IMPLEMENTATION OF OPEN DATA
IN HIGHER EDUCATION: A REVIEW

SITI NURASYIQIN ISMAEL[1],*, OTHMAN MOHD[2], YAHAYA ABD RAHIM[2]

[1]Faculty of Information & Communication Technology,
Universiti Teknikal Malaysia Melaka
[2]Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
*Corresponding Author: syiqin08@gmail.com

Abstract

The phenomenon of having an Open Data initiatives becomes pervasive in every country. In Malaysia, it started with Open Government Data Framework approach from The Malaysian Administrative Modernization and Management Planning Unit (MAMPU) and currently the initiative has been extended to the ministries and agencies level. In addition, it is important to implement Higher Education Open Data Framework as the agency plays a major role as contributors to support the needs of citizen in current and future education world. Every country has its own governance structure and open data policies. In implementing Open Data Initiatives, there are some rules need to be followed; principles and models. Three (3) countries have been reviewed relating to Higher Education Open Data which are United Kingdom (UK), Saudi Arabia and South Africa. These countries are chosen based on its country classification. UK presented as a developed country. Meanwhile South Africa and Saudi Arabia acted as developed country in which the latter one is more significant to Malaysia that falls under sub category of Asia. This paper discussed about the principles, models and themes used in determining data openness for selected Higher Education Countries; United Kingdom, South Africa and Saudi Arabia.

Keywords: Data openness, Higher education, Open data, Open data developing country, Ten principles.

1. Introduction

An open data phenomenon is expected to encourage public engagement with the government. According to Reale [1], the aim of Open Data Implementation is to access the data produced by the government through an open format datasets form. Tran and Scholtes [2] stated that the implementation of Open Data would address some existing legal challenges such as the scope of the access to information and the data ownership. The researcher is advised to be well conceptualized and understand on how the data is collected, why the data is collected and who will benefit from Open Data in order to ensure the data sharing is efficient and accurate. Open Knowledge International [3] identified that open data is the data that freely can be used, redistributed and reused by anyone. In addition, the characteristics of Open data are easily to access and always be available, can be reused, redistributed and involve universal participation.

The Availability and Access means that the data must be available and can be downloaded via internet. Readiness of the data should be expedient and in amendable form. The Re-use and Redistribution is referred to licensing term that must be cleared where the data can be re-use, redistribution and intermixing with other data. While the Universal Participation means that there should not be any restriction or any discrimination towards the group, field endeavour or any domain (e.g., only can be used in education cluster) [3].

Furthermore, the meaning of Open Data is well-defined as machine-readable information, predominantly the government data that can be shown, available and used to the others [4]. This new concept has been introduced and has create a phenomenon around the world for its potential capability to improve delivery of public service, encourage the participation of the citizens towards having a transparent government [4].

Since every country has its own governance structure and policies on open data [5], thus this paper categorized the factors/criteria that have influenced the openness of the data based selected Higher Education countries.

2. Related Works

Referring to the reports of open data barometer in year 2015 [6], Malaysia has ranked 41 upon 86 countries that towards the open data initiatives. Clustered into a group of one sided initiatives together with UAE and Ukraine. Based on the report shows that Malaysia is the weakest fit in this cluster as the researchers note that there is very little outreach in engaging users with the data [6].

According to Higher Education Commission [7], there is a high demand on open data from public institutions. It was suggested that the open data agenda for Higher Education (HE) sector is observed by the institutions and government.

In the Open Government Data Initiatives there are several challenges that have been identified. One of the challenges is converting raw data into useful representations in transforming passive data sets towards open data culture [7].

Three (3) terms has been introduced in Open Data Framework; civic start-ups, open data services and infomediaries. The civic startup is focused on developing civic-minded digital solutions that implies between government and the citizens. The open data services are included on top of open data sets released by

government. While, infomediaries reflects to the concept of the consumers who take the ownership of information about themselves [8].

