FREE OF CHARGE

AVAILABLE ONLINE

OPEN FORMAT

TIMELY

PRIMARY

MACHINE-READABLE

PERMANENT

METADATA

IN BULK/API

OPEN LICENSE

HUMAN-READABLE

IDENTIFIER

HONG KONG
OPEN DATA
INDEX
2019/20

HONG KONG
OPEN DATA INDEX

A project by

Internet Society
Hong Kong Chapter

ABOUT THE HONG KONG OPEN DATA INDEX

The Hong Kong Open Data Index is a project initiated by the Internet Society Hong Kong Chapter to assess and monitor the progress of releasing public data in Hong Kong against international standards. Built on 19 sets of established open data principles and assessment tools at international and regional levels, the Index and its assessment results are aimed to reveal achievements and identify challenges in Hong Kong, and introduce international best practices to the city. For more about the Index, please visit: opendata.isoc.hk

Advisory Committee

Ben Cheng
Edmon Chung
Ho Wa Wong
Himphen Hui

Project Staff

Benjamin Zhou
Sherry Shek
Daphne Ip

Contributing Researchers

Paula Lau, Chloe Ng, Bingzi Yu, Charlotte Yu, Lansie Lan, Candice Lok Yee Chow, Iqra Abbasi, Wipada Nganthavee

Report Production

Author: Benjamin Zhou
Editor: Bill Leverett
Designer: DESIGNORM

Publication Date

14 May 2020

Funder

WYNG Foundation

CONTENTS

HONG KONG OPEN DATA INDEX 2019/20

DATASETS PERFORMANCE

	OVERALL	DATA.GOV.HK	OTHER PUBLIC WEBSITES
Census & Statistics	90	73	90
Legislature & Advisory Bodies	82	47	70
Recreation & Culture	80	70	56
Welfare	78	78	56
Transport	77	77	43
Environment & Energy	76	76	62
Health	76	76	60
Science & Technology	75	72	60
Education	75	69	57
Public Finance	71	49	71
Climate & Weather	70	37	66
Government Operations	63	31	56
Housing	58	48	45
Justice & Safety	56	31	45
Land	47	43	25
Business Registration	47	47	28
Overall	70	58	55

INDICATORS COMPARISON

	OVERALL	DATA.GOV.HK	OTHER PUBLIC WEBSITES

Indicator Comparison	1. Available online	2. Free of charge	3. In Bulk/ API	4. Open license	5. Machine-readable	6. Open format	7. Primary	8. Timely	9. Metadata	10. Permanent	(11) Identifier	(12) Human-readable
Overall	9.13	9.24	5.03	4.07	6.90	8.51	7.52	8.04	5.08	6.38	4.15	6.53
data.gov.hk	7.27	7.35	4.83	3.65	6.75	7.13	5.87	6.07	3.52	5.18	3.01	4.07
other public sources	9.02	8.93	2.43	2.38	1.98	6.23	7.38	7.73	5.51	3.88	3.74	9.02

HONG KONG OPEN DATA INDEX

	Overall	1 Available online	2 Free of charge	3 In Bulk/API	4 Open license	5 Machine-readable	6 Open format	7 Primary	8 Timely	9 Metadata	10 Permanent	(11) Identifier	(12) Human-readable
Average		9	9	5	4	7	8	8	8	5	6	4	7
A Census & Statistics	90	10	10	5	5	10	10	10	10	10	10	10	10
		10 10	10 10	6 5	5 5	7 10	9 10	9 10	10 10	3 10	5 10	1 10	9 10
B Public Finance	71	9	10	3	4	5	10	9	10	6	6	6	10
		6 9	6 10	3 3	3 4	5 5	6 10	6 9	6 10	3 6	5 6	3 6	1 10
C Legislature & Advisory Bodies	82	10	10	6	4	10	10	9	9	5	10	3	7
		6 10	6 10	4 3	2 4	6 4	6 5	5 10	5 9	2 6	6 9	3 5	3 10
D Government Operations	63	10	10	3	4	4	6	9	10	4	3	1	9
		4 10	4 10	2 2	2 4	4 1	4 6	3 9	4 10	3 4	3 0	0 2	1 10
E Justice & Safety	56	9	9	3	4	5	9	5	6	4	3	3	8
		5 9	5 8	3 2	3 2	5 0	5 5	1 5	3 6	1 5	1 4	0 5	3 9
F Business Registration	47	5	5	4	5	5	5	5	5	3	5	5	3
		5 6	5 5	4 1	5 1	5 0	5 3	5 5	5 3	3 3	5 0	5 5	3 5
G Land	47	7	7	3	3	5	7	5	6	4	1	6	4
		6 4	6 4	3 1	3 0	5 3	6 4	4 3	4 3	4 3	2 0	6 2	3 3
H Housing	58	9	9	6	3	5	6	4	5	6	6	3	5
		6 8	6 8	5 3	3 1	5 1	5 5	4 5	4 6	4 6	6 3	0 4	3 10
I Transport	77	9	9	7	3	9	9	9	9	7	8	5	4
		9 9	9 9	7 2	3 0	9 0	9 3	9 6	9 9	7 3	8 3	3 4	4 9
J Welfare	78	9	10	5	5	10	10	9	9	3	8	5	5
		9 10	10 10	5 3	5 5	10 1	10 3	9 6	9 9	3 8	8 1	5 3	5 10
K Education	75	10	10	6	4	8	9	8	8	4	10	0	8
		9 10	9 10	6 3	4 1	9 0	9 6	8 8	6 10	4 4	9 5	0 0	6 10
L Health	76	10	10	6	5	9	9	6	9	3	9	0	7
		10 10	10 10	6 2	5 4	9 2	9 9	6 9	9 8	3 4	9 3	0 0	7 10
M Recreation & Culture	80	10	10	8	5	8	10	10	8	7	4	8	6
		8 10	8 10	8 0	4 2	8 0	8 10	8 10	8 7	6 7	4 0	8 6	4 10
N Environment & Energy	76	10	10	6	5	9	10	8	8	4	8	3	6
		10 10	10 10	6 5	5 0	9 4	10 5	8 8	8 5	4 8	8 3	3 0	6 10
O Climate & Weather	70	10	10	6	3	3	9	8	10	6	6	3	6
		5 10	5 10	5 4	3 3	5 0	5 9	4 8	3 10	3 6	0 6	5 0	0 9
P Science & Technology	75	10	10	6	5	8	8	9	8	4	7	7	7
		9 10	9 10	6 2	5 5	9 0	9 7	7 9	6 8	4 5	6 5	5 6	6 10
Overall	70	9	9	5	4	7	8	8	8	5	6	4	7
		7 9	7 9	5 2	4 2	7 2	7 7	6 7	6 8	3 6	5 4	3 4	4 9

■ overall
■ data.gov.hk
■ other public websites

SUMMARY

Open data is data that can be freely used, shared, and built-on by anyone for any purpose. Government data is the most important source of open data. It is regarded as an instrument to improve government transparency and openness, and a source to better public service delivery, business development and civic innovation, among others. [1]

The Hong Kong Open Data Index is developed by the Internet Society Hong Kong Chapter as an assessment tool to test the openness of public data. Built on 19 established open data standards at international and regional levels, the Index consists of 12 indicators (including two non-scoring indicators) to evaluate data in the city across 16 categories covering 69 types of datasets. The assessment results and analyses are aimed to inform the advocacy work by civil society and the business sector, and to help the government improve the open data policies and practices.

KEY FINDINGS

Highlights:

- The Hong Kong government's open data portal data.gov.hk is not more "open" than other public websites, mainly because it maintains fewer datasets than the Index is expecting compared to average governmental websites. The types of datasets that are not found on the portal but available on other websites are: calls for tender, bills & legislation, Legislative Council meetings, government contact points, declarations of interest, judiciary & judgements, travel alerts, public housing (including waiting time and allocation status), hiking trails, wind, rainfall distribution.

- Individual government departments adopt inconsistent standards for publishing data across their websites and in the open data portal, confusing users and potentially discouraging data reuse.

- Individual government departments adopt inconsistent standards for publishing data across their websites and in the open data portal, confusing users and potentially discouraging data reuse.

