Openness · Transparency · Participation · Collaboration · Citizen Engagement

UNITED NATIONS

الاسكوا

ESCWA

Open government for greater public sector transparency and accountability in Arab countries

Capacity development material on open data

Economic and Social Commission for Western Asia (ESCWA)

Open Government for Greater Public Sector Transparency and Accountability in Arab Countries

Capacity Development Material on Open Data

United Nations

Beirut, 2019

Acknowledgements

CONTENTS

LIST OF TABLES

CONTENTS (*continued*)

LIST OF FIGURES

LIST OF BOXES

Abbreviations and acronyms

4IR	Fourth Industrial Revolution
API	application programming interface
BI	business intelligence
CIAT	Centro Internacional de Agricultura Tropical
CKAN	Comprehensive Knowledge Archive Network
CMS	content management system
CSV	comma-separated value
DFID	Department for International Development
EFTA	European Free Trade Association
ESCAP	Economic and Social Commission for Asia and the Pacific
ESCWA	Economic and Social Commission for Western Asia
EU	European Union
FOI	Freedom of Information
GCC	Gulf Cooperation Council
GIS	Geospatial Information System
GML	geography mark-up language
HDX	Humanitarian Data Exchange
ICT	information and communications technology
IT	information technology
JSON	JavaScript Object Notation
NECTA	National Examinations Council of Tanzania
NGO	non-governmental organization
ODI	Open Data Institute
OECD	Organization for Economic Cooperation and Development
OGP	Open Government Partnership
PDF	portable document format

RDF	resource description framework
SDG	Sustainable Development Goal
SMART	specific, measurable, attainable, relevant, and timely
TfL	Transport for London
TSV	tab-separated value
UN-DESA	United Nations Department for Economic and Social Affairs
UNDP	United Nations Development Programme
URI	uniform resource identifier
URL	uniform resource locator
WHO	World Health Organization
XML	extensible mark-up language

Introduction

Introduction

This document aims to provide open data practitioners in Arab States with practical guidelines on the initiation and management of open government data programmes. The guidelines were developed while taking into account the following considerations concerning the state of open government data in the Arab region:

- Approximately half of the Arab States have yet to launch formal open data programmes, and existing programmes are still in an emerging state without much political, economic or social impact;

- While low impact is common in all Arab countries with open data programmes, readiness to implement such programmes into mainstream practice and culture varies from one country to another. This applies to both the supply side of open data (mainly government agencies) and its demand side (civil society and businesses).

Therefore, to maximize benefit, readers should take into account the local context specific to their government organizations or open data programme and adjust the recommended methods and tools accordingly.

I. Open government data: concepts and benefits

I. OPEN GOVERNMENT DATA: CONCEPTS AND BENEFITS

A. INTRODUCTION: GOVERNANCE IN THE AGE OF DATA

In 2018, the number of people connected to the Internet surpassed the 3 billion threshold, compared with only 2.3 million in 1990.[1] This is an unprecedented milestone in the history of humanity. In the Arab region, more than 125 million individuals are connected to the Internet with an annual average growth rate close to 30 per cent.[2]

The Fourth Industrial Revolution (4IR) is promoting the connection of "things" to the Internet in addition to people. All sorts of devices, from subway trains and wind turbines to toilet seats and toasters, are becoming sources of data. It is projected that the total number of "things" connected to the Internet will reach 25.1 billion by 2021, 64 per cent of which being consumer applications.[3]

As a result, the production of data has reached an unprecedented volume and pace. The size of the "digital universe" (the data we create and copy every year) is doubling every two years. By 2020, it is expected to reach 44 zettabytes or 44 trillion gigabytes.[4]

Data is now being referred to as the "new oil". It is to this century what oil was to the last one: a driver of growth and change. Flows of data have created new infrastructure, new businesses, new monopolies, new economics, and new models for the public sector.

Private companies of all sizes and in all sectors are becoming "data companies", and they are exploiting the so-called "data-network effect" – a powerful economic engine that uses data to attract more users, who then generate more data, which helps to improve services and attract more users.

Uber, for example, is best known as a transport company. The firm is worth an estimated $68 billion, this in part is because it owns the biggest pool of personal transportation data. The company operates in more than 600 cities around the world, and, by June 2018, it completed 10 billion trips across rides and deliveries.[5] For each trip, Uber collects a wide range of data points including the geographic coordinates of both the trip's starting point and destination, date and time, fare, and data about the riders.

By analysing millions of trips in a city, Uber can generate insights about the state and pattern of transport in that city, such as how long it takes to get from one area to another and travel patterns in different times of the day, days of the week or months of the year, in addition to insights on how road closures and other events might impact transport across the city.

Uber Movement, a platform launched by the company in 2017, offers access to such data to help urban planners make informed decisions about their cities, and the platform offers data about cities all over the world, whether it is London, Nairobi or Delhi.[6] These data and insights can be very valuable to city authorities and planners in both developing and developed countries.

While a significant portion of the world's data is owned by the private sector, especially through social media websites and applications, the remainder of global data lies in government hands. Governments own and continue to produce huge amounts of data, most of which is still stored in paper formats or legacy systems. This is a potential source of wealth and public value creation, provided that the methods of data collection and storage are updated, and that it is published and shared so it can be readily and easily consulted and reused (Jarrar, 2017). Many governments all over the world are trying to adopt the practices of private sector companies, such as Uber, to leverage the data they have for better public policies and services. However, these efforts are faced with the lack of skills and expertise required to that effect. According to Gartner,[7] only 39 per cent of government organizations agree that their technical staff have the competencies and expertise needed to support the desired analytics and business intelligence (BI) initiatives, compared to 56 per cent in other

industries. With regards to business analysts, only 22 per cent of government organizations agree that they have the necessary skills, compared to 36 per cent in other industries.

There is an increasing awareness of this gap in data skills, and serious measures are being taken to bridge it. In 2017, the Government of Singapore launched a four-year programme to train 10,000 government officials in digital and data skills.[8]

This gap, which may hinder progress in achieving national developments goals, is also widespread in the Arab region. Several Governments in the region have launched initiatives to remedy this issue, but the successful implementation of open data programmes and their tangible impact remain to be seen. This document aims to contribute to these on-going efforts by providing a reference to the essential concepts of open data.

<center>B. GOOD GOVERNANCE AND OPEN GOVERNMENT</center>

In September 2015, the General Assembly of the United Nations adopted resolution 70/1, entitled "Transforming our world: the 2030 Agenda for Sustainable Development", which identified 17 Sustainable Development Goals (SDGs) and 169 targets to achieve this transformation.

The agenda emphasized the building of "effective, accountable and inclusive institutions at all levels" (Goal 16), and identified a set of targets to that effect, including the following:

- Developing effective, accountable and transparent institutions at all levels (target 16.6);

- Ensuring responsive, inclusive, participatory and representative decision-making at all levels (target 16.7);

- Ensuring public access to information and protect fundamental freedoms, in accordance with national legislation and international agreements (target 16.10).

Achieving this goal clearly needs good governance. Governance can be defined as the "process whereby societies or organizations make their important decisions, determine whom they involve in the process and how they render account".[9] Another definition that focuses on governance at the country level states that it is "the exercise of economic, political and administrative authority to manage a country's affairs at all levels. It comprises of the mechanisms, processes and institutions through which citizens and groups articulate their interests, exercise their legal rights, meet their obligations and mediate their differences".[10] It is also defined more concisely in the same source as "the method through which power is exercised in the management of a country's political, economic and social resources for development". However, a simpler definition of governance states that it is "the process of decision-making and the process by which decisions are implemented (or not implemented)".[11]

Evidently, governance is applicable to different contexts and situations where decisions need to be made and implemented, whether in the public sector or outside it, at a local, national or even international level.

The focus of this report is on the public sector, and therefore the government is a key stakeholder in the governance process but is by no means the only one. Depending on the society discussed and focused on, other stakeholders may include non-profit organizations, private-sector businesses, the general public, and leaders of various social groups, such as tribal and religious leaders.

Due to the wide range of contexts and the variety of the stakeholders involved, good governance can be defined in a number of ways. One definition states that good governance is "sound public-sector management (efficiency, effectiveness and economy), accountability, exchange and free flow of information (transparency), and a legal framework for development (justice, respect for human rights and liberties)".[12]

Another definition, the one adopted in this report, was made by the United Nations Economic and Social Commission for Asia and the Pacific (ESCAP).[13] This definition identifies good governance through the following eight characteristics: It is participatory; consensus oriented; accountable; transparent; responsive; effective and efficient; equitable and inclusive; and it follows the rule of law.

The adoption of the SDGs coincided with an increased focus on the concept of "open government", especially for the achievement of better governance and greater transparency. Throughout the last decade, this concept gained growing attention, and many countries initiated programmes and efforts towards adopting it. Like governance, open government can be defined differently in different contexts. However, some regional and international bodies have mutually agreed on common definitions.

The Organization for Economic Cooperation and Development (OECD) defines open government as the "culture of governance based on innovative and sustainable public policies and practices inspired by the principles of transparency, accountability, and participation that fosters democracy and inclusive growth".[14]

Another definition states that open government is the "governing doctrine which holds that citizens have the right to access".[15]

The United Nations Economic and Social Commission for Western Asia (ESCWA) identified the following five dimensions for open government: accountability, transparency, inclusiveness, effectiveness, and contestability.[16]

Moreover, the Open Government Partnership (OGP), a "multilateral initiative that aims to secure concrete commitments from governments to promote transparency, empower citizens, fight corruption, and harness new technologies to strengthen governance",[17] identified the following four commitments for more than 70 member Governments:[18]

- Increase the availability of information about governmental activities;
- Support civic participation;
- Implement the highest standards of professional integrity throughout administrations;
- Increase access to new technologies for openness and accountability.

These commitments share with the above-mentioned definitions a common focus on transparency and accountability, availability of government data and information and the citizen's right to access them. These are essential characteristics of good governance, in addition to being crucial targets for the SDGs.

Open government is, therefore, increasingly considered an essential tool for fostering good governance and moving forward towards achieving the SDGs. The OGP Declaration states that "we embrace principles of transparency and open government with a view toward achieving greater prosperity, well-being, and human dignity in our own countries and in an increasingly interconnected world".[19]

To transform these principles and commitments of open government to successful standard practices, Governments can leverage many frameworks and models that have emerged in the past few years.

ESCWA recently developed a framework for the application of open government in Arab countries. The framework is consistent with the context of the countries, specifically the level of information and communications technology (ICT) investment, e-government status, and the legislative and regulatory aspects, Arab culture and the level of interaction between citizens and the public sector. The framework can be used as a guideline for a transformation towards open government (figure 1).

Figure 1. ESCWA policy framework on open government for Arab countries
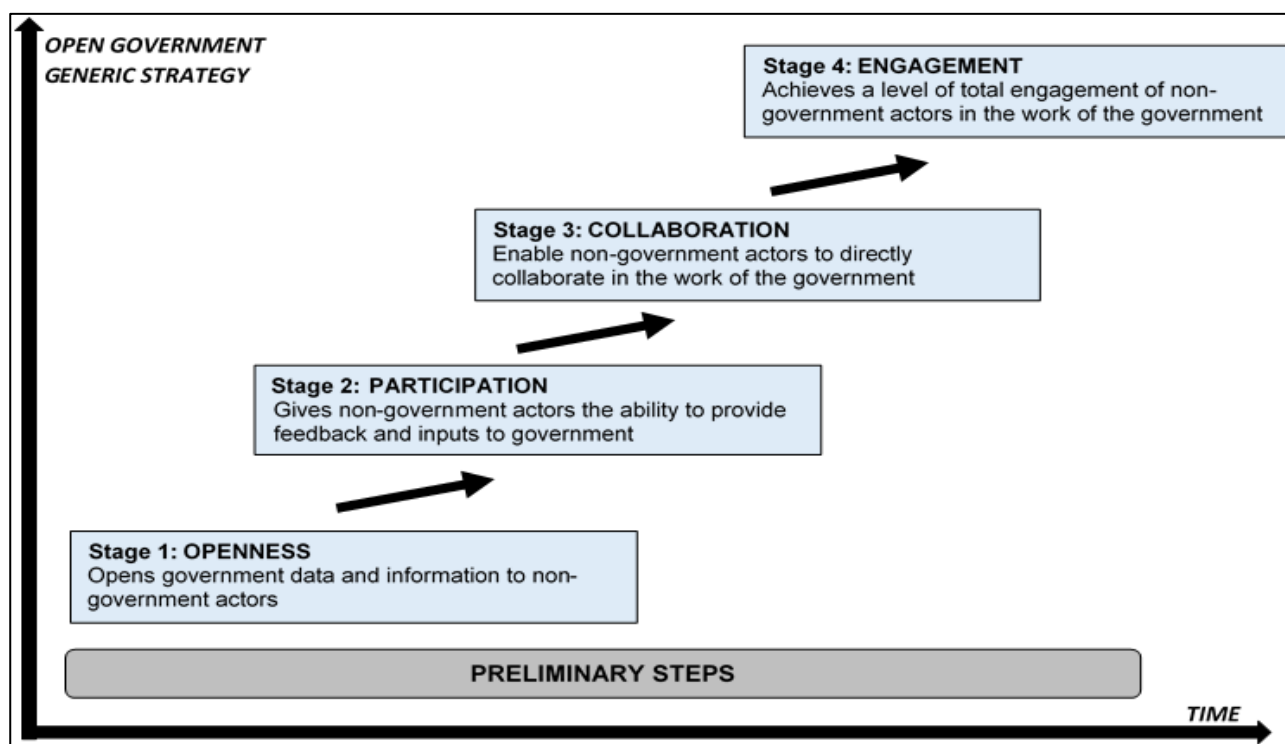
The ESCWA framework for open government consists of four stages, in addition to preliminary steps needed for the Arab countries.

Preliminary steps

Governments in Arab countries are encouraged to take several preliminary measures to prepare procedures for every phase of the proposed framework, including the following:

- Drafting a concept paper for wide dissemination that explains the meaning of open government;

- Organizing training, especially for government workers, on technology and the professional use of social media;

- Developing plans to improve ICT networks and reduce the cost of using them;

- Preparing a policy document that sets out a national framework for open government's vision and principals.

Stage 1: Openness

Stage 1 aims to achieve openness to enhance transparency, which is the main channel for implementing open government. This phase forms the cornerstone of open participation and cooperation between Government, citizens and other stakeholders. During this phase, it is necessary to do the following:

- Focus on data dissemination through ICT, while ensuring data quality and periodic updates;

- Raise awareness among the public and government workers on the importance of open data and government openness and accountability;

- Encourage innovation through open data investment to develop key public services.

Stage 2: Participation

Stage 2 aims to strengthen the culture and practices of open government by promoting interaction between Government departments and citizens. It also enhances inclusiveness and citizen involvement in Government work and decision-making. This could be achieved by the following:

- encourage citizens to present ideas, knowledge, comments and proposals to the Government;
- disseminate and share citizen feedback;
- Respond to citizens by providing feedback on steps taken.

Stage 3: Collaboration

Stage 3 aims to enhance open cooperation between Government departments, citizens, civil society organizations, and the private sector. It focuses on the following:

- Promote dialogue between Government and citizens on public policies and decisions;

- Provide services according to beneficiary needs;

- Collaborate with citizens for the development of tools and services using collaborative tools and applications such as "blog" and "wiki";

- Involve citizens in service design as this is expected to increase quality and innovation in Government services and reduce costs.

Stage 4: Citizen engagement

Stage 4 aims to achieve effective citizen participation by promoting the previous three phases to ensure full citizen involvement in Government work, where citizens and other stakeholders (civil society organizations, the private sector and Government actors) participate in formulating policies and in decision-making. This phase also aims to build an accountable government that places citizens at the centre of its concerns, so as to achieve the SDGs.

C. OPEN DATA: CONCEPT AND DEFINITION

1. *Concept*

Increasing transparency by publishing open data is the first stage in the framework for fostering open government we referred to previously, which makes it the foundation of any successful open government initiative. OECD describes open data as a "philosophy – and increasingly a set of policies – that promotes transparency, accountability and value creation by making government data available to all".[20]

The World Bank described open data itself as a very new concept that originated with the belief that the enormous amount of information routinely collected by Government entities should be available to all citizens. In the late 2000s, Governments and entities began to allow a greater number of users access to these resources.[21]

A major milestone in the global open data movement took place in 2009, when President Barack Obama issued the United States Open Government Directive, which required all Government agencies to post at least three high-value data sets online.[22] This directive lead to the launch of data.gov, the online repository on which Government data are published in open format, in addition to the related tools and resources.

2. *Definition*

As the adoption of open data spreads globally, different organizations and institutions have developed various versions of definitions that aim to explain the term and practice in various ways.

The United Nations Department of Economic and Social Affairs (UN-DESA) in its e-Government Survey defines open data as "government information proactively disclosed and made available online for everyone's access, reuse and redistribution without restriction".[23]

The same definition is also used by the World Bank.[24]

The Open Knowledge Foundation defines open data as the "data that can be freely used, shared and built-on by anyone, anywhere, for any purpose".[25] This definition is derived from the famous Open Definition which defines the term "open" with respect to knowledge, stating that "knowledge is open if anyone is free to access, use, modify, and share it".[26] Under this definition, data must satisfy the following three requirements to be considered "open data":

Open licence: Data must be published under a legal licence that enables anyone to access it, use it and share it.

Open access: Data should be accessible by anyone without any technical or financial barriers. For example, publishing data in portable document format (PDF) makes it extremely difficult for users to process.

Open format: This includes requirements such as publishing data in bulk and in machine-readable format.[27]

This report adopts the definition of the Open Data Institute, which states that open data is "data that anyone can access, use or share".[28] It is a simple definition that captures the key elements that are included in the other previously cited definitions.

For example, the Victoria State Government in Australia published on its open data portal a dataset titled "Post school destinations of year 12 or equivalent completers, Victorian schools, 2017".[29] The dataset has the following features that match the above definition of open data:

- It has an open legal licence (Creative Commons Attribution 4.0 International) that enables anyone to access it, share it and adapt it for any purpose, even a commercial one;
- It is published in comma-separated value (CSV) format, which is machine-readable and non-proprietary;
- It has other features that increase its value to the users, such as metadata.

These features will be explained in greater detail in other chapters of this report.

D. OPEN DATA: VALUE AND BENEFITS

As shown in figure 2, an increasing number of Governments across the globe are opening up their data for public information and scrutiny.[30] This provides Governments and societies with a number of benefits and helps them accomplish targets that are linked to their own national agenda, as well as the international agenda as laid out in the United Nations' SDGs.

Figure 2. Number of countries with open government data portal and/or catalogues, 2014, 2016 and 2018
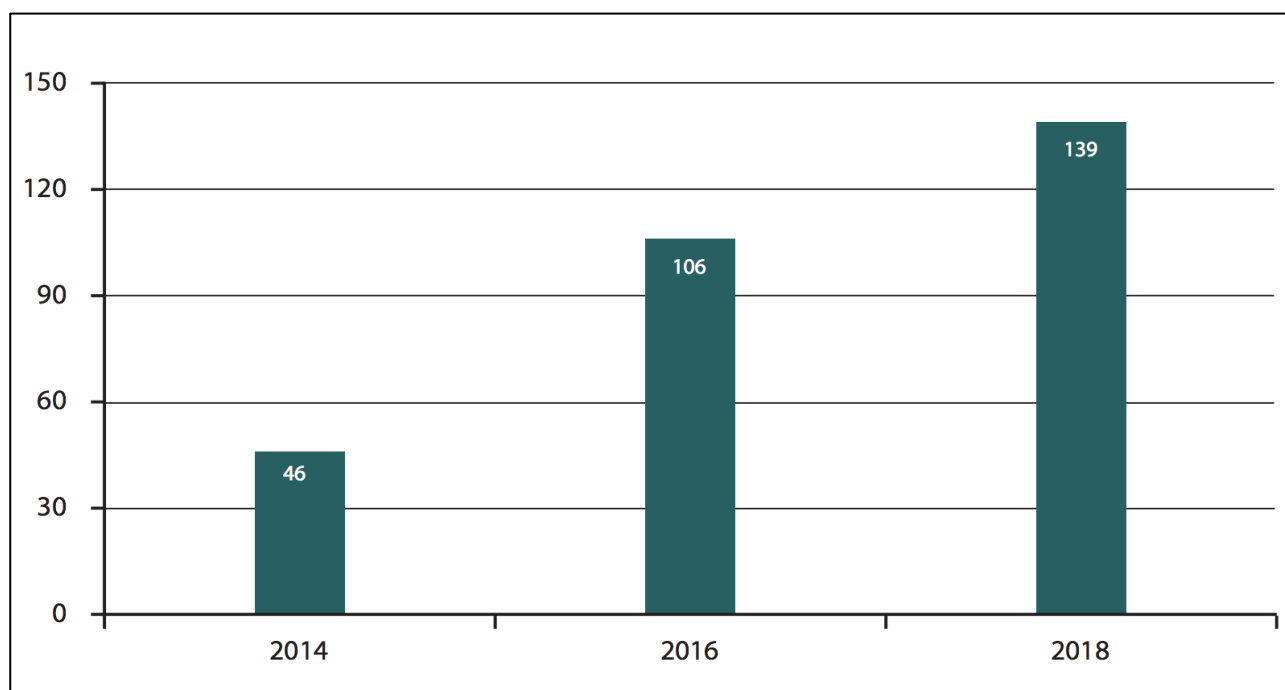
1. *Economic value of open data*

The economic value of open data can be shown in many areas such as the market size, number of created jobs and economic growth. In 2015, the direct market size of open data in EU 28+ (the 28 member States in the European Union plus the four States in the European Free Trade Association (EFTA)) was 51 billion euros, with an increase of 82 per cent compared to 2005.[31] As shown in figure 3, the direct market size is expected to increase by another 36.9 per cent over the period between 2016 and 2020 (from 55.3 billion euros in 2016 to 75.7 billion euros in 2020). The cumulative direct market size for the period 2016-2020 is estimated to reach 325 billion euros.

Figure 3. Direct market size of open data, 2015-2020
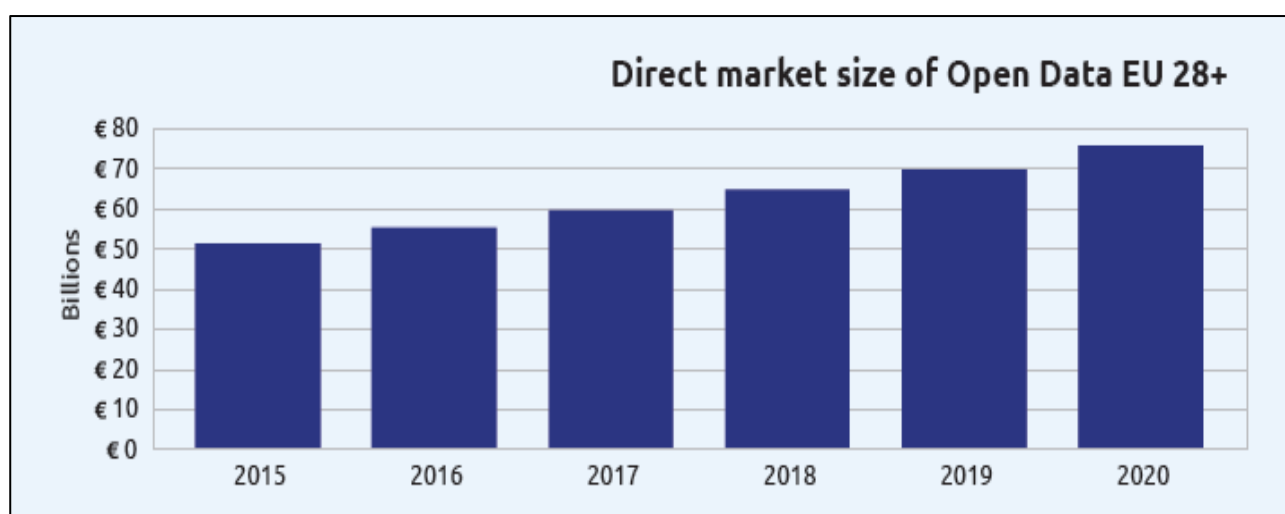
The European Data Portal extrapolated the direct market size per sector in 2020 and the number of jobs created by open data, as can be seen in figure 4 and figure 5. Although the sectors most influenced by open data are public administration, industry, trades, and transport, open data has a significant impact on other sectors as well. Open data jobs are created every year. In the period 2016-2020, more than 5,000 open data jobs are being created within the EU 28+ per year.

Figure 4. Direct market size of open data per market sector, 2020
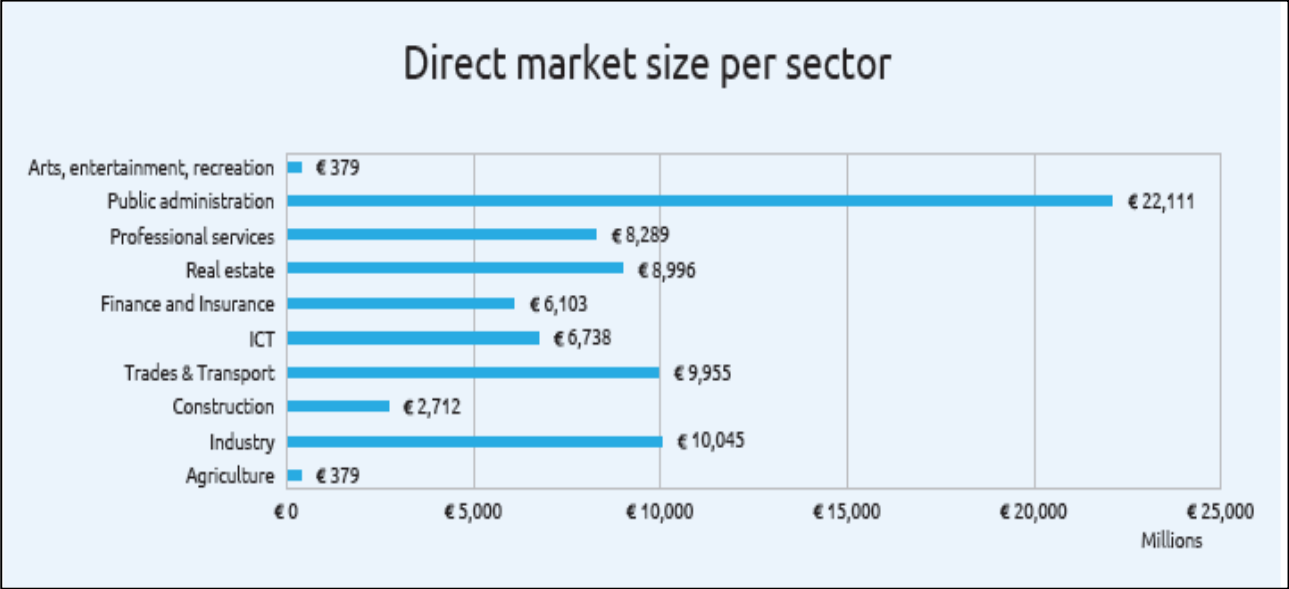
Source: European Data Portal.

From the aspect of economic growth, opening up government data can lead to significant economic gains at both the national and sectorial levels. It was estimated that open data can help unlock 3 to 5 trillion United States dollars in economic value for the global economy per year across seven sectors (figure 6).[32]

Figure 5. Direct open data jobs created, 2016-2020

Source: European Data Portal.

14

Education	890–1,180
Transportation	720–920
Consumer products	520–1,470
Electricity	340–580
Oil and gas	240–510
Health care¹	300–450
Consumer finance	210–280
Total	**3,220–5,390**

The values here are drawn from examples of open-data potential and are not a comprehensive sizing of potential value across the two sectors.

¹Includes US values only.

Source: Manyika et al., 2013.

2. *Impact of open data and linkages with the SDGs*

Open data has several dimensions of impact that include improving governance, empowering citizens, solving public problems and creating opportunities. Open data also contributes to improving Governments in tackling corruption through increased transparency and accountability, and in enhancing public service delivery and resource allocation.[33]

Publishing a wide range of datasets encourages economic growth and fosters innovation in many sectors, and it empowers citizen engagement by enabling more informed decision-making.

A paper published by GovLab in 2017 demonstrates how open data may contribute, directly and indirectly, to many sectors, especially in developing economies. Table 1 contains a summary of what is mentioned in the paper to that effect.

Table 1. How open data can create an impact[34]

	How open data can create an impact			
Sector	**Improving governance**	**Empowering citizens**	**Innovation\creating opportunities**	**Solving public problems**
Health	Increasing transparency and accountability of Government expenditure on the health sector helps in ensuring that resources efficiently and adequately target public health needs.	Accessing health data can bolster the ability of citizens to make informed choices regarding their service providers.	The health sector is increasingly data-driven. Based on data availability on both the supply and demand sides, open data can spur job creation and the establishment of new service models.	Accessing data on the availability and location of health resources and on emergent health outcomes can play an important role in addressing major epidemics or public health concerns.
Humanitarian aid	Tracking how money is being used can help root out corruption	By opening aid allocation data, citizens can provide valuable feedback to	Open humanitarian aid data allows NGOs and other civil society organizations to create innovative strategies based on	Open data can help organizations better identify where and how to invest

How open data can create an impact				
Sector	Improving governance	Empowering citizens	Innovation\creating opportunities	Solving public problems
	and catalyse better spending practices in government.	Governments on how aid is being used, and become active co-partners, rather than mere recipients.	new and accurate information, for example, to help communities in conflict areas or recovering from natural disasters.	humanitarian aid to most effectively solve social problems and identify sectors that most urgently require this aid.
Agriculture and nutrition	Open data can make agricultural agencies and organizations more accountable, by making information accessible on whether or not financial resources provided were used according to contractual obligations and whether they serve and support people and farmers.	Farmers can be full participants in defining, implementing and evaluating agricultural projects. Citizens can lobby Government institutions for more equitable food policies.	An increased awareness of weather trends, models of crop yields and other relevant datasets can help inform more strategic and evidence-based agricultural decision-making and increase the viability of individual farms.	Open data can play an important role in predicting potentially damaging conditions for crops, and informing more strategic planting choices following, for example, a catastrophic weather event.
Poverty alleviation	Identifying and addressing corruption through open data can lead to the reallocation of resources toward public services that are better suited for addressing systemic poverty.	An improved understanding of how the Government allocates resources can enable public mobilization around issues such as poverty.	Whether enabling job creation, frugal innovation efforts or more systemic economic growth, as developing economies begin to leverage data as an economic asset, poverty alleviation can accelerate as a result.	Open data can improve intervention programmes that seek to alleviate poverty and improve the quality of life by enhancing the understanding of cities, organizations and donors as to where the needs are the biggest.
Energy	Improving the transparency of a country's energy budget can potentially identify and prevent corruption. Moreover, open datasets such as open address data, can improve service delivery.	Open data on utility services and pricing can be used to identify the best-priced service, and citizens can make decisions regarding their usage and decrease their energy spending.	The use of open energy data to bolster predictive capabilities and reel in energy expenditures can have wide-ranging economic impacts for both the public and private sectors.	Open data from various sources can help decision makers prioritize investments in energy production and delivery.
Education	Opening education data can encourage Governments to weed out vested interests that may overdraw the public educational fund.	Open data on expenditures and performance of schools can help parents make more informed decisions about school choice, and mobilize citizens to identify deficiencies.	Education is increasingly driven by data and technology. More accessible data on schools and on the subjects taught in schools can spur the creation of a data-driven "Ed-Tech" industry.	Analysing open data through learning analytics can improve the often poor quality of the education sector.