At national government level, Open Government Data (OGD) is one of the effort introduced. Some observations in terms of policy and practical suggestion on OGD portal development has been prepared. This policy let the OGD publishes all data collected by the government agencies. However, in term of privacy, confidentiality and security principles, the OGD can only publishes the data that is agreed by law, in addition, accessing to OGD can be done from single data portal where according to Sayogo et al. [9], ten (10) principles have been introduced by Sunlight foundation's open government working group as shown in Fig. 1.

Principle	Indicator
Completeness	• Data set should be as complete as possible and includes metadata. • Datasets do not compromise with federal laws on privacy, security of privilege.
Primacy	• Datasets should be primary source and include original information. • Includes details of collection process.
Timeliness	• Available in timely fashion and be released as quickly as possible. • Release priority depends on the time sensitivity of data.
Ease of Physical and Electronic Access	• Datasets should be accessible and easy to obtained by any means. • Pay attention to the "findability" of datasets.
Machine Readability	• Information should be stored in widely-used file formats and that enable machine processing. • Documentation related to the format and how to use it is provided.
Non-discrimination	• Any person who can access the data at any time and without having to identify themselves or provide any justification for doing so.
Use of Commonly Owned Standards	• Used freely available alternative formats • There is no software license needed for accessed and used.
Licensing	• Clear labeling public information as a work of the government. • No subject: copyright, patent, trademark or trade secret regulation.
Permanence	• Online stickiness – information is available online in perpetuity.
Usage Costs	• No cost for accessing the data, or cost is as minimum as possible.

Fig. 1. Ten principles of OGD introduced by
Sunlight foundation's open government working group [9].

The rapid changes of information and communication technology nowadays required the transformation in working culture of public administration. A closed and self-referential system that was introduced before has to be an open data system. It able to adapt the bottom-up requests of transparency, participation and collaboration. This transformation is required in order to change the concept of top down approach in open data implementation that leads to the effectiveness of Open data concept [1].

Jannsen et al. [10] stated that open data concept will exclude the data with the type of private, confidential and classified as the data is inappropriate to publicize.

According to Malaysian Administrative Modernisation and Management Planning Unit [11], in Malaysia Government document, there are five (5) types of data has been categorized; Top Secret *(Rahsia Besar),* Secret (*Rahsia*), Confidential (*Sulit*), Limited (*Terhad*) and Open (*Terbuka*). The licensing terms are introduced as follows: -

- Data's owner gives the permission of using the data without any charge.
- The data can be copied, used, reused, redistributed, commercialized and combined the data in a product or applications.
- If number (2) applies, user need to cite the resources, give a link to licensing terms, give a consent of data intellect, exceptional data such as personal data, name logo and official stamp of data's owner.
- No warranty on readiness of the data is coming from the data's owner.
- Law in which the data's owner is the primary owner of respective data.

The impacts of having higher education open data framework that can benefit economy misleading and accountability of universities operations. The data shared will be very valuable in producing innovations to be a better university in the future [12].

3. Open Data in Higher Education

Higher Education Commission [7] acted as centralized body for data collection, and other organizations that need HE data will be able to obtain the necessary data from Higher Education Statistics Agency (HESA).

According to Manyika et al. [4] higher education performance and requirements data that have been categorized as Open might benefit the community especially school teachers, students and parents in order to ensure that they can make a wise decision when it comes to their future in term of having an updated choices of schools and academic or vocational concentrations.

University benchmarking is practiced to measure a wide range of aspects especially in term of how benchmarking is gathering, the availability and update frequency of the data [13].

3.1. Malaysia Higher Education Institution

In Malaysia, higher education institution is defined as Public Higher Education or Private Higher Education that registered with Ministry of Higher Education (MOHE) [14].

The higher education sector is responsible for the operation of Higher Education Institutions (HEIs) in Malaysia and is under the authority of MOHE. Malaysia's HEIs can be categorized as Public Universities, Private Higher Educational Institutions, Polytechnics and Community Colleges [15].