Good practices observed:

- **Availability** – The Hong Kong government and other public organisations have released most types of data online that are set out by major international open data initiatives and expected by local stakeholders, so as the online availability is scored 9.13 out 10.

- **Best practices** – A few organisations and data platforms— such as Census and Statistics Department, Legislative Council, and a newly launched platform Hong Kong GeoData Store—have set good examples of how open data should be managed and presented.

- **Timely and Primary** – The published data are fairly timely and primary as the datasets in Hong Kong score 8.04 and 7.52 out of 10 on these two indicators respectively.

- **Machine-readability** – The datasets in Hong Kong perform moderately in machine-readability, scoring 6.9 out of 10 overall, thanks to the fact that most datasets published via the portal data.gov.hk are machine-readable.

Some gaps identified:

- **Availability** – Data in the five categories of land, business registration, justice & security, housing, and government operations are significantly less open than other datasets. Datasets regarding ownership of lands and businesses are not freely available.

- **Open license** – The copyright licenses are not only inconsistent across datasets and platforms, but also lack explicitness and clarity about the extent that the public are authorised to use the data.

- **Metadata** – The metadata that accompany the datasets to summarise their basic information are not comprehensive enough, as the majority of the datasets are not denoted with their initial publication date, coverage period, theme and key words, etc.

- **Historical records** – The historical records/archives of published data are not properly maintained and presented online.

- **API** – Very few datasets come with an Application Programming Interface (API).

RECOMMENDATIONS

Immediate action

We also have developed a few steps that government departments in Hong Kong can take immediately to improve data openness.

- **Machine-readability.** Make the data readily available in both human- and machine-readable forms.

- **Core metadata.** Supplement metadata elements including the publication date, coverage period, theme and key words, data owners and maintainers etc.

- **Context.** Prepare an abstract description for each dataset published on the open data portal, to explain the contents, context, analytical limitations etc.

- **Primary**. Audit the published datasets to ensure their quality and granularity.

- **Historical records**. Archive dated copies when a dataset is updated, and make the archives available at a stable and searchable online location.

Comprehensive measures

We have identified more comprehensive measures that the government may consider to better manage the data opening process and thereby boosting availability and usability of the data.

- Establish an interactive public participation platform on the public sector information (PSI) portal data.gov.hk to facilitate a user-centric and demand-driven process to enhance data availability.

- Set up an interdepartmental task force to coordinate and standardise the data publication practices across official websites and the open data portal, in particular to identify and tackle issues that may jeopardise data usability.

INTRODUCTION

Open data is data that can be freely used, modified, and shared by anyone for any purpose.[2] Government data is the most important source of open data, and therefore it is also referred to as open government data.

The Hong Kong Open Data Index is a research-based initiative founded by the Internet Society Hong Kong Chapter to evaluate and promote open data in Hong Kong. The Index is developed as an assessment tool to test the openness of public data, including data provided by the government and other public organisations as well as data owned by private businesses but of public interest such as public transport data. Built on 19 established open data standards at international and regional levels, the Index consists of 12 indicators (including two non-scoring indicators) to evaluate data in the city across 16 categories covering 69 types of datasets. The assessment results and analyses are aimed to inform the advocacy work by civil society and business sector, and to help the government improve open data policies and practices.

WHY OPEN DATA IN HONG KONG

The concept of open data has been prospering in the world for at least a decade since a group of influential advocates gathered in California, US to develop a set of principles of open government data in 2007.[3] A lot of initiatives have been set up to promote and monitor open data around the globe ever since.

In Hong Kong, the first open data portal was launched by the government in 2011, and the civil society and the business sector have been calling for opening up more datasets of better quality. There is a consensus among the stakeholders including the government that open data matters for the city, and we have observed three major reasons:

Transparency and openness in governance

Open data enhances government transparency. All stakeholders can monitor government operations—such as public service delivery, public spending, and resource distribution—thus holding the government accountable. Open data also facilitates openness in the public decision-making process. Armed with clearer knowledge of public matters, all stakeholders can participate in policy consultation and formulation more effectively. Reciprocally, the transparency and openness in governance will build trust between government and the general public.

Digital transformation into a truly smart city

Open data enables "smart governance" and "smart operation", achieving which would make Hong Kong a truly smart city. Seamless integration of data across sectors can facilitate solutions responding to social issues, if the data is maintained and released to a uniform standard to enhance interoperability. Furthermore, interdepartmental and cross-sector resources could be better coordinated and distributed to streamline the city's daily operation.

Empower innovation to unleash social and economic values

"Data is the new oil." When data is truly accessible and usable in society at large, it will unleash the next wave of immense potentials. People can obtain insights from the data that is opened up, thus stimulating innovation in developing products and services for the community. Advancing what new technology does, open data helps trigger more business opportunities, boosts productivities and improves the quality of life.

WHAT IS THE HONG KONG OPEN DATA INDEX

Since the Hong Kong government launched a pilot of the open data portal titled "Data.One" in 2011,[4] the business sector and civil society together with the authorities have been acting to improve open data, while a number of civil groups formed and activities such as hackathons were frequently organised. Years later in 2017, the government levelled open data with its innovation and technology strategy by absorbing it into the Smart City Blueprint,[5] and subsequently firmed up a new open data policy in 2018.[6]

With all these efforts made by the stakeholders in the past years, we want to know the outcome: how open is the public data in Hong Kong now? Unfortunately, there was no such evaluation about Hong Kong on its own, except that the city

was ranked in the Global Open Data Index at 24th—after Taiwan, Japan and Singapore—in its last round in 2016 and thereafter the ranking was suspended without a stated reason. You cannot manage what you cannot measure. We need a system to measure the data openness and track the progress in Hong Kong in order to better manage it.

Therefore, the Internet Society Hong Kong Chapter has initiated the project Hong Kong Open Data Index, an assessment tool built on 19 established standard-setting initiatives at international or regional levels and consultations with local stakeholders, to inform people about the city's data openness. Among the established initiatives are 10 set of standard-setting principles and nine assessment tools.

The Index is aimed to be a stepping stone to a series of evidence-based research to reveal achievements, identify challenges, and introduce international best practices to Hong Kong. By showcasing the evidence, the Index will also become a vehicle for dialogues between stakeholders that could contribute to shaping consensus for new policies and measures.

The Hong Kong Open Data Index consists of 12 indicators (including two non-scoring indicators) to evaluate data in the city across 16 categories that cover 69 types of datasets.

FIGURE 1   **19 Established Standards that the Hong Kong Open Data Index is built on**

10 Standard-setting principles

Open Definition	2005-
Eight Open Government Data Principles	2007
Ten Principles for Opening Up Government Information	2010
Five-Star Open Data	2010
Open Data Policy Guidelines	2012/13/14
Open Government Data: The Book	2012/14
Open Data Handbook	2012/15
G8 Open Data Charter	2013
International Open Data Charter	2015
Internet Universality	2018

9 Assessment tools

Global Open Data Index	2013-2017
Open Data Barometer	2013-2018
ePSI Platform Scoreboard	2013
G20 Anti-corruption Open Data Principles Assessment	2015
Open Data Inventory	2015-
OECD OURdata Index on Open Government Data	2017
Open Standards Directory	2017
Taiwan Open Government Report	2017
China Open Data Index	2017-

FIGURE 2   **12 Indicators of the Hong Kong Open Data Index**

1	Available online to anyone
2	Free of charge
3	Downloadable in bulk and with API
4	Machine-readable
5	Open license
6	Open format
7	Primary
8	Timely
9	Metadata
10	Permanent
[11]	Identifier
[12]	Human-readable

FIGURE 3   **16 Data Categories for the Hong Kong Open Data Index**

A	Census & Statistics
B	Public finance
C	Legislature & Advisory bodies
D	Government operations
E	Justice & Security
F	Business registration
G	Land
H	Housing
I	Transport
J	Welfare
K	Education
L	Health
M	Recreation & Culture
N	Environment & Energy
O	Climate & Weather
P	Science & Technology

KEY FINDINGS

The Hong Kong Open Data Index 2019/20 evaluates the openness of public data in Hong Kong across 16 categories on 12 indicators (including two non-scoring indicators). The assessment results of the Index reveal:

HIGHLIGHTS

The Hong Kong government's open data portal data.gov.hk is not more "open" than other public websites, mainly because it maintains fewer datasets than the Index is expecting compared to average governmental websites.

