

## RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

### **Metadata quality matters in Open Government Data (OGD) evaluation! An empirical investigation of OGD portals of the GCC constituents**

Chokki, Abiola Paterne; Alexopoulos, Charalampos; Saxena, Stuti; Frénay, Benoît; Vanderose, Benoît; Ali, Mohsan

*Published in:*

Transforming Government: People, Process and Policy

*DOI:*

[10.1108/TG-09-2022-0118](https://doi.org/10.1108/TG-09-2022-0118)

*Publication date:*

2023

*Document Version*

Peer reviewed version

[Link to publication](#)

*Citation for pulished version (HARVARD):*

Chokki, AP, Alexopoulos, C, Saxena, S, Frénay, B, Vanderose, B & Ali, M 2023, 'Metadata quality matters in Open Government Data (OGD) evaluation! An empirical investigation of OGD portals of the GCC constituents: An empirical investigation of OGD portals of the GCC constituents', *Transforming Government: People, Process and Policy*, vol. 17, no. 3, pp. 303-316. <https://doi.org/10.1108/TG-09-2022-0118>

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



**Metadata quality matters in Open Government Data (OGD) evaluation! An empirical investigation of OGD portals of the GCC constituents**

Journal:	<i>Transforming Government: People, Process and Policy</i>
Manuscript ID	TG-09-2022-0118.R1
Manuscript Type:	Research Paper
Keywords:	Open government data, GCC, Metadata, Quality

SCHOLARONE™  
Manuscripts

**Metadata quality matters in Open Government Data (OGD) evaluation! An empirical investigation of OGD portals of the GCC constituents**

**Abstract**

**Purpose:** The study seeks to investigate the quality of metadata associated with the Open Government Data (OGD) portals of the six Gulf Cooperation Council (GCC) constituents- Bahrain, Kuwait, Oman, Qatar, Kingdom of Saudi Arabia (KSA), and United Arab Emirates (UAE).

**Design/methodology/approach:** A quantitative framework, supported by extant literature, is adopted to assess the metadata quality of the six OGD portals of the GCC constituents.

**Findings:** Among the six GCC countries, Qatar has the most advanced OGD metadata quality followed by KSA, UAE, Oman, Bahrain and Kuwait. Furthermore, the OGD metadata quality of UAE and Oman stand at the same pedestal whereas Bahrain and Kuwait OGD portals are lagging behind.

**Originality:** Whereas the OGD quality has been investigated in extant literature, the metadata quality of the OGD portals for the GCC countries has not been investigated so far- the present study seeks to plug this gap.

**Keywords:** Open Government Data, OGD, Open data, Metadata, Evaluation, Gulf Cooperation Council

**1. Introduction**

Open Government Data (OGD) is the resultant of the Open Government initiatives undertaken by the governments across the globe in response to the call made by the then-US President, Barack Obama, for bringing about transparency in administration besides furthering collaboration, participation and trust of the citizens in administration (Janssen, Charalabidis and Zuiderwijk, 2012). It is equally anticipated that with the engagement of the citizens in policy-making, economy, efficiency and effectiveness of administration would also be realized (Alexopoulos, Spiliotopoulou and Charalabidis, 2013). Therefore, OGD refers to the

publicly-available datasets related with the different administrative sectors that were hitherto hidden from public view at the pretext of confidentiality and secrecy (Attard et al., 2015). OGD is license-free and is accessible in machine-friendly formats without the intervention of any specialized software platforms (Safarov, Meijer, and Grimmelikhuijsen, 2017). Thus, there are dedicated web portals which provide datasets linked with dimensions like energy, transport, climate, education, infrastructure and the like (Ubaldi, 2013). OGD is amenable to statistical interpretation and analysis and the datasets are re-used by the diverse set of stakeholders (citizens, journalists, businesses, non-profit sector, etc.) (Gonzalez-Zapata and Heeks, 2015; Wirtz, Becker and Langer, 2022) for drawing inferences thereby leading to value creation and innovation (Jetzek, Avital and Bjorn-Andersen, 2013). Implicitly, OGD initiatives lead to economic growth as well (Zeleti, Ojo and Curry, 2016). Extant research focuses on the need for ensuring that the OGD is high-quality which also implies that the metadata-information about the data itself that lends credibility to the same- associated with the same is complete, accurate and updated. In specific terms, metadata addresses the following questions: "what the data are about as well as when, where, by whom, how and why the data are collected" (Bargh et al., 2022; pp. 17). Furthermore, metadata facilitates the discoverability, searchability and usability of the datasets (Neumaier, Umbrich and Polleres, 2016).

Ample research is available on the quality status and assessment of OGD portals across the globe (for instance, Alexopoulos et al., 2018; Lourenco, 2015) however, studies focused on metadata quality assessment are few and far between (Slibar and Mu, 2022). Retaining the six countries of the GCC (Gulf Cooperation Council) as the context given the scant research on the OGD progression in these countries and authors' familiarity with the region as a whole, the present study seeks to investigate the metadata quality of the OGD portals of Bahrain, Kuwait, Oman, Qatar, Kingdom of Saudi Arabia (KSA), and United Arab Emirates (UAE). The present study addresses the call made by Charalabidis and his colleagues that further research is warranted to deepen our understanding of "OGD interoperability" with a focus on the "various OGD metadata-related subtopics: data models, schemata, taxonomies, codelists, and ontology-based extended metadata sets for OGD" (Charalabidis, Alexopoulos and Loukis, 2016: pp. 52).