As for the current state of Malaysia's Open Data in Higher Education, there is no framework regarding this matter yet. Malaysia had launched an Open Data initiative in public sector in year 2014 by MAMPU in collaboration with Malaysia Digital Economy Corporation (MDEC). Hence, an Open Data framework for

public sector had been developed ever since where it only involves government agency, public and business community.

Selection of these 3 HE countries (UK, Saudi Arabia and South Africa) are based on the country classification by World Economic Situation and Prospects report (WESP). DJIA [16] has categorized UK as developed country where South Africa and Saudi Arabia has fallen into a group of developing countries in which the latter one specifically belongs to the same region Asia as Malaysia.

3.2. Open Data in UK Higher Education

According to Universities UK [17], creating value from Open Data is a project led by Universities UK in conjunction with Open Data Institute (ODI) to gain a value of UK Higher Education and promote an understanding on how the universities will get benefit of it.

Openness of the data

The first step of Data Openness is having identified the dataset to be released. This step delivers a questions like how the data is collected, who owns the data and more. By answering all these questions, it will help the universities to develop an organizational strategy for Open Data [17].

ODI has developed two (2) online tools on how to have a strategic approach towards the implementation of Open Data. The tools known as Open Data Certificate (ODC) and Open Data Pathway (ODP). The ODC tools is used to assess the data to be published for the purpose of re-use by third parties. The ODP tools used to assess the overall open data practice across a number of themes. The latter tool is developed based on Open Data maturity model. The results can be used to create a strategic plan that can guide the development of open data practice in the organization. Fig. 2 shows a themes component to be assessed in Open Data Maturity Model [17].

Themes	Description
Data management processes	Identifies the key business processes that underpin data management and publication including quality control, publication workflows and adoption of technical standards
Knowledge and skills	Highlights the steps required to create a culture of open data within an organization by identifying the knowledge sharing, training and learning required to embed an understanding of the benefits of open data
Customer support and engagement	Addresses the need for an organization to engage with both of their data sources and their data re-users to provide sufficient support and feedback to make open data successful
Investment and financial performance	Covers the need for organizations to have insight into the value of their datasets and the appropriate budgetary and financial oversight required to support their publication. In term of data consumption, organizations will need to understand the costs and value associated with their re-use of third-party datasets
Strategic oversight	Highlights the need for an organization to have a clear strategy around data sharing and re-use and an unidentified leadership with responsibility and capacity to deliver that strategy

Fig. 2. Themes as a component to be assessed in open data maturity model [17].

3.3. South Africa Higher Education Open Data Framework

In South Africa, the Department of Higher Education and Training (DHET) is the government department that responsible for public higher education sector and also is responsible to integrate the data as well as to develop responsive data driven intelligence system. Centre for Higher Education Transformation's (CHET) Open Data Portal has been developed to supply the data from DHET for public perusal. CHET drew on data from Higher Education Management Information System (HEMIS). CHET acted as intermediaries between DHET's HEMIS database and the end-users [18].

Openness of the data

Van Schalkwyk et al. [18] has stated that the first step in identifying data to be categorized as open is merely depend on these two (2) interest groups which are Open Government Data Principles which was developed by a working group of Open Government Advocates and second set principles by Open Data in Developing Countries (ODDC).

Each dataset will be assessed according to ten (10) criteria by ODDC and eight (8) criteria by OGD. If the datasets met about 80% of these criteria using each of the evaluator frameworks, then the dataset is considered to be opened [18].

According to the ODDC, many datasets failed these criteria; machine readability, data availability, licensing is not clearly defined and no linked data Uniform Resource Identifiers (URIs). However, CHET meet 80% of ODDC criteria that only not compromise on linked data URIs criterion [17].

In compared to OGD criteria that focuses more to transparency and accountability, three (3) criterions were not met by the dataset which are machine readability, unclear licensing of data and non-discriminatory access to data.