Source: GovLab, 2017.

In the following sections, the value and benefits of open data are demonstrated through a number of selected case studies, showing how it can make an impact.

(a) *Improving governance*

The trend towards openly publishing data to enhance transparency and accountability can be observed in the initiatives, policies and commitments of many Governments and international organizations all over the world. The Open Data Charter is a collaboration between Governments and experts that aims to help Governments practice openness through opening up data; 69 Governments joined the charter and 50 organizations endorsed it to date. A version of the charter, entitled "G8 Open Data Charter", was signed by the leaders of the Group of Eight (G8) in July 2013.[35]

Several Arab countries linked their open data initiatives with promoting the transparency and accountability of the public sector. Qatar's Open Data Policy of 2014 states that "Open data supports the Qatar National Development Strategy 2011-2016 and the Digital Government 2020 Strategy's call for transparency, efficiency and participation of Qatar's people".[36] Saudi Arabia has an open data portal that aims "to implement a public data hub and strategy to enable transparency, promote e-participation and inspire innovation".[37]

The United Kingdom, for example, is ranked first among 115 countries and jurisdictions on the 2016 version of the World Wide Web Foundation's Open Data Barometer of 2016, an index that analyses global open data trends and provides comparative data on countries and regions.[38] The United Kingdom also ranked second among 94 countries on the Open Knowledge Network's Global Open Data Index 2016/2017, which annually measures the openness of government data based on the open definition elaborated previously.[39]

In December 2017, Prime Minister Theresa May issued a letter to the Cabinet on government transparency and open data. In the letter, the Prime Minister stated that "since 2010, the Government has been at the forefront of opening up data to allow the public and press to hold public bodies to account. Such online transparency is crucial to delivering value for money, to cutting waste and inefficiency, and to ensuring every pound of taxpayers' money is spent in the best possible way".[40]

Moreover, the Prime Minister listed the information and data that the Government has committed to publish. Examples include civil service sickness absence data; ministerial salary data; central government contracts; tender opportunities and contract award notices, over £10,000; monthly payment card data over £500; gender pay gap data; and senior officials' business expenses, hospitality and meetings.

Similarly, local governments in the United Kingdom follow the Transparency Code which aims to meet the "Government's desire to place more power into citizens' hands to increase democratic accountability and make it easier for local people to contribute to the local decision-making process and help shape public services".[41] The Transparency Code emphasizes that "transparency is the foundation of local accountability and the key that gives people the tools and information they need to enable them to play a bigger role in society".[42] More importantly, the Transparency Code sets out the minimum data that local authorities are required to publish. Those include the details of each individual item of expenditure that exceeds £500 which might be in the form of individual invoices, grant payments or any other form of expenditure.[43]

(b) *Empowering citizens*

When Governments share open data about the various aspects of their work, they create a channel for their citizens to engage in a dialogue on Government performance and actively participate in the design of public policies and programmes. UN-DESA considers access to information collected and generated by Governments "an important prerequisite to the exercise of other rights, including the right to fully participate in the political process".[44]

In Ghana, for example, the Government started publishing open data about 25 Government-funded infrastructure projects, and connected communities with data on projects most relevant to them. This step helped in driving political participation across the country and engaged 400,000 citizens both online and offline.[45]

Box 1. Open data for reforming education in Tanzania

In 2012, six out of every ten students in Tanzania failed the standardized national secondary-level examination. This crisis led to demands for reform in the education sector. The poor results were the product of recent changes in the country's education system, where tuition fees for governmental primary schools were eliminated in an effort to boost the country's school enrolment and literacy rates. The move triggered a rapid increase in the net primary enrolment rate from 66 per cent in 2001 to 90 per cent in 2004. The increase, however, was not matched by a proportional increase in school funding.

In response to the crisis, two open data driven initiatives emerged. The first was the Education Open Data Dashboard (see figure in the box) launched by the National Examinations Council of Tanzania (NECTA). The dashboard is connected to the country's open data portal and offers data on both primary and secondary schools. The published datasets cover pupil-teacher ratios, annual and average pass rates, national rankings of school performance, and changes in pass rates since 2012. The users of the portal can browse and search the catalogue, download the datasets and perform visualizations.

This project was supported by the open data initiative in Tanzania, the World Bank and the United Kingdom's Department for International Development (DFID).

The second project was Shule, a portal with an objective similar to the first. However, this portal was launched by the individual effort of Arnold Minde, a programmer, entrepreneur and open data enthusiast.

At the time of writing this report, the data on both portals seems to be out of date, which probably indicates that both of them have reached the end of their lifespans. When launched, both projects were considered promising, and they had to deal with the challenge of low Internet penetration rates in Tanzania in addition to the low level of open data awareness in the country.

Education Open Data Dashboard

Source: http://educationdashboard.org/#/dashboard/primary.
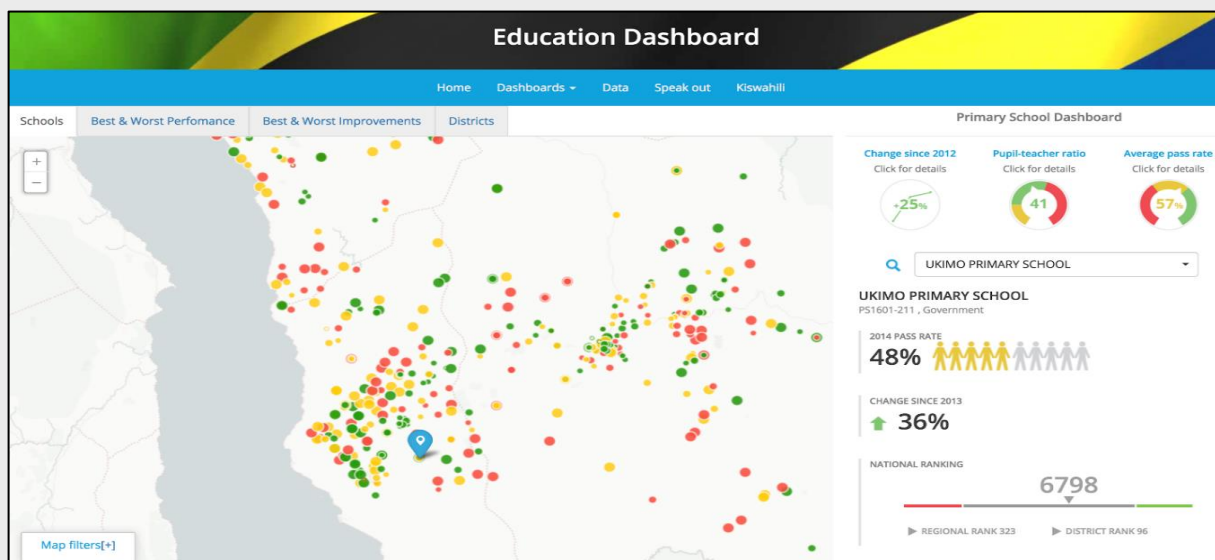
Key lesson: The innovative use of open data can be expected not only from well-established organizations, but also start-ups and individuals. Government agencies and other open data stakeholders should give priority to enabling the growth and sustainability of such initiatives.

_____

Source: http://educationdashboard.org/#/.

Innovation and entrepreneurship

When Government data is published openly, individuals, businesses and non-profit organizations can tap into it to generate insights, design solutions for the problem facing their societies, save time, and create economic value.

Transport for London (TfL), the Government agency in charge of the day-to-day operations of the London's public transport network, has opted not to develop mobile apps for its services. Instead, TfL focuses on publishing high-quality open data that enables entrepreneurs to innovate such mobile apps. As a result, the city has now more than 500 mobile apps developed by more than 8,000 developers.[46]

(d) *Solving public problems*

Open data is playing an important role in solving public problems and can directly and indirectly contribute to addressing local challenges.

Box 2. Columbia: Saving \$3.6 million in drought damage through open data

In Colombia, rice represents the most consumed staple food, particularly among poor people, who make up 30 per cent of the population. Rice farmers, however, struggle with climate change which has caused a decrease in rice production each year since 2007.

To tackle this challenge, a mix of open and private data (data owned by companies) is being used to help farmers take precautions to avoid drought damage.

Between 2007 and 2013, an association of farmers (the National Federation of Rice Growers Fedearroz), which represents more than 50,000 farmers, joined forces with the international research centre Centro Internacional de Agricultura Tropical (CIAT) and Colombia's Ministry of Agriculture to identify the issues behind yearly reductions in rice crop yields.

The coalition used data-mining techniques to analyse data from resources that included annual rice surveys, harvest monitoring data, experiments on rice sowing dates as well as weather data from the national institute of hydrology, meteorology and environmental studies.

This data-driven research provided recommendations to Colombian farmers on planting times so as to avoid droughts, improve total yields and incomes and increase efficient use of resources.

In Córdoba, one of the country's five rice producing areas, 170 farmers followed these recommendations and avoided crop loss in 1,800 hectares of irrigated rice, as planting times were set based on the seasonal forecasts from the data analysis results. This decision alone saved \$3.5 million in input costs.

In addition to forecasting droughts, the research team recommended the adoption of rice varieties that were less sensitive to sunlight.

Key lesson: When the high-value datasets are released in good quality, they could be used to create insights that directly impact the lives of citizens.

_____

Source: https://ciat.cgiar.org/outcome/using-big-data-to-understand-declining-rice-crop-yields-in-colombia/; and https://www.theguardian.com/global-development/2014/sep/30/colombia-rice-growers-climate-change.

Box 3. Open data for battling Ebola outbreaks in West Africa

In 2014, the largest Ebola outbreak in history occurred in West Africa. More than 28,000 confirmed, probable and suspected cases were reported in Guinea, Liberia and Sierra Leone, with 11,310 deaths.

One key challenge facing the World Health Organization (WHO) and other governmental organizations and NGOs that were involved in the efforts to battle the virus disease was that information on Ebola cases and response efforts were dispersed across a diversity of data collectors.

Data was key in the battle against Ebola. However, limited data sharing was leading to disrupted attempts to combat the epidemic. It was difficult to get hold even of the most basic data, such as the number of cases or the number of deaths, making it difficult to assess the severity of the epidemic and target interventions. Recognizing the importance of data and information, Governments, aid agencies and international organizations soon began designing and implementing a number of data tools.

One key tool was the Humanitarian Data Exchange (HDX), an open platform for sharing data which was launched in July 2014 with the goal of "making humanitarian data easy to find and use for analysis".

Through HDX, aid agencies and authorities were able to tackle the disease and manage its spread through improving the quality of humanitarian data available to people working to address the crisis.

HDX defines humanitarian data as one of the following three types:

- Data about the context in which a humanitarian crisis is occurring, such as damage assessments and geospatial data;

- Data about the people affected by the crisis and their needs, such as the dataset on the number of Ebola health-care worker deaths;

- Data about the response by organizations and people seeking to help those who need assistance.

HDX also offers maps and visualizations of the published dataset and other tools that help the users of the platform use the data and generate insights.

Since 2014, HDX has expanded beyond Ebola in terms of the data published in its catalogue and users who access it. Today, the platform has more than 6,400 datasets accessed by users in over 200 countries and territories.

Key lesson: Open data can enable collaboration across different organizations that work in the same sector or share the same mission. This can be achieved by offering a single location for a wide range of datasets under that sector.

_____

Source: https://data.humdata.org/.

(e) *Addressing SDGs*

Open data is a tool for helping Governments to engage with NGOs, the private sector, citizens, and all other stakeholders in making new collaborations and improving public goods and services. As shown in the previous section, open data has an impact on a wide range of sectors, such as health, agriculture and education. This impact, overlapping with many different sectors, can demonstrate the important role of open data in contributing to achieve SDGs. A note prepared by the World Bank Open Data Team, under the title "Open Data for Sustainable Development", addressed how open data can contribute to each SDG. Table 2 was extracted from it.

Table 2. Open data linkages with the SDGs

SDG	Open data contribution
1. No poverty	• Publishing data about public goods and services provides people living in poverty with access to basic care; • Using open data drives business growth and thus the creation of new job opportunities; • Open data that can be used to match job-seekers with employers can help increase national employment levels overall.
2. Zero hunger	• Releasing agricultural and nutrition data promotes better education of farmers and ordinary consumers; • Using open data can help counter local problems that interfere with agriculture, such as unfair subsidies, disease or theft; • Open data can help correct erratic food prices, a potentially serious problem in low-income countries where food can be scarce.
3. Good health and well-being	• Open data can be structured into an effective tool for connecting potential patients to the optimal providers of health care; • Practitioners can use open data to make health care more efficient and reduce the costs; • Open government data combined and supplemented with crowd-sourced and other non-governmental data can also be critical in tracking infectious diseases.
4. Quality education	• Many countries are now using open government data and other data to show the availability and improve the quality of public schools; • Open data about basic school operations and facilities helps to inform policymakers about failures in schools and students' education.
5. Gender equality	• Analysis of open data can highlight disparities in the ways that education and health systems cater to women and girls; • Open data on health-care and education facilities can provide young women and girls with resources in sexual and maternal health.
6. Clean water and sanitation	• GPS and satellite imagery open data can be used to map out communities and infrastructures to plan the placement of water and sanitation points; • Open data on water quality can help for quicker clean-up/sanitation service responses; • Open data can also be used to identify areas that are undergoing a water supply crisis; • Open data about urban low-income areas can be used to create development and sanitation programmes aimed at the reduction of water and sanitation-related public health risks.
7. Affordable and clean energy	• Open data on competing electric utility services can be used to provide the best price points to consumers; • Open data from household energy surveys, satellite imagery and other sources can help Governments and private companies prioritize investments in energy generation and grid extension; • Open data can support the development of renewable energy sources strategically and efficiently.
8. Decent work and economic growth	• The provision of open data can be the raw material for innovative new data-driven businesses; • Open data can help existing companies optimize their businesses; • Open data can be a means to helping individuals receive the initial financing needed to launch a company.
9. Industry, innovation and infrastructure	• Open data opens up prospects for rethinking urban infrastructure through "smart cities". The concept of smart cities involves combining Government-provided open data with extensive, diverse, and timely data collected from sensors around the city that measure traffic, air quality, and other factors;

SDG	Open data contribution
	• Open data is being used to understand urban issues and improve urban planning; • Open data can improve transportation systems/infrastructure.
10. Reduced inequalities	• Open data promotes greater transparency in governance to reveal inequality and support action to correct it; • Open data informs domestic economic regulations to best provide for those living on social welfare or for those living near or below the poverty line.
11. Sustainable cities and communities	• Some of the most effective applications of open data, not only data collected by the Government, but also by citizens, helps in managing disaster risk and relief efforts.
12. Responsible consumption and production	• Open data can be used to track consumer prices on a global scale and spot in real time early signs of food shortages and inflationary pressures, and to detect other important trends and anomalies for better decision-making.
13. Climate action	• Open data programmes can be designed to help cities and countries become more resilient in the face of climate change and the floods, droughts and other extreme events that may ensue.
14. Life below water	• Open data can prove to be a useful way to monitor volume of fishing intake in national/supranational waters to ensure proper compliance with sustainable sourcing; • Open data on various marine activities can potentially lead to the discovery and usage of marine energy sources, such as geothermal vents and others.
15. Life on land	• Open data on mining, drilling and energy industry practices helps promote transparency that will prevent bad practices out of fear of public reaction.
16. Peace, justice and strong institutions	• Open data has helped several countries to launch spending transparency initiatives; • Open data on budgeting and elections promotes greater transparency in political institutions and can consequently encourage greater political involvement; • Open contracting – making Government contracts available for public review – deters favouritism and hidden deals, and simultaneously benefits Government, businesses and investors.
17. Partnerships for the Goals	• Open data can be used to collaborate on and implement solutions that have global benefits; • Open data on aid to developing countries can allow for more efficient allocation of foreign aid resources and support for sustainable development in less developed countries.

Source: World Bank Group, 2015.

Two initiatives were selected as examples on how open data can contribute to achieving the SDGs. The first example (box 4) is an initiative that contributes to gender equality (SDG 5), and the second example is on an initiative that contributes to sustainable cities and communities (SDG 11) (box 5).

Box 4. Noceilings.org initiative

In 1995, at the Fourth United Nations World Conference on Women in Beijing, leaders from Governments and the civil society around the world came together and committed to ensuring that women and girls have the opportunity to participate fully in all aspects of life. Twenty years later, in 2015, the Bill and Melinda Gates Foundation and the No Ceilings initiative of the Bill, Hillary and Chelsea Clinton Foundation have joined forces to gather data and analyse the gains made for women and girls, as well as the gaps that remain. Those efforts resulted in the launching of the No Ceilings portal and report.

The portal is home to 850,000 data points, spanning more than 20 years, from over 190 countries, including many Arab States. In addition to its comprehensive dataset, the portal offers data visualizations and presents stories on the achievements made by girls and women around the world and the challenges facing them.

Source: http://www.noceilings.org/map/.

- Examples of the cross-country data featured on the portal include the following: The number of female entrepreneurs is rising. In Nigeria, for example, 41 per cent of women are entrepreneurs;

- Girls' rights to education are cut short. While 71 per cent of national constitutions protect girls' access to primary education, few countries protect girls' rights to attend middle and high school;

- One in three women suffers physical or sexual violence.

Another helpful feature of the portal is country snapshots, which provides a glance into the state of women's participation at the country level. The figure in the box shows the snapshots of Saudi Arabia and Morocco.

The portal offers rich methods for users to navigate and view the data. In addition to the ability to download the entire row dataset, there are plenty of interactive charts, maps and videos.

Country snapshots of Saudi Arabia and Morocco
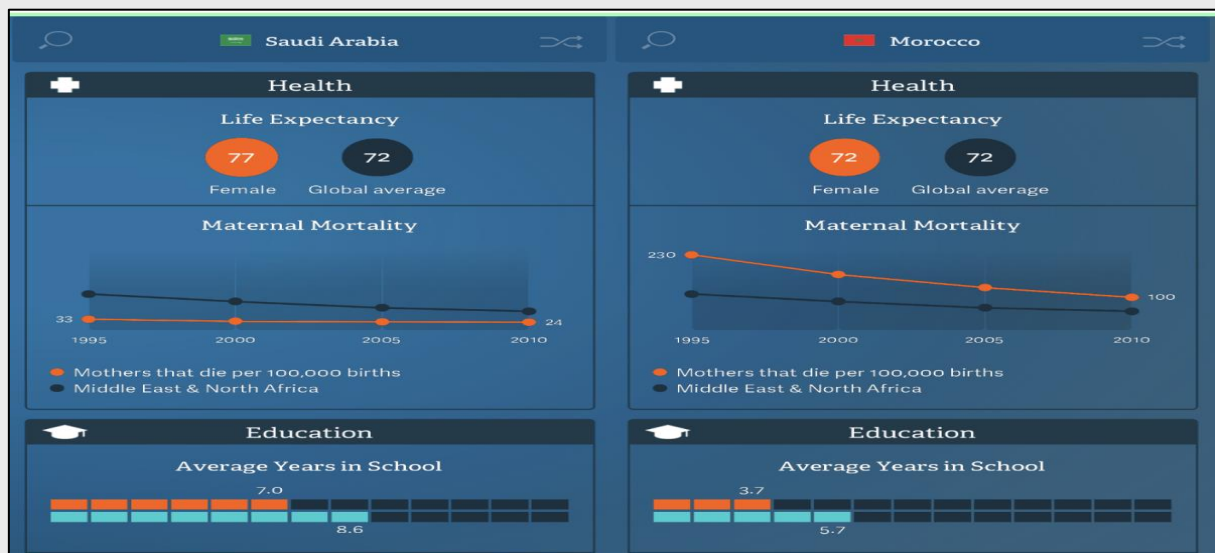
Key lesson: Historical data enables open data users to study and detect patterns and generate sound insights and decisions. In addition, the use of visualization tools can maximize the value of the published data.

_____

Source: http://www.noceilings.org/.

Box 5. National cities performance framework dashboard in Australia

The Australian Government has launched a portal to track the progress and performance across the country's largest cities. The portal uses the open data published by the federal and State Governments to measure the progress in more than 20 cities against the following six policy priorities that are linked to the objectives of the country's Smart Cities Plan:

Jobs and skills: Featured indicators include the percentage of each city's population who completed year 12 of school education, and who completed a bachelor degree or higher;

Housing: With several indicators, including data on the average number of persons per dwelling and the average housing construction costs;

Infrastructure and investment: The data covers such indicators as jobs accessible by car within 30 minutes and peak travel delay;

Liveability and sustainability: Indicators include dwellings with access to green spaces and percentage of population able to get crisis support;

Innovation and digital opportunities: Data includes trademark applications per 100k people and LinkedIn contacts by location;

Governance, planning and regulation: Indicators include data on population density and local governments per 100k people.

In addition, the portal offers data about the context in each city. Context data explains the characteristics of a city based on the measurement framework. By offering open data that covers these indicators across many cities, the portal creates an opportunity for these cities to learn from each other and cross-share good practices and expertise. Data in the portal is updated annually.

The LinkedIn contacts by location data page

Source: https://smart-cities.dashboard.gov.au/all-cities/innovation.
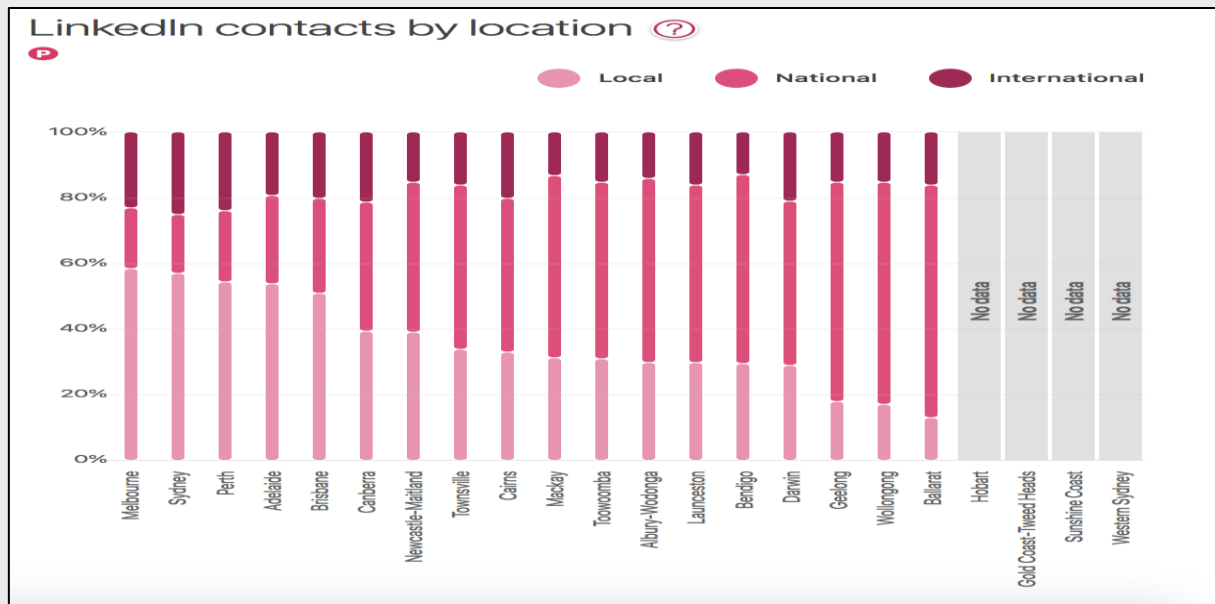
Key lesson: Open data can be a powerful tool for planning inclusive, safe, resilient, and sustainable cities. When using open data for this purpose, it is essential to select and publish the datasets that are directly linked to the strategic policy agenda.

II. Planning and managing open data initiatives

II. PLANNING AND MANAGING OPEN DATA INITIATIVES

A. INTRODUCTION

Having a well-designed plan is essential for successfully implementing a sustainable open data programme that creates the desired impact. In addition to the usual challenges and risks that face public sector projects in the Arab region, open data is considered a new domain in many Arab States and a totally unchartered territory in others.

When executing open data programmes or projects, a one-size-fits-all approach will not be effective. Each country, city or even a Government organization has its unique context that can be defined by many factors, including the objectives of launching such programmes; the surrounding political and socioeconomic environment; and the human and financial resources available.

Some Governments in the Arab region have already launched open data programmes at the city level (such as Dubai) or the national level (such as Tunisia and Saudi Arabia). When reviewing the performance and results of these programmes, certain observations can be drawn about the design and execution of successful and sustainable open data programmes.

Figure 7. Arab States in the Open Data Barometer

Source: Calculated from https://opendatabarometer.org/.
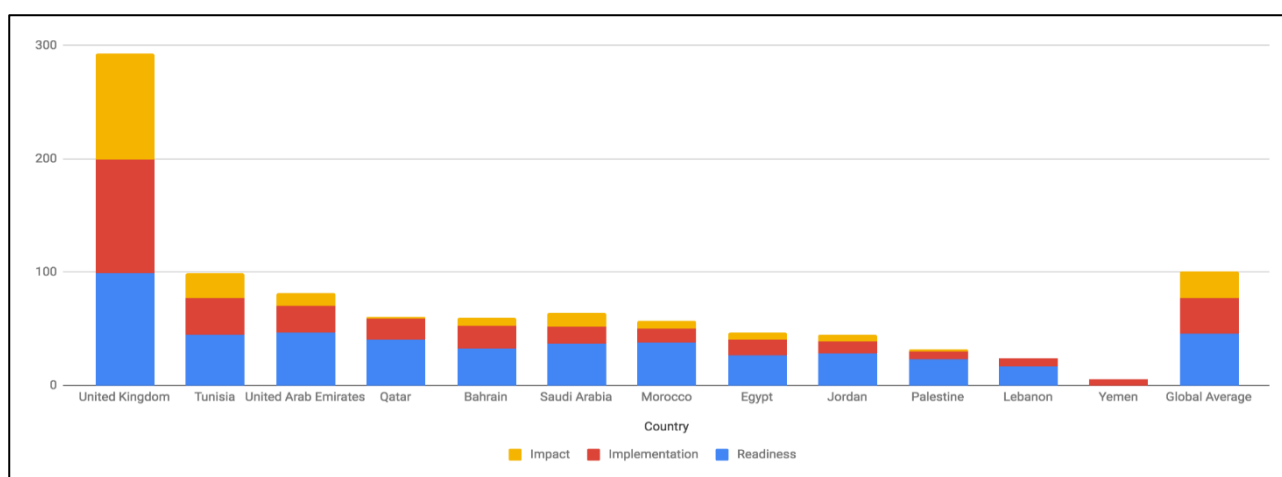
Figure 7 features an analysis of the performance of the 11 Arab countries covered by the Open Data Barometer and compares it with that of the United Kingdom, which tops the barometer's ranking. The global average was also included in the chart for benchmarking purposes. Based on the chart, the following key observations about open data programmes in the Arab region can be made:

- The total score of all Arab countries is below average, and very low when compared with the United Kingdom. This indicates that all open data programmes in the region are still at an early stage and that much work is still needed to improve the level of maturity of these programmes and to create the aspired value;

- Almost all of these programmes share the feature of achieving their best scores in the readiness subindex, which measures the readiness of Governments, citizens and entrepreneurs to secure the benefits of open data. This is followed by the implementation subindex, which measures the extent to which accessible and timely open data is published. The lowest scores were invariably in the impact subindex.[47] This may be an indicator that the programmes lack long-term sustainable design and implementation. A high level of readiness to release open data does not necessarily guarantee creating value or making an impact on the ground;

- It is of no less importance to take in consideration that the even scores in the readiness subindex can vary greatly between countries in the region. This reflects differences in local contexts. In the Gulf Cooperation Council (GCC) countries, for example, the readiness of the public sector (the data publisher) is generally higher than that of the civil society which constitutes a very important segment of the data users. In other Arab states (such as Tunisia), the situation is reversed. These observations indicate the importance of understanding local contexts and of designing and implementing plans according to that context.

B. PLANNING AND MANAGING OPEN DATA INITIATIVES: A STANDARD PLAN

In the following section, a standard plan is recommended for open data initiatives. The plan was developed based on global references that include a study carried out jointly by the United Nations Development Programme (UNDP) and the Partnership for Open Data(World Bank, Open Data Institute and Open Knowledge),[48] and the European Data Portal's Open Data Goldbook for Data Managers and Data Holders,[49] in addition to successful international open data programmes such as those of the United Kingdom, Australia and other countries. The plan was supported with examples from the Arab region and around the world.