- The portal is designed to publish open data, but its overall performance (scored 58 out of 100) is almost the same as the average for other public websites (55);

- It records a lower score on the indicator of data availability than other online sources (7.27 vs. 9.02 out of 10), driving down scores on other indicators, and it provides fewer datasets that are qualified for the assessment than other public websites do in 11 out of 16 data categories;

- The types of datasets that are not found on the portal but available on other websites are: calls for tender, bills & legislation, Legislative Council meetings, government contact points, declarations of interest, judiciary & judgements, travel alerts, public housing (including waiting time and allocation status), hiking trails, wind, rainfall distribution.

- A number of datasets in the portal come with no description or context in their metadata (3.52 vs. 5.51 for other public websites), leaving users without information about the strengths and limitations of the data;

- The open data portal outperforms other websites in providing machine-readable data (6.75 vs. 1.98), and so a lower data availability of the portal pulls down the overall data machine-readability in HK.

Individual government departments adopt inconsistent standards for publishing data across their websites and on the open data portal, confusing users and potentially discouraging data reuse.

- The copyright licenses vary across individual department websites. Some use open licenses that permit users to access, reuse, and redistribute the datasets, while some other datasets are restricted to downloading for internal use only, posing legal risks for users especially those who are not familiar with copyright law and hence potentially discouraging data reuse;

OVERALL · DATA.GOV.HK · OTHER PUBLIC WEBSITES
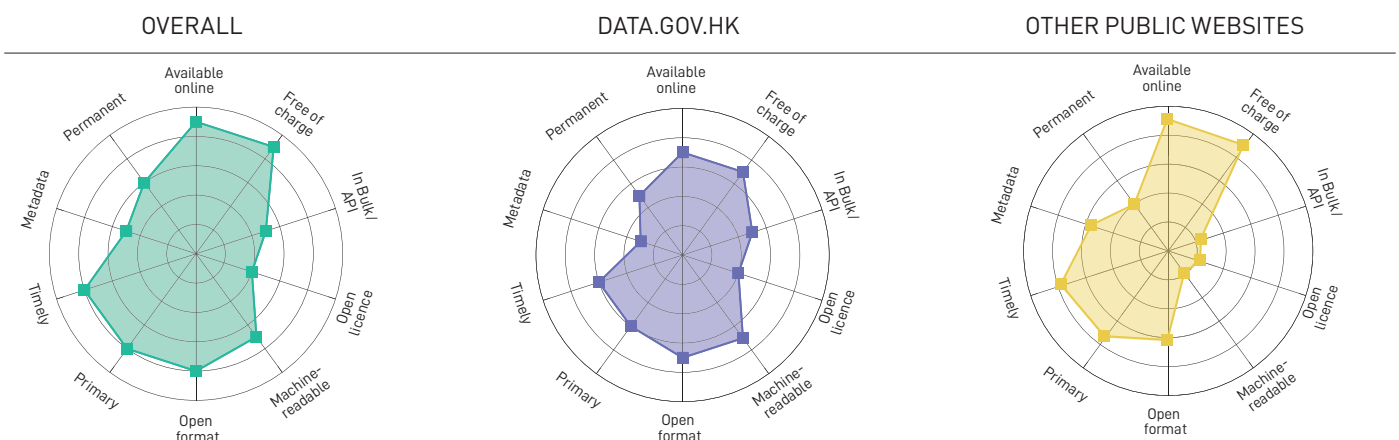
FIGURE 4   Scores of the 12 Indicators of the Hong Kong Open Data Index

Indicator Comparison	1 Available online	2 Free of charge	3 In Bulk/ API	4 Open license	5 Machine-readable	6 Open format	7 Primary	8 Timely	9 Metadata	10 Permanent	(11) Identifier	(12) Human-readable
Overall	9.13	9.24	5.03	4.07	6.90	8.51	7.52	8.04	5.08	6.38	4.15	6.53
data.gov.hk	7.27	7.35	4.83	3.65	6.75	7.13	5.87	6.07	3.52	5.18	3.01	4.07
other public sources	9.02	8.93	2.43	2.38	1.98	6.23	7.38	7.73	5.51	3.88	3.74	9.02

- Only a few government departments retain the records of data that have been previously published on their websites or the data portal. Nevertheless, the few datasets with historical records are not stored in a uniform manner—some merge the historical data into the updated version, some create new datasets every time they have updates and maintain the old ones in their original form, while some others provide tools such as APIs to retrieve the archives;

- The metadata schemes and download/API options also vary across datasets and platforms.

GOOD PRACTICES OBSERVED

The Hong Kong government and other public organisations have released online most types of data set out by major international open data initiatives and expected by local stakeholders, so the online availability score is 9.13 out of 10.

A majority of datasets on the list for the Index to examine—built on various checklists of established open data assessment tools in the world and a local stakeholder consultation—can be found online in Hong Kong, with a better availability on average government websites than on the open data portal. Therefore, the priority for these readily available datasets in the near future should be to improve their openness, i.e., to publish them on the portal in machine-readable forms with proper metadata and archives.

A few organisations and data platforms in Hong Kong have set good examples of how open data should be managed and presented.

Examples include the websites of the Census and Statistics Department (scoring 90) and the Legislative Council (88). The Hong Kong GeoData Store (87)—a digital platform managed by the Lands Department to facilitate the discovery and dissemination of spatial data such as locations of Wi-Fi spots and approved/rejected/pending residential development projects—has showcased how essential features of an open data portal could be properly designed and arranged. Hence, these leading organisations should contribute more as a part of interdepartmental and coordinated efforts to improve the open data practices in Hong Kong.

The published data are fairly timely and primary as the datasets score 8.04 and 7.52 out of 10 on the two indicators respectively.

The Index researchers observed that most datasets had been updated within a few months of the moment of their examination, suggesting that a majority of data providers have fulfilled their maintenance duty once their datasets are made available publicly. Meanwhile, update frequency is another aspect of timeliness but there is no universal standard because it varies with the nature and value of the data. It is noteworthy that three bus companies released their real-time arrival data through APIs in 2019, as well as four rail lines of the MTR. This is a breakthrough after years of advocacy from civil society and business start-ups, as well as the government's coordination. In addition, real-time data also fulfils the criteria set out by the indicator of "primary", which advocates that data should be published in its original and unmodified form without aggregation. In the event that aggregation is inevitable—e.g. for privacy and confidentiality concerns—the data should be disaggregated at the finest level of granularity when possible.

Overall, Hong Kong performs moderately in machine-readability as it scores 6.9 out of 10 on this indicator, because most datasets published via the portal are machine-readable.

Data machine-readability is a principle that is frequently mentioned in discussions about open data around the world including Hong Kong. A machine-readable document must fulfil two criteria: first, the document format must be one fits to be processed by computers, such as CSV, XML, and JSON; second, data in the document must be structured and standardised, for example, the numbers should be aligned with column and row headers in a CSV file, or it cannot be easily "read" by any "machine". Due to its significance, machine-readability is the element that incurs most criticism from open data advocates, as can be seen from news reports in the past few years. The Index assessment reveals that machine-readable is not the worst- nor the best performing indicator for Hong Kong—it is at a moderate position at 6.9, because the datasets released in the portal largely fulfil the criteria.

FIGURE 5   Ranking and Scores of the 16 Data Categories of the Hong Kong Open Data Index

	OVERALL	DATA.GOV.HK	OTHER PUBLIC WEBSITES
Census & Statistics	90	73	90
Legislature & Advisory Bodies	82	47	70
Recreation & Culture	80	70	56
Welfare	78	78	56
Transport	77	77	43
Environment & Energy	76	76	62
Health	76	76	60
Science & Technology	75	72	60
Education	75	69	57
Public Finance	71	49	71
Climate & Weather	70	37	66
Government Operations	63	31	56
Housing	58	48	45
Justice & Safety	56	31	45
Land	47	43	25
Business Registration	47	47	28
Overall	70	58	55

SOME GAPS IDENTIFIED

Data in the five categories of land, business registration, justice & security, housing and government operations are significantly less open than other datasets.