A quick glance on evaluation of Openness based on two (2) datasets as shown in Tables 1 and 2.

From this exercise, it was suggested that South Africa Higher Education Open Data have an additional effort to in line with protocol and open standards.

Table 1. Evaluation of openness of two datasets using the Exploring the emerging Impacts of ODDC 10-point evaluation [17].

ODDC Open Data Criteria	DHET	CHET
Does the data exist?	√	√
Is it available online in digital form?	√	√
Is the data machine readable?	√	√
Is the data available in bulk?	√	√
Is the dataset available free of charge?	√	√
Is the data openly licensed?	x	√
Is the dataset up-to-date?	√	√
Is the publication of the dataset sustainable?	√	√
Was it easy to find information on the dataset?	x	√
Are linked data URIs provided?	x	x
TOTAL SCORE	7	9

**Table 2. Evaluation of openness of two
datasets using the 8 Principles of OGD [17].**

Open Government Data Criteria	DHET	CHET
Data must be complete	√	x
Data must be primary	√	x
Data must be timely	√	x
Data must be accessible	√	√
Data must be machine process-able	√	√
Access must be non-discriminatory	√	√
Data formats must be non-proprietary	x	√
Data must be license-free	x	√
TOTAL SCORE	6	5

3.4. Saudi Arabia Higher Education Open Data

According to AlRushaid et al. [5], Saudi Arabia has million records of data that has been collected and need to turn it into information to be useful to others. Open data initiatives always focused on the accessibility and the technology without giving a glance from user perspectives. Citizen awareness is essential to be measured because without citizen engagement, the initiatives will be impractical.

Openness of the data

There are some of evaluations' assessment need to be done in order to measure the openness level of the data. A five-stage model has been used to measure the data availability. The maximum number of score is 4 when the data fulfilled all the requirement. In measuring the extension of public data availability; Sir-Berners-Lee proposed a star rating. The maximum number of score is 5 when the data fulfilled all the requirements [5].

However, it is not adequate to just measure the data availability. There are a lot of indicators need to be measured which can't be measured by the above models. Table 3 shows a models used for measuring OGD [5].

Table 3. Models used for measuring OGD [5].

Model's Name	Indicators/Aspects to be measured
Four-Stage Model of Data Availability (David Osimo,2008)	√ Availability
Five-Star Model of Data Availability (Sir-Berners-Lee,2010)	√ Availability
(European Commission Model,2011)	√ Number of Open Datasets Available √ Timeliness √ Reuse Conditions √ Pricing √ Accessibility √ Take-Up by App Developers √ Number of Application Developed Based On Open Data
Open Data Benchmark (Socrata,2011)	√ Accessibility √ Availability
Scoring Model by Open Knowledge Foundation	√ Data Exist √ Data in Digital Format √ Publicly Available √ Free of Charge √ Available Online √ Openly Licensed √ Machine-Readable Available in Bulk √ Updated

Scoring model contains a well-defines open government data principles. In fact, as of now, this model has been used to measure openness level of 22 government portal around the world.