While presenting the plan and its phases, the work needed in each phase will be described and supported by methods, tools and examples. While studying this recommended approach, it is important to consider the following:

- Though detailed in many of its part, the recommended approach remains general and mainly for reference purposes. It should be modified according to the needs and contexts of different organizations;

- While it is important to develop or improve the various elements explained throughout the phases of the recommended approach in localized open data programmes, they do not have to be developed in the same described order. For example, open data policy was introduced within the second phase of this approach, but certain Government organizations may find it more suitable to develop their open data policies in the first phase;

- The barriers that may face the programme, the need for innovation and key success factors presented at the end of this section should be analysed, and lessons should be extracted and applied.

The plan consists of the following four phases: preliminary phase; development phase; roll-out phase; and standard practice phase.

Table 3 provides a description of each phase and its elements.

Table 3. Four phases of the recommended open data programme plan

No.	Phase	Key tasks
1	Preliminary phase	- Running the baseline data maturity assessment; - Developing the open data strategy.
2	Development phase	- Developing the open data policy; - Creating the open data portal; - Prioritizing and publishing the datasets that should be released first; - Recruiting the central team for coordinating the open data initiative/programme; - Building up the human capacity.
3	Roll-out phase	- Publishing more high-value datasets; - Improving organizational capacity; - Promoting the use of the published open data.
4	Standard practice phase	Maintaining the sustainability of the open data programme.

C. PHASE 1: PRELIMINARY PHASE

The preliminary phase is quite significant considering that it lays down the foundations for the entire open data programme. In this phase, the organization should aim to secure the political and administrative support required to successfully implement the plan and to develop the business case for the open data initiative.

This phase usually starts with a senior decision maker authorizing exploratory work and normally ends with a political declaration to launch the open data initiative.[50] That was the case in Dubai, where an open data committee was formed in December 2014 by decree of the city's crown prince and chairman of the executive council.[51] The committee initiated the Government efforts in this domain which later evolved and lead to the establishment of the Dubai Data portal.[52]

The following key tasks are included during this phase:

1. *Baseline data maturity assessment*

This assessment provides the organization with a holistic and deep understanding of where it stands in terms of its readiness to launch and manage an open data initiative. For this purpose, the Open Data Maturity Model[53] is recommended as it enables Government organizations to assess the effectiveness of their operational and strategic activities in relation to the context and objective of the open data initiative. This multidimensional model also offers guidance on the potential areas for improvement.

In addition, if the organization (or country) has been evaluated by a national, regional or international open data assessment programme (such as the Open Data Barometer), such evaluation results should also be taken into consideration in the maturity assessment.

The outcomes and findings of the maturity assessment exercise form the baseline for carrying out the following tasks, and they should be used as benchmarks in future programme reviews and assessments.

2. *Developing the open data strategy*

As is the case with any strategy, an open data strategy is mainly concerned with defining the destination of the efforts to be made and designing the path towards it from the current situation. If the organization has a standard methodology or framework for designing its strategies, it can be used here.

In all cases, it is recommended to consider the following when developing an open data strategy:

- Making the process open and collaborative by communicating the strategy design plans to all stakeholders, including the public, and offering them continuous updates;

- Making it engaging by following an approach that allows the public to contribute to the design of the strategy through public consultation;

- Aligning the strategic destination with the higher organizational or national objectives and agendas.

An example of such a process is when the Irish Government, in June 2017, launched a public consultation on the draft of the country's open data strategy for the period 2017-2022.[54]

3. *Outline of the open data strategy*

In principle, the outline of a typical open data strategy document is no different than other strategy documents developed by the organization. However, considering certain special aspects of open data and the fact that different organizations might have different approaches and expectations of what is to be covered in the strategy, a standard outline is presented in the following with examples from the United Kingdom's "London's Open Data Strategy";[55] the United States' "Open Data Plan of New York City";[56] Australia's "Open

Data Strategy of Department of Health in Queensland";[57] and Canada's "Open Data Strategy of Edmonton City".[58]

(a)   *Context and background*

This section should offer a review of the local environment in which this strategy was designed and will be implemented. This provides a reference regarding the events that triggered the strategy and the circumstances that influenced its design and implementation. This should cover, at minimum, other related plans and activities of the organization, but optimally it should also cover other political and socioeconomic circumstances in the city or country. Context is a component of the strategy's starting point.

The data strategy for London (Data for London)[59] covers the context under the "Policy Context" and "Strategic Overview" sections. The two sections describe the data landscape in London prior to the strategy, and the strategy's linkages to other policy areas, including the energy policy and the broader smart cities agenda.

(b)   *Stakeholders and governance*

Identifying the stakeholders of the open data programme can be very helpful in understanding its context. Table 4 provides a list of some of the main stakeholders. It should be noted that the weight and significance of each stakeholder may differ between Arab countries depending on their local political and socioeconomic contexts.

Understanding the contributions, roles and responsibilities of different stakeholders is essential for designing the appropriate governance model and structure for the programme. For example, open data programmes in Dubai and Qatar adopted a governance model where a central Government organization leads the programme and coordinates with other involved organizations through a coordinator or champion.

Table 4. Key stakeholders in the open data programme

Stakeholder	Expected role
Political leadership	Securing support and sponsorship for the open data programme, and helping to overcome the barriers or challenges that it may face.
Public-sector organizations	Engaging continuously in the open data programme, and making sure that open data becomes an integral part of their day-to-day work.
Legislatures	Supporting the development of any enabling legislation needed at any stage of the open data programme.
Businesses/startups	Using the published data effectively and demonstrating cases of that use, and providing feedback to data publishers on their datasets, in addition to recommendations for new datasets.
Civil society/citizens	Engaging with the data publishers to guarantee that open data supports the practice of maximizing transparency and accountability.
Information professionals and librarians	Contributing with their knowledge and expertise in data management, curatorship and archiving.
Researchers and scientists	Identifying economic, social and other opportunities of open data. This leads to promoting and fostering a data-driven research culture.
Media	Engaging with the open data programme to promote it and deliver its message to all the stakeholders.

Source: https://www.per.gov.ie/wp-content/uploads/Draft-Open-Data-Strategy-2017-2022.pdf.

(c) *Strategic direction*

This can be shaped as the "vision" or the "mission" of the programme, and offers a clear description of the situation the programme aspires to achieve. In the case of Edmonton City in Canada, the open data strategy has the vision of "leading the way to an open, smart and resilient city".[60] The keywords in this statement (such as resilient) should shape the remaining sections of the strategy, especially the objectives and principles.

(d) *Objectives*

The objectives provide a quantifiable version of the strategic direction. Consequently, they have to bespecific, measurable, attainable, relevant, and timely (SMART).

In Australia's "Open Data Strategy for the Department of Health in Queensland", objectives grouped into themes. The theme of improving data quality has the following two objectives:

- Ensure data is released within set standards and in accordance with legislative and other protections;
- Review currently available data sets to ensure links are active.[61]

(e) *Actions*

Actions are the steps that need to be implemented to achieve the defined objectives. As a whole, the actions should move us from the current situation to the aspired one. In London's data strategy, an action states that "we will assess, plan and build a pan-London Centre of Data Innovation".[62]

(f) *Metrics for measuring progress*

In some organizations, these are defined as key performance indicators. Notwithstanding the name, they should be carefully selected and frequently measured (for instance, quarterly or annually). There are, for example, three key measures in London's Data Strategy, classified by theme into institutional capability, social capability and innovation capability.[63]

(g) *Open data release*

At least one of the objectives should define the organization's target for its open data inventory and catalogue. This should explain the number of targeted datasets and their domains or departments/sections.

In the Open Data Plan of New York City, a detailed list of the datasets scheduled for release between 2015 and 2018 was identified, with the name of the agency that owns each dataset among other details.[64]

(h) *Principles*

Principles are the set of commitments that guide the overall implementation of the strategy. They are usually aligned with the organization's overall values and the general principles of openness. This section enables the open data team to emphasize the key principles that are essential for promoting the culture of openness and participation.

Edmonton City, Canada, has set a number of principles, one of which states that "the City actively engages Edmontonians, non-profit organizations, businesses and the community to design and deliver programmes and services in an integrated and effective manner".[65]

(i) *Budget*

Budget estimates are necessary for building the business case and securing the support of decision makers. The most common factors that contribute to the cost of open data programmes include the following:[66]

- Training and capacity-building activities which continue throughout all the programme phases;

- External consultants who often have a higher day-rate than in-house staff;

- Conversely, external information technology (IT) resources, such as cloud computing, may be cheaper than in-house services.

(j)   *Other elements*

The strategy document may contain other sections, such as key success factors and themes under which the objectives and actions may be grouped.

D. PHASE 2: DEVELOPMENT PHASE

Upon completion of the preliminary phase and the development of the open data strategy, the second phase focuses on defining further details of the open data programme. The key tasks in this phase were described in table 3, the following is a detailed discussion of each task.

1. *Developing the open data policy*

A well written policy is an essential component for running a mature and sustainable open data practice. The policy should address both internal and external stakeholders with the following aims:

- Outline the details of the organization's commitment to publishing, sharing and consuming data;

- Help the internal stakeholders (such as the data owners) identify and prioritize releases of open data according to a standard process;

- Help the external stakeholders (such as developers) understand how the organization will be releasing its data and ways in which they can be involved;

- Offer a central reference for all stakeholders on how the various open data activities should be carried out and how the different stakeholders communicate and collaborate;

- Offer a common educational content of the essentials of open data (for instance, definition) so that all stakeholders have the same understanding and expectations.

Similar to the strategy, the open data policy should be developed through an open and engaging process. An example of this is the open consultation launched by the Ministry of Transport and Communications in Qatar on the draft version of the open data policy document. The objective of the consultation was to "seek feedback and comments on the proposed policy provisions from the Government agencies, the private industry and general public".[67]

2. *Outline of the open data policy*

Similar to the open data strategy document, different organizations may design their open data policies in different ways. A standard outline of a standard open data policy is compiled in the following, and explained through sample policies from the Arab region and beyond.

- Jordan: The Open Data Policy released by the Ministry of Information and Communications Technology (national level);

- Qatar: The Open Data Policy issued by the Ministry of Transport and Communications (national level);

- United States: City of Seattle Open Data Policy (city level);

- Australia: The Transport Open Data Policy released by the Transport for the New South Wales Department (department level).

The samples cover different countries and different levels (national, city and department levels) to demonstrate how some of the policy elements can be approached differently according to the governance level. The following presents the core elements recommended to be covered in the open data policy.

(a)   *Context and background*

At minimum, the context should cover the following:

- Reference to the strategy or any other programme that led to the releasization of the policy;

- Reference to any other legislations, policies or guidelines relevant to this policy or the practice of sharing Government data and information across the Government or with other stakeholders.

In Jordan, the open data policy released by the Ministry of Information and Communications Technology refers to Jordan's Third National Plan 2016-2018, under the Open Government Partnership (OGP) Initiative+ among other references.[68]

(b)   *Basics of open data*

This should, at minimum, cover the definitions of closed, shared and open data. Qatar's open data policy, for example, dedicates a section to definitions and acronyms that covers many open data concepts. In addition, the policy explains the 5-Star Linked Open Data model.[69]

(c)   *Purpose and scope*

This section should include an explanation of the objective of releasing the policy, and whether it was released to promote open data as a practice and culture, or to seek the commitment of target stakeholders to the delivery of a specific open data programme with clear targets and timelines.

The scope of the policy can be identified by defining the target Government organizations and NGOs expected to adhere to the policy (for instance, all Government department in a city) and by defining the types of data targeted by the policy (for instance, all Government data owned by these organizations). Clarifying the purpose and scope of the policy is essential for setting expectations and helping the target stakeholder cooperate and engage with the policy and the wider open data programme.

The Open Data Policy of the City of Seattle states that the policy "covers data created, collected and maintained by the City of Seattle or by contractors or third parties on behalf of the City of Seattle".[70] Another example is Australia's Transport Open Data Policy, released by the Transport for New South Wales Department identified five objectives for the policy, including the simplification and facilitation of the release of appropriate data.[71]

(d)   *Policy provisions or directives*

This is the core part of the policy and should cover the organization's key messages on how to run the open data programme. The following lists the common elements that a good policy should cover.

Governance: Defines the various stakeholders targeted by the policy and their roles and responsibilities in applying it. Also, the interrelations between stakeholders should be clearly explained.

In Australia, theOpen Data Policy developed by Transport for New South Wales lists the following departments as the parties to which this policy is applicable: Transport for New South Wales, Roads and Maritime Services of Sydney Trains, New South Wales Trains, and the State Transit Authority. Furthermore, the policy refers to the New South Wales Government Data and Information Custodianship Policy for specifics on how data and information should be managed.

In Qatar, considering that the scope of the policy is national, a different approach is followed. The policy requires every Government agency to assign responsibility to a senior officer to oversee the administration and implementation of the open data activities and the various responsibilities detailed in the policy. Moreover, the policy has detailed the roles and responsibilities of the central agency (which is the Ministry of Transport and Communications) and other agencies in a separate appendix.

Data selection and release: This sets the approach for identifying and prioritizing data for release, in other words, how the open data inventory and catalogue will be created and maintained.

In Qatar's policy, for example, agencies are encouraged to develop their own criteria for "prioritizing the opening of data assets, accounting for a range of factors, such as the volume and quality of datasets, user demand, internal management priorities, and agency mission relevance, usefulness to the public, etc.". To help agencies in this endeavour, the policy recommends using the key datasets identified by the Open Data Charter as a reference point.

Data licensing: This should cover the description of the licence under which the datasets will be published, in addition to the process of clearing rights during data collection.

The open data policies of Qatar, Jordan and Seattle require the use of open licence and define that licence. However, the open data policy of Transport for New South Wales goes a step further and recommends Creative Commons as a standard licence to be used.

Privacy and security considerations: Defining what is considered personal data that can be used to identify an individual, and thus should not be included in the published datasets. Open data programmes in the Arab States must give special attention to this issue considering that many countries still lack the legislation to regulate it.

Qatar's open data policy has a detailed definition of personal data, in addition to a clear article stating that "data must be free of any personal information that may directly or indirectly identify an individual. Data should be anonymized or pseudonymized before being published". Moreover, other articles across the policy addressed the issue of personal data.

The open data policy developed by Transport for New South Wales, however, reflects the legal framework that addresses the issue. It requires the removal of "personal or identifying information from datasets in line with the Government Information Public Access Act, the Privacy and Personal Information Protection Act 1998 (NSW) or other applicable legislation".

Data quality and publishing standards: This includes the recommended standard formats, and the use of metadata among other standards. As the majority of the open data programmes in the Arab region are still in emerging phases, it is recommended to set clear and specific recommendations that the data publishers can adopt without confusion.

All the policies referred to in this section address this topic in varying levels of details. In Jordan, the policy addresses the issue at a broader level under the "Open Data Principles" section, and requires publishers to create the needed processes to ensure adherence to the stated general principles. The policy in Qatar, however, provides more detailed measures, and requires all datasets to be published in an open format, listing the formats that are considered open (for instance, CSV, Jason). The policy also defines metadata and requires it to be maintained without providing specifications on what metadata should include.

Review and measuring progress: This entails deciding on the metrics that will be used to measure the successful implementation of the policy and should be accompanied by a clear timeline and review process.

The open data policy in Qatar contains a clear list of targets that all Government agencies need to accomplish and a clear timeline (in months) for each target. One target, for example, states that each Government agency is required to develop its own open data plan within five months.

Engaging with the data users: The users in this context are the external stakeholders (researchers, developers and the general public). This engagement improves the quality of the published data, helps in guiding its release and ensures its ease of use.

The open data policy of the City of Seattle allocates a section for the stakeholder engagement where it offers recommendations on how this engagement should be carried out.

(e)   *Other elements*

The policy document may offer some standard managerial or technical templates or tools.

3.   *The open data portal*

A typical open data programme requires a portal, which is a central online destination where datasets are published so they can be accessed.

Further details on the open data portal and what to consider in launching it are provided in chapter 4 of this guide.

4.   *Prioritizing and publishing the first wave of the datasets*

The first version of the portal should be populated with high value datasets that are limited in numbers but highly credible. It is recommended to pick the datasets that are relatively easy to prepare and publish, and to avoid, at this stage, the publication of those that are controversial or that raise difficult policy questions.

Chapter 3 of this guide offers full details on the process of prioritizing and publishing datasets.

5.   *The open data team*

Most successful open data initiatives have a small central team, a task force that leads and coordinates the work within the organization and also across other involved Government organizations and NGOs if the programme was at the city or national levels. Therefore, the hierarchy and size of the team may differ from one programme to another. However, there is a common range of managerial and technical responsibilities that should be handled by all open data teams and that require a specific set of skills and experiences.

Table 5 lists the recommended job titles and their responsibilities. While viewing the roles suggested in the list, the following should be taken in consideration:

- The list is not inclusive. It merely provides a starting point reference for deciding on the formation of the open data team;
- There are no one-to-one relationships between roles and team members. For example, the same member can act as the portal manager and data publisher. In other cases where the work load is high, there may be a need for more than one data publisher. It all depends on the organization's needs and availability of resources, and both can change over time;
- Certain roles may be more significant and crucial than others depending on the needs of the organization and the maturity of the open data programme. For instance, when launching a new nation-wide programme, the coordinator's role is crucial to the successful collaboration between the central team and other organizations involved;
- The central team is usually overseen by a steering committee or advisory board within the Government.

Table 5. Suggested roles for the open data team

Role/title	Suggested responsibilities
Programme sponsor	Secures political and legislative support and approvals as necessary.
Chief data officer	Leads the entire programme with its various tracks, including the development of the open data strategy, policy and data prioritization and release.
Legal advisor	• In charge of the process of drafting the suitable open licence; • Leads the development of any necessary legislation.
Portal/IT manager	• Sponsors the development of the open data portal; • Provides advice and guidance on agency-specific technological issues.
Subject matter expert	• Provides the detailed technical knowledge of open data; • Participates in the regular review activities to ensure that datasets are published and updated as required.
Coordinator	• Leads the communication work of the initiative within the Government, including briefing political sponsors and raising external awareness; • Leads the communications with stakeholders and partners outside the Government, including the civic community; • Manages the process and participation in activities; • Arranges suitable training and capacity-building activities; • Acts as the programme's focal point for partners and stakeholders.
Data publisher	• Publishes data on the data portal as required by the approved process and quality standards.

Sources: https://www.gartner.com/smarterwithgartner/understanding-the-chief-data-officer-role/; https://theodi.org/article/how-to-plan-and-budget-an-open-data-initiative/; and https://data.sa.gov.au/sites/default/files/Toolkit/Open-Data-Process-Guide.pdf.

Box 6. Allocation of roles and responsibilities in Dubai and Qatar

The open data programmes in Qatar (where the programme is run at the country level) and Dubai (at the city level) follow a model where a central open data team leads the work and coordinates with all Government organizations. Each stakeholder has a certain well-defined role in advancing the programme's strategy.

According to this model, each Government organization appoints one of its staff members as the open data lead or coordinator to oversee open data activities within the organization, and to coordinate with the central team.

Examples of the tasks assigned to each side in both programmes include the following:

Central open data team:

• To provide leadership and management for the whole open data programme;

• To update the open data strategy and policy from time to time to address new or emerging needs and trends;

• To develop the open licensing scheme for the entire open data programme;

• To lead and manage the communications strategy and awareness campaigns and activities across governmental organizations and other stakeholders.

Organization's open data team:

• To prepare and execute the open data plan in accordance with the guidelines and targets set out by the central team;

• To create and maintain the open data inventory and catalogue as per the provided guidelines and targets;

• To establish the needed internal accountabilities and governance processes;

• To report periodically the progress made and the challenges faced by the central team.

_____

Sources: http://dubaidata.ae/pdf/Dubai-Data-Manual-20160612.pdf; and http://www.motc.gov.qa/sites/default/files/open_data_policy_0.pdf.

6. *Capacity-building*

Training and capacity-building should continue throughout the entire lifetime of the open data programme. While some initial open data training and orientation activities may be carried out during the preliminary phase, it is important to design a comprehensive and ongoing training programme for the open data team and public servants to drive the initiative forward and to spread an open data culture.

The actual skill gap and training needs should be identified in the data maturity assessment conducted in the preliminary phase. Nevertheless, examples of the skills needed may include the following:[72]

- Basic concepts and practices of open data;

- Managing the open data portal;

- Training relevant staff in the necessary skills to publish data and in the use of the portal software;

- Training data publishers in the skills necessary for finding high-value datasets and prioritizing their release;

- Skills related to legal, privacy and security aspects;

- Skills needed for successfully engaging with the public and stakeholders.

The training programme may include study visits to countries and organizations with advancement programmes and initiatives in the open data domain.

Box 7. Capacity-building initiatives

Different Governments around the world have launched several versions of capacity-building programmes that aim to provide public servants with the data skills they need and to spread the culture of open data.

In Singapore, the Government launched a Government-wide programme to provide 10,000 public servants with data science training to quicken the Smart Nation efforts. The programme which was announced in 2017 is being delivered in partnership with the National University of Singapore.

In Australia, the federal Government has launched a reference framework for the data skills and capability in the Australian public service. The framework identified certain roles across the public service that require more specific data skills, and offered guidance for public-sector organizations on how to upskill their employees in that regard.

In the United Kingdom, the Open Data Institute (ODI) has developed the Open Data Skills Framework, a simple, three-tier framework that describes the knowledge and skills needed for interacting with open data, from the beginner through to the expert level. The skills framework enables learners to identify where they stand in their learning journey and how to proceed in improving their skills.

In the United Arab Emirates, the Mohammed bin Rashid School of Government has collaborated with the Open Data Institute Dubai (ODI Dubai) in conducting a series of open data training programmes for public servants in the federal and local Governments.

_____

Source: https://www.straitstimes.com/singapore/10000-public-servants-to-receive-data-science-training-under-govtech-nus-tie-up; https://www.pmc.gov.au/sites/default/files/publications/data-skills-capability.pdf; https://theodi.org/article/open-data-skills-framework/; and http://www.mbrsg.ae/HOME/EXECUTIVE-EDUCATION/Open-Enrolments-Programs/Open-Data.aspx.

E. PHASE 3: ROLL-OUT PHASE

The main focus during this phase is on publishing increasing amounts of open datasets. This requires improving the internal capacity of the open data team and the organization(s) involved and promoting the use of the published data by the public and the target users.

Examples of the key areas organizations may need to focus on during this phase include the following:

- The processes surrounding the prioritization and release of open data need to be improved to cope with the increase in the number of published datasets and to respond properly to the needs and expectations of the data users;

- The open data team will need to handle some difficult (or precedent-setting) cases of data release, including privacy and anonymization, among other issues;

- The open data portal may need to be improved to add features that were not available in its initial version. This decision should be taken in the light of the feedback from the open data team and the users;

- Developer-oriented events may be carried out, such as hackathons, to promote the use of the published datasets. Hackathons are design sprint-like events that gather both data suppliers and teams of application developers in order to generate examples of data reuse;

- Data innovation challenges and competitions outside the developers' community may be carried out as well to promote the innovative use of data by different segments in the society.

F. PHASE 4: STANDARD PRACTICE PHASE

By this phase, the open data programme should have reached a stage where the publication and update of data are parts of the normal business and standard procedures of Government organizations. However, a smaller open data team or unit may still be needed to ensure that the policy and processes keep up with requirements and that the open data portal is working effectively to meet its objectives.[73]

The main focus should be on ensuring the long-term sustainability of the open data programme, and this is likely to require the following tasks:

- The frequent review and update of the open data standards and references (for instance, the policy);

- The update and maintenance of the open data portal;

- Continuing the publication of datasets in alignment with the prioritization plans and user needs;

- The frequent and continuous evaluation of the open data programme to assess the level of progress achieved and to identify areas for improvement;[74]

- Maintaining engagement with data users inside and outside the Government.

G. BARRIERS IN PUBLISHING OPEN DATA

Throughout the different phases of the open data plan, the open data team should expect to face a number of barriers that may hinder or challenge the process of data publications. Table 6 lists a number of common barriers and their recommended remedies.

Table 6. Barriers to publishing open data and their remedies

Barriers	Recommended remedies
Data is not interesting	Focus should be on publishing high-value datasets and then letting the users judge how interesting or useful they are. Even niche datasets have people that care about them. It should be remembered that "open by default" is one of the open data principles.
Questioning the purpose or benefit for our organization	The value of open data has been demonstrated and documented in many cases in different countries around the world.
Too many user requests on our data will be received	On the contrary, when raw data is provided, end-users will extract and combine the data according to their needs. This will potentially reduce the number of requests.
Users will draw superficial conclusions from the data	This risk exists in any form of publication. Generally, the data is accompanied by supplementary documentation which provides the context.
Data is not sufficiently accurate to be shared	Improving the data quality and the whole organization data maturity takes time. Publishing data can help through receiving feedback from the users and improving the quality of the data.
It is too costly to put the data into a standard format	With good planning, the production of or conversion to an open standard format requires little effort.
Will there be a risk of getting a negative reputation?	On the contrary, publishing open data can enhance the image of openness and transparency of the organization in addition to creating the desired public value.
Publishing low quality data will harm the image of the Government/organization	User feedback provides an opportunity to improve the quality of data. Open communication should be maintained while publishing data.

Source: https://www.europeandataportal.eu/sites/default/files/european_data_portal_-_open_data_goldbook.pdf.

H. SUCCESS FACTORS

While working on the design and implementation of the open data plan, the open data team should consider the following key success factors derived from international practices:[75]

- Strong and sustainable leadership and sponsorship for the programme;

- Appropriate governance structures in place within public sector bodies to ensure success. This includes the necessary support at senior level and the designation of an open data liaison officer in each organization;

- Implementation of the open data strategy will require collaboration and cooperation between data producers and data users to ensure that resources and efforts are targeted where most value will be derived;

- Focusing on creating impact and value to all stakeholders as per the goals defined in the open data strategy;

- Publication of high-quality datasets underpinned by the national quality standards and framework and aligned to international best practices to ensure cross-border inter-operability;

- Allocating resources for capacity-building, technical support, training support and to provide expertise as required to enable knowledge-sharing;

- Engagement with stakeholders to ensure maximum data use and reuse;

- Promotion of the culture of open data and the principles associated with it such as openness, collaboration and engagement;

- Commitment to tracking, understanding and sharing progress openly with all stakeholders, including the general public;

- Promoting innovation: It is essential for open data teams to think strategically about their own organizations and local contexts and find ways to innovate in designing and implementing the entire open data programme to create the desired impact on their organizations and communities. In addition, they need to innovate in solving the challenges that they might face (box 8).

Box 8. Open data for everyone

One of the common challenges facing open data programmes is the public perception that open data is mainly a technological subject. Therefore, people from outside the technology industry and community might feel less encouraged to engage with open data programme, although this can negatively affect the scope and depth of the impact of the programme.

One reason behind this perception is that open data programmes are usually incubated in e-government programmes, as is the case in almost all Arab countries. In Morocco, for example, the open data programme is launched under the e-Government Programme, and the same applies to Dubai, where Dubai Data is under Smart Dubai. Two cases from Singapore and the United Kingdom can provide solutions on how to get more people from outside the technological sphere to use open data, and thus maximize the impact and value being created.

In the United Kingdom, the Open Data Institute (ODI) has launched the Data as Culture programme back at an early stage of its foundation. The Data as Culture art programme "engages new and diverse audiences with work by artists who explore data critically and materially". Since 2012, the programme commissioned internationally renowned artists and artists in residence for ODI's headquarter in London, for other organizations and galleries, and public spaces. The art programmes and partnerships have reached hundreds of thousands of people. The data-driven artworks have included a semi-sentient vending machine, data collection performances, photographs, networked artworks, and pneumatic machines.

In Singapore, while the country was approaching its 50th anniversary in 2015, Infocom Development Authority launched the Data in the City Visualization Challenge. In this competition, all citizens were invited to use the open datasets to tell their Singapore story with creative use of this data. Many people from different walks of life participated in the challenge and created sophisticated data visualizations that reflect 50 years in their nation in very different ways.

_____

Sources: http://culture.theodi.org/; and https://e27.co/singapore-data-city-visualisation-challenge-20140918/.

III. Open data inventory and catalogue

III. OPEN DATA INVENTORY AND CATALOGUE

A. INTRODUCTION

As previously discussed in chapter 1, Government organizations own and continue to produce huge amounts of data. Although Governments have, in many instances, invested political and financial resources and efforts in data, they still often struggle with legacy databases and their incompatibility. Governments hold datasets relating to different sectors with diverse characteristics, some of which are personal and even sensitive. Furthermore, the utilization of Government data is often hampered by the constraints of traditional storage methods and old technology.[76]

Government organizations, therefore, face a number of challenges in the context of open data, including the selection and categorization of the right data to publish and prioritizing the publication process in alignment with their resources and public demand. This chapter aims to help open data teams in addressing these challenges through providing guidelines for creating open data inventories and catalogues, which is an essential practice in any successful open data programme.

B. DEFINITIONS

The terms "inventory" and "catalogue" are sometimes used interchangeably. They are, however, defined as follows for the purposes of this guide: An open data inventory of a Government organization is the master record of all high-value datasets owned by that organization and are candidate to be published as open data in a standard format. An open data catalogue contains datasets that have already published as open data; in other words, it forms only a subset of the inventory.

Figure 8 shows, as an example, the open data catalogue of the city of Toronto in Canada. The catalogue, as of its last update on 20 July 2018, contains 275 datasets although one can expect that the city has more datasets that are not published yet. Over time, more datasets are being selected from the open data inventory to be published online on the open data catalogue. The metadata of the catalogue last updated on the portal aims to inform users on the last time the catalogue was updated, which most likely means the last time new datasets were added to this catalogue from the inventory.

Figure 8. Open data catalogue of the city of Toronto

Source: https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/.

Similarly, Morocco's open data catalogue has only 136 datasets (figure 9), while it is expected that the Government has many more datasets that will be added in the future.

Figure 9. Open data catalogue of Morocco

Source: http://www.data.gov.ma/data/fr/dataset.

C. BENEFITS OF THE OPEN DATA INVENTORY AND CATALOGUE

When built and maintained properly, the open data inventory and catalogue can deliver the several benefits for both Government organizations, namely, the publishers, and the general public, namely, the users. The following list includes some of those benefits.