Both the land and business registration categories record the lowest score at 47, as the land and business ownership datasets are not openly available because users have to inquire through a search system requiring to knowledge of names and registration numbers beforehand. Also, detailed data regarding ownership is not for free. The data about justice & safety and government operations are largely not available on the open data portal but they can be found on the government department websites. This drags down the openness of these categories, because government department websites usually perform worse on technical indicators such as machine-readability compared to the open data portal. The housing sector fails to offer primary data at the finest granularity in the public domain—for example, transaction records of individual apartments are not freely available.

The copyright licenses are not only inconsistent across datasets and platforms, but also lack explicitness and clarity about the extent that the public are authorised to use the data.

Users have to parse the legal language on their own in the "Terms and Conditions" or "Important Notices"—hidden in the corner of government websites—to understand whether they can use the data for their own work and then publish it. It may prevent average people from making use of the data for fear of legal risks, and such risks do exist because a number of datasets only allow users to download them for internal use. In the international open culture community, Creative Commons (CC) sets a good example of a standardised licensing system that features a range of visual icons allowing users to grasp their rights to use the work at a glance. The Hong Kong government may consider introducing such a system.

The metadata that accompanies the datasets to summarise their basic information is not comprehensive enough, as the majority of the datasets are not noted with their initial publication date, coverage period, theme and key words, etc.

The International Open Data Charter (2015) advocates that core metadata, a type of metadata that provides important and fundamental information, may include the dataset title, source, publication date, format, and others that support discoverability.[7] The public data portal and websites in Hong Kong fall short of providing a complete set of core metadata, as publication dates, cover period and key words are rarely found. Also, they lack documentation to explain the context, data fields, data dictionaries, the purpose and method of data collection, etc. Such information would have helped users to utilise the data for suitable scenarios and to address respective issues.

The historical records (archives) of published data are not properly maintained and presented online.

Making the data that had been released in previous years available in stable online locations will ensure users can conduct comparative analyses over time, based on authentic documents rather than amended versions. However, many datasets in Hong Kong only provide the latest version without archiving the old copies or not showing them in a location easily accessible to the public. Even when some of the "outdated" data were retained and showcased, they were not exhibited in a way that gave users a clear expectation of how to find them and what contents would be included. For example, the data portal offers the function "Download Historical Data", but it does not necessarily allow users to retrieve the dated copies for every dataset. A few other data providers create new datasets every time for updates while leaving the "dated" ones paralleled to the new ones instead of archiving them. There is no answer yet about which method is the best, as long as they are managed uniformly without leaving users confused. It once again underscores the importance of consistency in data management across departments.

Very few datasets come with an Application Programming Interface (API), so it is scored 1.29 out of 5 on this sub-indicator.

Making datasets downloadable in bulk and accessible through an API are two common methods to open up data, while the former is more familiar to the public and data providers than the latter. API is a mechanism that allows users to get some slice of the data by placing specific queries, and make it possible for programmers to automate the data access process. It is particularly valuable for datasets that are updated as frequently as every few minutes (real-time). The provision of an API requires extra resources to upgrade the data management system and therefore it is usually a lower priority on the government's open data agenda than providing downloadable files, which can be executed immediately once the providers decide to publish them. Nevertheless, APIs indicate a better data management and sharing capability, and therefore the extremely low score implies that there is still a long way to go before Hong Kong becomes a "smart city".

RECOMMENDATIONS

Based on our findings, we have identified two major aspects that the government should consider to better manage the data opening process and thereby boost availability and usability of the data. We also have developed a few steps for the government departments and other data providers in Hong Kong to take immediately to improve their performance in data openness.

Establish an interactive public participation platform in the public sector information (PSI) portal data.gov. hk to facilitate a user-centric and demand-driven process to enhance data availability.

- The Hong Kong government started to announce annual open data plans including the lists of datasets to be released via the portal data.gov.hk over three years from the beginning of 2019.[8] The plans were formulated by over 80 bureaux and departments, and aimed to facilitate the public to provide feedback and suggestions on the types of data to be opened up and their potential applications and hence to help the government assess the priority of datasets for publication.[9]

- However, the Index assessment (as of the latest round of review in March 2020) reveals that the portal fails to provide enough of the datasets expected by both the international and local open data communities, and while a number of them are already available on the government websites they may not cater to open data criteria. There is a lack of transparency on what feedback and suggestions the authorities have received and what their priorities are. Therefore, we suggest setting up an online forum in the open data portal to open up the feedback, discussion and deliberation process.

- Opening up datasets is ultimately about helping people, and so the government should communicate openly with them to understand what data they need and in what forms. The proposed forum will not only help the authorities collect wishlists and respond to complaints from the public, but also help the public to discuss data availability and usability, hence holding dialogues between the government and citizens in order to improve the open data practices and maximise data value.

- This type of forum has become a common practice particularly in Asia. For example, the cities of Taipei[10] and Shanghai[11] have already set up such platforms on their open data portals in recently years. Government departments in Taipei have been constantly responding to the recommendations posted on the forum.

- In addition to addressing online comments, the authorities should also take the initiative to gather requirements and feedback from stakeholders, for example, by conducting surveys and workshops with specific data user groups.

Set up an interdepartmental task force to coordinate and standardise the data publication practices across official websites and the open data portal, in particular to identify and tackle issues that may jeopardise data usability.

- The governance of open data is more complex than managing information technology systems or implementing innovation and technology policies. It consists of multiple layers: management (decision-making), legal and policy, technical and standards, capacity building.[12] Therefore, good open data governance requires top level oversight as well as interdepartmental and cross-sector coordination at other layers to tackle the challenges that emerge in the data opening process.

The Open Data Governance Model
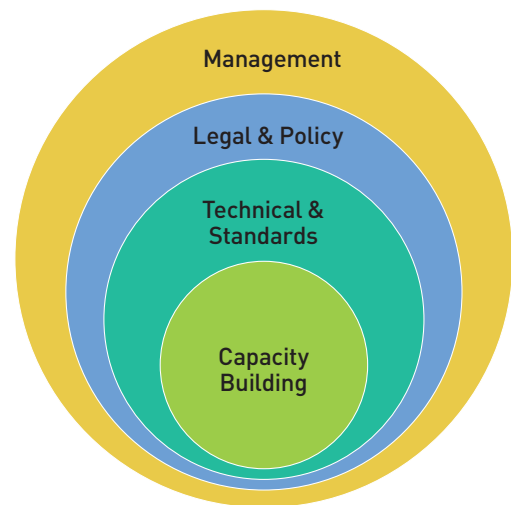
Source: World Wide Web Foundation, 2019

Hong Kong has established a steering committee overseeing the innovation and technology strategies including open data, chaired by the chief executive since 2017 to provide leadership.[13] However, the Index reveals that the departments still adopt nonuniform standards on various aspects of open data across online platforms, suggesting a lack of coordination when it comes to the implementation of strategies.

- We recommended setting up an interdepartmental task force comprised of experts in data management, information technology, law, administration, archives, communications and other disciplines. The task force will 1) ensure the standards are constantly being refined and implemented in a uniform manner, 2) tackle the issues that may affect data usability, e.g., copyright license, machine-readability, metadata schemes and historical records.

- Copyright license is one of the first issues that should be addressed as the Index has revealed its inconsistency and obscurity across online platforms. The same dataset released on the government website and data portal respectively should be under the same licensing terms and conditions, if not a complete open license. Ideally, all datasets owned by the government, published on any platforms, should come with a uniform open license in an easy-to-understand form, exemplified by the Creative Commons licenses, with few exceptions. The department overseeing the legal policy should play a significant role in unifying the licensing.

- A lack of enough metadata elements and dataset archives are two other major issues identified in the Index report. Improving these aspects should not be a challenge for the government as there are experts already in the establishment with rich experience in information management. These technocrats can contribute their expertise to open data through the interdepartmental coordination mechanism.