However, according to Open government maturity model that been used to measure the citizen empowerment, collaboration and information sharing, its clearly shows that scoring model has only reflected to the first and second stage of data maturity models. The scoring model and 5 level of data maturity as Fig. 3.
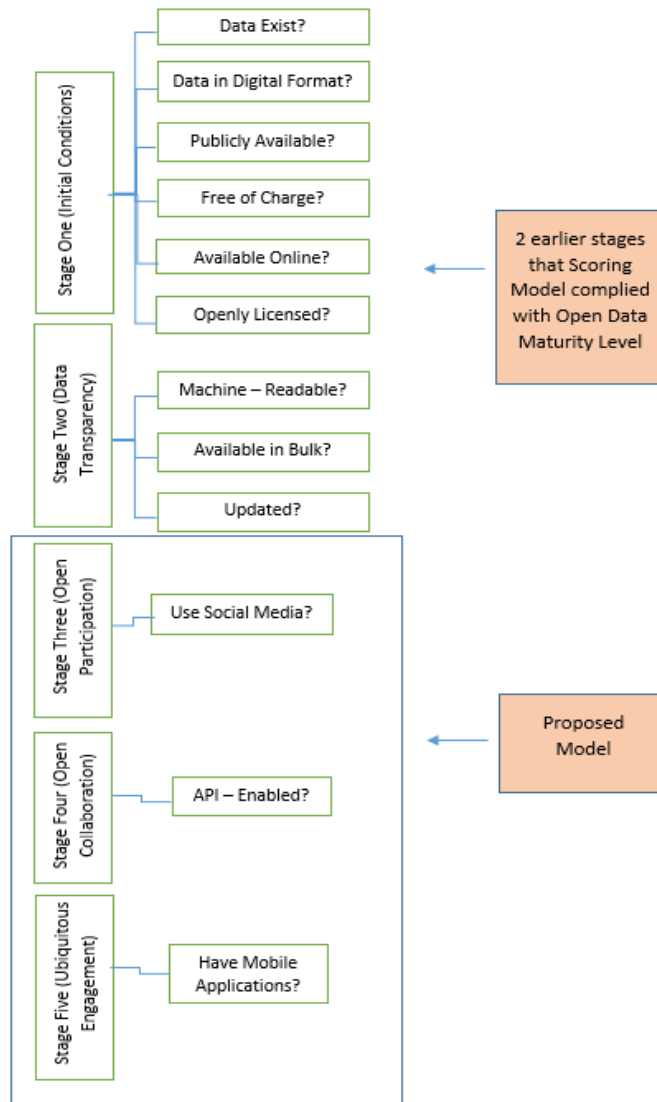
Fig. 3. The scoring model and 5 level of data maturity [5].

By having this proposed model from [5], the level of portal openness is evaluated. The scoring model indicators implies the Data Exist, Data in Digital Format, Publicly Available, Free of charge, Available Online, Openly Licensed,

Machine-Readable, Available in Bulk, Updated, Use Social Media tools, API-Enabled and have Mobile Applications; These indicators matched with both OGD Maturity Model and OGD Principles that are aligned with MAMPU framework.

4. Data Openness Approach in Higher Education

Based on these three (3) (UK, South Africa and Saudi Arabia) HE openness of data, it can be seen that they are using three (3) different approach as such UK use a tools like ODC and ODP in selecting a dataset and also in creating a strategic plan that can guide the development of open data practice in the organization. This tool is based on Open Data Maturity Model [5, 7, 18].

Meanwhile, South Africa use a different approach of Openness in selecting datasets as the categorization is depend on two (2) interest group which are ODDC and OGD. If 80% of these two (2) criterions respectively were met, then only the dataset can be declared as opened [18].

Saudi Arabia do a measurement level of openness based by scoring and open government data maturity models. This two (2) models are derived from OGD with nine (9) principles introduced. However, the addition of several elements like collaboration, participations and citizen engagement has led to the introduction of new proposed model as the current model on satisfied the earlier 2 stages of open government maturity model. This readiness and effectiveness of this new model has been evaluated according to the four-teen (14) experts from 14 different ODI [5].

While in Malaysia Government, the document has been governed to the licensing terms by MAMPU, where it was categorized into five (5) types of data that are Top Secret (*Rahsia Besar*), Secret (*Rahsia*), Confidential (*Sulit*), Limited (*Terhad*) and Open (*Terbuka*) [11].

In implementing HE Open Data, there are challenges need to addressed that included the data availability, licensing and machine readability [18]. Therefore, to overcome this issues, the interest groups/tools were established in identifying Data Openness used within selected of HE Countries.