The remainder of the paper is structured as follows: Section 2 presents a brief on the research on different dimensions associated with the OGD initiatives of the GCC context; Section 3 covers

research methodology adopted for the present study; Section 4 covers the results and analysis and the last two Sections cover the concluding observations and the academic and practitioner implications.

2. Related research

Thematic research veering around OGD in the GCC context is countable (Table 1). Five inferences may be drawn from this Table: first, OGD initiatives in the GCC region at a nascent stage and political will is required to institutionalize them; second, there are structural (for instance, the cultural and societal, information and communications technology (ICT) literacy, insularity of the people; risk-averse propensity of the countries, etc.) and functional constraints that deter the governments from rolling out the OGD initiatives in an efficacious manner; third, the quality of the OGD portals requires to be monitored and upgraded on a regular basis; fourth, efforts should be made by the governments to promote the OGD initiatives and the potential benefits to be reaped from the same via value generation and innovation and, finally, there should be a revamp of the government departments to facilitate inter- and intra-organizational coordination for the success of the OGD initiative at the national, regional and local levels.

Author/s (year)	Countries covered	Major insights
Al-Anazi and Chatfield (2012)	Middle East (inclusive of all the GCC constituents)	Only 3 out of the total of 13 countries investigated (UAE, Saudi Arabia and Bahrain) spearheaded OGD initiatives.
Sayogo, Pardo and Cook (2014)	35 countries including the United Arab Emirates (UAE)	UAE’s OGD portal lags behind in terms of the provision of conducting statistical analysis and user engagement facilities via the OGD portal are conspicuously missing.

Al-Kubaisi (2014)	Qatar	In terms of the citizen engagement with the national OGD portal, findings show that the engagement is relatively less.
Saxena (2016)	Oman	Infusion of Big Data analytics into OGD would be a potential advantage for the country's OGD initiative.
Al-Rushaid and Saudagar (2016)	Kingdom of Saudi Arabia (KSA)	Government has not pushed forth the OGD initiative for furthering citizen engagement in the OGD initiative and citizens are unable to appreciate the potential to be reaped from the same.
Saxena (2017a)	GCC	Given the shift from oil to non-oil economic progression, OGD initiatives might be a viable support mechanism for realising this objective.
Saxena (2017b)	GCC	GCC countries have been lagging behind in terms of their OGD initiatives and it is important to undertake drives to make them sustainable in the long run.
Saxena (2017c)	Oman	It is important for the OGD portal of Oman to be refurbished and improvised for the involvement of the stakeholders concerned.
Saxena (2017d)	GCC	The six countries of the GCC need to take up measures for the smooth execution and institutionalization of the OGD initiatives.
Tamimi, Hoshang and	United Arab Emirates (UAE)	Machine-readable formats of the datasets are not published via the OGD portals. Furthermore, in

Al-Blooshi (2017)		many instances, the API(Application Programming Interface) and guidelines are unavailable.
Saxena (2018a)	Kingdom of Saudi Arabia (KSA), Japan and The Netherlands	Given the high collectivist culture with an equally-high ranking on the “power distance” and “uncertainty avoidance” scores, Saudi Arabia has been on the backfoot in terms of publishing only the datasets that are non-controversial, insensitive and readily available to further citizen engagement.
Saxena (2018b)	Oman	The OGD portal of Oman poses difficulties for the users to engage with it given its non-user-friendly features and non-availability of the requisite datasets.
Al-Kubaisi (2018)	Qatar	It is important that the country furthers its OGD initiative with a strategic vision and direction.
Katbi & Al Ammary (2019)	Bahrain	Apart from the demand-side (i.e. the users) and supply-side (i.e. the government authorities) requirements for the publishing of datasets, it is important that the quality of OGD, inclusive of the metadata, should be maintained.
Al-Kubaisi (2019)	Qatar	The OGD initiative of Qatar is not as advanced in terms of its features and applications in contrast with that of the West-specifically the United Kingdom (UK), in this study.

Saxena (2019)	Kingdom of Saudi Arabia (KSA)	It is important that the datasets be published on a regular basis and there should be updation of the same accordingly.
Katbi (2020)	Bahrain	The government of Bahrain needs to be more proactive for ensuring that the OGD initiative is sustainable.
Al Sukhayri, et al. (2020)	Kingdom of Saudi Arabia (KSA)	There is low usability of the datasets owing to multiple formats of the datasets like Excel sheets, CSV files (Comma Separated Values), images, scanned documents and only one social media source, i.e. Twitter, is available for social media dissemination of the datasets.
Abu Samra, Mezher and Azar (2021)	United Arab Emirates (UAE)	While engaging with the OGD portal, the stakeholders concerned-case in point being the academia- encounter challenges in harnessing the portal for appropriate purposes.
Al-Hujaylan, Carr and Ryan (2022)	Kingdom of Saudi Arabia (KSA)	OGD is being re-used by only some stakeholders and not all owing to the lack of incentives and promotions by the government.
Al-Sulaimani and Ozuem (2022)	Oman	For furthering citizen trust in the OGD initiative of Oman, a collaborative landscape needs to be in place wherein the stakeholders may be co-parties for improvising the OGD value chain.