A few steps to take immediately to improve data openness:

- Make the readily available datasets in both human- and machine-readable forms. Review the data that have been published on the websites of individual institutions and develop an inventory of readily available datasets. Transform these datasets into machine-readable formats and ensure they are structured and standardised. Present the data in human-readable form on the websites while providing an option to download the machine-readable files that could be also re-published on and linked to the open data portal data.gov.hk.

- Supplement metadata elements including the publication date, cover period, theme and key words, data owners and maintainers etc., to the websites that publish the data, whether on the individual website or the portal.

- Prepare an abstract description for each dataset published on the open data portal, to explain the contents, context, analytical limitations and other information that could facilitate the usability, and provide external links for data users who want to know more about the background of a dataset.

- Audit the published datasets regularly to ensure they are error-free and at the possible granularity.

- Archive the dated copies every time a dataset is updated, and make available the archives available in a stable and searchable online location.

INDICATORS ANALYSIS

INDICATORS OVERVIEW

	OVERALL	DATA.GOV.HK	OTHER PUBLIC WEBSITES
Free of charge	9.24	7.35	8.93
Available online	9.13	7.27	9.02
Open Format	8.51	7.13	6.23
Timely	8.04	6.07	7.73
Primary	7.52	5.87	7.38
Machine-readable	6.9	6.75	1.98
Permanent	6.38	5.18	3.88
Metadata	5.08	3.52	5.51
In bulk/API	5.03	4.83	2.43
Open license	4.07	3.65	2.38
Human-readable	6.38	4.07	9.02
Identifier	4.15	3.01	3.74
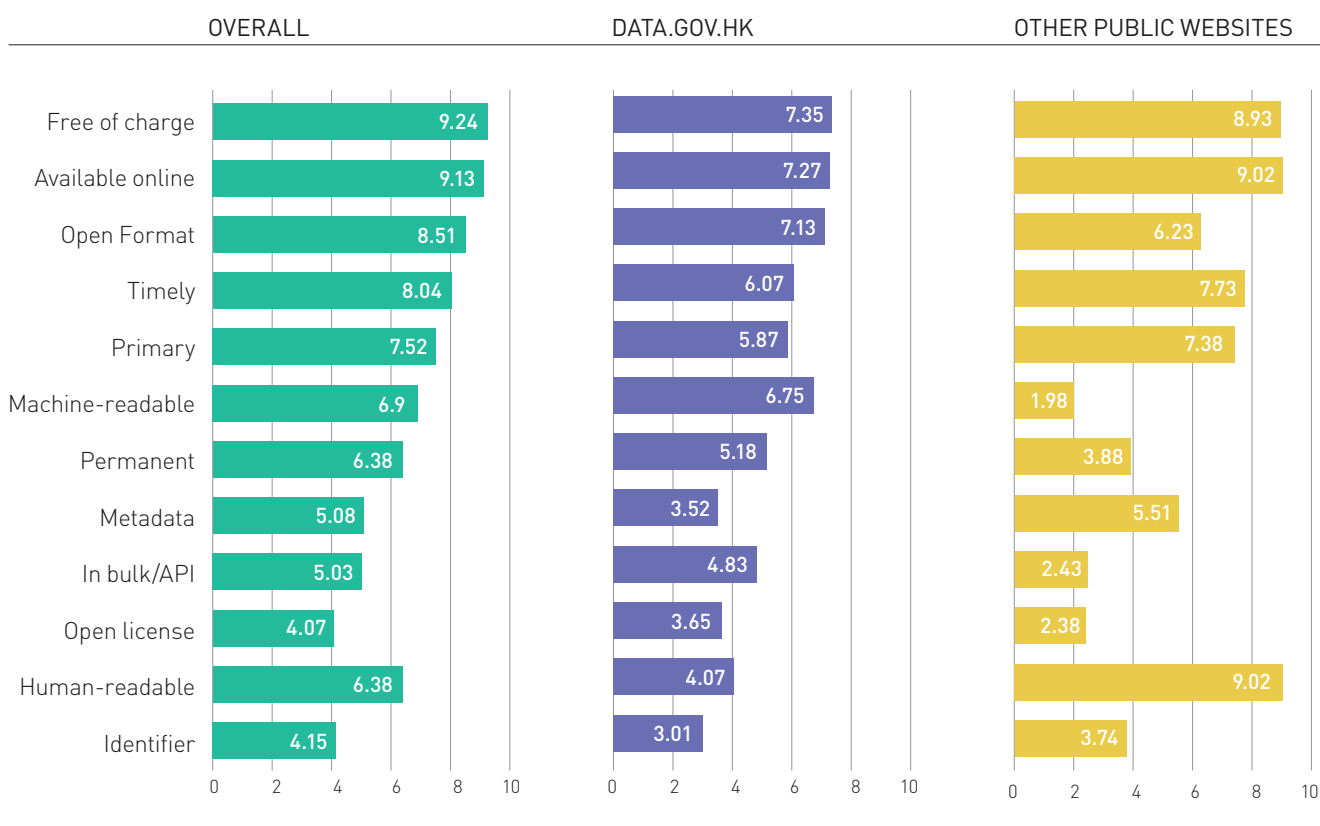
AVAILABLE ONLINE AND FREE OF CHARGE

		Overall	data.gov.hk	Other official websites	Full score
1	Available online	9.13	7.27	9.02	10
	1. Online without reqests	4.57	3.60	5.09	5
	2. No registration	4.68	3.71	5.13	5
2	Free of charge	9.24	7.35	8.93	10

Data is not meaningfully open if it is not available on the internet to the widest range of users, and therefore the availability indicator is the premise for others. Typical barriers to availability are requirements to submit requests and mandatory registration before you are allowed access to the data. Another hurdle for public members is the potential fee charged by data providers—usually the government departments—for the claimed administrative costs. Most of the open data standard-setting initiatives worldwide are upholding the "free of charge" principle (see Appendix IV: Indicator and dataset references).

The Hong Kong government and other public organisations in the city have made available most types of data online (scored 9.13 out of 10) without a charge (scored 9.24). A majority of datasets on the list for the Index to examine—an extensive checklists was compiled using established open data assessment tools worldwide and a local stakeholder consultation—can be found online in Hong Kong, either on the open data portal data.gov.hk and/or on the respective organisation's website.

AVAILABLE ONLINE

Category	other public websites	data.gov.hk	overall
Average	9.02	7.11	9.13
Science & Technology	10	9	10
Climate & Weather	10	5	10
Environment & Energy	10	10	10
Recreation & Culture	10	8	10
Health	10	10	10
Education	10	8.75	10
Welfare	10	9.38	9.38
Transport	8.75	8.75	8.75
Housing	8.13	6.25	8.75
Land	4.17	5.83	6.67
Business Registration	5.83	5	5
Justice & Safety	8.75	5	8.75
Government Operations	10	1.25	10
Legislature & Advisory Bodies	10	6	10
Public Finance	8.75	5.63	8.75
Census & Statistics	10	10	10

- other public websites
- data.gov.hk
- overall

The Index reveals that the data categories of business registration (scored 5 out of 10) and land (6.67) are significantly less available than others, because for the land and business ownership datasets users have to inquire through a search system requiring them to know specifics such as names and registration numbers beforehand. Such a measure fails the "no request" criterion. Also, detailed information regarding ownership is not available for free. For further analysis, please refer to the respective data section in the chapter Datasets Highlights.

In addition to the low availability of a few datasets, another major issue with this criterion is that the open data portal data.gov.hk registers a lower score on the indicator than other online sources (7.11 vs. 9.02), driving down scores on other indicators and hence weakening overall openness of the portal. When we compare the number of datasets, the open data portal provides fewer than other public websites that are qualified for the assessment in 11 data categories out of 16. Our assessment has revealed 11 types of datasets that are available on the websites of respective government departments or organisations but not on the open data portal. They are:

B.   Public Finance
B.3   Calls for tender

C.   Legislature & Advisory Bodies
C.2   Bills & Legislation
C.4   Legislative council (LegCo) meetings

D.   Government Operations
D.1   Government contact points
D.3   Declarations of interest

E.   Justice & Security
E.1   Judiciary & Judgement
E.4   Travel alerts

H.   Housing
H.4   Public housing
         (including waiting time and allocation status)

M.   Recreation & Culture
M.5   Hiking trails

O.   Climate & Weather
O.2   Wind
O.3   Rainfall distribution

The major issue in only publishing on the organisations' websites is that these websites usually did not consider open data criteria in the initial design. For example, we found that on average, departmental websites perform worse than the portal on the indicator of machine-readability (1.98 vs. 6.59) and historical records (3.88 vs. 5.03). Therefore, re-publishing the readily available data on the open data portal will immediately improve the openness of these datasets.