The open data inventory:

- Enables the Government organization to set the boundaries of the landscape of the high-value data that can be published as open data;
- Can be used to design a roadmap for publishing open datasets in alignment with the organization's priorities and resources;
- Can be used by the open data team to keep track of their progress in publishing the potential datasets as open data.

The open data catalogue:

- Facilitates user access to published datasets, especially when it has features like search and tags;
- Enables the open data team to automate some of the data publication and update tasks.

44

There are several approaches to creating open data inventories depending on the level of governance. Box 9 demonstrates this through examples from the Governments of Jordan and Qatar and the City of Seattle in the United States.

Box 9. Sample approaches for creating open data inventories and catalogues

Qatar's national open data policy targets all Government agencies in country and aims to provide them with specific actions to ensure that they all make their data available to the general public in accordance with the policy's provisions.

The policy does not clearly distinguish between an inventory and a catalogue and provides Government agencies with general guidelines on how to create them. It requires each Government agency to conduct a review and identify the datasets to be released and to maintain an open data inventory/catalogue. The policy adopts the key datasets identified by the G8 Open Data Charter as reference points for Government agencies to consider when creating their open data inventory/catalogue, while advising them to prioritize opening up their datasets at their discretion.

A similar general approach is adopted by Jordan's open data policy. The policy requires Government entities to "conduct an inventory of existing data early in the process of development of their respective open data policy in order for the Government and other stakeholders to be aware of the full potential dimensions of data release". It also requires Government entities to make the inventory public and to assign an individual or a group to ensure the ongoing maintenance and accuracy of the inventory.

The City of Seattle adopts a slightly more detailed approach in its open data playbook. While still recognizing the differences between departments in terms of how each one is managed, how much data it owns and the characteristics of that data, the playbook offers the departments the following guidelines to help them create their inventory:

- A general three-step process;
- A set of guiding questions to use in the process of dataset selection;
- A template for logging the datasets in the inventory.

_____

Sources: http://www.motc.gov.qa/en/documents/document/open-data-policy; http://www.mpwh.gov.jo/Pages/Open-Government-Data-policy.pdf; and
https://www.seattle.gov/Documents/Departments/SeattleGovPortals/CityServices/OpenDataPlaybook_Published_2016.08.pdf.

In this guide, the following four-phases process is recommended for creating the open data inventory and catalogue:

Phase 1: Preparing a master list of all high-value datasets owned by the organization.

Phase 2: Classifying each dataset in the master list to create the open data inventory.

Phase 3: Releasing the open datasets from the inventory to the open data catalogue.

Phase 4: Updating both the open data inventory and catalogue.

As highlighted in the introduction of this guide, the local context of the organization and country should be taken into account while adopting this process.

E. PHASE 1: PREPARING THE MASTER LIST

The open data team at the organization should run brainstorming sessions to identify the master list of high-value datasets. The high-value datasets are the ones that are more likely to make an impact once released as open data. This can be done through identifying the areas where data can be exploited for solving problems,

improving efficiency and effectiveness and collaborations, and draw a pathway for what data should be published.[77] The desired impact should be measured in alignment with the strategic agenda and objective of the Government organization as defined in its open data strategy.

The SDGs provide a universal reference for selecting and publishing high-value datasets. Table 7 lists potential high-value datasets that are recommended by international open data practices. Most Arab countries have yet to publish the majority of the datasets in the table, so they should represent a good starting point for them.

In addition to the datasets listed in the table, the central team may consider the datasets that fulfils any of the following criteria:

- Datasets that are aligned with the country's national priorities and agenda;
- Datasets that are aligned with the organization's strategic objectives and priorities;
- Datasets that are in high demand by the general public.

Table 7. Recommended list of high-value datasets

Sector/category	Subsector/subcategory	Sample datasets
Education	Education providers	- Performance of education providers; - Locations of education providers; - Education providers' sanitation data.
	Public libraries	- Bibliographies of national libraries; - Library locations.
	Budgets and finance	- Government spending on education; - Budgets per locality or school.
Health care	Health-care providers	- Address and geocoded locations of registered health-care providers (doctor practices, clinics, hospitals, pharmacies, etc.). This dataset can also contain contact details; - Health-care services offered by each provider; - Quality and performance level of each provider based on relevant performance measures and opinion surveys.
	Disease, vaccination and prescription data	Data on the prevalence, risk and treatment of diseases and illnesses, by locality where possible.
	Budget, spending and contracting data	Data on public-sector finances, which may include public-sector spending for each sector, by locality and by public body (at an individual payment level) and budget, grant and contract data by department.
Transport	Public transport timetables	- Timetables for various forms of public transportation (trains, buses, trams, ferries, etc.); - Prices for various forms of public transportation (trains, buses, trams, ferries, etc.).
	Public transport location data	- Geocoded location data for routes, stops and terminals for the various forms of public transport; - Accessible features available in the public transport facilities (e.g. ramps and tactile tiles at platform edges).
	Traffic flow information	- Flow rates of traffic at specific locations; - Geocoded locations of traffic incidents.
	Budget, spending and contracting data	Data on public sector finances, which may include public-sector spending for each sector, by locality and by public body (at an individual payment level), and budget, grant and contract data by department.

Sector/category	Subsector/subcategory	Sample datasets
Environment	Weather and pollution	• Real-time and historic data of temperatures, rainfall, wind speed and other observations; • Forecast data; • Data on various types of pollution including, but not limited to, a pollution inventory (regulated and otherwise) and pollutant concentration.
	Soil data	Data on the state of soil, such as contaminated land locational data and expected soil conditions and nutrients.
	Cadastral map	A map that shows the precise location (using GPS coordinates), dimensions, boundaries and ownership of land parcels, which may also include additional details such as unique identifying numbers, and certificate of title numbers.
Sanitation and waste	Waste disposal sites	Geocoded locations on waste disposal sites.
	Toilet facilities	Geocoded locations of toilet facilities which may include the type of toilet facilities available where possible.
Governance	Elections	• Geocoded locations of polling stations; • Election results.
	Finance and contracts	• Transaction spending; • Tenders and contracts.
Economy and business	National and local accounts	• National budget (planned and spent); • Local budget (planned and spent); • Company/business register.
Crime and justice	Crime	• Geocoded locations of crimes which may include the type of crime and an outcome (e.g. an arrest) where possible.

Sources: https://theodi.org/article/how-to-prioritise-open-data-to-drive-global-development/; and
https://www.diplomatie.gouv.fr/IMG/pdf/Open_Data_Charter_Annex_FINAL_13_June_2013_cle0ff8a3.pdf.

F. PHASE 2: CLASSIFYING DATASETS IN THE MASTER LIST

In principle, Government data should all be open by default. This implies that Government organizations should assume that all data and information is open, and should identify whether there is any pressing reason for not making data or information open. This is different from the current common practice, where the public has to ask Government officials for specific data or information to be released.

However, not all the datasets collected in phase 1 and listed in the master list are eligible for immediate release to the general public as open data. This can be due to any of the following reasons:

- Legal: Releasing the dataset might violate an existing law or regulation;

- Security: Releasing a dataset might threaten the national security;

- Privacy: A dataset may contain personal data that can be used to identify individuals once the dataset is released.

Therefore, the main purpose of this phase is classification. The open data team needs to run a classification exercise to check if any dataset in the master list should not be published as open data. The remaining datasets should be marked as open data and represent the open data inventory.

The data spectrum

The data spectrum is a tool that can aid Government organizations in the data classification and inventory-building exercise, and, thus, it is recommended to use it to better understand and implement this classification process.

As shown in figure 10, the data spectrum provides a visual representation of data classification. It emphasizes spectrums of data rather than absolute and rigid categorizations. This is done through classifying data based on its legal status which determines its location on the spectrum, so data can be 'closed' on the far left side, 'open' on the far-right side, or somewhere in between.

Figure 10. The data spectrum
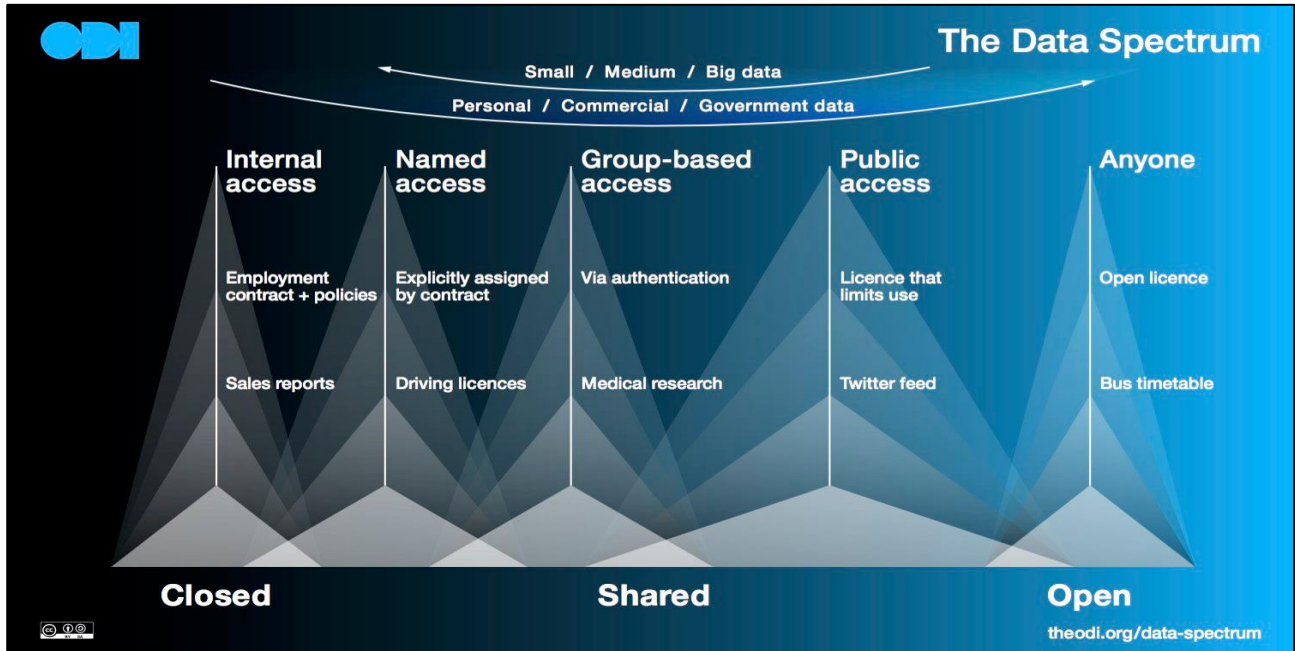
Given its simplicity and accessibility, the data spectrum is applicable to all data regardless of the type of organization that owns that data, and it serves as a reference point for Government agencies, businesses, citizens, and stakeholders alike to have informed conversations on how data can be used and the implications of different categorizations.

To elaborate further, the following provides brief definitions of the key classifications shown in the data spectrum in figure 10.

Internal access: This classification, on the far left of the spectrum, relates to closed data that can only be accessed by some staff in the Government organization and is never to be shared outside it. The reason is that the data contains personal elements A good example are employee contracts, which usually contain some personal information such as salaries, and thus should not be open to public access.

Named access: Moving from left to right on the spectrum, data begins to be shared outside the Government agency. Data labeled as "named access" can only be shared with named people or organizations outside the agency. An example of this are driving licences, which contain personal data about drivers that they alone should have access to.

Group-based access: This is a more relaxed status where the data is shared with a group of people outside the agency and not only with specific individuals, as is the case with named access. The group can be identified based on their job titles or the organizations they work for, or any other factors that necessitate access to the data.

An example of this are patient records in hospitals. Although these records contain personal information, they can be shared with insurance companies outside the hospital when needed or with pre-approved group of practitioners in the medical field.

Public access: Data that is available to anyone under terms and conditions, so it is not yet open data. An example of this is Twitter feed anyone can request access to for purposes such as the development of an application that uses the published tweets, but Twitter will only grant access under certain limitations.

Anyone: To the far right is the data that can be described as open data as defined in the first chapter of this report. This is data that anyone can access, use and share. An example is the crimes dataset on the United States data portal.[78]

For an easier and more practical implementation of the data spectrum, Government agencies can just use the three key categories: closed, shared and open. Closed data in this categorization refers to data that is available exclusively to a particular individual, team or organization. Shared data refers to named, group-based, or public access data. And finally, open data is that which anyone can access, use or share, as defined in chapter 1. Some examples from Arab countries are presented in box 10.

Box 10. Data classification in samples from Arab countries

In the United Arab Emirates, the Dubai Data Manual adopts an approach similar to that of the data spectrum, mainly with respect to open and shared data. The manual offers "a suite of standards and best practices that describes how Government entities should manage their data in accordance with the Dubai Data Law". Dubai Data is defined as any data relating to any aspect of the Government, economy, culture, and life within the Emirate of Dubai.

Open data is defined by the manual as "data provided by the Government or private sector entities to individuals, to be freely or subject to a minimum limit, used or exchanged with third parties". Shared data is also defined as "data owned by Government entities […] made available for sharing and reuse by other Government entities, with appropriate controls". A dataset can be classified as shared if it fits one of the following categories:

- Confidential: Shareable across Government entities according to professional responsibilities;
- Sensitive: Shareable within certain groups and subject to strict controls;
- Secret: Shareable in a limited way between certain individuals and under strict controls.

In Qatar, the open data policy provides guidelines for Government agencies on how to classify data before releasing it. According to the policy, agencies must conduct an open data assessment to ensure that the data "is free from law, policy, regulation or national security concerns, or because the data are subject to privacy, confidentiality, security, trade secret, contractual, or other valid restrictions to release".

In addition, the policy includes an open data assessment checklist which covers the following points:

- Privacy: The dataset must be free of any personal information that may directly or indirectly identify an individual. If this is not the case, then the data should be anonymized or pseudonymized before being published;

- Copyright: The dataset must include only content created and owned by the Government of Qatar, and there should be no exclusive licence for another party to use or access the materials;

- Legal, contractual or policy constraints: The release and use of the dataset should be permitted under the law and related policies. In addition, there should be no legal, contractual or policy restrictions or limitations to the public release or use of the dataset;

- Security controls: The release of the dataset should be compliant with the national information assurance, and the information security officer in the agency should be contacted for this purpose;

- Pricing frameworks: The dataset should be available to the public without collecting a fee;

- Data source: The agency should have the primary responsibility of the dataset. In addition, the data should be complete (meaning that no subset of the data was excluded) and provided in machine-readable format.

_____

Sources: http://dubaidata.ae/pdf/Dubai-Data-Manual-20160612.pdf; and
http://www.motc.gov.qa/en/documents/document/open-data-policy.

The size of the data inventory does not matter in this categorization as even the so-called big data can be examined using the data spectrum. Big data can be defined as datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse; however, it is worth noting that defining this concept is still subjective given the fluid nature of technological advancement, which changes over time what can be considered big data.[79]

G. PHASE 3: RELEASING DATASETS FROM THE INVENTORY TO THE OPEN DATA CATALOGUE

In the previous phase, the organization would have created an open data inventory containing a pool of datasets that are marked as open datasets. These are the datasets qualified for release to the public. In the current phase, the aim is to release the datasets to the public through publishing them online on the open data catalogue. To do that, it is recommended to implement the following three process on each dataset, namely selection, preparation and publication.

1. *Dataset selection*

While all the datasets in the open data inventory should be of high value, the open data team may still need to prioritize the release of these datasets in alignment with the current demand from both the general public and the Government, and also according to the organization's internal capacity for dataset release and maintenance.

For example, the city of Buffalo in the United States updates its crime incidents dataset on a daily basis.[80] This is an example of a dataset with a high frequency of updates, compared to low frequency datasets such as school geographic locations, which may not be updated more than once a year.

If the Government organization does not have the required resources or experience to commit to the daily update of a dataset, then it is recommended to give proirity to datasets that require high frequency of updates.

2. *Dataset preparation*

This is the step where the open data team needs to carry out most of the work. For each dataset, the team must at least prepare the following so that the dataset is ready to be published:

- Ensure the application of the appropriate open data licence;
- Use the appropriate open data format;
- Update the appropriate metadata;
- Check the dataset against the organization's standard quality criteria.

The following explains further each one of these requirements.

(a) *Open data licence*

As mentioned previously, the key pillar of the data spectrum is the legal status of the data and how explicitly that status is clarified. The legal status of data can be encapsulated in the licence of the dataset. The data licence centrally focuses on how permissions are granted to users for the use or extraction of the published and owned data of an agency or entity. In other words, it involves the transfer of ownership and the rights associated with it.

With regards to open data, and open government data in particular, licensing basically involves placing very few restrictions on how the content or data can be used. This entails processes such as republishing content or data on websites, deriving content from other entities, reaching untapped potentials of data and exploiting its potential and republishing content or data for a particular set fee.

According to different national and international rules and regulations, licensing may come under different categories and/or levels. It may either be a public domain licence with no restrictions, which technically waives rights to content or data, or an attribution licence that states explicitly that users must attribute data used to the owner or publisher of the data, or an attribution and share-alike licence which states that users must give attribution and share any derived content or data under the same licence.[81]

The following list shows the different ways a Government organization can acquire the open licence required for its open data:

- Self-developed licence: a licence that is developed by the organization's own team for the sole use on its open data page or portal;

- Government-wide licence: a licence that is developed by a central agency in the public sector and used by all agencies across the public sector;

- Creative Commons licence: The Government agency chooses and applies the suitable Creative Commons licence. Creative Commons is a global non-profit organization that enables sharing and reuse of creativity and knowledge through the provision of free legal tools, including a variety of licences.[82]

To demonstrate these three approaches, box 11 presents examples from the United Arab Emirates, the United Kingdom and Australia.

It is strongly recommended to avoid the self-developed licence approach when possible, and adopting one of the two other approaches instead. This is because when different Government organizations self-develop their open data licence, this may lead to inconsistency across the Government open data. This can cause confusion to the users and hinder the value of the datasets. Adopting a central licence for the entire Government can also be more cost-effective.

Box 11. Approaches to open data licence

The Ministry of Interior in the United Arab Emirates has developed its own licence for open data. The licence is presented on the open data policy page shown in the figure in the box. This is an example of a self-developed licence.

The Open Data Policy page at the website of the Ministry of Interior, United Arab Emirates

---

In the United Kingdom, the National Archive developed the Open Government Licence as a tool "to enable information providers in the public sector to license the use and reuse of their information under a common open licence".

When a dataset is published under this licence, the user is free to do the following:

- Copy, publish, distribute, and transmit the information;

- Adapt the information;

- Exploit the information commercially and non-commercially, for example by combining it with other information, or by including it in their products or applications.

The user always has to "acknowledge the source of the information in their product or application by including or linking to any attribution statement specified by the information provider(s) and, where possible, provide a link to this licence".

The Australian Government follows a similar central approach but has not created its own licence. Instead, the Government uses Creative Commons. The Creative Commons licence adopted in Australia specifies that users are free to do the following:

- Share, namely, copy and redistribute the material in any medium or format;
- Adapt, namely, remix, transform and build upon the material for any purpose, even commercially.

This should be done through ensuring attribution by giving appropriate credit and indicating if any changes are made. Additionally, it specifies ensuring that no additional technological measures are applied that would restrict others from using data as the licence permits.

_____

Sources: Creative Commons Australian Government Licence, (no date); http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/; and https://www.moi.gov.ae/en/Open.Data/genericcontent/open.data.policy.aspx.

(b)   *Dataset format*

The format of an open dataset refers to the way in which the data is structured and made available for humans and machines, which includes three elements: format, structure and delivery mechanism.[83]

Open data should be published in a format that enables anyone to access, use or share it. This means that the data should be non-proprietary and in a machine-readable format, regardless of which operating system or licensed software the user has access to.

There is a wide range of file formats that can be used for publishing datasets, table 8 lists the most common of them.

Table 8. Popular open data formats

No.	Format name	Definition	Type of data to use this for
1	Comma-separated values (CSV)	CSV is a great way for storing large amounts of data with just commas separating the data values. CSV files often contain a header with names describing what data populates the file.	Tabular data
2	Tab-separated values (TSV)	TSV is a very common form of text file format for sharing tabular data and is highly machine-readable.	Tabular data

No.	Format name	Definition	Type of data to use this for
3	JavaScript Object Notation (JSON)	JSON uses human-readable text to transmit data objects consisting of attribute-value pairs. It is used primarily to transmit data between a server and web application as an alternative to XML. JSON files are usually more compact than XML files.	Complex-structured, multidimensional or tabular data
4	Extensible mark-up language (XML)	XML is a widely known mark-up language that defines a set of rules for encoding documents in a format that is readable to both humans and machines. Users create and define their own tags.	Complex-structured, multidimensional or tabular data. e.g. database extract metadata
5	Resource description framework (RDF)	RDF is a standard model for data interchange on the web. RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed.	Used for events/announcements
6	Geography mark-up language (GML)	GML is the XML grammar defined by the Open Geospatial Consortium for expressing geographical features. GML serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the Internet.	Spatial/location data
7	GeoJSON	GeoJSON is an open standard format for encoding collections of simple geographical features along with their non-spatial attributes using JSON.	Spatial/location data

Source: https://data.sa.gov.au/sites/default/files/Toolkit/Open-Data-Process-Guide.pdf.

The structure of the dataset describes the way the data is presented, and the most common structures include the following:

- Tabular data: data is presented in columns and rows;
- Hierarchical data: data is presented in a tree-like structure;
- Network data: a flexible representation in which nodes are not restricted to a hierarchy.

The delivery mechanism focuses on two key features, namely:

- The frequency of updating the datasets;
- Whether to make the dataset downloadable or not. If it is to be downloadable then its size should be small.

When selecting the format for a dataset, the open data team should consider the following:

- Choose the format that is most likely to be used by the general public;
- Publish the same dataset in different formats to accommodate for the different user needs;
- Publish a CSV version of datasets, as it is the most common format, in addition to any another format when needed.

(c)   *Metadata*

Metadata is structured information that describes, explains, locates, or makes it easier otherwise to retrieve, use or manage an information resource. Metadata is often called data about data or information about information.[84] Metadata is useful because it provides information enabling the user to make sense of the published dataset.
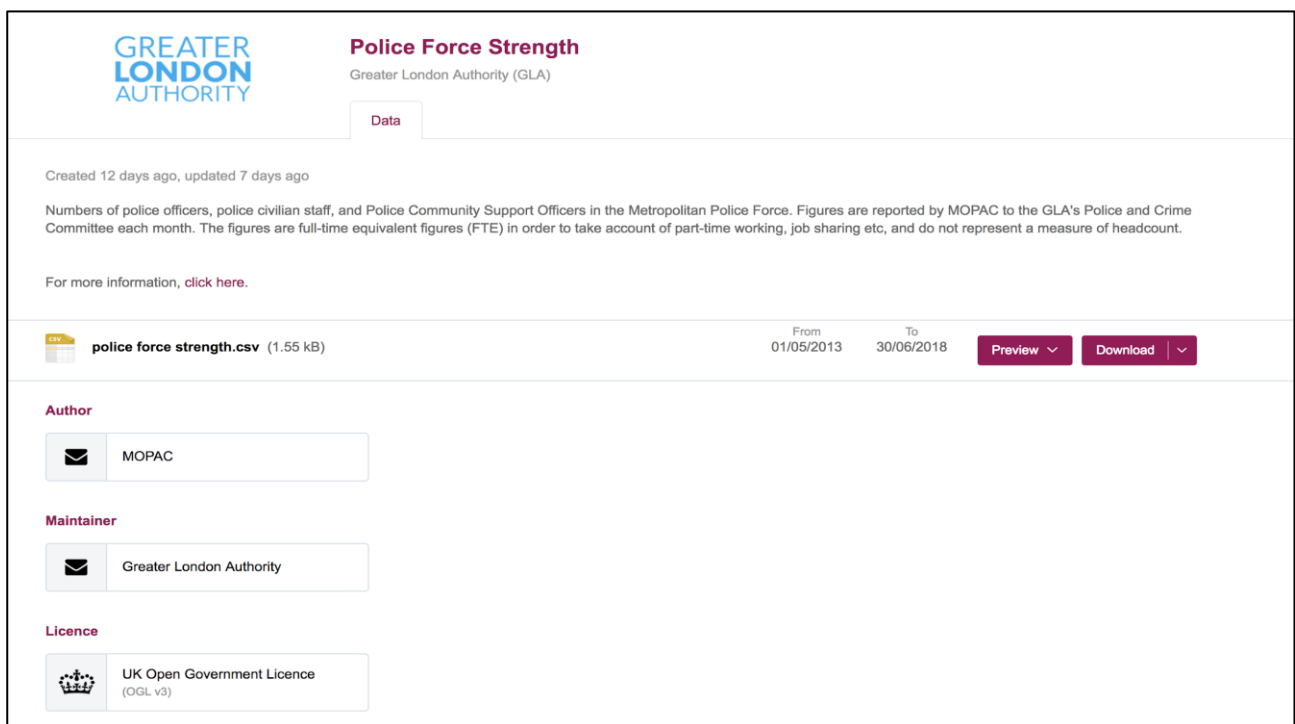
There are four general types of metadata:

- Administrative metadata is the most common and is produced in data collection, production, publication, and archiving. Most open data metadata is in this category;

- Structural metadata describes a structure of the dataset, including its format, organization and variable definitions. This is highest in demand by researchers and academics;

- Reference/descriptive metadata is a broad term that mostly involves descriptions of methodology, sampling and quality;

- Behavioural metadata records the reactions and behaviours of the dataset's users such as rating or user analytics.[85]

The Police Force Strengh data set of the Greater London Authority (figure 11), is published with a metadata set that includes the following:

- Name: Police Force Strength;
- Date published/created: 12 days ago;
- Date of last update: 7 days ago;
- Licence: UK Open Government Licence;
- Format: CSV.

Figure 11. Example of metadata

Source: https://data.london.gov.uk/dataset/police-force-strength.

Table 9 lists the recommended standard basic metadata.

It is useful to notice that the life cycle of metadata is longer than that of the data itself, which means that the metadata of a dataset may be created before the dataset itself and may be kept after the dataset has been removed.[86]

Table 9. List of standard metadata

No.	Metadata	Description
1	Title	Name of the dataset
2	Description	A free text
3	Keyword/tags	A keyword or tag describing the dataset
4	Publisher	Name of the Ggovernment organization responsible for creating the dataset online
5	Themes	Business domains under which the dataset is classified (e.g. health care or transportation). A dataset can have multiple themes.
6	Language	Language of the dataset (e.g. English or Arabic)
7	Licence	Type of legal licence the dataset is released under
8	Publish (or release) date	Date on which the dataset was published for the first time
9	Update date	Most recent date on which the dataset was changed, updated or modified
10	Update frequency	How often the dataset is updated
11	Contact	Contact details of the publishing organization including, email and phone number if possible
12	Format	N/A

Sources: https://www.w3.org/TR/vocab-dcat/; and https://data.gov.au/dataset/data-gov-au-metadata-and-other-schemas/resource/c79c5839-008e-4833-b469-3c550d38d05c.

(d) *Checking the dataset against the standard quality criteria of the organization*

Before publishing the dataset, it has to pass a check against all open data quality criteria approved by the organizations.

3. *Publication of the dataset*

After preparing the dataset, it is to be published on the online open data catalogue. This can be done in the following three different ways:

- Manually: This is generally used for datasets that are relatively small and uncomplicated, and where the refresh rate is quarterly or longer. This method is easy to implement but can be time-consuming if it is used for a large number of datasets;

- Automated: In this method, software is used to automatically extract the dataset from where it is stored (for instance, a database) and to publish it on the online catalogue;

- Programmatic with application programming interface (API): The use of APIs is necessarily for publishing and updating datasets with a live update frequency, namely, in real time or near-real time.

Depending on what technology the organization has, the automated method is the most recommended method to use.

H. PHASE 4: UPDATING OPEN DATA INVENTORY AND CATALOGUE

Once the open data inventory and catalogue have been created, the open data team needs to keep on maintaining and updating them. The time and effort needed for this depends on the size of both the inventory and the catalogue, in addition to the level of usage and engagement from the general public.

These continuous efforts include the following essential tasks:

- Adding more datasets to the open data inventory;

- Enriching the open data catalogue by selecting new datasets from the inventory and publishing them on the catalogue;

- Receiving and responding to any inquiries or requests from the public;

- Solving any issues, such as broken links, that may emerge;

- Updating the datasets as promised in the metadata.

IV. Open data platforms

IV. OPEN DATA PLATFORMS

A. INTRODUCTION

Publishing open data on the web makes it universally accessible, which satisfies an essential element of the definition of open data as discussed in chapter 1 of this guide. This makes the web the core channel for people to find and access open data. However, the web platforms used for publication can vary from one organization to another depending on their ambitions and their open data strategy. In general, there are three main options Government organzsations can choose from: web downloads, data portals and APIs.[87] This chapter discusses these options in greater detail.

B. OPTION 1: WEB DOWNLOAD

This is probably the simplest and least expensive way to publish open data. Datasets can be published on the organization's regular website by adding a dedicated page or section to the main website that contains links to the datasets. This option is ideal during the initiation phase of the data publication when the organization still has a limited number of datasets to publish.

From the technological point of view, the content management system (CMS) used by the organization, such as Drupal and Wordpress, can be used to publish the datasets in combination with an internal database.

The Ministry of Energy and Industry in the United Arab Emirates follows this option in publishing its data (figure 12). The data is published on a page within the Ministry's website where a list of datasets in various formats (such as PDF and XLS) that are labelled as open data can be found.[88] Clicking on any of these datasets will trigger the dataset to be downloaded.