**Table 1:** Extant literature pertaining to the OGD initiatives of the GCC

### 3. Research Methodology



A quantitative research method (Bloomfield and Fisher, 2019; Queirós *et al.*, 2017; Sukamolson, 2007) is used to address the research question of this paper. This method was primarily chosen because the measure to be analyzed in this study, metadata quality, is a numerical measure. In addition, it has been widely used and proven effective in the context of open data research (Chokki *et al.*, 2022; Machova *et al.*, 2018). The research design consists of four phases: (1) selection of the portals, (2) definition of metrics used for the metadata quality assessment, (3) collection of the portal datasets, (4) calculation of portal metadata quality, and (5) comparison of metadata quality among different countries.

**Selection of the portals.** Due the infeasibility of assessing the metadata quality of every national portal in the world given the huge number of national portals in the world (Slibar and Mu, 2022), and the lack of studies of metadata quality for the Gulf Cooperation Council (GCC) member states, we choose to fill this gap by assessing the metadata quality portals of the GCC member states: Bahrain, Kuwait, Oman, Qatar, Kingdom of Saudi Arabia (KSA), and United Arab Emirates (UAE). All the GCC countries, except Kuwait, have their national open data portal. For Kuwait, we have taken into account datasets published by the Central Statistical Bureau. The list of portals to be evaluated with information about their access link are as follows:

1. Bahrain (BH) | <https://www.data.gov.bh>
2. Kuwait (KW) | <https://www.csb.gov.kw>
3. Oman (OM) | <https://data.gov.om>
4. Qatar (QA) | <https://www.data.gov.qa>
5. Kingdom of Saudi Arabia (KSA) | <https://data.gov.sa>
6. United Arab Emirates (UAE) | <https://bayanat.ae>

It may be pertinent to note that in terms of the maturity of OGD infrastructures (Alexopoulos, Diamantopoulou and Charalabidis, 2017), the OGD portals of Qatar and Kingdom of Saudi Arabia (KSA) fall in the second generation OGD maturity rubric which encapsulates traditional OGD infrastructures based on Web 2.0 with advanced information and system quality. However, the others' OGD portals are best described in line with the characteristics of the first generation OGD maturity rubric which encompasses traditional OGD infrastructures based on Web 1.0 with limited information and system quality.

**Definition of metrics.** Multiple metrics have been identified in the literature to assess the metadata quality of the portal (Bargh *et al.*, 2022; Máchová and Lněnička, 2017; Neumaier *et al.*, 2016; Slibar and Mu, 2022). After cross-referencing portal metadata metrics with those in the literature, we finally chose to use the following 9 metrics: title, description, language, theme, keywords, license, publisher, references and release date. Table 2 shows the metrics as well as the rules used to check whether a dataset satisfies a specific metric.

Metric	Required	Rule
<b>m<sub>1</sub></b> . title	Yes	check if the title of the dataset is set
<b>m<sub>2</sub></b> . description	Yes	check if the description of the dataset is set
<b>m<sub>3</sub></b> . language	Yes	check if the language of the dataset is set
<b>m<sub>4</sub></b> . theme	Yes	check if the theme or topic associated with the dataset is defined
<b>m<sub>5</sub></b> . keywords	Yes	check if the keywords to identify the dataset are provided
<b>m<sub>6</sub></b> . license	Yes	check if the license of the dataset is set
<b>m<sub>7</sub></b> . publisher	Yes	check if the publisher name or email of the dataset is set. In case an email is provided, check if it is in email format
<b>m<sub>8</sub></b> . references	Yes	check if the source link or reference name of the dataset is set
<b>m<sub>9</sub></b> . release date	Yes	check if the release date of the dataset is set

**Table 2:** List of metrics used for metadata quality assessment

**Collection of the portal datasets.** In this step, we collect all available datasets from the evaluated portals. Since not all evaluated portals had an API (application programming interface), we used two additional methods: web scraping<sup>1</sup> and manual collection of datasets, to easily collect their datasets. Since the different portals are not implemented using the same technology, a python application (using requests<sup>2</sup> and BeautifulSoup<sup>3</sup> packages) is developed for each portal that requires the use of API or web scraping to collect datasets based on the structure of the portal datasets. The source code used to collect the different datasets is available at [https://github\\_url\\_available\\_after\\_review/](https://github_url_available_after_review/).

<sup>1</sup> [https://en.wikipedia.org/wiki/Web\\_scraping](https://en.wikipedia.org/wiki/Web_scraping)

<sup>2</sup> <https://requests.readthedocs.io/en/latest/>

<sup>3</sup> <https://beautiful-soup-4.readthedocs.io/en/latest/>

**Calculation of portal metadata quality.** The metadata quality of a specific portal is represented by the mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of the metadata qualities of all datasets within the specific portal. We use the formula (1) to define the Metadata Quality (MQ) for a dataset D. Similar to (Bargh *et al.*, 2022), we assume that all metrics are equivalent and thus do not apply a weight for the value associated with each metric.

$$MQ_D = \frac{\sum_{i \leq ||M||} |m_i|}{||M||} \quad (1)$$

where  $||M||$  is the total number of metrics (equal to 9 in this study).  $|m_i|$  is the assigned value of the metric  $m_i$  for the dataset D. The value of  $|m_i|$  is equal to 1 if the rule set for the metric  $m_i$  is met, otherwise it is equal to 0.