Meanwhile, we have found a couple types of dataset that are only freely available on the portal but not on the respective departments' websites:

G.   Land
G.1   Mapping
G.2   Boundaries

OPEN LICENSE

		Overall	data.gov.hk	Other official websites	Full score
4	Open license	**4.07**	**3.65**	**2.38**	**10**
	1. Open license	4.07	3.65	2.38	5
	2. Explicit	0.00	0.00	0.00	5

Open license

The published data, like any creative work, is usually subject to copyright, and therefore must seek its creators' permission for use. An open license is one which grants users permission to access, reuse and redistribute a work with few or no restriction (Open Definition v2.1, 2015). Almost every established open data initiative in the world demands that governments should adopt an open license to waive some of the major copyrights when they release data.

In Hong Kong, the HKODI's open license indicator is registered with the lowest score (4.07 out of 10) among all 12 indicators in the assessment, and we see two major issues. First, other than the open data portal data.gov.hk, the websites of government departments and public organisations adopt inconsistent terms and conditions regarding copyrights, as a few of them do not allow people to reuse and reshare the government data in the public sphere in the literal sense. Second, even though a number of datasets are released with open licenses, none of these licenses is presented explicitly to allow average people to quickly grasp the extent that they are authorised to use the data.

Below is a statement extracted from the copyright notice attached to the website of a government department, which publishes data regularly on both its website and the open data portal.

"Permission is granted for users to ... (ii) download the materials herein to store them in local computers, provided that this is solely for personal or non-commercial internal use, and provided further that this copyright notice is downloaded at the same time. "

—Copyright Notice, Electrical and Mechanical Services Department

The condition contradicts the open license definition that allows users to "access, reuse and redistribute a work with few or no restrictions." At the same time, the datasets published on this website are also available on the open data portal data.gov.hk, which adopts another license:

"You are allowed to browse, download, distribute, reproduce, hyperlink to, and print in their original format the Data for both commercial and non-commercial purposes on a free-of-charge basis..."

— Terms and Conditions of Use (Version 1.0), updated on 22 July 2019, DATA.GOV.HK

This license could be regarded as open, even though the terms of use still reserve a few conditions including a requirement for attribution to and indemnification of the government. The overall authorisation is enough for data reuse in most scenarios.

The above example of the contradictory copyright licenses for the same datasets published on two different platforms showcase how inconsistent the licenses are across departments and online platforms. It also means a lack of open license for a number of organisations. This together may discourage people from reusing data for fear of legal risks. Nevertheless, there is no evidence yet that anyone has been found liable for breaching a copyright license of open data in Hong Kong.

Explicitness

Another aspect of the copyright licensing issue is that they all lack explicitness and clarity. Users have to parse the legal language on their own in the "Terms and Conditions" or "Important Notices"—hidden in the corner of websites—to understand whether they can use the data for their own works and then publish it, exemplified as:
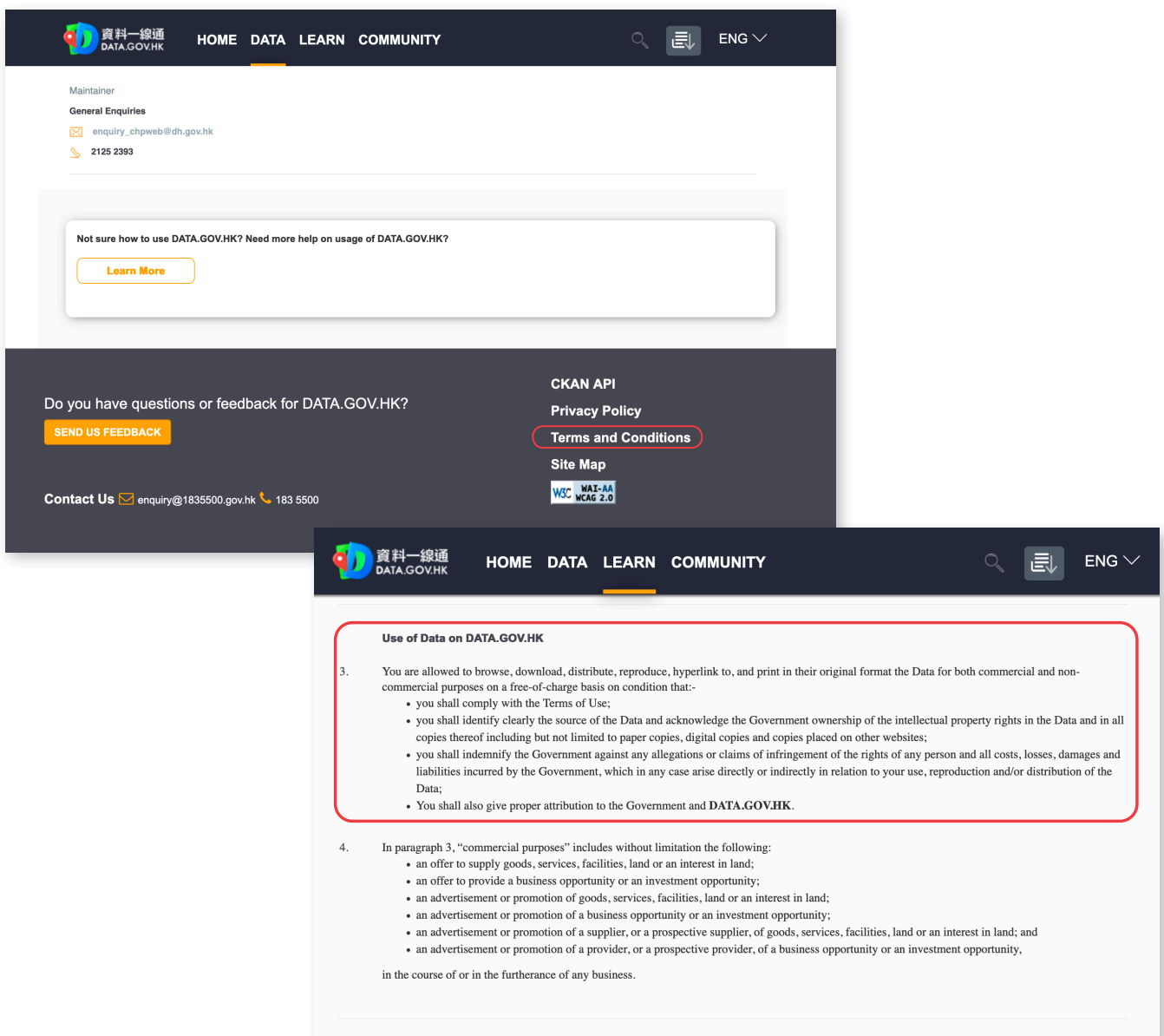
Source: The Hong Kong Public sector information portal DATA.GOV.HK

The concealment of open license may further prevent average people from making use of the data. A solution is to introduce a standardised license with a visual icon system to indicate which rights are reserved and which are not.

In the international open culture community, Creative Commons (CC) is a recognised licensing system and it features a range of visual icons that allow users to understand their rights to using the works at a glance. The Hong Kong government may consider introducing licenses as per CC.

Attribution

You let others copy, distribute, display, and perform your copyrighted work — and derivative works based upon it — but only if they give credit the way you request.

Noncommercial

You let others copy, distribute, display, and perform your work — and derivative works based upon it — but for noncommercial purposes only.

No Derivative Works

You let others copy, distribute, display, and perform only verbatim copies of your work, not derivative works based upon it.

Share Alike

You allow others to distribute derivative works only under a license identical to the license that governs your work.

Source: Creative Commons Hong Kong

Another common practice on governmental websites across advanced economies is to standardise their open licenses and present them uniformly. Here are examples from the open data portals of the US, UK and Singapore.