Figure 12. Open data page at the website of the Ministry of Energy and Industry, United Arab Emirates
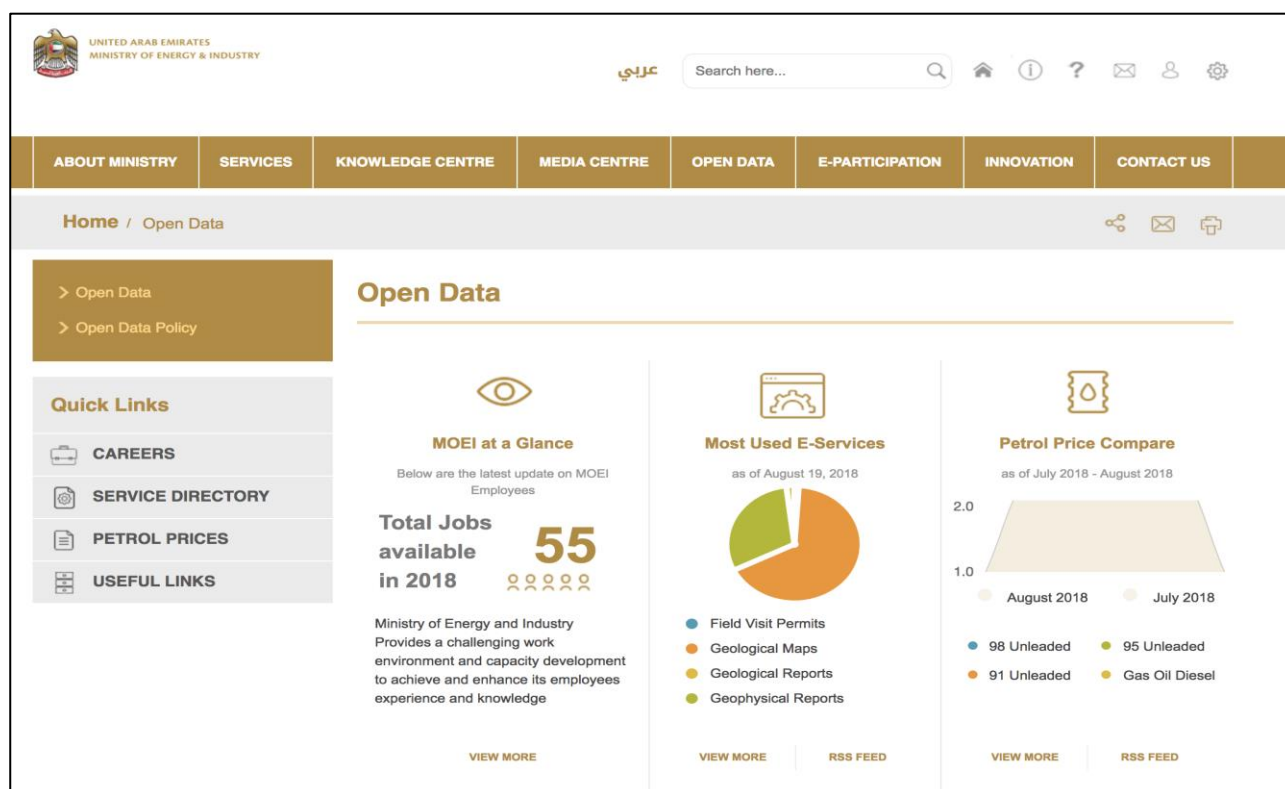
Source: Ministry of Energy and Industry, United Arab Emirates, n.d.

C. OPTION 2: DATA PORTAL

This is a more advanced way of publishing datasets where the organization dedicates a whole portal to open data independent from its corporate website. This option is particularly useful if the publishing organization has large amounts of datasets to publish that need to be updated regularly. Publishing datasets on a data platform can offer both to the publisher and the users a set of benefits explained later in this chapter, but it requires more financial, human and technical resources as opposed to web downloads.

To implement the data portal option, organizations can choose from two common scenarios:

- Scenario 1: To create their platform, own it and run it;

- Scenario 2: To use another organization's platform. Usually, that other organization is the one in charge of the central national or city-wide open data programme.

In some cases, where the organization has a very strong digital presence and direct communication with its data users, it can adopt a combination of both scenarios. The Department for Transport in New South Wales, Australia, offers a good example that demonstrates both options. The Department has its own open data platform (figure 13) which, at the time of writing this report, contained 89 datasets.

Figure 13. Open data platform of the Department for Transport, New South Wales, Australia
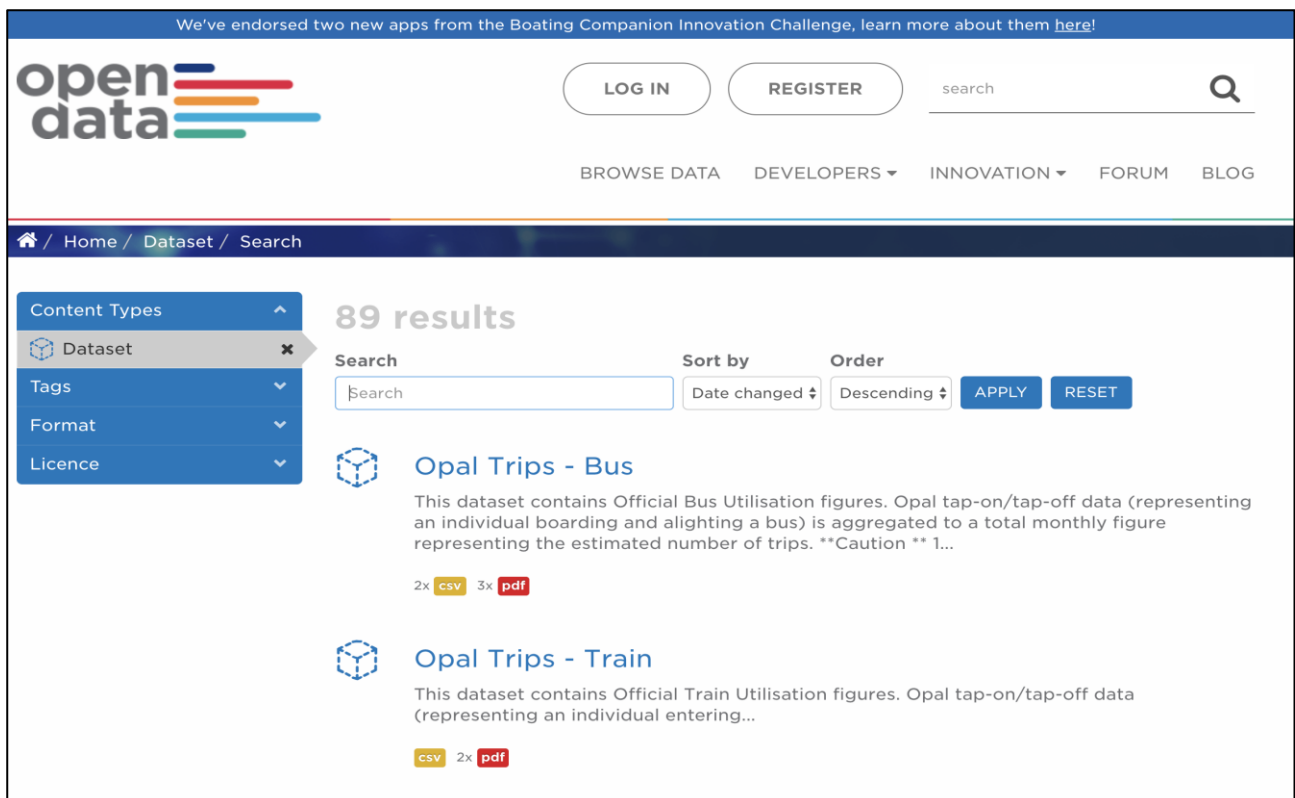
Source: Transport for New South Wales, Australia, n.d.

In addition to this, the Department adopts the second scenario as it publishes some of its datasets on another portal run by the state Government and represents the official data portal for the whole Government of New South Wales (figure 14).

In Tunisia, the central open data platform contains datasets collected from different ministries and Government organizations (figure 15).

Figure 14. Open data platform of the Government of New South Wales, Australia

From a technological point of view, setting up an open data portal requires specialized software as explained later in this chapter.

Figure 15. Open data platform of the Government of Tunisia
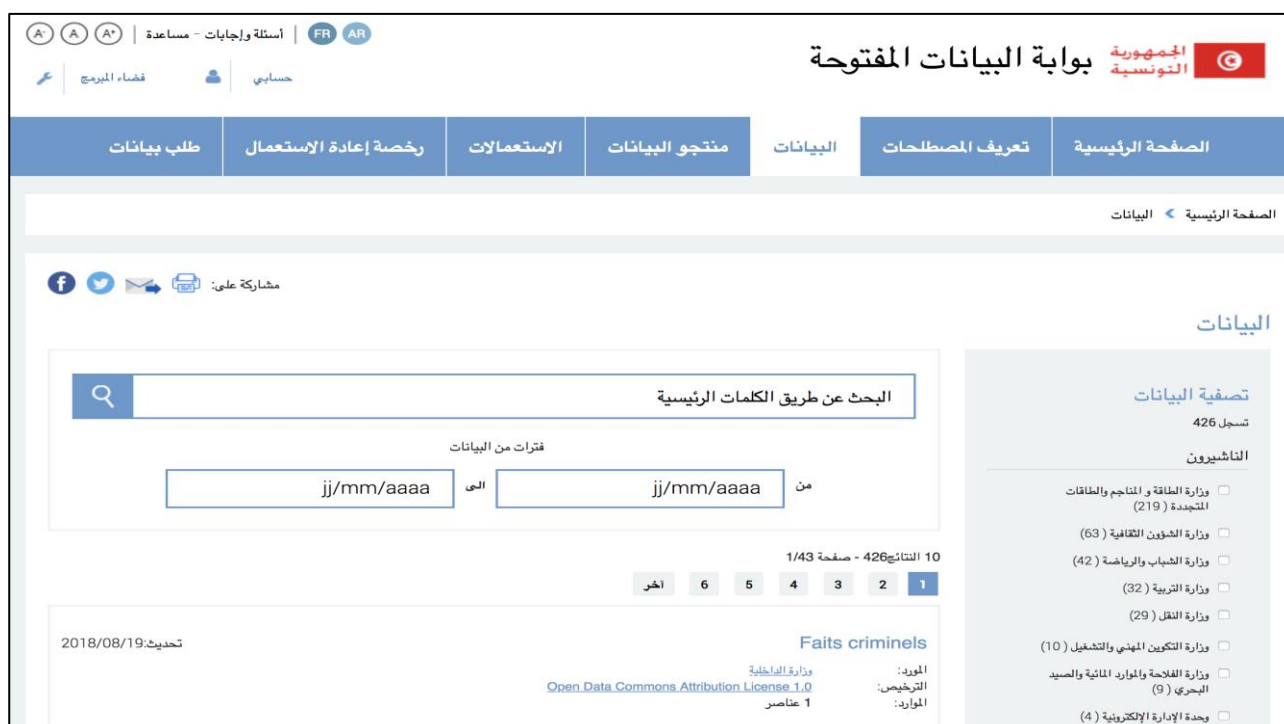
D. OPTION 3: APPLICATION PROGRAMMING INTERFACES

APIs, rather than being treated as a stand-alone option, are widely adopted as an advanced feature on top of any of the previous two options.

Figure 16. API of Transport for London

Source: Transport for London n.d.

Figure 17. CityMapper app using the API of Transport for London
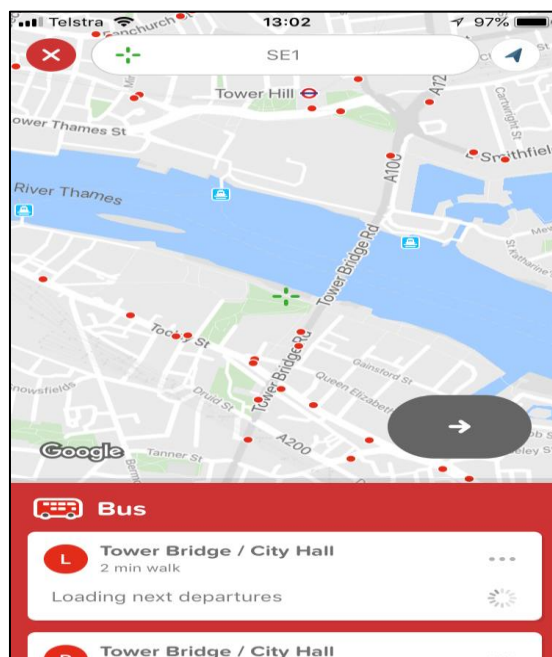
Source: CityMapper, n.d.

An API is an advanced way for connecting distinct systems or applications to one another. In the context of open data, an API is a software that provides a system with direct access to data from another system. The API grants access to the use of the data catalogues and their functionalities. Without the interference of any end-user interfaces, such as portals or web pages, a third-party application can load the data by means of a request protocol. In short, it is particularly useful if the data should be up to date, directly accessed and reused by third parties, and the application using the data needs direct access to the database without any interference.

TfL has an API that enables developers to connect their software and access the data in a simple way (figure 16). As a result, thousands of developers have been using this API to build hundreds of |mobile applications including CityMapper, which is one of the most popular transport applications in the city (figure 17).

E. PLATFORM DEVELOPMENT

Using any of the three options discussed previously, Government organizations can develop their own open data platform using one of the following three broad approaches:

Custom web development: **I**n this approach, the Government organization builds a custom platform using its own internal software and technology team or through hiring an external vendor;

Open source: There are several open source packages for open data portals. The organization can use one of them and customize it for its own data portal;

Software as a service: These are cloud-based open data portals that can be used for a recurrent fee calculated by the data volume among other parameters.

Each approach is associated with a number of cost elements summarized in table 10.

Table 10. Cost centres platform development approaches

Cost centre	Custom web development	Open source	Software as a service
Design and development/customization	X	X	-
Hosting	X	X	-
Vendor contracting/software selection	X	X	X
Software licence	Possible	-	X
Installation	X	X	
Initial design/branding customization	-	-	Possible

Source: Open Data Institute, n.d. See https://theodi.org.au/tools/.

F. KEY FEATURES OF OPEN DATA PLATFORMS

A well-designed open data platform helps the organization to achieve its goal of enabling anyone to access and use the published data. Although the detailed features of an open data platform may vary depending on the used technology and development approach, there are some common features that include the following:

Usability and accessibility: Portals should be designed and developed in alignment with the international usability and accessibility standards such as the ones developed by the World Wide Web

Consortium. Users should be able to access and use the datasets anonymously without user registration. Data.gov and almost all open data portals featured in this guide have no registration or sign-in requirements;

Data catalogue: When provided as a built-in feature, it facilitates the process of publishing managing datasets;

Metadata: A built-in support for metadata can enable the publisher to assign the required metadata to dataset(s) in an easy and quick way;

Data licence: Similarly, some data platforms enable the publisher to define one central licence and apply it to all or some of the datasets;

Dataset download option: The user should be able to download the published datasets in a simple and easy way;

Data format: The portal should support a wide range of data formats (for instance, CSV, XML and JSON);

Data preview: This feature enables users to preview the data prior to downloading it;

Data visualization: This feature uses built-in graphing or mapping tools to enable the user to visualize the data. The data portal of the Australian Capital Territory offers preview and built-in visualization and mapping feature for its datasets, such as cyclist crashes heat map;

Search function: The user should be able to easily search the data catalogue by a number of parameters such as subject, organization or data format. On Data.gov.uk, users can browse data available within the catalogue by theme (for instance, education and environment), by publisher (for instance, Government agencies) or by format (for instance, CSV), among other options. In addition, the catalogue has a search engine with advanced features and result-filtering options;

Engagement: Open data portals should offer a space for dialogue between users and the data publisher. This space can offer some resources and user guides in addition to e-participation tools, such as discussion forums and social media communications;

Application programming interface (API): As explained previously, APIs can enable software developers to build creative solutions using the published datasets;

Security: Open data catalogues should implement security measures to protect data and metadata from being changed by unauthorized users. Data.gov provides details on the applied security measures to protect the portal's data and content on its Privacy and Policies page. An example states that "commercially reasonable tools and techniques are used to protect against unauthorized access to Data.gov systems";[89]

Monitoring and analytics: This feature provides the data publisher and portal administrator with insights into the usage of the portal and the datasets.

G. OPEN DATA PLATFORM SOLUTIONS

The following provides an overview of the most common open data platform solutions that are used by governmental organizations and NGOs around the world to publish and manage open data. A brief summary is provided about each solution, its key features, and whether it supports the Arabic language or not, in addition to three examples of open data portals.

This list is not inclusive, and it does not indicate any endorsement for any of these platforms. Government organizations may use this list as an indicative guidance but must perform their own research exercise to choose the platform that suits their needs.

The Comprehensive Knowledge Archive Network (CKAN): An open source for creating and managing open data websites.I It can be used by national and local Governments, research institutions and other organizations who collect a lot of data.

Key features:

- Provides a wealth of features and has over 200 community extensions which can fill almost any feature gap;
- Provides a streamlined way to make data discoverable and presentable. Each dataset is given its own page for the listing of data resources and metadata, which adds to the value of the data catalogue and facilitates searching in it.

Supports Arabic language: Yes

Examples of government open data portals that use CKAN:

- https://www.data.gov/ (US);
- https://data.gov.au/ (Australia);
- https://www.europeandataportal.eu/ (EU).

DKAN: A Drupal-based community-driven open source open data platform that offers a full suite of cataloguing, publishing and visualization features.

Key features:

- Community-driven feature development;
- Built-in visualization features;
- Manages diverse data sets.

Supports Arabic language: Yes

Examples of government open data portals that use DKAN:

- https://data.ca.gov/ (California, US);
- http://data.gov.sa/ (Saudi Arabia);
- http://data.gov.bd/ (Bangladesh).

OpenDataSoft: A turn-key solution fully accessible from a web browser that can be used by organizations to publish datasets and by users to visualize data in maps and graphs.

Key features:

- Built with an API-first approach, each dataset has an API with real-time data processing, querying and aggregation capabilities;
- Intuitive data visualization tools;
- Data processing and enrichment features.

Supports Arabic language: Yes

Examples of government open data portals that use OpenDataSoft:

- https://opendata.paris.fr/page/home/ (Paris, France);
- http://spms.min-saude.pt/ (Ministry of Health, Portugal);
- https://www.kapsarc.org/ (King Abdullah Petroleum Studies and Research Centre, Saudi Arabia).

Socrata: A cloud-based solution that allows Government organizations to publish their data online, make data-driven decisions, operate more efficientl,y and share insights with citizens.

Key features:

- Access a single source of trusted data in a self-service environment and break down silos and encourage collaboration;

- Transform data into insights with visualizations and performance measurements;

- Enhance the flow and quality of open data;

- Increase transparency and resident engagement.

Supports Arabic Language: No

E*xamples of g*overnment open data portals that use Socrata*:

- https://opendata.cityofnewyork.us/ (New York City, US);
- https://opendata.cityofnewyork.us/ (Chicago, US);
- https://www.dallasopendata.com/ (Dallas, US).

V. Evaluation of open data programmes
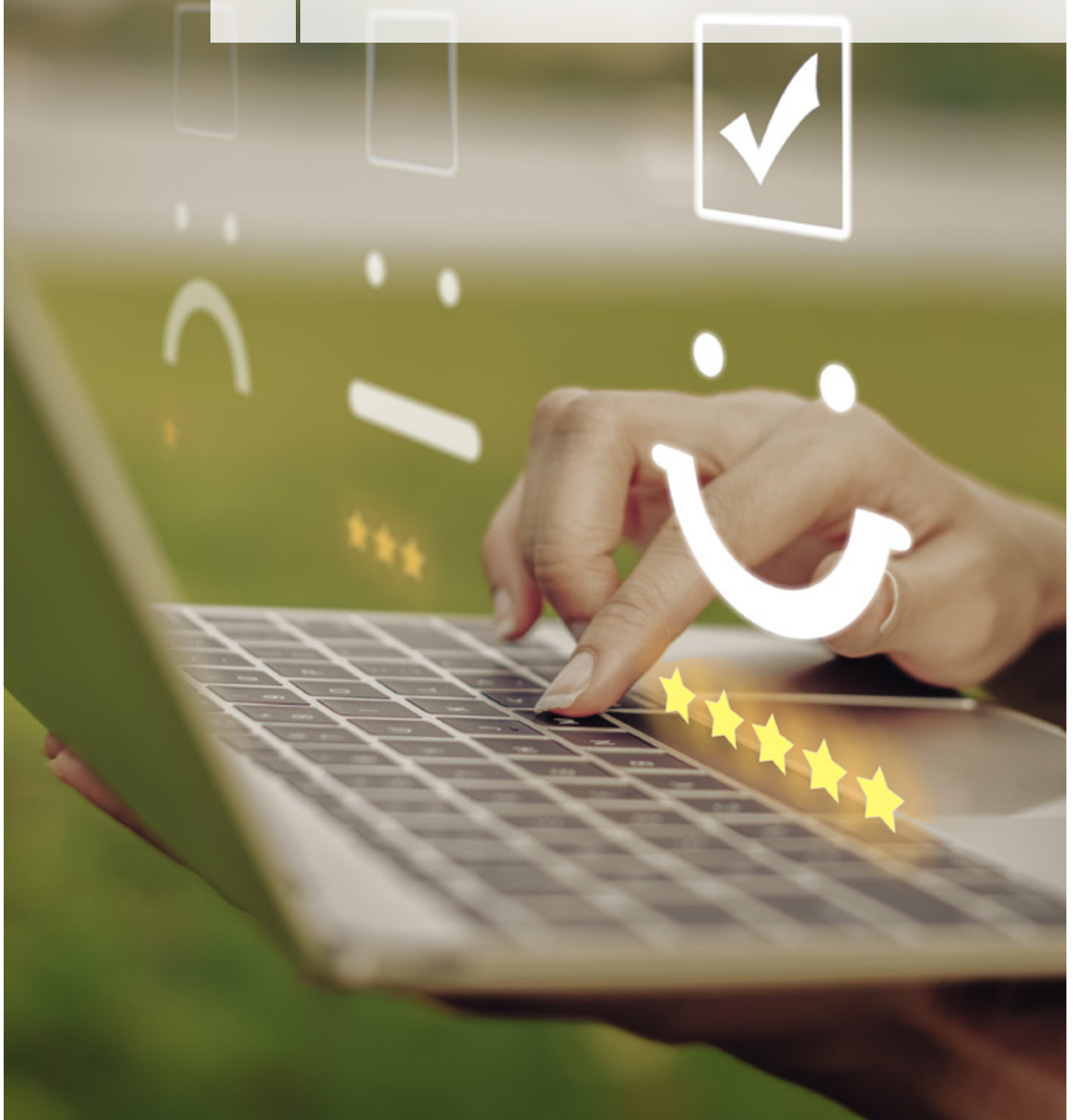
V. EVALUATION OF OPEN DATA PROGRAMMES

A. INTRODUCTION

Previous chapters of this guide presented case studies that demonstrate how governmental organizations and NGOs can leverage open data to create growth and change through several forms of impact, including improving governance and empowering citizens. However, the task of capturing and evaluating the impact of open data programmes and the value generated by them can be challenging due to several factors.

Open data attracts new types of users, mainly small and medium enterprises and certain new business models such as advertising models, rather than the end-user subscription model. These new users and business models can challenge the traditional methods for calculating the economic value of Government programmes. In the Netherlands, for example, the mobile app Buienradar[90] enables users to find out if it is going to rain in their current location over the next few hours. The app provides this forecast by using the open data published by the Royal Netherlands Meteorological Institute. The app generated around 300 million hits throughout Europe in 2010. Attracted by this high traffic, advertisers were greatly interested in investing in the service, and it is paid for through advertising revenues while being provided completely free of charge to end-users.[91]

Another challenge is that it is apparent now that the value is not in the data by itself, but in the combination of the data with factors such as innovative ideas, the use of new technologies (such as the Internet of Things) and the availability of new techniques to process data (such as "big data analytics" methods). PlumeLabs, a startup based in Paris, offers its customers a mobile app that shows them the level of pollution in the air both indoors and outdoors and provides recommendations on whether certain activities (such as exercising) are suitable or not, based on the air quality of the user's location. To do that, PlumeLabs uses the open data provided by Government authorities in the cities covered by the app. In addition to open data, the company also uses a mix of technologies to provide the desired output. One of these technologies is the smart air quality tracker: a hardware sensor that measures the quality of air around the user. The company uses the data generated trackers of its users around the world. This is a direct application of Internet of Things technology and the crowdsourcing technique.[92]

This chapter presents an evaluation process that is not confined to impact assessment, but to all the development stages of the open data programme.

B. EVALUATING OPEN DATA PROGRAMMES

Despite the challenges associated with the assessment of open data programmes, many governmental organizations and NGOs have launched programmes and research projects that aim to assess open data programmes and to measure their value, focusing on a range of angles (such as economic versus social aspects) and using a variety of methods and tools. Table 11 provides examples of these programmes and studies.

The programmes listed in the table vary in their aims in assessing open data programmes, as well as the methods they apply. Overall, assessing open data programmes may focus on:

- Technical assessment of datasets;
- Evaluation and ranking of open government data initiatives;
- Providing quantitative metrics of open data outcomes and impacts;
- Providing qualitative judgements on performance of an open data initiative;
- Developing qualitative case studies about open data use and impacts.[93]

No.	Study	Brief description	Run by
1	United Nations e-Government Survey	A global survey that offers a snapshot of trends in the development of e-government in countries across the globe and assesses their progress. The survey contains open data issues.	UN-DESA
2	The Open Data Barometer	The barometer aims to uncover the true prevalence and impact of open data initiatives around the world.	World Wide Web Foundation
3	Open Data Certificate	A digital platform that uses a questionnaire-style assessment to recognize well-published open data with digital badges.	Open Data Institute
4	OpenData500	An assessment of the value of open data in the private sector that covers six countries including the United States, Australia and Canada.	Governance Lab (GovLab)
5	OpenDataMonitor	A research project that provides an overview of available open data resources and undertakes analysis and visualization of existing data catalogues using innovative technologies.	A consortium of European organizations funded by the EU

Source: GovLab, 2018.

The open data in developing economies logic model (figure 18) offers an overall review of how different actors in open data programmes work together to create an impact. The model suggests that open data released by Government organizations (supply), when analysed and leveraged by the actors from both governmental and non-government sectors (demand), can be used in a variety of ways (actions and outputs), within the parameters established by certain enabling conditions (and disabling factors) to create the desired impact.

Figure 18. Open data in developing economies logic model
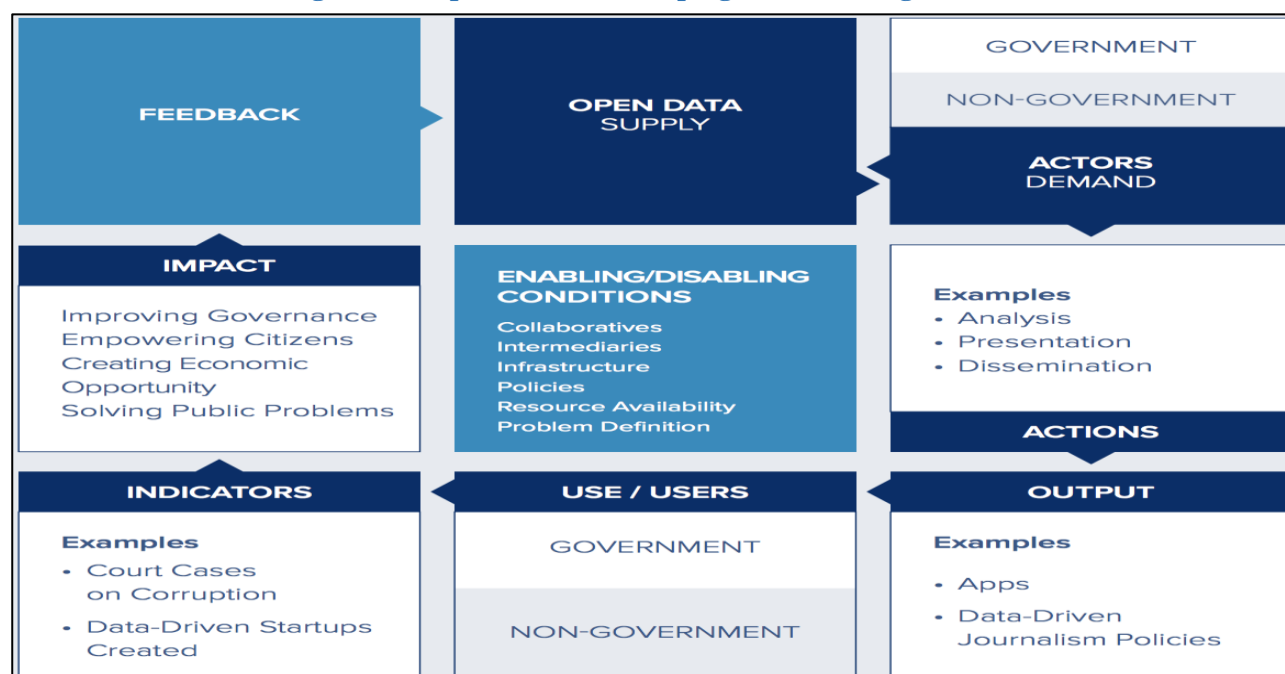
Source: Verhulst and Young, 2017.

This model can be helpful in understanding ways to create impact in developing countries and the importance of the local enabling and disabling conditions. This is an essential foundation for the evaluation model introduced in the next section.

The following introduces a framework for evaluating open data programmes in the Arab countries. The framework has been designed in accordance with the following two guidelines:

- Adopting a variety of reliable open data evaluation frameworks and methods, such as the ones listed in table 11;

- Fitting the current state of open data programmes across the Arab countries where the progress being made can vary greatly from one country to another.

The aim is to make the evaluation exercise as simple as possible, with clear steps and outputs. However, organiations should expect to go through a steep learning curve while building their human capacity and adjusting their organizational practices.

Before presenting the framework itself, it is important to highlight the following:

- Framework and evaluation process do not comprise a stand-alone exercise that is isolated from the open data programme. Instead, it should be considered as an integral part of the programme;

- The evaluation process should be carried out continuously and on a regular basis. This would provide the organization with experience in the process and also accumulate time-series data that can be very helpful in monitoring output trends of the evaluation process.

The framework shown in figure 19 consists of five steps connected to the entire open data programme.