Once the metadata quality of each dataset with a specific portal is calculated, we later use these values to calculate the mean and standard deviation in order to assess the overall metadata quality of the portal. Let  $\mu_P$  and  $\sigma_P$  be respectively the mean and standard deviation of the metadata quality for a specific portal P. Formulas (2) and (3) represent how they can be calculated.

$$\mu_P = \frac{\sum_{D \in P} MQ_D}{||D||_P} \quad (2)$$

$$\sigma_P = \sqrt{\frac{\sum_{D \in P} (MQ_D - \mu_P)^2}{||D||_P}} \quad (3)$$

where  $||D||_P$  is the total number of datasets collected from the portal P.

**Comparison of metadata quality among different countries.** We performed a T-test<sup>4</sup> (specifically the p-value generated by the T-test) in order to determine if there is a significant difference between the metadata qualities of two portals. If  $p\text{-value} \leq 0.05$ , then we conclude that there is a statically significant difference between the metadata qualities of two portals.

**Analysis of the technological portfolio**

<sup>4</sup> <https://rb.gv/cdw5so>

Finally, we proceeded with the analysis of the technological aspects of these portals in order to provide a more comprehensive view of the identified portals. The technological aspects that are examined are coming from Alexopoulos et al. (2018) and include: (a) OD platform used (in case an already made open data platform applied like CKAN or DKAN); (b) CMS: Content Management System used; (c) Analytics: whether and which analytics platform is applied; (d) Web frameworks used for the development; (e) Web servers and (f) the specific JavaScript libraries that were applied.

#### 4. Results and Analysis

Among the 6 portals evaluated, API was used to collect datasets from the Qatar (QA) portal, web scraping was used to collect datasets from the Oman (OM), Saudi Arabia (KSA) and United Arab Emirates (UAE) portals, and datasets were manually collected from the Bahrain (BH) and Kuwait (KW) portals. Table 3 presents the number of datasets collected from each portal. It also presents for each portal the percentage of datasets that meet the rules defined for each metric in Table 3. Apparently, KSA has the highest number of datasets, followed by the United Arab Emirates, Kuwait, Qatar, Bahrain, and Oman.

Table 4 presents the mean and standard deviation of metadata quality of each portal. Based on the results, it appears that the Qatar portal has the highest metadata quality ( $\mu_p=0.99$ ), followed by the Saudi Arabia portal ( $\mu_p=0.57$ ), the United Arab Emirates and Oman portals ( $\mu_p \approx 0.52$ ), the Bahrain portal ( $\mu_p=0.28$ ) and the Kuwait portal with the lowest metadata quality ( $\mu_p=0.22$ ). The higher metadata quality of the Qatar portal can be explained by the fact that the metrics were correctly set for most of their datasets, as evidenced in Table 3 where the % datasets that meet some metrics are equal to 100% or other are greater than 88%. The Kuwait portal has the lowest metadata quality, as only the "title" and "release date" metrics have been correctly set for its datasets, with the other metrics not being set.

In order to verify whether there is a statistically significant difference between the metadata qualities of any pair of portals, we have performed a T-test. Table 5 summarizes the p-value results between each pair of portals. Referring to the results presented in Table 5, we can say that

there is a statistically significant difference ( $*p \leq 0.05$ ) between each pair of portals except the pair United Arab Emirates and Oman where the  $p$ -value = 0.31.

Based on these observations, we can conclude that the Qatar portal has the best metadata quality among the 6 portals evaluated, followed by Saudi Arabia. The metadata quality of the United Arab Emirates (UAE) and Oman portals is quite similar. Bahrain and Kuwait portals have the lowest metadata quality.

Country	Number of datasets	Percent of datasets that meet the metric (%)								
		title	description	language	theme	keywords	license	publisher	references	release date
UAE	2610	99.96	99.62	0	74.79	98.39	1	3.75	0	100
BH	78	100	57.69	0	100	0	0	0	0	0
KW	289	99.65	0	0	0	0	0	0	0	100
OM	54	100	70.37	0	100	0	0	100	0	100
QA	137	100	100	100	100	100	100	98.54	88.32	100
KSA	6549	100	100	0	94.43	9.05	0	100	11.77	100

Table 3: Number of datasets and percent of datasets that meet each metric for the evaluated portals

Country	Metadata quality of portal	
	$\mu_p$	$\sigma_p$
UAE	0.53	0.06
BH	0.28	0.06
KW	0.22	0.01
OM	0.52	0.05
QA	0.99	0.04
KSA	0.57	0.05

