–621, 2018, doi: <https://doi.org/10.1016/j.giq.2018.10.004>.
- [28] A. M. Alsukhayri, M. A. Aslam, I. H. Khan, R. A. Abbasi, and A. Babour, “Toward Building a Linked Open Data Cloud to Predict and Regulate Social Relations in the Saudi Society,” *IEEE Access*, vol. 10, pp. 50548–50561, 2022, doi: 10.1109/ACCESS.2022.3174090.
- [29] R. E. Sieber and P. A. Johnson, “Civic open data at a crossroads: Dominant models and current challenges,” *Gov. Inf. Q.*, vol. 32, no. 3, pp. 308–315, 2015, doi: <https://doi.org/10.1016/j.giq.2015.05.003>.
- [30] S. Saxena, “Drivers and barriers towards re-using open government data (OGD): a case study of open data initiative in Oman,” *foresight*, vol. 20, no. 2, pp. 206–218, Jan. 2018, doi: 10.1108/FS-10-2017-0060.
- [31] “National Centre for Statistics and Information - Sultanate of Oman. Sultanate of Oman - DATA PORTAL,” 2017. <https://data.gov.om/>.
- [32] “Information Technology Authority of Oman (ITA). Open Data - Omanuna Portal,” 2022. <https://www.oman.om/wps/portal/index/aboutportal>.
- [33] J. W. Creswell and J. D. Creswell, *Research Design Qualitative, Quantitative, and Mixed Methods Approaches*, 5th ed. SAGE Publications, Inc, 2018.
- [34] A. Bryman and D. Cramer, *Quantitative data analysis with IBM SPSS 17, 18 and 19: A guide for social scientists*. United Kingdom: Routledge-Cavendish/Taylor & Francis Group, 2011.
- [35] M. L. Mchugh, “The Chi-square test of independence Lessons in biostatistics,” *Biochem. Medica*, vol. 23, no. 2, pp. 143–9, 2013, [Online]. Available: <http://dx.doi.org/10.11613/BM.2013.018>.