Figure 19. Open data evaluation framework
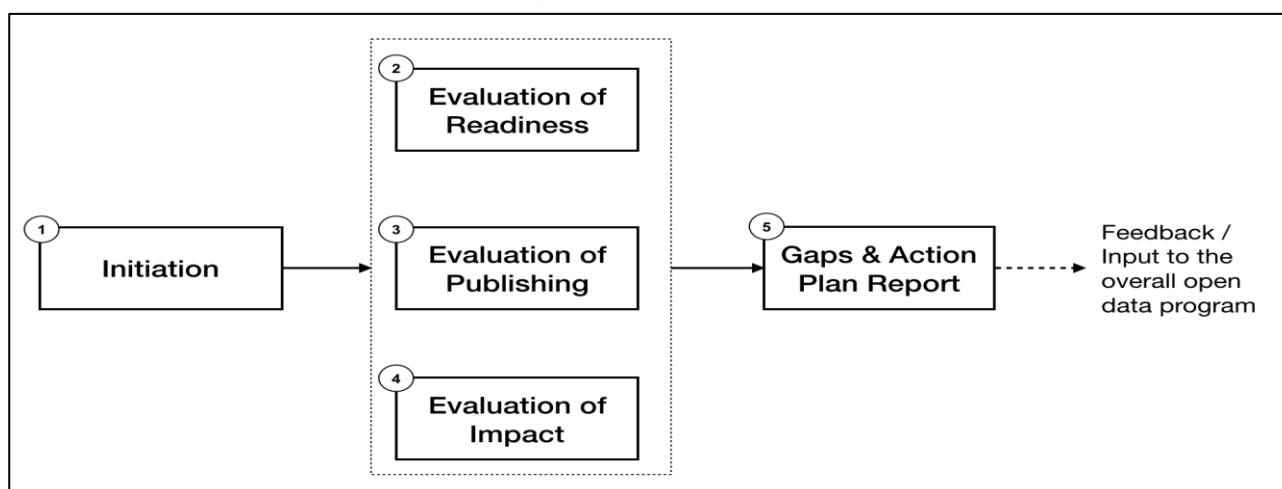
The core evaluation process occurs in steps 2, 3 and 4, and the headlines of readiness, publication and impact, addressed in those steps, are widely adopted in similar open data evaluation models and methodologies, including that of the Open Data Barometer.

The following offers a description of each individual framework step.

1. *Step 1: Initiation*

At the beginning of the evaluation work, it is recommended to carry out the following steps:

- Securing the support of the organization's leadership; this will increase the chances of conducting a smooth, honest and effective evaluation;

- Identifying the evaluation team and team leader; usually, this includes the team in charge of the open data programme, in addition to other stakeholders from inside and outside the organization such as representatives of the data owners and external vendors who are in charge of some of the ongoing open data activities;

- Defining the scope of the evaluation; the evaluation should cover the entire open data programme. However, this scope might be too large for some organizations to handle all at once. In this case, they may decide to start by focusing on a part of this scope;

- Allocating the human and financial resources needed;

- Ensuring that the evaluation process is well aligned with the open data strategy.

2. *Step 2*: *Evaluation of readiness*

Objective: To evaluate the extent to which the Government organization has the required organizational and human capabilities to carry out a successful open data programme.

Output: A list of readiness gaps under the five key themes.

Methods and tools: Open Data Maturity Model[94] and Open Data Pathway.[95]

The Open Data Maturity Model was designed to help organizations in carrying out their open data programmes and activities through the following:

- Supporting assessment of the effectiveness of an organization in its operational and strategic activities around open data;

- Providing guidance to organizations on potential areas of improvement;

- Comparing organizations to highlight their respective strengths and weaknesses;

- Supporting wider adoption of best practices and helping to improve processes.[96]

To achieve that, the model covers the following five themes that represent key areas of open data activities within the organization:

Strategic oversight: Highlights the organization's need for a clear strategy concerning data sharing and reuse, and an identified leadership with responsibility and capacity to deliver that strategy.

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 the organization by identifying the knowledge-sharing, training and learning required to embed an understanding of the benefits of open data.

User support and engagement: Addresses the organization's need to engage with both their data sources and their data reusers to provide sufficient support and feedback to make open data successful.

Investment and financial performance: Covers the organization's need for an insight into the value of its datasets and the appropriate budgetary and financial oversight required to support their publication. In terms of data consumption, the organization has to understand the costs and value associated with reuse of third-party datasets.

For each one of these five themes, the model contains a set of organizational activities. There are 15 activities in total, shown in table 12.

Table 12. Themes and activities of the Open Data Maturity Model

Strategic oversight	Investment and financial performance	Data management processes	Knowledge and skills	User support and engagement
Open data strategy	Financial oversight	Data release process	Open data expertise	Engagement process
Asset catalogue	Dataset valuation process	Standards development and adoption	Knowledge management	Open data documentation
	Open data in procurement	Data governance		Reuser support processes
		Data desensitizing		Community norms

Source: Open Data Institute, 2015.

Process: To apply the Open Data Maturity Model, the organization should carry out an assessment exercise in which it reviews its current practices in each of the five themes. This exercise can be done internally (self-assessment) or with the help of external independent assessors. If the organization opts for self-assessment, it is recommended to use Open Data Pathway,[97] a free online assessment tool built to help organizations in assessments using the Open Data Maturity Model.

In either case, the following eight-step approach is recommended for a successful and smooth assessment process:

1. Identifying an organizational lead to coordinate the collection of information from across the organization and to lead the entire assessment process.

2. Identifying the scope which can cover the entire organization or focus on some individual departments that are directly involved in the open data programme.

3. Identifying key participants; those are the people in the organization who may be needed to help answer specific questions or support the evaluation.

4. Assessing and scoring each activity, which is the main step in the exercise. The organization needs to assess each of its activities (table 12) and thereby identify the level of maturity achieved by the organization. The activities with low scores should be incorporated in the gap report to be produced in a later stage of the evaluation exercise.

When conducting the assessment, the organization needs to assign an assessment result (score) for each activity. The following is a recommended scale which divides activity assessment into five levels:

Level 1: **Initial**: The desirable processes are non-existent or ad hoc, with no organizational oversight;

Level 2: **Repeatable**: Processes are becoming refined and repeatable, but only within the scope of individual teams or projects. There are no organizational standards;

Level 3: **Defined**: Processes are standardized within the organization based on best practices identified internally or from external sources. Knowledge and best practices start to be shared internally. However, the processes may still not be widely adopted;

Level 4: **Managed**: The organization has widely adopted the standard processes and started to monitor them using defined metrics;

Level 5: **Optimizing**: The organization is attempting to optimize and refine its process to increase efficiency within the organization and, more widely, within its business sector.[98]

The Open Data Maturity Model offers a complete list of the five themes and the activities within each of them, in addition to a description of each of the five possible levels for each activity, as follows:

5. Setting appropriate targets; after conducting a baseline assessment, appropriate targets for improvement should be identified. This will involve either maintaining or improving the score for specific activities.

6. Developing an action plan based on the results and the identified targets. The action items should be included in the plan, which is to be designed at a later stage of the evaluation.

7. Circulating results, targets and action plans within the organization, including among those involved in supporting the assessment. Sharing this information more widely outside the organization can be taken into consideration too.

8. Setting a date for the next assessment; the action plan should set a date for a further assessment. This will allow the organization to monitor its progress. It is recommended to maintain regular annual assessments.[99]

By the end of evaluation of readiness, the organization should have identified a list of the activities that need to be improved, and also should have generated recommendations on how to improve them. At a later phase in the exercise, both the identified activities and recommendations will be incorporated in the gap report and action plan respectively.

The following presents an example to elaborate the steps mentioned above:

Organization: Department of Transport and Main Roads in Queensland, Australia.

Scope: The whole open data programme; the department had 182 datasets published on its websites at the time of the assessment.

Lead: The open data team which is part of the information technology branch in the department.

Tools and methods: The department ran a self-assessment exercise through a series of internal workshops attended by the open data team members and used the Open Data Pathway assessment tool. In addition, the department published all the results of this assessment on its website.[100]

Results: The summary of the assessment results are provided in table 13. The summary, in addition to the full and detailed results that are published on the department website, can be useful to other Government organizations in several ways, including the following:

• They offer a useful list of the detailed questions asked under each of the five themes and their activities;

• The answers provided in response to these questions can be a great source of learning and benchmarking for other organisations carrying out open data programmes. They can offer real-life examples of the issues and challenges that may face any open data programme and how to tackle them.

Table 13. Summary of the Open Data Maturity Assessment results, Department of Transport and Main Roads, Queensland, Australia

Theme	Score	Maximum score
Data management process	12	20
Knowledge and skills	2	10
Customer support and engagement	6	20
Investment and financial performance	3	15
Strategic oversight	4	10
Total	27	75

Source: Department of Transport and Main Roads, Australia, n.d.

The detailed results, in addition to the future targets set by the department leadership, can be downloaded from the department's website.[101]

3. *Step 3*: *Evaluation of publishing*

Objective: To evaluate to what extent the Government organization successfully publishes its datasets according to the predefined quality criteria.

Output: A list of gaps between the current status of publishing and the planned one.

Methods and tools:

- The organization's own data inventory and catalogue plans (as discussed in chapter 3);
- The organization's quality criteria (quality criteria will be discussed in detail in chapter 6);
- The Open Data Barometer might be used in some cases.[102]

Process: Similar to the previous step, evaluation of publishing can be carried out either internally via a self-assessment exercise, or with the help of an external independent assessor. The following steps are recommended for this assessment:

- Capturing the current version of the open data catalogue that contains the latest list of open datasets published on the organization's open data page or portal;

- Using the list to evaluate dataset quantity and quality in the following manner:

To evaluate the quantity of the published open datasets, the organization needs to simply compare the number of the already published datasets on the catalogue to the number of the datasets it planned to publish as per the open data strategy. The outcome should be a percentage.

To evaluate the quality of these published datasets, the organization needs to assess each individual dataset against the organization's quality criteria. For example, for each dataset, the format should be examined to confirm that it matches the list of formats approved by the organization as part of the quality criteria. The same test should be conducted for metadata and the remaining quality criteria.

At the end of the process, the evaluation team should compile a list of gaps and issues and relevant notes about them, in addition to recommendations on how to fix them. These outputs should be incorporated in the gaps and action plan report to be produced at a later stage of the evolution exercise.

The city of San Francisco, for example, runs such a quality and quantity evaluation of the datasets published by 52 government departments. The results of the evaluation are published on the city's data portal. The latest version of the evaluation shows that the city has published 437 datasets, which represents 52 per cent of the inventoried datasets (figure 20).

Figure 20. Evaluation results for the city of San Francisco: Quantity of datasets
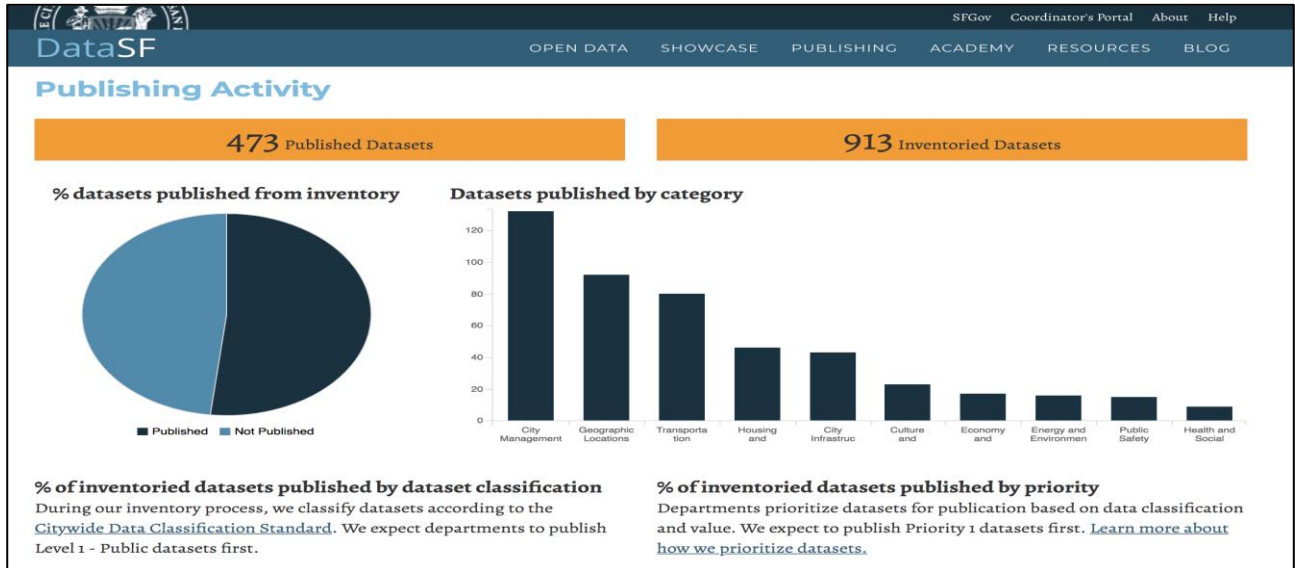
Source: City of San Francisco, n.d.

Figure 21. Evaluation results for the city of San Francisco: Quality of datasets
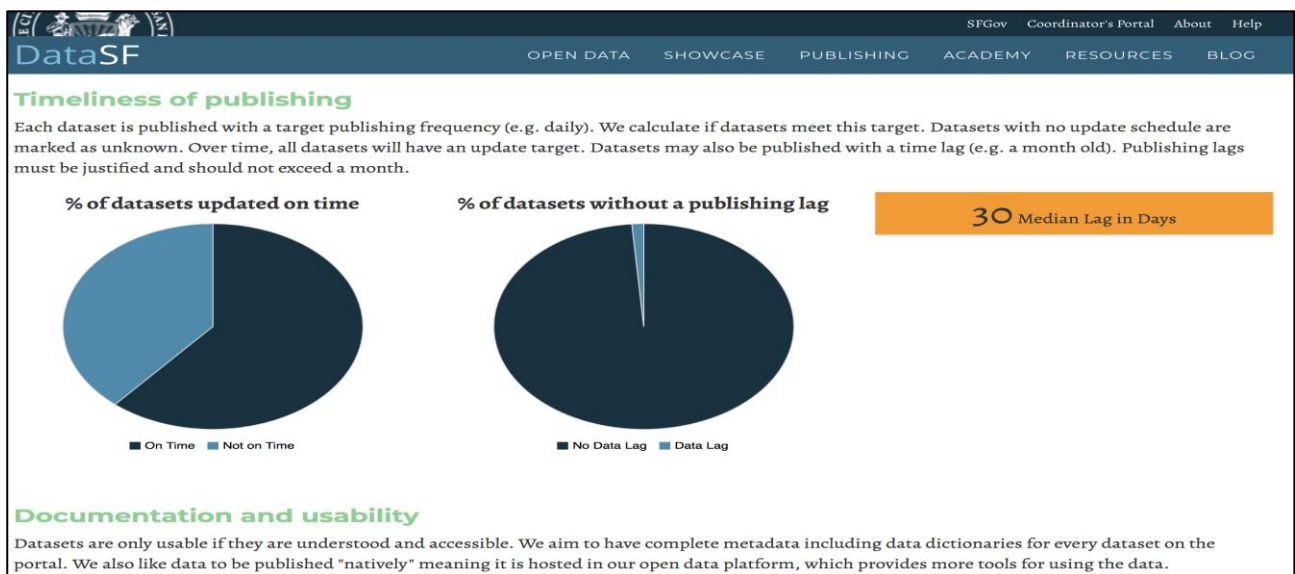
Source: City of San Francisco, n.d.

The quality part of the evaluation covers the criteria identified by the city. For instance, the timeliness criterion indicates that 61 per cent of the published datasets are updated on time (figure 21).

4. *Step 4*: *Evaluation of impact*

As pointed out at the beginning of this chapter, evaluating the impacct is the most challenging step in the evaluation framework and exercise.

Objective: To evaluate the extent to which the open data programme successfully achieved its objectives as defined in the open data strategy. Objectives can vary from one organization to another depending on its mandate and what value it aspires to create through the open data programme.

Output: A review of the current gap between what has been achieved and the objectives stated initially in the open data plan, in addition to recommended action steps to bridge this gap.

Methods and tools: A mix of quantitative and qualitative methods that the evaluation team has to identify. Generally speaking, the relevant literature refers to two main approaches to measure the impact and benefits of open data:[103]

- Top-down (macroeconomic): This approach considers the value of open government data through the resources devoted to generating it or using it;

- Bottom-up (microeconomic): This approach seeks to find an aggregate figure by adding up various components using business surveys, local and international case studies and consultations.

In addition, the evaluation can cover assessing the progress made in implementing the actions identified in the open data plan, such as the delivery of training programmes or the launch of a new data portal.

Three examples are provided in the following to demonstrate how different levels of the impact of open data programmes are being evaluated.

Figure 22. Progress report for New York City's open data programme
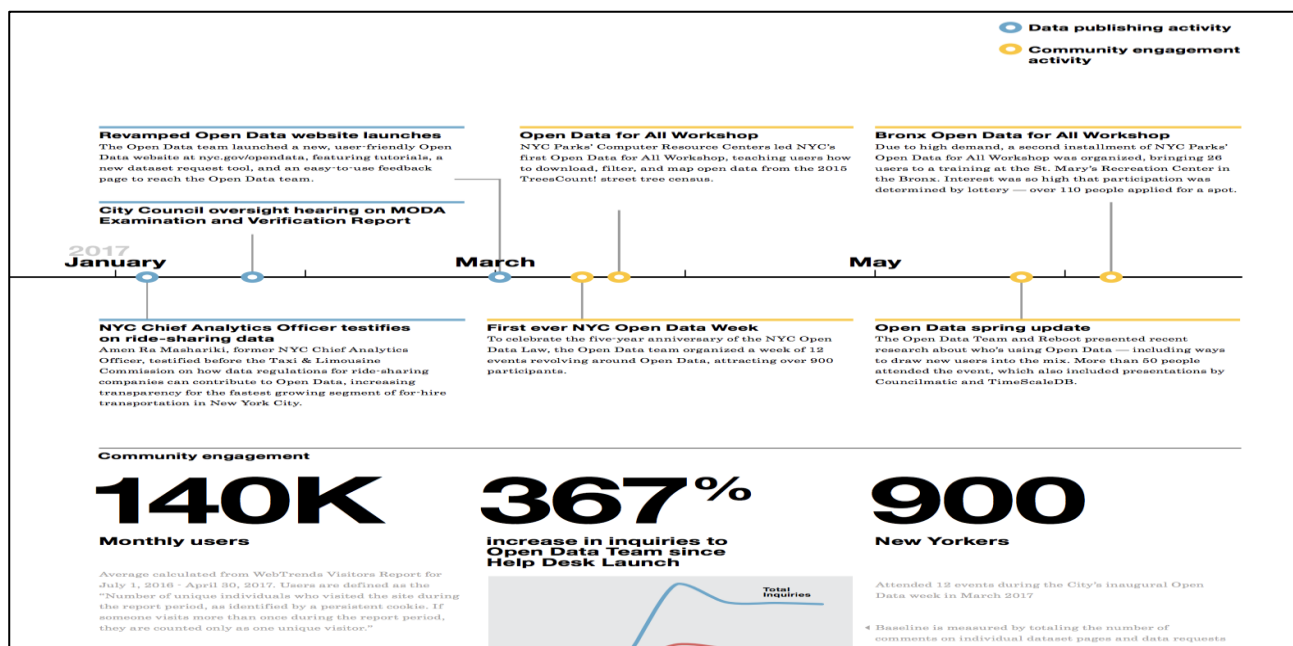
Source: New York City, 2017.

New York City's 2017 progress report for its open data programme (figure 22) covers the previous 12 months. The report addressed several aspects of the city's open data plan, including data publications and community engagement.

The United States' federal data portal (Data.gov) also has a page dedicated to impact. The page presents a collection of case studies to demonstrate how the open data published on this portal has been used by startups and companies to create new businesses and jobs. For each company in the list, the actual datasets used by that company are identified, in addition other important information such as the number of jobs created.[104]

The EU has conducted a study to collect, assess and aggregate all economic evidence to forecast the benefits of open data for all 28 European member States and EFTA countries for the period 2016-2020. The study followed a mix of top-down and bottom-up approaches where the economic impact of open data was assessed using several tools including the market size of open data (in both euros and as a percentage of the gross domestic product), the cost savings that can be reached in the public sector and efficiency gains for individual citizens.[105]

5. *Step 5*: *Gaps and action plan report*

This step represents the output of the evaluation team's work in previous steps and should be allocated a minimum time compared to the previous three phases.

Objective: To compile and document all the issues captured in steps 2, 3 and 4 in a report, in addition to a set of recommended action steps on how the organization should handle these issues to improve the overall open data programme.

Output: A comprehensive document of all the gaps and recommended actions.

Methods and tools: To create this report, the team should follow the standard policies and guidelines applied in the organization (such as communication guidelines). In addition, the team should consider publishing a version of the report online on the organization's website. The team can also use some of the digital collaboration tools (such as Google Docs) to facilitate co-working among the team members.

Steps: The following suggests generic steps to create the report. These steps can be amended depending on factors such as the organization's scope, team size and available resources:

- Appointing a moderator (or key author) for the report. The moderator's main job is to lead the work of the report production and facilitate the contribution of different team members;

- Agreeing on the overall structure (outline) of the report during the initiation phase, this should help in capturing all the needed information and minimize the risk of revisiting some of the evaluation steps to recapture information. At minimum, the report outline should include the following:

  - Objective of the evaluation;
  - Scope;
  - Stakeholders from inside and outside the organization;
  - Evaluation of readiness: gaps and recommended corrective action steps;
  - Evaluation of publishing: gaps and recommended corrective action steps;
  - Evaluation of impact: gaps and recommended corrective action steps;
  - Action plan: full list of action steps provided with targets and timeline;
  - The planned date for the next evaluation.

- Capturing all the issues and recommended actions throughout the evaluation phases: this should happen throughout the meetings and workshops conducted by the evaluation team;

- It may be helpful to use some of the digital collaboration tools (such as Google Docs) to enable those involved in the evaluation exercise to co-edit the report on an ongoing basis;

- Reviewing the draft version of the report and gaining the approval of the organization's leadership;

- Publishing the report online: It is good practice to consider publishing the report (or parts of it) in different formats. For example, it may be useful to publish the section related to the evaluation of readiness that contains the actual scoring in spreadsheet format to facilitate its access and use.

The evaluation reports from New York City and the city of San Francisco, highlighted previously, offer good examples for transparent evaluation reports.

VI. Open data quality

VI. OPEN DATA QUALITY

A. Introduction

As discussed in chapter 1, there is a continuous growth in the number of countries with open data initiatives and the number of datasets openly available. However, this increase in available open data does not necessarily imply that it is of a high quality. Publishing high-quality open datasets is crucial because open data of a poor quality can reduce user trust and negatively affect their use of the data.

This chapter provides helpful standards, tools and administrative advice on how to assess and improve the quality of open data. However, and as in the case of many other parts of this guide, the local specific context of organisztions, cities or countries should be taken into account, and recommendations should be adapted accordingly.

B. What is open data quality?

There are many different interpretations and ways of measuring data quality, including timeliness of publication, reliability and completeness.[106]

Simply put, high-quality open data is data that is usable, and open data becomes usable "when a human can understand it and a machine can manipulate it".[107] It is quite possible for open data publishers to focus on humans as their target users and forget that datasets can also be used by machines (such as mobile applications). The human-focused approach can lead to poor data quality expressed, for example, in publishing datasets in formats that suit humans but not machines. The case study presented in box 12 provides a good illustration of this.

This issue can become critical for Governments and cities that have smart-city and artificial Intelligence (AI) initiatives, examples of which are the United Arab Emirates, in particular Dubai, and should therefore be given special attention.

Box 12. Quality data is needed for both humans and machines

Our daily life is empowered by abundant software applications and increasingly growing artificial intelligence solutions. Open government data is increasingly becoming an important source for the data that feeds these solutions.

Air pollution, for example, is an issue that directly affects how people work and live on a daily basis. It is becoming a serious environmental and economic risk factor in major cities around the world, including the capital of India, New Delhi.

As part of its efforts to make the public aware of the level of pollution in the city and support air cleaning research and activities, the Government of India publishes an open dataset that provides data on the air quality in New Delhi and many other locations and that is updated on an hourly basis (as shown in the figure). This dataset satisfies many requirements expected from a high-quality dataset as explained later in this chapter, including offering an open licence, a machine-readable format (XML in this case) and the use of metadata. This dataset can be used by both humans and machines in different ways. Researchers and scientists, for example, can download the entire dataset to conduct research. Software developers can access the dataset using the provided application programming interface (API) and create mobile applications that use the data feed to offer its content in a user interface that is more accessible and usable to the average citizen as compared to the sophisticated open data portal.

A good example of these applications is the Plume Air app which offers the residents of New Delhi a visualization of live and historic data from the same dataset. Moreover, the app provides helpful tips and recommendations on the activities (for instance, exercising) that can be carried out in the current level of air pollution. This AI-powered app offers forecasts for the expected pollution levels during the following hours and days. Average citizens can use this app instead of the original dataset, which contains way more data than they need.

Real-time Air Quality Index dataset

Source: Government of India, n.d.

New Delhi's air quality report on Plume Air mobile app
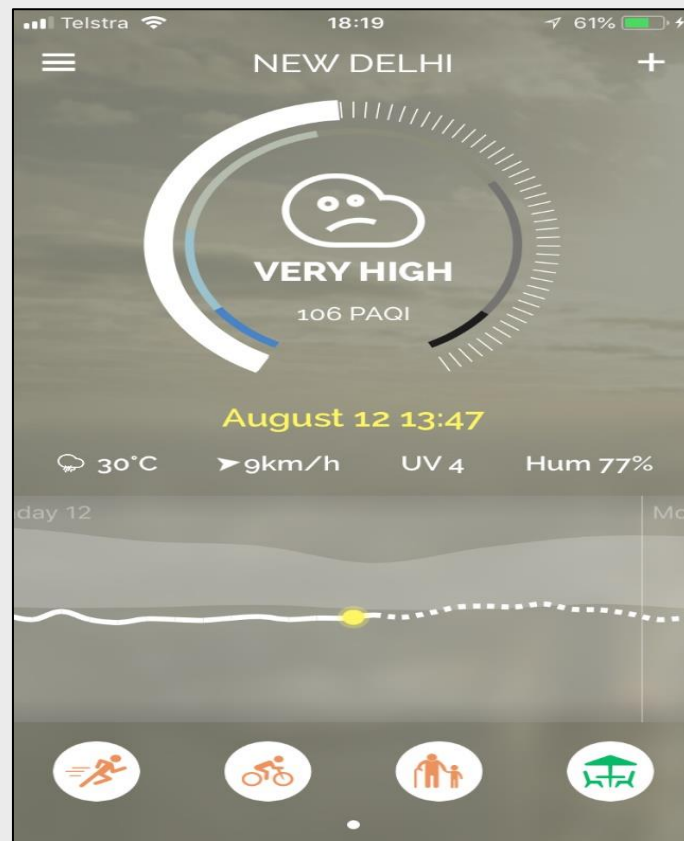
Source: PlumeLabs, 2018.

Defining an inclusive list of detailed criteria for open data quality can be challenging considering that quality can be interpreted in different ways. However, there is a common consent on the outlines of open data quality. The European Commission[108] has compiled these outlines in dimensions that are greatly aligned with the principles of the Open Data Charter and that also overlap with the quality parameters highlighted in the United Nations e-Government Survey 2016.[109] These outlines are as follows:

- Accuracy: Is the data correctly representing the real-world entity or event?

- Consistency: Is the data free from contradictions?

- Availability: Can the data be accessed now and over time?

- Completeness: Does the data include all the items that represent the entity or the event?

- Conformance: Is the data following accepted standards?

- Credibility: Is the data based on trustworthy sources?

- Processability: Is the data machine-readable?

- Relevance: Is the amount of data included appropriate?

- Timeliness: Does the data represent the actual situation and is it published soon enough?

Translating these dimensions into a specific set of applicable quality criteria can be challenging, and, therefore, the Open Data Certificate may be easier to understand and apply. Open Data Certificates were developed by the Open Data Institute and are used or endorsed by governmental organizations and NGOs, including the European Commission.[110] The certificates are used to assess and recognize the sustainable publication of quality open data and based on a set of criteria that are grouped under four categories: legal, practical, technical, and social.[111] A complete list of these criteria under each category is provided in the following.

1. *Legal criteria*

The criteria in this category cover issues related to rights, licensing and privacy; table 14 describes each of them.

Table 14. Open data quality criteria: Legal category

Criteria	Description
Open licence	The statement that describes to people their right to use this dataset and explains what they can and cannot do with the data.
Clear rights statement, detailing any copyrights	Needed if the data or part of it was created by the organization's intellectual effort, for example, by writing text that is within the data, or deciding whether particular data is to be included. There is no copyright if the data only contains facts where no judgements were made about whether to include them or not.
Privacy issues	How does the organization protect people's privacy in the published data? Can individuals be identified from this data? If there is such a risk, then the dataset needs to be de-identified before being published.
Machine-readable rights statement	It is good practice to embed information about rights in machine-readable formats so people can automatically attribute this data back to the author when they use it.

Source: Open Data Certificate, 2017.

2. *Practical criteria*

The criteria under this category (table 15) ensure that open data can be found and relied upon by users.

Table 15. Open data quality criteria: Practical category

Criteria	Description
Accessibility	Data is available on the Internet for the widest range of users including people with disabilities. The data should be in a format that enables its reuse.
Discoverability	Data can be found more easily if, for example, it is linked to the organization's main website or to its publications, or if the dataset is listed somewhere else (e.g. as part of a collection of related datasets).
Data is time-stamped	The date of publishing and update should be part of the metadata associated with the dataset. This ensures that people using the data do not unintentionally use out-of-date information.
Data availability	Will the dataset remain published for a long time? Or could it disappear at any time?
Data timeliness	Data should always be up to date.
Data backup	Backup can be done automatically depending on the technology used by the organization. In all cases, people need to know that they can rely on the availability of the data.
Quality documentation	Each dataset should have a uniform resource locator (URL) (can be part of the metadata) where there is documentation of the quality of data and of any issues with that quality, so as to clarify the reliability of the data.