Table 4: Mean and standard deviation metadata quality

Country	AE	BH	KW	OM	QA	SA
UAE	X	Sig* ( $p < 0.00001$ )	Sig* ( $p < 0.00001$ )	Not Sig* ( $p=0.31$ )	Sig* ( $p < 0.00001$ )	Sig* ( $p < 0.00001$ )
BH		X	Sig* ( $p < 0.00001$ )	Sig* ( $p < 0.00001$ )	Sig* ( $p < 0.00001$ )	Sig* ( $p < 0.00001$ )

			0.00001 )		0.00001 )	0.00001 )
<b>KW</b>			X	Sig*( $p < 0.00001$ )	Sig*( $p < 0.00001$ )	Sig*( $p < 0.00001$ )
<b>OM</b>				X	Sig*( $p < 0.00001$ )	Sig*( $p < 0.00001$ )
<b>QA</b>					X	Sig*( $p < 0.00001$ )
<b>KSA</b>						X

\* $p \leq 0.05$ , Sig: significant difference

**Table 5:** Statistical test results

Table 6 presents the results of the technological analysis of these open data portals. It seems that most of the portals (except Qatar) are using analytics platforms which is quite important for their improvement. Most of the open data portals in GCC countries are using a custom implementation of the portal without applying any already made open data platform. Only two of them applied already tested technologies regarding open data portal development: Qatar is using openDataSoft and Kingdom of Saudi Arabia is using CKAN. No commonalities (or standard way of development) exist in Content Management System, Web servers and Web frameworks used.

Country	Portal Access Link	Technology Portfolio
Bahrain (BH)	<a href="https://www.data.gov.bh">https://www.data.gov.bh</a>	OD platform: Custom CMS: N/A Analytics: Google Analytics Web frameworks: N/A Web servers: N/A JavaScript libraries: jQuery 1.6.2
Kuwait (KW)	<a href="https://www.csb.gov.kw">https://www.csb.gov.kw</a>	OD platform: Custom CMS: N/A Analytics: Google Analytics Web frameworks: Microsoft ASP.NET 4.0.30319

		Web servers: IIS 10.0 [+ Operating system: Windows Server] JavaScript libraries: Modernizr 2.7.1; OWL Carousel; jQuery 2.1.4
Oman (OM)	<a href="https://data.gov.om">https://data.gov.om</a>	OD platform: Custom CMS: Microsoft SharePoint 15.0.0.5109 Analytics: Errorception; Google Analytics Web frameworks: Microsoft ASP.NET Web servers: N/A JavaScript libraries: Modernizr 2.8.3; Lodash 4.17.21; jQuery UI; jQuery 1.9.1
Qatar (QA)	<a href="https://www.data.gov.qa">https://www.data.gov.qa</a>	OD platform: OpenDataSoft CMS: N/A Analytics: N/A Web frameworks: AngularJS 1.8.0 [+ Search engines: Algolia] Web servers: Nginx [+PaaS: Amazon Web Services] JavaScript libraries: Moment.js 2.22.2; Modernizr 2.8.3; jQuery 2.2.4
Kingdom of Saudi Arabia (KSA)	<a href="https://data.gov.sa">https://data.gov.sa</a>	OD platform: CKAN (Comprehensive Knowledge Archive Network) CMS: Drupal 7 Analytics: Google Analytics Web frameworks: N/A Web servers: Nginx 1.14.0 JavaScript libraries: Isotope; jQuery UI 1.12.1; jQuery 3.3.1
United Arab Emirates (UAE)	<a href="https://bayanat.ae">https://bayanat.ae</a>	OD platform: Custom CMS: Sitecore Analytics: Google Analytics

		Web frameworks: Microsoft ASP.NET
		Web servers: Nginx
		JavaScript libraries: Isotope; core-js 2.5.7; OWL
		Carousel; jQuery 1.11.1.

**Table 6:** Technical characteristics of open data portals

## 5. Conclusion

Given that the OGD phenomenon is emerging (Kalampokis, Tambouris and Tarabanis, 2011), there are a lot of considerations that need to be considered to ensure that the high-quality datasets are provisioned via the web portals. Furthermore, while factoring into account the quality of OGD, it is important that the metadata associated with the same is complete, accurate and updated (Charalabidis, Loukis and Alexopoulos, 2014). Given the fact that academic interest in the OGD initiatives is emerging over the years in tandem with the gradual progression of the OGD initiatives in the GCC region, the present study sought to contribute to the existing literature by focusing on the micro-dimension associated with the OGD phenomenon-that is, the metadata. This became pertinent because research on metadata quality is few and far between (Nogueras-Iso et al., 2021) notwithstanding the fact that metadata constitutes an intrinsic component of a dataset and without meeting the proper metadata quality standards in terms of its accuracy, completeness, coherence, consistency and updation on a regular basis, some of the other issues related with OGD re-use remain untackled, especially in terms of interoperability and further value creation and innovation.

Furthermore, there is no coordinated technological development effort in GCC countries. Every country is applying its own technologies even if there are already tailored-made solutions in the open data portals development. The same applies in the provided metadata. There is no common metadata standard that would benefit the cross-country development of added-value services.