Source: Wiki Commons

Source: DATA.GOV hosted by the U.S. General Services Administration

Source: Data.gov.sg managed by the Government Technology Agency of Singapore

Source: data.gov.uk built by the UK Government Digital Service

In terms of standardisation, the Hong Kong government may refer to a list of licenses conforming with the Open Definition.[14] The list includes a few governmental licenses that have been authenticated through an approval process,[15] such as the ones developed by Canada, Germany, the UK and Taiwan. The Open Definition approval process will ensure the interoperability of the licenses across the world, and the Hong Kong government may also consider doing thish.

ACCESSIBLE VIA BULK DOWNLOAD AND API

		Overall	data.gov.hk	Other official websites	Full score
3	In Bulk/API	5.03	4.83	2.43	10
	1. Download in bulk	3.74	3.67	2.03	5
	2. API	1.29	1.15	0.40	5

Providing bulk download and Application Programming Interface (API) are two common methods to open up datasets for the public, the former is more familiar to the public and data providers than the latter. The Index reveals that more than half of the datasets examined are downloadable in bulk (scored 3.67 out of 5), but few of them are provided with APIs (1.22 out of 5).

API is a mechanism that allows users to get some slice of the data by placing specific queries, and makes it possible for programmers to automate the data access process. It is particularly valuable for datasets that are updated as frequently as every few minutes (real-time). The provision of an API requires extra resources to upgrade the data management system and therefore it is usually a lower priority on the government's open data agenda than providing downloadable files, which can be executed immediately instead. Nevertheless, an API is a sign of better data management and sharing capability of a government, and therefore its lower score for Hong Kong implies that there is still a long way to go before the city becomes a "smart city" and its government becomes a "smart government".

METADATA

		Overall	data.gov.hk	Other official websites	Full score
9	Metadata	5.08	3.52	5.51	10
	1. Core metadata	2.57	2.12	2.24	5
	2. Context description	2.66	1.56	3.27	5

Metadata was originally used in the catalogues of libraries to enable users to find books. For open data, metadata provides several kinds of information that defines and explains datasets, so users can find data by searching key words on the internet. Users can also quickly identify whether the dataset is usable for them by browsing information in the metadata columns on the respective webpage. The most common elements for a metadata scheme are the dataset title, source, publication date, and format, etc. These are also called "core metadata".

For the Index assessment, we reviewed the metadata column on the webpage of each dataset for usability, but we did not examine the metatags in the website's HTML codes that are for searchability such as by using Google. This arrangement was due to the constraints of our research capacity, but a sample test indicated that if a metadata element was missed on the human-readable webpage then it is unlikely to exist as a meta tag in its codes. Therefore our simplified method can still reflect the metadata condition at large.

We found that the majority of the datasets in Hong Kong are annotated with some core metadata elements including the dataset title, source (government departments), last revision date, and format; but fall short of publication date, cover period and key words or theme. Also, they were found to lack documentations to explain the context, data fields, data dictionaries, the purpose and method of data collection, etc. Such information would have helped users to fit the data into suitable scenarios and to address proper issues.

Below is a comparison of the metadata columns on the open data portals in Hong Kong and Singapore, and it shows that Singapore has a more complete metadata scheme in place than Hong Kong.
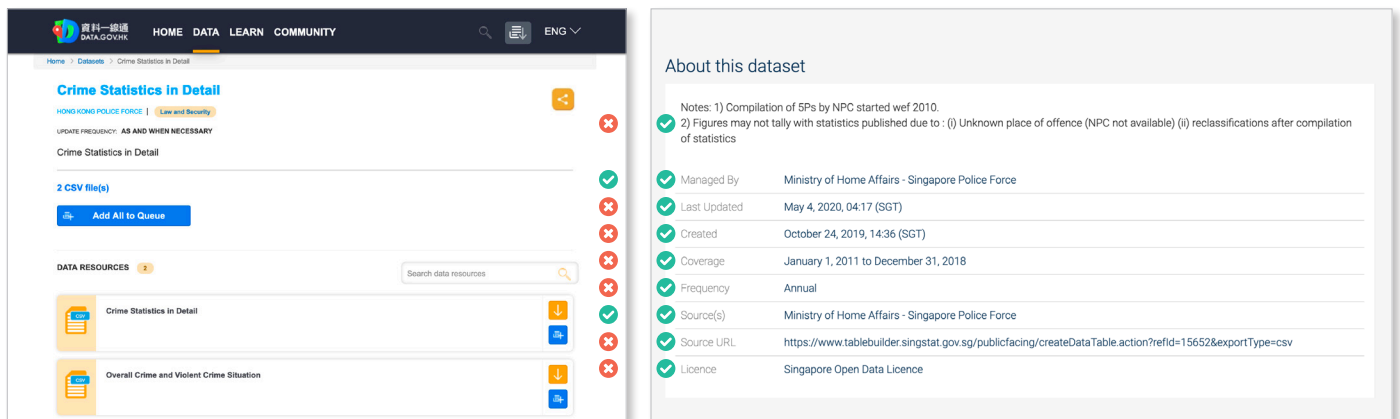
Source: Hong Kong Public Sector Information portal DATA.GOV.HK (left) and Singapore government's one-stop portal data.gov.sg (right)

Also, we have examples of metadata columns extracted from the open data portals of Shanghai and the United States.

Source: Shanghai Open Data Platform data.sh.gov.cn

Source: Data.gov managed and hosted by the U.S. General Services Administration

PERMANENT

		Overall	data.gov.hk	Other official websites	Full score
10	Permanent	6.38	5.18	3.88	10

The "permanent" indicator concerns archiving published datasets at stable online locations. Making the old data permanently available will ensure users can conduct comparative analyses over time, based on authentic documents rather than amended versions.

The Index assessment found that a number of datasets in Hong Kong only provide the latest version without archiving, or fail to make the online archives accessible to the public: they are maintained without navigation and description.

For example, the open data portal offers the function "Download Historical Data", but it is not functioning consistently for every dataset. The first dataset below showcases that from Jan 2018 to March 2020, there is only one file in the "historical data", which is actually the current data. Nevertheless, the "Download Historical Data" is not available for the second dataset below.
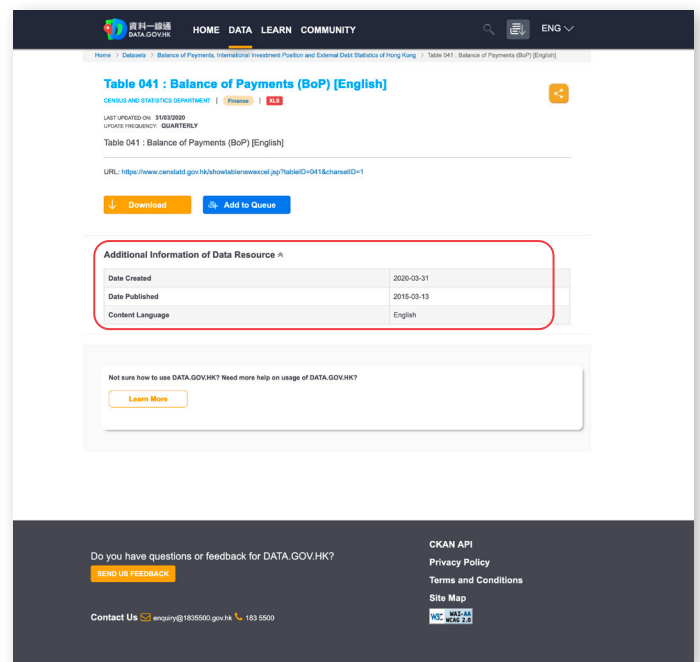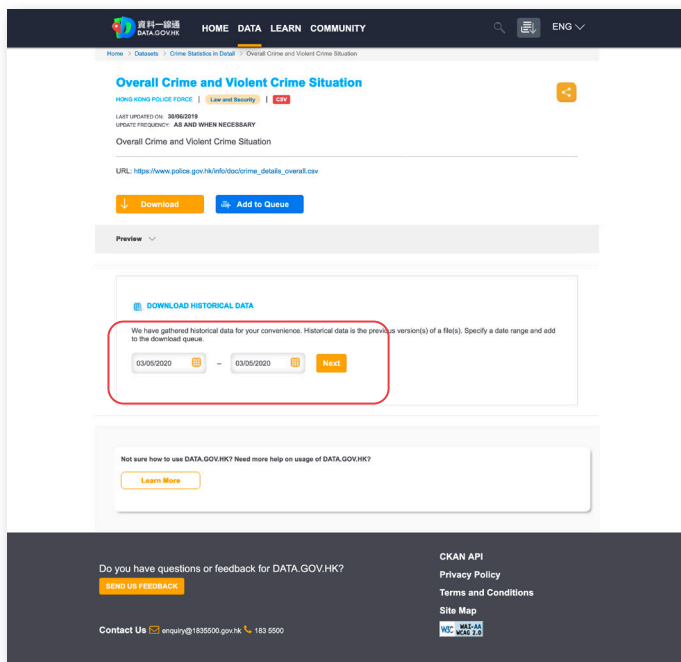
Source: Hong Kong Public Sector Information portal DATA.GOV.HK