Source: Open Data Certificate, 2017.

3. *Technical criteria*

These criteria cover technical issues, such as the format in which the data is published.

Table 16. Open data quality criteria: Technical category

Criteria	Description
Machine-readable format	Structured format that can be automatically read and processed by a computer, such as CSV, JSON, XML, etc. PDF is not a machine-readable format.
Open standard machine-readable formats	Open standards enable anyone to implement them and are widely supported, so it would easier to share data with more people. XML, CSV and JSON are all open standards.
Machine-readable provenance documentation	Clarifies the origins of how the data was created and processed before it was published. It builds trust in the data because it enables people to trace back how it has been handled.
Content-appropriate formats	Different data formats suit different datasets depending on the dataset content. For example, geographical data can be plotted on a map as points or boundaries, whereas a tabular format can be more suitable for statistical or numeric data (e.g. census results).
Single consistent URL	A URL for accessing and downloading the data.
URLs as identifiers	Data is usually about real things, such as schools or hospitals, or uses a coding scheme. If data from different sources uses the same persistent and unique identifier to refer to the same things, people can combine sources easily to create more useful data.

Source: Open Data Certificate, 2017.

These criteria aim to make the datasets more usable by human consumers.

Table 17. Open data quality criteria: Social category

Criteria	Description
Data is documented	Documentation about the data helps users understand the data's context and content. For example, documentation about the data API can help developers understand and use it.
Contact details provided	Providing a link (URL) to the webpage containing the contact details of the person(s) responsible for, or could respond to questions about, the data.
Machine-readable metadata (documentation)	Enables users to easily access and use the metadata.
Social media accounts	Social media can be used to promote the data and connect with its users. If the organization uses social media, the accounts should be provided as part of the contact details.
Discussion page or forum	This could simply be the webpage for the dataset itself where people can post a comment to have a discussion, or a dedicated forum on the data portal or the publisher's website.
Dedicated communication team	This team will engage through social media and blogging and arrange "hackathons" or competitions to encourage people to use the data.

Source: Open Data Certificate, 2017.

While applying all four categories of the quality criteria, the open data team should likely notice that the level of implementation of most of these criteria may vary from one organization to another, depending on the organization's needs and capacities. For example, a certain organization may not have the required human capacity or financial resources to frequently update a dataset.

Therefore, it is recommended that Government organizations design their own version of the quality criteria, based on open data principles, but tailored to the organization's business domain, level of readiness, user expectations, and other factors particular to the organization. The organization must observe a high level of commitment to the quality and also to keep improving them over time to boost the quality level of their data.

Box 13. Sample quality criteria from different organizations

Different Government organizations at the department, city and national levels adopt open data quality standards that are consistent with the previously highlighted dimensions, but their criteria vary in their level of detail, focus and strictness.

In Canada, the Open Data Directive of the Government of Ontario has set a number of general quality principles that apply to all Ontario ministries and provincial agencies. The principles are very strictly defined and allow data publishers some space to adapt the principles to their needs. For example, on the timeliness dimension, the directives states that datasets should be as up-to-date "as possible, without compromising accuracy".

In the United Kingdom, the Transport for London's Open Data Policy focuses on providing data for "engaging developers to deliver new products, apps and services for our customers". Transport for London's business model refrains from developing apps and focuses instead on releasing high-quality open datasets that enable more than 11,000 registered developers to develop different mobile apps. Therefore, the brief policy requires publishing data in three ways, namely:

Static data files: Data files which rarely change.

Feeds: Data files refreshed at regular intervals.

D. PROCESS AND TOOLS FOR OPEN DATA QUALITY ASSURANCE

To help data publishers in the Arab region adopt and develop their own customized versions of the quality criteria presented previously, it is recommended to use the following standard process, supported by a set of tools that can facilitate its application.

1. *The quality assurance process*

The process consists of the following six steps (figure 23):

(a) **Define** a standard set of criteria;
(b) **Share** with stakeholders;
(c) **Inform and train** stakeholders;
(d) **Apply** on new datasets;
(e) **Check** the quality of the published datasets;
(f) **Review** and update the quality criteria.

Figure 23. Recommended open data quality assurance process
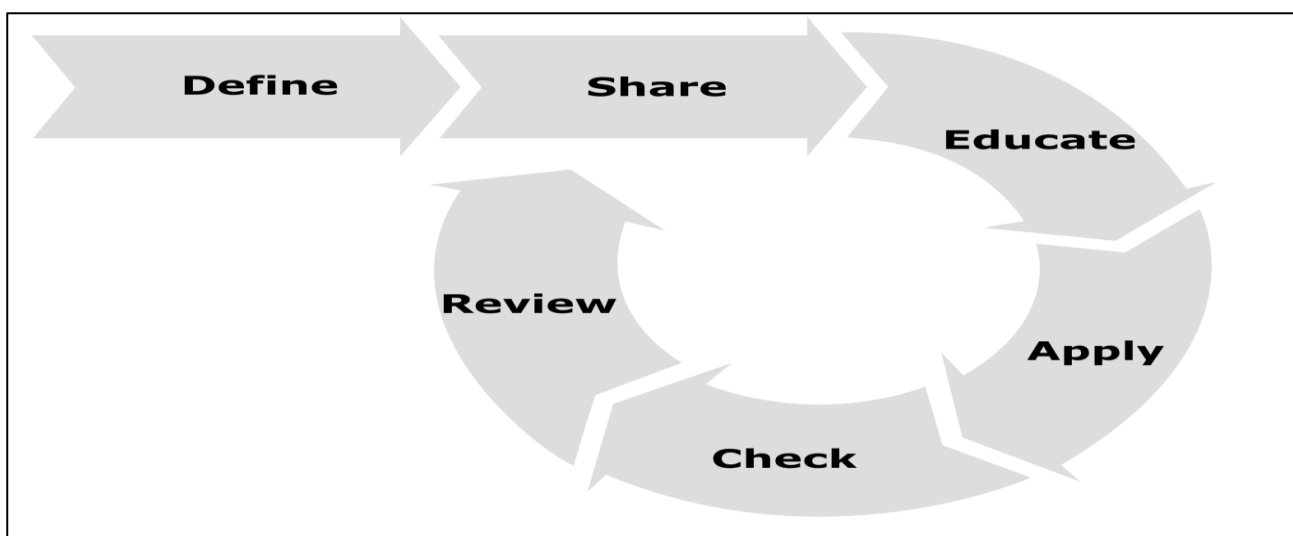
(a)   *Defining the criteria*

The open data team should create a standard set that clarifies the set of criteria approved by the organization. The set will be applicable to all the datasets to be published under the open data programme. This process can be completed as part of the open data policy design process (as highlighted in chapter 2).

Recommended steps to that end may include the following:

- Studying the list of four categories of criteria discussed in the previous section;

- Identifying the quality needs of the organization as determined by its business domain or the range of its datasets, in addition to user quality expectations. These needs and expectations should then be matched with the appropriate level of quality. For example, if the organization's core business produces a lot of Geospatial Information System (GIS) data, then it is essential to use open machine-readable geographic data formats. In all cases, the organization should not go below the minimum criteria that satisfy the definition of open data. Satisfying the criteria of open licence, for instance, is compulsory in this context;

- As recommended in many parts of this guide, the process of defining the quality criteria should be open and collaborative. The open data team should engage the stakeholders in the process to develop the most suitable set of criteria possible, and to promote implementing the criteria.

(b)   *Sharing the criteria with stakeholders*

Once the quality criteria have been defined and approved, the team should share them with all stakeholders, including the public, to promote their usage. The quality criteria can be shared through several channels; coordinators and communication officers in the team have to decide which of those channels are the most suitable. For example, after publishing the policy on the organization's open data portal or on a page on its main website, it could be promoted through social media. Moreover, the policy should be referred to, when appropriate, in all the events and publications created by the organization.

(c)   *Informing and training stakeholders*

Similarly, the open data team should offer all stakeholders the training and information needed to explain the process and prepare them for its application and to utilize its outcomes. It is good practice to offer these training activities via both online and offline channels to accommodate the needs of different stakeholders.

(d)   *Applying on datasets*

As discussed in chapter 3, every dataset has to pass a quality check before being published. This quality check will rely on the quality criteria designed in this process.

(e)   *Checking datasets for quality*

The open data team should carry out a periodic and ongoing quality control procedure to review the published datasets and fix any possible quality issues. For example, the links to some datasets may be broken or the publishers may forget to update certain datasets on time as planned.

(f)   *Reviewing the quality criteria*

The quality criteria themselves need to be reviewed and periodically updated. This may take place once a year. The feedback and suggestions for changes in the criteria can be collected from the following two main resources:

Internally: from the open data team and internal stakeholders. The team should look for suggestions for improving the quality criteria to reflect the development in the organization's readiness (for instance, human or technological) or some changes in the datasets (for instance, need to publish live data);

Externally: from all segments of data users (for instance, researchers or software developers) who may have new requests (for instance, to offer API for certain datasets). Every time a new version of the quality criteria is developed, the team should go through the process of sharing and educating.

2. *Tools*

Several tools can help in applying the different dimensions of the data quality criteria. This section focuses on the following four widely-adopted tools:

(a) Open Data Certificate;
(b) 5-Star Open Data scheme;
(c) International open data indexes.

(a) *Open Data Certificate*

This tool has already been discussed in this chapter in the context of the four-category classification of open data quality criteria. This online tool can be used to assess each published dataset and receive the assessment result, which is given in the form of a digital badge.

The assessment consists of a questionnaire for the data publisher to evaluate the level of support and the steps taken to make data reusable and discoverable. Based on the answers, a digital badge is awarded which can be embedded on the publisher's website to signify the quality level of their publication process.

Four levels of badges are awarded by the certificate:

  (i)   **Bronze**: Data is openly licensed, available with no restrictions, accessible and legally reusable;

  (ii) **Silver**: Satisfies the bronze requirements, data is documented in a machine-readable format, reliable and offers ongoing support from the publisher via a dedicated communication channel;

  (iii) **Gold**: Satisfies the silver requirements, is published in an open standard machine-readable format, has guaranteed regular updates, offers greater support, documentation, and includes a machine-readable rights statement;

  (iv) **Platinum**: Satisfies the gold requirements, has machine-readable provenance documentation, uses unique identifiers in the data, the publisher has a communications team offering support. This is an exceptional example of an information infrastructure.[112]

The certificate is freely available online and can be accessed and used by any agency; it is currently used by many agencies around the world. Examples include the dataset on schools in the Borough of Barnet in North London. Through satisfying most assessment criteria of the practical, legal, social, and technical information, the dataset was awarded a silver badge.[113]

It is important to note that the survey used in the certificate contains some questions that are context-sensitive, and therefore need to be reviewed and possibly updated before being applied in Arab countries. The survey, however, still represents a potential opportunity that can be utilized by Government agencies in Arab countries to develop their own local versions and pioneer efforts for improving the quality of open data in the region.

5-Star Open Data scheme

This famous model can help in carrying out a quick and immediate quality assurance action that can help the agency identify issues such as absence of the open licence.[114] In addition, it can be used as a reference model in the standard process of releasing high-quality data. The model is a rating system for open data proposed by Tim Berners-Lee, founder of the World Wide Web, and it mainly addresses the legal and technical aspects of datasets, mainly, the licence and the data format. Its rating system consists of five stars, and for the dataset to secure a number of the stars, it must fulfil the associated requirements (table 18).

Table 18. 5-Star Open Data model

Number of stars	Requirement
1	Making the dataset available on the Web (in any format) under an open licence.
2	Making the dataset available as structured data (e.g. Excel instead of image scan of a table)
3	Making the dataset available in a non-proprietary open format (e.g. CSV instead of Excel)
4	Using uniform resource identifiers (URIs) to denote data entities so that users can refer to the data easily
5	Linking the data to other data to provide context

Source: 5-Star Open Data, 2012.

It is obvious that applying this 5-star scheme can increase the quality of the published datasets. However, it is also noticeable that it mainly focuses on dataset format and disregards other non-technical factors. Therefore, the scheme "is arguably best used as a technical roadmap and a short-hand assessment of the technical aspects of data publishing".[115]

(c) *International open data indexes*

As part of the global open data movement, several organizations and research institutions have focused their efforts on establishing international indexes for assessing open data programmes around the world. The following are some of the best-known and widely adopted indexes:

Open Data Barometer: from the World Wide Web Foundation in partnership with other organizations. The Barometer assesses national open data programmes according to a set of standard criteria that are grouped under three tiers: readiness, implementation and impact. In its latest edition in 2016, the Barometer included the following 11 Arab countries: Tunisia, United Arab Emirates, Qatar, Bahrain, Saudi Arabia, Morocco, Egypt, Jordan, Palestine, Lebanon, and Yemen. Detailed examples are explained in box 14.

Global Open Data Index: run by the Open Knowledge Network. This index uses the Open Definition to measure and benchmark the publication of open government data. The index was launched in 2013, and in its latest edition (2016) covered 91 countries, including two Arab countries, namely, Tunisia and Oman.

Considering international open data indexes in the organization's approach in addressing open data quality can be helpful because they can offer a global reference on which datasets should be published and which criteria define a high-quality open dataset, and a benchmark with other open data programmes from around the world with varying levels of maturity and progress.

Box 14. Using the Open Data Barometer for quality check

The implementation tier of the Open Data Barometer quantitatively and qualitatively measures the datasets published by a given country. Upon reviewing the performance of Arab countries (as shown in the figure in the box), the report concluded that that none of the 180 datasets analysed from the region qualifies as truly open. Even Tunisia, the highest-ranking Arab country, stands at a modest 50th global rank. This evident low performance is worth noting.

Bahrain is an example of the countries listed in the middle of the table. If Bahrain's profile is further scrutinized and the level of implementation of its open data initiative evaluated, it is noticeable (second figure in the box) that the list of achieved criteria overlaps greatly with the principles developed in the UN e-Government Survey.[a]

Performance of Arab countries in the Open Data Barometer

GLOBAL RANK	SCORE	COUNTRY	READINESS	IMPLEMENTATION	IMPACT
50 ▼	32	Tunisia	45	32	22
59 ▼	26	United Arab Emirates	47	23	12
74 ▼	19	Qatar	41	18	2
74 ▼	19	Bahrain	33	20	7
74 ▼	19	Saudio Arabia	37	15	12
79 ▼	17	Morocco	38	12	7
85 ▼	14	Egypt	27	14	6
87 ▼	13	Jordan	28	11	6
100 NEW	8	Palestine	23	7	2
104 NEW	6	Lebanon	17	7	0
114 ▼	0	Yemen	0	6	0

Source: Open Data Barometer, 2016.

As indicated by the full row of green circles answering the first criteria question ("Does the data exist?"), Bahrain has been successful in publishing all the datasets targeted by the evaluation. Moving down the list to more advanced criteria, however, this gradually changes as the green circles begin to be replaced by yellow and red ones.

When considering the question "Is the data openly licensed?", it is noticeable that all datasets are marked red, which indicates that none of them has an open licence that enables the users to utilize them.

Bahrain profile (implementation dimension)

DATASET SCORED	IMPLEMENTATION														
🔴 Bahrain	15	5	60	80	5	15	15	15	45	15	5	40	5	15	15
Does the data exist?															
Is it available online from government in any form?															
Is the dataset provided in machine-readable and reusable formats?															
Is the machine-readable and reusable data available as a whole?															
Is the dataset available free of charge?															
Is the data openly licensed?															
Is the dataset up to date?															
Is the dataset being kept regularly updated?															
Was it easy to find information about this dataset?															
Are data identifiers provided for key elements in the dataset?															

Source: Open Data Barometer, 2016.

This can be further confirmed by visiting Bahrain Open Data Portal,[b] where no open licence can be found, but a "Copyright © 2015 Information & e-Government Authority" label at the footer instead.

As reference for comparison, the example of the United Kingdom's Open Government Licence may be considered.[116] The licence clearly states that the user is free to do the following:

- Copy, publish, distribute and transmit the information;

- Adapt the information;

- Exploit the information commercially and non-commercially, for example, by combining it with other information, or by including it in products or applications.

The profile of Bahrain is common across Arab countries, where most of the resources are allocated to publishing as much data as possible, and much less attention is given to maintaining the quality of the dataset.

_____

[a] Open Data Barometer, 2016.
[b] Bahrain Open Data Portal. Available at http://www.data.gov.bh/.

E. USER ENGAGEMENT

As discussed in chapter 1, full engagement between Governments, citizens, the private sector, civil society, and other stakeholders is the objective of open government framework. Throughout the steps of the open data quality assurance process, as previously explained, the importance of engaging data users was emphasized since they are the consumers of the data. Moreover, actively engaging the users can ensure that information and data is used in more diverse ways.

It is recommended for open data teams to use the Open Data Engagement Model to engage their users in a planned and structured way. Similar to the 5-Star Open Data scheme, this is a five-star model where a star

represents a key step that can be taken to engage with data users. Each star includes a set of questions that open data teams have to address to achieve the optimal level of engagement (table 19).

Table 19. Open Data Engagement Model

Rating on the engagement scale	Description
★ **1 star** **Be demand driven**	Are your choices about the data you release, how it is structured, and the tools and support provided around it based on community needs and demands? Have you got ways of listening to people's requests for data, and responding with open data?
★ ★ **2 stars** **Put data in context**	Do you provide clear information to describe the data you provide, including information about frequency of updates, data formats and data quality? Do you include qualitative information alongside datasets such as details of how the data was created, or manuals for working with the data? Do you link pages from data catalogue to the analysis of the data that your organization or third parties have already carried out with it, or to third-party tools for working with the data?
★ ★ ★ **3 stars** **Support conversation around data**	Can people comment on datasets or create a structured conversation around data to network with other data users? Do you join the conversations? Are there easy ways to contact the individual 'data owner' in your organization to ask them questions about the data or to get them to join the conversation? Are there offline opportunities to have conversations that involve your data?
★ ★ ★ ★ **4 stars** **Build capacity, skills and networks**	Do you provide or link to tools for people to work with your datasets? Do you provide or link to how to guide on using open data analysis tools so people can build their capacity and skills to interpret and use data in the ways they want to? Do you go out into the community to run skill-building sessions on using data in particular ways, or using particular datasets? Do you sponsor or engage with capacity-building to help the community work with open data?
★ ★ ★ ★ ★ **5 stars** **Collaborate on data as a common resource**	Do you have feedback loops so people can help you improve your datasets? Do you collaborate with the community to create new data resources (e.g. derived datasets)? Do you broker or provide support to people to build and sustain useful tools and services that work with your data? Do you work with other organizations to connect up your data sources.

Source: Open Data Engagement n.d.

# VII.	Case studies and success stories

VII. CASE STUDIES AND SUCCESS STORIES

A. OPEN DATA STRATEGY IN IRELAND

Ireland's open data programme[117] represents a successful experience that can be leveraged from because it includes a comprehensive strategy and a clear vision starting with the development of open data policy supported by the necessary laws and legislation. Despite differences in terms of context between Ireland and most Arab countries, the Irish experience can be considered and benefitted from in developing open data policies in the Arab region.

Ireland's open data strategy[118] aims to encourage and drive the uptake and use of open data and add value to the economy by increasing transparency, stimulating new business applications, building trust in the Government and improving the lives of citizens by delivering better services. This strategy is aligned with reforms and key Government priorities such as the ICT strategy, the public service reform programme and the civil service renewal plan, which increase the efficiency and effectiveness of the public service. It is also aligned with the Open Government Partnership Action Plan, the proposed National Data Infrastructure and the reuse of Public Service Information directive and regulations.

1. *Legislations*

As explained earlier, the access to information law represents the legislative backbone of the open government and open data initiatives in order to increase transparency, participation and full citizen engagement in decision-making. The Irish Government realized the importance of this law; therefore, it adopted the Freedom of Information Act 1997 (FOI) (amended in 2003 and again in October 2014). The FOI Act is defined as n Act "to enable members of the public to obtain access, to the greatest extent possible, consistent with the public interest and the right to privacy, to information in the possession of public bodies, other bodies in receipt of funding from the State and certain other bodies and to enable persons to have personal information relating to them in the possession of such bodies corrected and, accordingly, to provide for a right of access to records held by such bodies".

The FOI Act obliges the public sector and a wide range of Government entities to publish data related to their work and activities and make them available to citizens.[119] It provides citizens with the right to access data in governmental entities, to amend official information relating to themselves where it is incomplete, incorrect or misleading, and to obtain reasons for decisions affecting them.

2. *Governance of the strategy*

This strategy is being implemented and overseen by the Open Data Governance Board supported by the Public Bodies Working Group and the Open Data Unit in the Department of Public Expenditure and Reform.

3. *Principles and strategic themes*

Ireland's open data strategy includes the essential and important principles of open data, previously discussed in this document, such as the openness of data by default (except in sensitive cases related to the national security or privacy). It also encourages the publication of high-quality datasets that are accurate, authoritative, interoperable, regularly updated, and linked to the national open data portal (data.gov.ie) while following a uniform technical framework such as standards, formats and metadata. These principles are supported by life-cycle approach to maintain and manage datasets starting from creation and publishing to ensure continuity.

The strategy has set out seven strategic themes which will guide the open data initiative over the lifetime of the strategy (2017-2022) and which are listed in table 20.

Table 20. Themes of Ireland's open data strategy

Theme 1	Broaden the range of public bodies actively engaged in the open data initiative (open data providers)
Theme 2	Broaden the scope and improve the quality, quantity and range of open data and associated metadata (to be used internally and made available to the public); improve the quality and range of services provided through the national open data portal
Theme 3	Continue to engage with all stakeholders and encourage use of open data
Theme 4	Support and encourage various groups of open data users
Theme 5	Provide a framework to support and train all data providers and build capacity in the management and use of open data
Theme 6	Evaluate the impact, benefits and risks of the open data initiative and benchmark against other jurisdictions
Theme 7	Ensure that effective governance structures are in place to implement the strategy

Source: Open Data Unit.

4. *Outcomes*

The intended benefits from applying an open data strategy in Ireland include increasing transparency and accountability, enhancing innovation in businesses, attracting international and local IT firms, and facilitating citizen participation and engagement in policy developments and decision-making.

The key output of Ireland's open data initiative is the national open data portal which provides access to official non-personal Government data in open format. This portal was originally launched in July 2014 and has been significantly enhanced since then. By mid-2018, the portal linked to some 8,686 datasets from 102 publishers. Table 21 provides some of the important open data initiative portals presented from a variety of entities and sectors, in addition to the governmental portal.

Table 21. Samples of open data portals in Ireland

Portal	Description[a]
Ireland's Open Data Portal https://data.gov.ie/	Provides easy access to datasets that are free to use, reuse and redistribute. The portal is operated by the Government Reform Unit of the Department of Public Expenditure and Reform.
Dublin Dashboard http://www.dublindashboard.ie	Provides real-time information, time-series indicator data, and interactive maps about all aspects of the city. It enables users to gain detailed, up-to-date intelligence about the city that aids everyday decision-making and fosters evidence-informed analysis.
All-Island Research Observatory http://airo.maynoothuniversity.ie/	Undertakes academic and applied mapping research and produces spatial datasets and specialist tools to aid in their analysis.
Transport for Ireland https://www.transportforireland.ie/	Real-time passenger information helps members of the general public plan their journeys better. It provides data including timetables and travel information from all licensed public transport providers across Ireland, including information on train, bus, tram, ferry, and taxi services.

Portal	Description[a]
Myplan.ie http://myplan.ie/index.html	Myplan.ie is a Web map portal providing spatial information relevant to the planning process in Ireland. This site is an initiative of the Department of Housing, Planning, Community and Local Government in conjunction with Irish local authorities.
Heritagemaps.ie http://heritagemaps.ie/	The Heritage Maps viewer is a Web-based spatial data viewer which focuses on the built, cultural and natural heritage around Ireland and off-shore.Heritage Maps allows users to look at a wide range of built and natural heritage data sets in map form, much of which has never been accessible to the public before.
Beaches.ie https://www.beaches.ie/	The national bathing water information website developed by the Environmental Protection Agency. The bathing water monitoring results and information available from 190 beaches are provided by 10 local authorities. Data about tides from the Marine Institute is also included.
Smart Dublin https://smartdublin.ie/	Smart Dublin is an initiative of the four Dublin local authorities to engage with smart technology providers, researchers and citizens to solve challenges and improve city life.
Cork Smart Gateway http://www.corksmartgateway.ie/	This initiative will compliment and support the already large number of organizations in Cork who are working on smart solutions to improve the physical realm, from energy conservation to open data.

Source: Open Data Unit.

Note: Most descriptions were extracted from the websites.

B. CASE STUDIES ON OPEN DATA

Although most of the government open data programmes worldwide have been launched only recently, some of them are well advanced and have shown signs of success; such programmes are the open data programmes of France and the Republic of Korea.

The open data programme of France was only launched in 2014. However, massive investment was made in establishing this programme, which was placed under the supervision of the prime minister. This gave the programme strong political support from the start. The French Government succeeded in convincing public institutions and ministries to publish data at the national and local levels. In addition, the Government succeeded in launching a dynamic around the creation of applications based on open data, and even supported start-ups in launching value-added services using open data. The governance of the open data programme is based on two newly established entities under the prime ministry. One is the Chief Data Officer who leads the programme and champions it among relevant organizations. The second is Etalab, a unit at the Directorate of Inter-ministerial Information System in charge of implementing actions in the area of open data and maintaining and updating the open data portal. Both entities have succeeded in working with the civil society and Government entities mainly through a network of ministerial data officers. The open data programme in France is one of the most dynamic and efficient in the world, and it benefited from two strong laws related to the right to access information and the digital law, where provisions about open data were included.

Another world leader in government open data is the Republic of Korea, a country that established a strong national open data programme with supportive laws on open data and the right to access information and a powerful fully linked and interactive open data platform. Like France, the Republic of Korea scaled very

fast in this area and succeeded in creating a dynamic around released data and in developing partnerships with the private sector, especially start-ups.

Some Arab countries have started their open data programmes in parallel with those countries but have not achieved the same level of success. Tunisia was one of the first countries to launch a national open data portal in 2012 in addition to other sectorial open data portals, based on the Open Data Decree Law of 2011 and the Organic Law on the Right of Access to Information of 2016. Although the latter has no provisions on data reuse, Tunisia is still well positioned to build a strong programme with a number of running actions such as increasing the coverage of its national data inventory, developing a data reuse law and updating its open data decree.

The United Arab Emirates enrolled in the trend of opening data and established a strong legal and organizational framework to accelerate the publication of Government agency data. To that end, data owned by the Government was classified as open or shared data, where open data was to be published by the Government or private-sector entities and used or exchanged with individuals or third parties openly. While shared data is made available for sharing and reusing among Government entities, subject to certain terms and conditions. A strongly linked open data national platform was also developed. However, a dedicated entity in charge of open data programme development still needs to be created at a high Government level to boost the release and exploration of open government data.

The Government of Jordan has shown strong commitment to enhancing transparency and freedom of access to public information. Since 2007, it has promulgated the Right to Information Law and the Law on Securing the Right to Information Access (No. 47 of 2007). The latter was the first of its kind in Arab countries.

Table 22 summarizes the state of open data programmes in the five countries discussed in this section.