To drive home the arguments, an empirical framework was conceptualized for quantifying the metadata characteristics and the same was applied to the OGD portals of the six GCC countries. Metadata quality was assessed in terms of metrics identified in extant research, viz., title, description, language, theme, keywords, license, publisher, references and release date. Findings from the analysis show that whereas Qatar leads in terms of the OGD metadata quality, Oman



and UAE are roughly parallel in terms of the metadata quality metrics. The metadata quality of the OGD portals of Bahrain and Kuwait need to be refurbished, however.

Findings from the metadata quality metrics in the GCC countries raise concerns: case in point being the OGD portal of Kuwait which is rudimentary and is lagging behind across all the quality metrics or the OGD portals of Bahrain and Kuwait which do not provide the publisher information. Likewise, the language of the metadata across all but Qatar are problematic and the absence of keywords in the OGD portals of Bahrain, Kuwait and Oman pose substantial bottlenecks in ensuring effective re-use of the datasets. Likewise, on the one hand, the license information is missing across Bahrain, Kuwait, Oman and Kingdom of Saudi Arabia, on the other hand, the references are missing for the OGD metadata of the portals of UAE, Bahrain, Kuwait and Oman. In line with the OGD maturity frameworks, it may be further deduced that the metadata quality of UAE, Bahrain, Kuwait and Oman is flat or closed being based on traditional OGD infrastructures whereas that of Qatar is open, contextual and detailed and is based on advanced OGD infrastructures which facilitates linking datasets (Alexopoulos, Diamantopoulou and Charalabidis, 2017). This concurs with previous findings pertaining to the OGD maturity of GCC OGD portals wherein the need to improvise the OGD portals in terms of their quality was being emphasized to further data linkage possibilities (Alromaih, Albassam and Al-Khalifa, 2016).

If these results are viewed in the context of the extant research findings on the OGD portals, the findings from the present study seek to bolster and validate the latter given that the quality of OGD and its metadata are apparently congruent in the GCC OGD portals' cases.

**6. Academic and practitioner implications**

Policy-makers and politicians need to appreciate the value of OGD initiative for ensuring that the same is institutionalized and sustainable because the potential of OGD initiative is huge. Given that the GCC countries have chalked out their “Vision” statements already, it would also be in place to allow the permeability of the digital governments via OGD initiatives. Lastly, without the integration of the organizations and departments with the involvement of all the personnel concerned, the success of the OGD initiative is a cul-de-sac, if nothing else. Therefore, apart from factoring into consideration the manpower training and development requirements, other issues pertaining to incentivization of those involved and the resource arrangement and allocation should also be taken into account. There should be a more coordinated effort to align

these open data portals so they can be used for research and development of added-value services.

The present study leaves academic insights for further research. For instance, a comparative approach may be adopted in further studies wherein the key lessons from the OGD experiences of the developed countries may be derived for being imbibed in the GCC settings. Likewise, further research is required to appreciate the manner in which the stakeholders perceive the overall utility and implications of the OGD initiative and their expectations from the same. Another line of research may relate to the short-term and long-term objectives of the OGD initiatives with a focus on the smart cities of the region. From the methodological standpoint, further studies may look into the metadata quality in terms of the spam datasets and the associated metadata (Assaf, Senart and Troncy, 2015) as also the use of Analytic Hierarchy Process (Kubler et al., 2018). Also, further research is warranted to understand how interoperability of datasets may be facilitated via the OGD portals of the GCC given the implications of data linkage for value creation, transparency and user collaboration (Aryan et al., 2014). Finally, further research is mandated to appreciate the extent to which OGD initiatives of the GCC region may be fortified for furthering value creation and innovation by the stakeholders.

References:

Abu Samra, A., Mezher, T., and Azar, E. (2021), “Public Sector Data for academic research: The case of the UAE”, In: Azar, E., and Haddad, A.N. (Eds.), “*Artificial Intelligence in the Gulf*”, Palgrave Macmillan, Singapore.

Al-Anazi, J. M. and Chatfield, A. T. (2012), "Sharing government-owned data with the public: A cross-country analysis of open data practice in the Middle East", 18th Americas Conference on Information Systems (AMCIS) 2012 Proceedings (pp. 1-10), *AIS eLibrary*, <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1150&context=amcis2012>.

Alromaih, N., Albassam, H. and Al-Khalifa, H. (2016), "A proposed checklist for the technical maturity of open government data: An application on GCC countries", *Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services*, pp. 494-499.

Assaf, A., Senart, A. and Troncy, R. (2015), "Roomba: Automatic validation, correction, and generation of dataset metadata enhancing dataset search and spam detection," in *WWW 2015 Companion - Proceedings of the 24th International Conference on World Wide Web*, pp. 159-162.

Alexopoulos, C., Diamantopoulou, V., and Charalabidis, Y. (2017), "Tracking the evolution of OGD portals: A maturity model", *Electronic Government (EGOV 2017), Lecture Notes in Computer Science*, Vol. 10428. Springer, Cham. [https://doi.org/10.1007/978-3-319-64677-0\\_24](https://doi.org/10.1007/978-3-319-64677-0_24).

Alexopoulos, C., Loukis, E., Mouzakis, S., Petychakis, M., and Charalabidis, Y. (2018), "Analysing the characteristics of open government data sources in Greece", *Journal of Knowledge Economy*, Vol. 9, pp. 721-753.