5. Conclusion

Higher Education Open Data is one of the government initiatives that are currently has undertook the sensation in order to share the data throughout the public universities thus can benefit the HEIs and public as well. Due to this initiatives, there are a lot of things to consider towards the openness of the data. This paper discussed about the Openness of Higher Education data towards Open Data implementation. Further research will be conducted to identify other possible characteristics in developing Malaysia Technical University Network (MTUN) Open Data Framework.

Acknowledgement

Abbreviations	
CHET	Centre for Higher Education Transformation
DHET	Department of Higher Education and Training
HE	Higher Education
HEIs	Higher Education Institutions
HEMIS	Higher Education Management Information System
HESA	Higher Education Statistics Agency
JESTEC	Journal of Engineering Science and Technology
JPA	Public Service Department
MAMPU	Malaysian Administrative Modernization and Management Planning Unit
MDEC	Malaysia Digital Economy Corporation
MOHE	Ministry of Higher Education
MTUN	Malaysia Technical University Network
ODC	Open Data Certificate
ODDC	Open Data in Developing Countries
ODI	Open Data Institute
ODP	Open Data Pathway
OGD	Open Government Data
UK	United Kingdom
URIs	Uniform Resource Identifiers
UTeM	Universiti Teknikal Malaysia Melaka
WESP	World Economic Situation and Prospects report

References

1. Reale, G. (2014). Opportunities and differences of open government data policies in Europe. *Athens Journal of Social Sciences*, 1(3), 195-206.

2. Tran, E.; and Scholtes, G. (2015). Open data literature review. *The 19th Annual BCLT/BTLJ Symposium: Open Data: Addressing Privacy, Security, and Civil Rights Challenges 2*. Budapest, 1-33.

3. Open Knowledge International (2012). What is open data? Retrieved January 5, 2018, from http://opendatahandbook.org

4. Manyika, J.; Chui, M.; Groves, P.; Farrell, D.; Van Kuiken, S.; Groves, P.; and Doshi, E.A. (2013). *Open data: Unlocking innovation and performance with liquid information*. McKinsey & Company.

5. AlRushaid, 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*, 7(12), 113-122.

6. OpenDataBarometer, (2015). The open data barometer. Retrieved February 14, 2018, from http://opendatabarometer.org

7. Higher Education Commission. (2016). *From bricks to clicks: The potential of data analytics in higher education.* Policy Connect.

8. Magalhaes, G.; Roseira, C.; and Strover, S. (2013). Open government data intermediaries. *ICEGOV 2013 Proceedings - 7th International Conference on Theory and Practice of Electronic Governance*. Seoul, Republic of Korea, 330-333.

9. Sayogo, D.S.; Pardo, T.A.; and Cook, M. (2014). A framework for benchmarking open government data efforts. *Proceedings of the 47th Hawaii International Conference on System Sciences*. Hawaii, 1896-1905.

10. Janssen, M.; Charalabidis, Y.; and Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Journal of Information Systems Management*, 29(4), 258-268.

11. Am, P.; and Menteri, J.P. (2015). *Pelaksanaan data terbuka.*MAMPU.

12. Krumova, M. (2017). Higher education 2.0 and open data: A framework for university openness and co-creation performance. *ICEGOV 2017 Proceedings*. New Delhi, India, 562-563.

13. Krumova, M.Y. (2016). Open data benchmarking for higher education: Management and technology perspectives. *Proceedings of the 4th Smart Cities and Regional Development Journal,* (02-2017)*.

14. Jabatan Perkhidmatan Awam. (2009). *Surat Pekeliling Perkhidmatan Bilangan 7*.

15. Study Malaysia. (2015). The Malaysian higher education system - An overview. Retrieved March 14, 2018, from https://studymalaysia.com.

16. DJIA Indexes. (2014). Country Classification. *Dow Jones Indexed Journal*, 1-3.

17. Universities UK. (2015). *Open data in higher education: An introductory guide*. Universities UK and Open Data Institute.

18. Van Schalkwyk, F.; Willmers, M.; and Czerniewwicz, L. (2014). *Open data in the governance of South African higher education*. OpenUCT Initiative.