Table 22. Summary of the five country case studies

Category	Indicator	Description	France	Korea	Tunisia	United Arab Emirates	Jordan
	Political support level	Political level supporting the policy	Prime Ministry	Ministry of the Interior and Safety	Prime Ministry	Prime Minister	Minister of ICT
Policy	Key motivation	Key motivation behind establishing the policy	Produce essential data, spread around, encourage exploring data[a]	- Guarantee citizens' right to access public data; - Contribute to improving their quality of life; - Utilize such public open data in the private sector.	- Enhance transparency of public action; - Utilize public open data for creation of value-added services.	- Fulfil Dubai Smart City; - Enhance transparency; - Increase efficiency of services.	- Ensure transparency; - Encourage innovation; - Increase participation; - Provide income opportunities; - Enhance decision-making.
	Policy document	Text of the policy	- Vademecum on Open Data;[b] - Road map.[c]	Government 3.0 Policy	Action plan prepared jointly with World Bank	Dubai Open and Shared Data Framework[d]	Open government data policy[e]
	Date of publication	Date policy approved	February 2013 and updated in 2017	2013	October 2017	February 2016	August 2017
Legislation	Right to information (RTI) law		CADA Law (loi 78-753 du 17 juillet 1978)	Act on Disclosure of Information by Public Agencies, Act No. 5242, 1996[f]	Loi organique n° 2016-22 du 24 mars 2016, relative au droit d'accès à l'information[g]	--	Law on Securing the Right to Information Access Law No.47/2007
	RTI law ranking[h]		94	61	11	--	105
	Open data legislation		Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique[i]	Act on Promotion of the Provision and Use of Public Data, Act No. 11956, 2013[j]	Decree being prepared	Law No. 26 of 2015 on Regulating Data Dissemination and Exchange in the Emirate of Dubai[k]	--

Category	Indicator	Description	France	Korea	Tunisia	United Arab Emirates	Jordan
	Political support level	Political level supporting the policy	Prime Ministry	Ministry of the Interior and Safety	Prime Ministry	Prime Minister	Minister of ICT
	Privacy law		Loi informatique et libertés	Information and Communications Networks Act No. 6360, January 16, 2001[l]	Organic Law No. 2004-63 of 24 July 2004 on Protection of Personal Data[m]	--	Draft version, yet to be concluded officially
Governance	**Responsible Central body**		Prime Ministry - Administrateur général des données; - Etalab.	Ministry of the Interior and Safety (MOIS)	Prime Ministry E-government Unit	Smart Dubai	Ministry of ICT
	Open data board	Existing board or committee governing the open data programme	--	- Public Open Data Strategy Committee (consisting of chairpersons, one of whom is the prime minister, secretary is the minister of the interior); - Working committee.	--	Open Data Committee	e-Government Steering Committee Chief Information Officer Council
	Others	Any other committee or group	Ministry data administrator	Committee on Mediation of Disputes	Expert group	--	Joint Committee on Open Government Data

Category	Indicator	Description	France	Korea	Tunisia	United Arab Emirates	Jordan
	Political support level	Political level supporting the policy	Prime Ministry	Ministry of the Interior and Safety	Prime Ministry	Prime Minister	Minister of ICT
Data & portal	URL	URL of central OD portal	data.gouv.fr	www.open.go.kr	Data.gov.tn	bayanat.ae/	jordan.gov.jo
	Year open government data portal launched		2013	2011	2012	2018	2018
	Managing agency	Agency that built and maintains the portal	Prime ministry (Etalab)	MOIS, National Information Agency	Prime ministry (e-Government Unit)	Federal Competitiveness and Statistics Authority	Ministry of ICT; Policies and Strategies Directorate and e-Government Programme
	Services and features	Existing services in the open government data portal aside from offering data	Search engine, open API, data visualizations, applications, Q&A, forum, start-up space, community activities	Search engine, open API, data visualizations, requesting open data, use cases, Q&A, forum, start-up space, dispute mediation	Search engine, open API, data visualizations, requesting open data	Search engine, selected visualizations, geo data, request data, example of OD applications	Request data, send complaint
	Licensing	Data licence published on the portal	Two licences authorized: Open Licences of Etalab and Open Data Base Licence (ODbL-1.0)	Creative Commons Licence Korean	Open Tunisian Licence, Creative Commons Attribution 1.0	No licence was found on the portal	No licence used, refer users to policy
	Restricted data	Data that cannot be accessed	Defined by law - Anonymization; - Intellectual property; - Commercial and industrial secrets.	Defined by law[n]	Defined by law[o]	Data is classified into open data and shared data, in accordance with adopted data manual[p]	Defined by law[q]

Category	Indicator	Description	France	Korea	Tunisia	United Arab Emirates	Jordan
	Political support level	Political level supporting the policy	Prime Ministry	Ministry of the Interior and Safety	Prime Ministry	Prime Minister	Minister of ICT
	Fees for data	Any data that can be charged for	Gratuity generalized for all sectors, data of industrial and commercial public services stay excluded[r]	Indicated in metadata	By law, all public data is free	No fees applied	Policy specifies that open data can be used without any fees or costs. Additional requests for open government data can be charged to cover costs for preparation and customization.
	Number of datasets published	Number of datasets on the portal	21,420	22,024 datasets, 2,540 API	1,174	784	80
	Interoperability between data	Any linked data	API	Open API	API	--	--
	Data reuse	Number of reuse cases	1,793[s]	148	4	29	--
Start-ups	Number of start-ups[t]	Number of start-ups in the country	1,356	295	35	235	77
Open government	OGP member		Yes (2014)	Yes (2011)	Yes (2014)	No	Yes (2011)
Funding	Existing funds		Government	Government	Government and World Bank	Government	Government
E-Government Readiness Score (rank)[u]	E-Government Dev. Index		0.8790 (9)	0.9010 (3)	0.6254 (80)	0.8295 (21)	0.5575 (98)
	Online Service Index		0.9792	0.9792	0.8056	0.9444	0.4931
	Human Capital Index		0.8598	0.8743	0.6640	0.6877	0.7387

Category	Indicator	Description	France	Korea	Tunisia	United Arab Emirates	Jordan
	Political support level	Political level supporting the policy	Prime Ministry	Ministry of the Interior and Safety	Prime Ministry	Prime Minister	Minister of ICT
	Telecom. Infra. Index		0.7992	0.8496	0.4066	0.8564	0.4406
	E-Participation Index		0.9663 (13)	1.000 (1)	0.7978 (53)	0.9438 (17)	0.4831 (117)
Index Score (ank)	**OD Barometer (2016)[v]**		85.13 (3)	81.16 (5)	32.17 (50)	26.17 (59)	12.58 (87)
	Statistic OD Inventory (2017)[w]		57 (33)	66 (17)	37 (97)	51 (51)	38 (87)

a. See: https://www.etalab.gouv.fr/wp-content/uploads/2018/04/RapportAGD_2016-2017_web.pdf.
b. See: http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/vademecum-ouverture.pdf.
c. See: https://www.etalab.gouv.fr/qui-sommes-nous.
d. See: http://dubaidata.ae/pdf/Dubai-Open-Shared-Data-Framework-20170427.pdf.
e. See: http://moict.gov.jo/uploads/Policies-and-Strategies-Directorate/Policies/Open-Government-Data-policy.pdf.
f. See: http://www.rti-rating.org/wp-content/themes/twentytwelve/files/pdf/South%20Korea.pdf.
g. See: http://www.rti-rating.org/wp-content/uploads/Tunisia.pdf.
h. See: http://www.rti-rating.org/country-data/.
i. See: https://www.legifrance.gouv.fr/eli/loi/2016/10/7/ECFI1524250L/jo.
j. See: http://elaw.klri.re.kr/eng_mobile/viewer.do?hseq=30365&type=part&key=4.
k. See: http://ogp.dubai.gov.ae/documants/pdf/ltiwndiymjezmdk.pdf.
l. See: http://www.koreanlii.or.kr/w/images/d/df/DPAct2014_ext.pdf.
m. See: http://www.inpdp.nat.tn/ressources/loi_2004.pdf.
n. See: http://elaw.klri.re.kr/eng_mobile/viewer.do?hseq=30365&type=part&key=4.
o. See: http://www.rti-rating.org/wp-content/uploads/Tunisia.pdf.
p. See: http://dubaidata.ae/pdf/Dubai-Open-Shared-Data-Framework-20170427.pdf.
q. Jordan Right to Access to Information Law No.47 of 2007.
r. Loi relative à la gratuité et aux modalités de la réutilisation des informations du secteur public n° 2015-1779 du 28 décembre 2015, http://frama.link/Loi_OD_gratuite.
s. See: https://www.data.gouv.fr/fr/reuses/.
t. See: https://www.startupranking.com/countries.
u. See: https://publicadministration.un.org/egovkb/Data-Center.
v. See: https://opendatabarometer.org/?_year=2016&indicator=ODB.
w. See: http://odin.opendatawatch.com/.

VIII.   Recommendations

VIII.   RECOMMENDATIONS

Throughout the chapters of this guide, the importance of open data and the processes of planning and implementing open data strategies and policies has been explained in detail. Some Arab countries have made concrete steps toward establishing open data programmes, yet several shortcomings are noticeable in these programmes. Consequently, a number of measures and actions need to be taken to adjust and boost them. In this regard, some points and actions that may be more suitable to the status of open data programmes in Arab countries are recommended to be taken into consideration while developing open data initiatives or improving them:

(a)   The success of open data programmes requires a clearly announced high-level political support. In this context, it is recommended that high-level political announcements be made in support of the open data strategy or policy and for them to include a number of clear actions to boost the open data ecosystem. It is also recommended to instruct public institutions to work on identifying datasets to be published as open data and release them within well-defined timelines;

(b)   It is necessary to set up a permanent mechanism for follow-up, evaluation and monitoring of the open data programme. An annual report could be prepared by the open data unit or team and submitted to the political entity responsible for the programme (such as the prime minister) for review. It is important to set up an evaluation and monitoring action plan as part of the national strategy in the area of open data;

(c)   A law that clearly defines the right to access information has a fundamental impact on strengthening accountability and transparency through the extension of oversight to the performance of governments, both at the central and local levels. In the Arab region, some countries (Jordan, Yemen, Lebanon, Morocco and Tunisia) have passed legislation on access to government data, while others are still working on it. These provisions fall within the context of open government activities in these countries. It is highly recommended that other Arab countries cover this legislative aspect while developing their open government programmes;

(d)   An open data programme also requires the development of other specific open data legislations, such as an open data law. An open data law is a powerful tool that can quickly help establish a strong open data ecosystem. Opening data should be enforced and clearly defined by law. If the access to information law clearly defines what data may be disclosed, then it is possible to opt for supplementing it with a decree/by-law that includes all aspects related to the development of an open government data programme such as obligations to release data, type of data to be released, types of licences attached to datasets, open data portal, data inventory, and more;

(e)   To insure successful coordination between ministries and public institutions involved in the open data programme, it is important to designate an open data coordinator in each government entity. The open data coordinator should be given enough resources to implement the programme and collaborate with the responsible open data unit or team;

(f)   Capacity-building is very crucial for the enrolment of public services in the dynamics of open data. It is recommended that an intensive training programme be defined and established to introduce open data. Special training should be given to Government employees working in the area such as on formats of open data, metadata, open data portals, and legal frame work, among others. Training of trainers is another important approach in spreading knowledge about open data in the public sector;

(g)   Raising awareness among the private sector, civil society and the general public about the advantages of open data is important. It is recommended that a communication plan be established to educate people about open data, the utility of releasing it and the forms of reusing it. Furthermore, it is important to show efforts made by governments to establish such programmes;

(h)   In order to provide technical assistance to different ministries and public institutions in the area of open data, it is recommended that a technical unit be established within the ministry of ICT, or a similar Government entity, to provide other open data teams with technical assistance in releasing data, data formats, metadata, and more;

(i)   Certain sectors, such as transport, agriculture or education, have better understanding of the benefits of open data and are usually more prepared to launch an open data programme in their departments. It is recommended that priority be given to those departments that have shown interest in open data;

(j)   As discussed in chapter 4, it is recommended to develop a new dedicated open data portal that follows the international standards and includes functions such as a search engine, open API, data visualizations, requesting open data, use cases, questions and answers, a forum, start-up space, dispute mediation, and more;

(k)   Technical aspects are very important in open data programmes. It is recommended that all public agencies coordinate in that respect to insure at least minimum conformity in produced datasets, and to set up standards in data exchange and formats. For this, it is necessary to define a common standard of metadata for all open datasets. This should facilitate exchange of data and link of datasets;

(l)   One of the pillars of the open data programme is the establishment of a national open data inventory. It is important that such project be undertaken with the help of enrolled structures in order to identify potential datasets that could be released. The inventory should be conducted according to a predefined methodology and based on a unified format (as discussed in chapter 3).

(m)   In parallel to the data inventory, it is important to develop internal use cases of open data. These are projects undertaken by Government services to develop applications based on some examples of open government data. Those examples should be published on the open data portal to demonstrate possible ways of open data exploration;

(n)   In order to trigger the use of open datasets, it is recommended to organize data exploration events called hackathons. Both Government representatives and representatives from private sector and NGOs can participate in such events to boost exploration of published data. Hackathons could be organized based on a partnership between data publishers and users in the Government, the private sector, universities, and the civil society;

(o)   The success of an open data programme depends greatly on the level of interaction with all stakeholders (as discussed in chapter 2), especially the civil society, either through consultation on the policy and standards adapted, or through partnership in developing examples of data reuse or the organization of hackathons and other open data events. A mixed committee for the open data programme is highly needed to gather support for the programme and obtain the necessary support for it;

(p)   It is important that a dedicated budget be allocated for the open data programme. The budget should include components such as development and maintenance of the data portal, developing open data regulatory framework, organizing hackathons, training sessions, seminars and workshops, conducting data inventories, and others;

(q)   International collaboration programmes can be useful in launching open data programmes since they provide expertise and know-how.

References

5-Star Open Data (2015). Available at http://5stardata.info/en/.

Al Munjid, Mohammed (2016), Open Government for Increasing Transparency and Accountability in the Arab Region. Paper prepared for the United Nations Economic and Social Commission for Western Asia.

Australian Capital Territory, Australia (n.d.). Open Data Portal DataACT. Available at https://www.data.act.gov.au/.

Australian Government, Australia (n.d.). data.gov.au. Available at https://data.gov.au/.

Australian Government (2015). Public Data Policy Statement. Available at https://www.pmc.gov.au/sites/default/files/publications/aust_govt_public_data_policy_statement_1.pdf.

Australian Government (2016). Open Government Data and Why It Matters. Available at https://www.communications.gov.au/file/14716/download?token=KSwxZ_B6.

Buienradar (2018). Available at https://www.buienradar.nl/.

City of San Francisco (n.d.). Datasf.org. Available at https://datasf.org/progress/.

CityMapper (n.d.). Available at https://citymapper.com/london.

Creative Commons (n.d.). Australian Government Licence. Available at https://creativecommons.org/licenses/by/3.0/au/deed.en.

Data.Gov (2018). Crimes: 2001 – present. Available at https://catalog.data.gov/dataset/crimes-2001-to-present-398a4.

Department of Transport and Main Roads, Australia (n.d.). Open Data Maturity Assessments. Available at https://data.qld.gov.au/dataset/department-of-transport-and-main-roads-open-data-maturity-assessments.

Dodds, Leigh (2015). Comparing the 5-Star Scheme with the Open Data Certificates, Open Data Institute. Available at https://oldsite.theodi.org/blog/5-star-open-data-certificates-tim-berners-lee.

European Commission (2011). Digital Agenda: Commission's Open Data Strategy. Available at http://europa.eu/rapid/press-release_MEMO-11-891_en.htm.

European Commission (2015). Creating Value through Open Data. Available at https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf.

European Commission, (n.d.). Achieving Impact with Open Data. Available at https://www.europeandataportal.eu/elearning/en/module16/#/id/co-01.

European Data Portal. Creating Value through Open Data.

Government of India (n.d.). Data.gov.in. Available at https://data.gov.in/catalog/real-time-air-quality-index.

GovLab (2018). OpenData500. Available at http://www.opendata500.com/.

Gov.uk (2017). Letter from the Prime Minister on government transparency and open data. Available at https://www.gov.uk/government/publications/letter-from-the-prime-minister-on-government-transparency-and-open-data.

Humanitarian Data Exchange (n.d.). Available at https://data.humdata.org/.

Institute on Governance (2003). Principles for Good Governance in the 21st Century. Policy Brief No.15. Available at http://unpan1.un.org/intradoc/groups/public/documents/UNPAN/UNPAN011842.pdf.

Jarrar, Yasar (2017). What is the role of government digital age data? *World Economic Forum Agenda*. Available at https://www.weforum.org/agenda/2017/02/role-of-government-digital-age-data/.

Jitzek, Thorhildur, Michel Avital and Niels Bjorn-Anderson (2013). Generating Value from Open Government Data, Thirty Fourth International Conference on Information Systems.

Lathrop, Daniel and Laurel Ruma (2010). Open Government: Transparency, Collaboration and Participation in Practice. O'Reilly Media.

Manyika, James and others (2013). Open data: Unlocking innovation and performance with liquid information. McKinsey Global Institute.

Ministry of Energy and Industry, United Arab Emirates (n.d.). Open Data Page. Available at https://www.moei.gov.ae/en/open-data.aspx.

Ministry of Transport and Communications, Qatar (2014). Open Data Policy. Available at http://www.motc.gov.qa/en/documents/document/open-data-policy.

National Archives (n.o d.). Open Government License. Available at http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/.

New York City (2017). Open Data for All 2017 Progress Report. Available at https://www1.nyc.gov/assets/doitt/downloads/pdf/OD4A-report_2017.pdf.

NSW Government, Australia (n.d.). Data.NSW. Available at https://data.nsw.gov.au/data/dataset.

Open Data Barometer (2016). Available at http://opendatabarometer.org/?_year=2016&indicator=ODB.

Open Data Certificate (2017). Available at https://certificates.theodi.org/en/datasets/215670/certificate.

Open Data Certificate (n.d.). Available at https://certificates.theodi.org/en. Accessed on 21 January 2018.

Open Data Charter (2013). Available at https://opendatacharter.net/.

Open Data Charter (2015). Principles. Available at https://opendatacharter.net/principles/.

Open Data Engagement (n.d.). Available at http://www.opendataimpacts.net/engagement/.

Open Data Impacts (n.d.). Open Data Engagement. Available at http://www.opendataimpacts.net/engagement/.

Open Data Institute (2015). Open Data Maturity Model. Available at https://theodi.org/guides/maturity-model.

Open Data Institute (n.d.). How to plan and budget an open data initiative. Available at https://theodi.org/article/how-to-plan-and-budget-an-open-data-initiative/.

Open Data Institute (n.d.). Publisher's Guide to Open Data Licensing. Available at https://theodi.org/article/publishers-guide-to-open-data-licensing/.

Open Data Institute (n.d.). The Data Spectrum. Available at https://theodi.org/data-spectrum.

Open Data Institute (n.d.). What is open data? Available at https://theodi.org/what-is-open-data.

Open Data Institute Certificate (n.d.). Available at https://certificates.theodi.org/en/about/badgelevels.

Open Data Monitor (2016). Available at https://project.opendatamonitor.eu/.

Open Definition (n.d.). Open Definition 2.0. Available at http://opendefinition.org/od/2.0/en/.

Open Data Portal, Saudi Arabia (n.d.). Available at http://www.data.gov.sa/en/about.

Open Data Unit. Open Data Strategy 2017-2022, Ireland.

Open Government Partnership (2011). Open Government Declaration. Available at https://www.opengovpartnership.org/open-government-declaration.

Open Government Partnership (n.d.). About OGP. Available at https://www.opengovpartnership.org/about/about-ogp.

Open Knowledge Foundation (2013). Defining Open Data. Available at https://blog.okfn.org/2013/10/03/defining-open-data/.

Open Knowledge Foundation (2017). Open Data Index. Available at https://index.okfn.org/.

Open Knowledge International (2017a). Open Data Handbook. Available at http://opendatahandbook.org/.

Open Knowledge International (2017b). Open Data Quality: the Next Shift in Open Data? Available at https://blog.okfn.org/2017/05/31/open-data-quality-the-next-shift-in-open-data/.

Organization for Economic Cooperation and Development (2015). Rebooting Public Service Delivery – How Can Open Government Data Help Drive Innovation?. Available at https://www.oecd.org/gov/Rebooting-Public-Service-Delivery-How-can-Open-Government-Data-help-to-drive-Innovation.pdf.

Organization for Economic Cooperation and Development (2016). Open Government: The Global Context and the Way Forward.

Organization for Economic Cooperation and Development (2017). Open Government Data. Available at http://www.oecd.org/gov/digital-government/open-government-data.htm.

Parisdata, France (n.d.). Available at https://opendata.paris.fr/page/home/.

Plume Labs (2018). Available at https://plumelabs.com/en/.

Presidency of the Government, Tunisia (n.d.). Open Data Portal. Available at http://www.data.gov.tn/.

Press, Gil (2014). 12 Big Data Definitions. Available at
https://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-
yours/#5ccd573113ae.

Sollazzo, Giuseppe (2015). Four Challenges for the Future of Open Data. Available at
https://data.blog.gov.uk/2015/03/23/four-challenges-for-the-future-of-open-data/.

Supreme Legislation Committee in the Emirate of Dubai (2016). Dubai Data Law. Available at
http://ogp.dubai.gov.ae/documants/pdf/ltiwndiymjezmdk.pdf.

Transport for London (2014). Why a new website but no app? Part 1. Available at
https://blog.tfl.gov.uk/2014/03/12/why-a-new-website-but-no-app-part-1/.

Transport for London (2016). More than 2,000 new developers sign up for TfL's open data in last six
months. Available at https://tfl.gov.uk/info-for/media/press-releases/2016/april/more-than-2-000-new-
developers-sign-up-for-tfl-s-open-data-in-last-six-months.

Transport for London (n.d.). Our Open Data. Available at https://tfl.gov.uk/info-for/open-data-users/our-
open-data#on-this-page-0.

Transport for London (n.d.). Unified API. Available at https://tfl.gov.uk/info-for/open-data-users/unified-api.

Transport for NSW, Australia (n.d.). Open Data. Available at https://opendata.transport.nsw.gov.au/.

United Kingdom's Cabinet Office (n.d.). Data.Gov.UK. Available at https://data.gov.uk/.

United Kingdom Department for Communities and Local Government (2015). Local Government
Transparency Code. Available at https://www.gov.uk/government/publications/letter-from-the-prime-
minister-on-government-transparency-and-open-data.

United Kingdom Government Digital Service (2017). Transformation Strategy: Better use of data. Available
at https://www.gov.uk/government/publications/government-transformation-strategy-2017-to-
2020/government-transformation-strategy-better-use-of-data.

United Nations Economic and Social Commission for Asia and the Pacific (n.d.). What is Good
Governance? Available at https://www.unescap.org/sites/default/files/good-governance.pdf. Accessed
28 December 2017.

United Nations Economic and Social Commission for Western Asia (2018). Fostering Open Government in
the Arab Region. Available at
https://www.unescwa.org/sites/www.unescwa.org/files/page_attachments/study-fostering-open-
government-arab-region-ar.pdf.

United Nations (2014). A World that Counts. Available at http://www.undatarevolution.org/wp-
content/uploads/2014/11/A-World-That-Counts.pdf.

United Nations (2015). United Nations General Assembly. Available at
http://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A
_RES_70_1_E.pdf.

United Nations (2016). United Nations e-Government Survey. Available at
https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2016.

United Nations (2018). E-Government Survey.

Verhulst, Stefaan G. (2017). Open Data in Developing Economies. Toward Building an Evidence Base on What Works and How. African Minds.

Verhulst, Stefaan G. and Andrew Young (2017). Open data in developing economies: Toward building an evidence base on what works and how. Available at http://odimpact.org/developingeconomies.htm. Accessed 13 August 2018.

White House (2009). Open Government Directive. Available at https://obamawhitehouse.archives.gov/open/documents/open-government-directive.

World Bank (2014). Open Data for Economic Growth. Available at http://www.worldbank.org/content/dam/Worldbank/document/Open-Data-for-Economic-Growth.pdf.

World Bank (2017). The Open Data Toolkit. Available at http://opendatatoolkit.worldbank.org/en/open-data-in-60-seconds.html.

Endnotes

1. See: https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx.

2. See: http://www.mbrsg.ae/getattachment/5147489c-d2f8-43c1-b54d-9bb29690341f/The-Arab-World-Online-Trends-in-Internet-Usage-in.

3. See: https://www.gartner.com/doc/3840665/forecast-internet-things--endpoints.

4. See: https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm.

5. See: https://marketrealist.com/2017/08/why-uber-could-struggle-to-stay-on-top-of-the-ride-hailing-market.

6. See: https://movement.uber.com/?lang=en-AU.

7. See: https://www.gartner.com/doc/3880666?ref=SiteSearch&sthkw=government%20%22open%20data%22&fnl=search&srcId=1-3478922254#dv_3_itscore_assessments.

8. See: https://govinsider.asia/digital-gov/singapore-to-train-10000-officials-with-data-skills/.

9. Institute on Governance, 2003.

10. United Nations Economic and Social Council, 2006.

11. United Nations Economic and Social Commission for Asia and the Pacific, n.d.

12. United Nations Economic and Social Council, 2006.

13. United Nations Economic and Social Commission for Asia and the Pacific, n.d.

14. Organization for Economic Cooperation and Development, 2016.

15. Lathrop and Ruma, 2010.

16. Al-Munjid, 2016.

17. Open Government Partnership, n.d.

18. Open Government Partnership, 2011.

19. Open Government Partnership, 2011.

20. OECD, 2017.

21. World Bank, 2017.

22. White House, 2009.

23. United Nations e-Government Survey, 2016.

24. World Bank, 2017.

25. Open Knowledge Foundation, 2013.

26. See: http://opendefinition.org/od/2.0/en/.

27. Further details on these requirements are discussed in chapter 3.

28. See: https://theodi.org/what-is-open-data.

29. See: https://www.data.vic.gov.au/data/dataset/post-school-destinations-of-year-12-or-equivalent-completers-victorian-schools-2017.

30. United Nations e-Government Survey, 2016.

31. European Commission, 2015.

32. Manyika et al., 2013.

33. Young and Verhulst, 2016, p. 55.

34. This table is summarized based on data in GovLab, 2017.

35. For more information on the charter see: https://opendatacharter.net/history/. For the full text of the version signed by G8 leaders see: https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex.

36. See: http://www.motc.gov.qa/sites/default/files/open_data_policy_0.pdf.

37. Open Data Portal Saudi Arabia, n.d. See: http://www.data.gov.sa/en/about.

38. Open Data Barometer, 2016.

39. Open Knowledge Foundation, 2017.

40. See: https://www.gov.uk/government/publications/letter-from-the-prime-minister-on-government-transparency-and-open-data.

41. United Kingdom Department for Communities and Local Government, 2015, p. 4.

42. Ibid.

43. Ibid., p. 23.

44. United Nations e-Government Survey, 2016.

45. See: https://apolitical.co/solution_article/ghana-makes-public-works-data-accessible-boost-political-participation/.

46. Transport for London, 2016.

47. The sub-indexes are not shown in the figure.

48. See: https://theodi.org/article/how-to-plan-and-budget-an-open-data-initiative.

49. See: https://www.europeandataportal.eu/sites/default/files/european_data_portal_-_open_data_goldbook.pdf.

50. See: https://theodi.org/article/how-to-plan-and-budget-an-open-data-initiative/.

51. See: https://gulfnews.com/news/uae/government/dubai-open-data-committee-formed-1.1429600.

52. Dubaidata.ae, n.d.

53. Presented in chapter 5 of this guide.

54. See: https://data.gov.ie/blog/ministerdonohoelaunchespublicconsultationprocessforthenationalopendatastrategy2017-2022.

55. See: https://data.gov.uk/dataset/83215e9c-a55e-4f15-810c-8d3b59bce6bc/data-for-london-a-city-data-strategy.

56. See: https://www1.nyc.gov/assets/home/downloads/pdf/reports/2015/NYC-Open-Data-Plan-2015.pdf.

57. See: https://www.health.qld.gov.au/__data/assets/pdf_file/0030/436809/open-data-strategy.pdf.

58. See: https://www.edmonton.ca/city_government/documents/PDF/CityofEdmonton_OpenDataStrategy.pdf.

59. See: https://data.gov.uk/dataset/83215e9c-a55e-4f15-810c-8d3b59bce6bc/data-for-london-a-city-data-strategy.

60. See: https://www.edmonton.ca/city_government/documents/PDF/CityofEdmonton_OpenDataStrategy.pdf.

61. See: https://www.health.qld.gov.au/__data/assets/pdf_file/0030/436809/open-data-strategy.pdf.

62. See: https://data.gov.uk/dataset/83215e9c-a55e-4f15-810c-8d3b59bce6bc/data-for-london-a-city-data-strategy.

63. See: https://data.gov.uk/dataset/83215e9c-a55e-4f15-810c-8d3b59bce6bc/data-for-london-a-city-data-strategy.

64. See: https://www1.nyc.gov/assets/home/downloads/pdf/reports/2015/NYC-Open-Data-Plan-2015.pdf.

65. See: https://www.edmonton.ca/city_government/documents/PDF/CityofEdmonton_OpenDataStrategy.pdf.

66. See: https://theodi.org/article/how-to-plan-and-budget-an-open-data-initiative/.

67. See: http://www.motc.gov.qa/en/documents/document/public-consultation-draft-open-data-policy.

68. See: http://www.mpwh.gov.jo/Pages/Open-Government-Data-policy.pdf.

69. See: http://www.motc.gov.qa/en/documents/document/open-data-policy.

70. See: http://www.seattle.gov/Documents/Departments/SeattleGovPortals/CityServices/OpenDataPolicyV1.pdf.

71. See: https://www.transport.nsw.gov.au/system/files/media/documents/2017/open-data-policy.pdf.

72. See: https://theodi.org/article/how-to-plan-and-budget-an-open-data-initiative/.

73. Ibid.

74. More details on this can be found in chapter 5.

75. See: https://webfoundation.org/2015/08/three-critical-success-factors-to-turn-an-african-open-data-revolution-from-vision-to-reality/; https://www.per.gov.ie/wp-content/uploads/Draft-Open-Data-Strategy-2017-2022.pdf.

76. United Kingdom Digital Government Service, 2017.

77. IEAG, 2014.

78. Data.Gov, 2018.

79. Press, 2014.

80. See: https://data.buffalony.gov/Public-Safety/Crime-Incidents/d6g9-xbgu.

81. Open Data Institute, n.d.

82. See: https://creativecommons.org/.

83. See: https://accelerate.theodi.org/.

84. See: http://www.niso.org/publications/press/UnderstandingMetadata.pdf.

85. See: https://support.socrata.com/hc/en-us/articles/115008609707-Best-Practices-for-Metadata-Management.

86. See: https://www.seattle.gov/Documents/Departments/SeattleGovPortals/CityServices/OpenDataPlaybook_Published_2016.08.pdf.

87. See: https://www.europeandataportal.eu/sites/default/files/european_data_portal_-_open_data_goldbook.pdf.

88. Labeling these dataset as open is questionable because they do not have an open legal licence, which violates the definition of open data. However, the page is still a good example of Web download from a technological point of view.

89. See: https://www.data.gov/privacy-policy.

90. Buienradar, 2018.

91. World Bank, 2014.

92. Plume Labs, 2018.

93. GovLab, 2018.

94. Open Data Institute, 2015.

95. Open Data Institute. See: http://pathway.theodi.org/.

96. Ibid.

97. Ibid.

98. Ibid.

99. Ibid.

100. Department of Transport and Main Roads, Australia, n.d.

101. Ibid.

102. World Wide Web Foundation, 2016.

103. Australian Government, 2016.

104. United States. For the page on impact, see: https://www.data.gov/impact/.

105. European Commission, 2015.

106. Open Knowledge International, 2017.

107. European Union Data Portal.

108. See: https://joinup.ec.europa.eu/sites/default/files/document/2015-05/d2.1.2_training_module_2.2_open_data_quality_v1.00_en.pdf.

109. United Nations e-Government Survey, 2016.

110. See: https://www.europeandataportal.eu/elearning/en/module5/#/id/co-01.

111. Open Data Certificate, 2017.

112. See: https://certificates.theodi.org/en.

113. See: https://certificates.theodi.org/en/datasets/215670/certificate.

114. 5-Star Open Data, 2015.

115. Dodds, (2015).

116. National Archives, n.d.

117. Most of the content extracted from Open Data Unit.

118. Ibid.

119. See: http://www.citizensinformation.ie.