Alexopoulos, C., Spiliotopoulou, L. and Charalabidis, Y. (2013), "Open data movement in Greece: A case study on open government sources", *Proceedings of the 17th Panhellenic Conference on Informatics*, pp. 279-286.

Al-Hujaylan, A., Carr, L., and Ryan, M. (2022), “Towards an integrated conceptual model for Open Government Data in Saudi Arabia”, In: Nagar, A.K., Jat, D.S., Marin-Raventos, G., & Mishra, D.K. (Eds.), “*Intelligent Sustainable Systems*”, Lecture Notes in Networks and Systems, 334, Springer, Singapore.

Al-Kubaisi, A.S. (2014), “Improving the transparency, openness and efficiency of e-government in Qatar in the era of Open Government Data, and beyond”, *PhD thesis submitted to the Faculty*

of Law, Queensland University of Technology, available at [https://eprints.qut.edu.au/75540/2/Ali\\_Al-Kubaisi\\_Thesis.pdf](https://eprints.qut.edu.au/75540/2/Ali_Al-Kubaisi_Thesis.pdf), accessed on 20th August, 2022.

Al-Kubaisi, A.S. (2018), "Enhancing the adoption of e-government systems through Open Government and Open Government Data (OGD) initiatives in Qatar", *Journal of Information Engineering and Applications*, Vol. 8 No. 1. <https://www.iiste.org/Journals/index.php/JIEA/article/view/40706>.

Al-Kubaisi, A.S. (2019), "An overview of open government data in developed countries: A comparison study between the state of Qatar & UK", *Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)*, No. 25, available at [https://qspace.qu.edu.qa/bitstream/handle/10576/12493/An\\_overview\\_of\\_open\\_government\\_data\\_in\\_developed\\_countries.pdf?sequence=1&isAllowed=y](https://qspace.qu.edu.qa/bitstream/handle/10576/12493/An_overview_of_open_government_data_in_developed_countries.pdf?sequence=1&isAllowed=y), accessed on 25 August, 2022.

Al-Rushaid, M.W. and Saudagar, A.K.J. (2016), "Measuring the data openness for the Open Data in Saudi Arabia e-government-A case study", *International Journal of Advanced Computer Science and Applications*, Vol. 7 No. 12, pp. 113-122.

Al-Sukhayri, A.M., Aslam, M.A., Saeedi, K., and Malik, M.S.A. (2020), "A linked open data-oriented sustainable system for transparency and open access to government data: A case study of the public's response to women's driving in Saudi Arabia", *Sustainability*, Vol. 12 No. 20, 8608. Available at <https://doi.org/10.3390/su12208608>.

Al-Sulaimani, A.H.A. and Ozuem, W. (2022), "Understanding the role of transparency, participation, and collaboration for achieving open digital government goals in Oman", *Transforming Government: People, Process and Policy*, In press, <https://doi.org/10.1108/TG-04-2022-0044>.

Aryan, P.R., Ekaputra, F.J., Sunindyo, W.D. and Akbar, S. (2014), "Fostering government transparency and public participation through linked open government data: Case study: Indonesian public information service", *International Conference on Data and Software Engineering (ICODSE)*, 2014, pp. 1-6.

Attard, J., Orlandi, F., Scerri, S., and Auer, S. (2015), "A systematic review of open government data initiatives", *Government Information Quarterly*, Vol. 32 No. 4, pp. 399-418.

Bargh, M.S., Choenni, S., Meijer, R. and Choenni, S. (2022), "A method for assessing the degree of openness of Semi-Open Data initiatives: applied to the justice domain", *International Journal of Electronic Governance*, Vol. 14 No. 1-2, pp. 207-235.

Bloomfield, J. and Fisher, M.J. (2019), “Quantitative research design”, *Journal of the Australasian Rehabilitation Nurses Association*, Vol. 22 No. 2, pp. 27-30.

Charalabidis, Y., Alexopoulos, C. and Loukis, E. (2016), "A taxonomy of open government data research areas and topics", *Journal of Organizational Computing and Electronic Commerce*, Vol. 26 No. 1-2, pp. 41-63.

Charalabidis, Y., Loukis, E. and Alexopoulos, C. (2014), "Evaluating second generation open government data infrastructures using value models," *47th Hawaii International Conference on System Sciences*, pp. 2114-2126.

Chokki, A.P., Simonofski, A., Frénay, B. and Vanderose, B. (2022), “Open government data awareness: eliciting citizens’ requirements for application design”, *Transforming Government: People, Process and Policy*, available at:<https://doi.org/10.1108/TG-04-2022-0057>.

Gonzalez-Zapata, F., and Heeks, R. (2015), “The multiple meanings of open government data: Understanding different stakeholders and their perspectives”, *Government Information Quarterly*, Vol. 32 No. 4, pp. 441-452.

Janssen, M., Charalabidis, Y. and Zuiderwijk, A. (2012), "Benefits, adoption barriers and myths of open data and open government", *Information Systems Management*, Vol. 29 No. 4, pp. 258-268.

Jetzek, T., Avital, M. and Bjorn-Andersen, N. (2013), "The generative mechanisms of open government data", In: *Proceedings of the 21st European Conference on Information Systems (ECIS)*, No. 156, Utrecht, The Netherlands.

Kalampokis, E., Tambouris, E., and Tarabanis, K. (2011), “Open government data: A stage model”, Berlin Heidelberg, Springer.

Katbi, A. (2020), “The demand side of Open Government Data: A case study of Kingdom of Bahrain”, *International Journal of Managing Information Technology (IJMIT)*, Vol. 12 No. 2. Available at <https://ssrn.com/abstract=3635928>, accessed on 25 August, 2022.

Kubler, S., Robert, J., Neumaier, S., Umbrich, J., and Le Traon, Y. (2018), "Comparison of metadata quality in open data portals using the Analytic Hierarchy Process," *Government Information Quarterly*, Vol. 35 No. 1, pp. 13-29.

Lourenco, R.P. (2015), "An analysis of open government portals: A perspective of transparency for accountability", *Government Information Quarterly*, Vol. 32 No. 3, pp. 323-332.

Máchová, R., Hub, M. and Lnenicka, M. (2018), "Usability evaluation of open data portals: Evaluating data discoverability, accessibility, and reusability from a stakeholders' perspective", *Aslib Journal of Information Management*, Vol. 70 No. 2, pp. 252-268.

Máchová, R. and Lněnička, M. (2017), "Evaluating the quality of open data portals on the national level", *Journal of Theoretical and Applied Electronic Commerce Research*, Vol. 12 No. 1, pp. 21-41.

Neumaier, S., Umbrich, J. and Polleres, A. (2016), "Automated quality assessment of metadata across open data portals", *Journal of Data and Information Quality*, Vol. 8 No. 1, p. 1-29.

Nogueras-Iso, J., Lacasta, J., Urena-Camara, M.A. and Ariza-Lopez, F.J. (2021), "Quality of Metadata in Open Data Portals," *IEEE Access*, Vol. 9, pp. 60364-60382.

Queirós, A., Faria, D. and Almeida, F. (2017), "Strengths and limitations of qualitative and quantitative research methods", *European Journal of Education Studies*, Vol. 3 No. 9, pp. 369-387.

Safarov, I., Meijer, A., and Grimmelikhuijsen, S. (2017), "Utilization of open government data: A systematic literature review of types, conditions, effects and users", *Information Polity*, Vol. 22, pp. 1-24.

Saxena, S. (2016), "Integrating open and big data via 'e-Oman': Prospects and issues", *Contemporary Arab Affairs*, Vol. 9 No. 4, pp. 607-621.

Saxena, S. (2017a), "Prospects of open government data (OGD) in facilitating the economic diversification of GCC region", *Information and Learning Sciences*, Vol. 118 No. 5-6, pp. 214-234.

Saxena, S. (2017b), "Open public data (OPD) and the Gulf Cooperation Council (GCC): Challenges and prospects", *Contemporary Arab Affairs*, Vol. 10 No. 2, pp. 228-240.

Saxena, S. (2017c), "Nature and scope of Open Government Data (OGD) in Oman", *Journal of Social and Administrative Sciences*, Vol. 4 No. 2, pp. 199-204.

Saxena, S. (2017d), "Significance of open government data in the GCC countries", *Digital Policy, Regulation and Governance*, Vol. 19 No. 3, pp. 251-263.

Saxena, S. (2018a), “National open data frames across Japan, The Netherlands and Saudi Arabia: Role of culture”, *Foresight*, Vol. 20 No. 1, pp. 123-134.

Saxena, S. (2018b), “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.

Saxena, S. (2019), “Evaluation of the National Open Government Data (OGD) Portal of Saudi Arabia”, In: Visvizi, A. and Lytras, M.D. (Ed.), “Politics and Technology in the Post-Truth Era”, *Emerald Studies in Politics and Technology*, Emerald Publishing Limited, Bingley, pp. 221-235.

Sayogo, D.S., Pardo, T.A. and Cook, M. (2014), “A framework for benchmarking Open Government Data efforts”, *47th Hawaii International Conference on System Sciences*, pp. 1896-1905, available at <https://ieeexplore.ieee.org/abstract/document/6758838>, accessed on 4 September, 2022.

Slibar, B. and Mu, E. (2022), “OGD metadata country portal publishing guidelines compliance : A multi-case study search for completeness and consistency”, *Government Information Quarterly*, <https://doi.org/10.1016/j.giq.2022.101756>.

Sukamolson, S. (2007), “Fundamentals of quantitative research”, *Language Institute Chulalongkorn University*, Vol. 1 No. 3, pp. 1-20.

Tamimi, H., Hoshang, S.A. and Al -Blooshi, E.J. (2017), “Analysis of UAE open government data usability within mobile application development”, *IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, pp. 437-441.

Ubaldi, B. (2013), "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives", *OECD Working Papers on Public Governance*, No. 22, OECD Publishing, Paris.

Wirtz, B.W., Becker, M., and Langer, P.F. (2022), “An integrated model of digital open government”, *International Journal of Public Administration*, <https://doi.org/10.1080/01900>.

Zeleti, F.A., Ojo, A., and Curry, E. (2016), "Exploring the economic value of open government data", *Government Information Quarterly*, Vol. 33 No. 3, pp. 535-551.