Another example is the datasets concerning the population census, which is conducted every five years. The picture below shows that the latest census datasets are listed together with previous ones that are not in any order, you cannot tell how these datasets correspond to each other. They should have been grouped in clusters by years and theme.
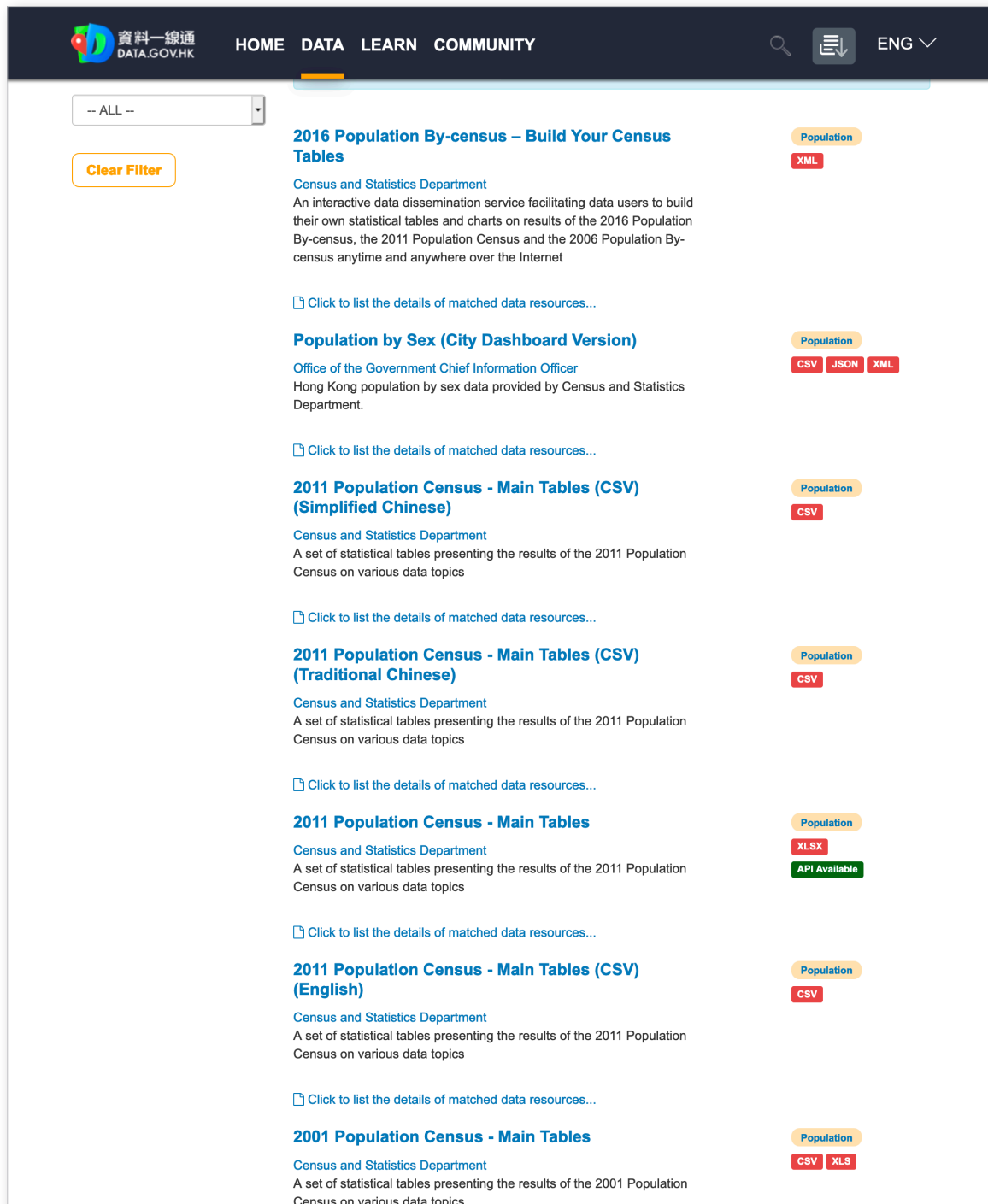
Source: Hong Kong Public Sector Information portal DATA.GOV.HK

There is no answer yet about what is the best method to present historical records, as long as they are managed uniformly without leaving users confused. It once again underscores the importance of consistency in data management.

MACHINE-READABLE AND OPEN FORMAT

		Overall	data.gov.hk	Other official websites	Full score
5	Machine-readable	6.90	6.75	1.98	10
	1. Machine-readable format	3.60	3.63	1.05	5
	2. Structured and standardised	3.53	3.43	0.93	5
6	Open format	8.51	7.13	6.23	10

Machine-readable

Data machine-readability is a principle that is frequently mentioned in discussions of open data around the world including Hong Kong. A machine-readable document must fulfil two criteria: first, the document format must be one that is fit to be processed by computer, such as CSV, XML or JSON; second, data in the document must be structured and standardised, for example, the numbers should be aligned with column and row headers in the CSV file.

Due to its significance, machine-readability is the subject of much criticism by open data advocates directed at data providers. Hong Kong was no exception to this as we have seen a number of articles exposing a lack of machine-readability in government documents published online in the past few years.

However, the Index assessment reveals that Hong Kong overall perform moderately in data machine-readability as it scores 6.9 out of 10 on the indicator. A majority of datasets published via the open data portal are in machine-readable formats (3.6 out of 5). Other official websites receive a lower average score of 1.98 out of 10, but we understand that they were not designed for open data and it would take additional resources to improve their machine-readability, which would need a cost/benefit analysis. Compared to other indicators, machine-readability is neither the worst nor the best.

The structure and standardisation aspect of the machine-readability—usually less prominent than the file formats—is also scored moderately at 3.53 out of 5. The example below is a typical unstructured data table in which each column has two or three different headers that would "confuse" the computer.

Source: A CSV file downloaded from the Telephone Directory of the HKSAR Government and Related Organisations

Open format

Open format is an independent indicator for the Index but correlated to machine-readable formats. It is a format that can be processed by at least one free or open-source software tool. No one possesses exclusive ownership over an open format, and therefore they are also called non-proprietary formats. Typical proprietary formats are those for commercial software, such as XLS and DOC

developed by Microsoft, and PDF by Adobe. The most common open formats are CSV, XML, HTML, PNG, and ODF. Most machine-readable formats are open formats as well because they were mainly created by open-source advocates.

Hong Kong scores at 8.51 out of 10 on the open format indicator, as most datasets are presented on the webpages (HTML format) if not downloadable.

TIMELY AND PRIMARY

		Overall	data.gov.hk	Other official websites	Full score
7	Timely	7.52	5.87	7.38	10
8	Primary	8.04	6.07	7.73	10

Timely

The published data are fairly up-to-date as the datasets cumulatively score 8.04 out of 10 on the indicator for timeliness. The Index researchers observed that most datasets had been updated within the last few months from the moment of their examinations, suggesting that a majority of data providers have fulfilled their maintenance duty once their datasets are made available publicly.

Update frequency is another aspect of timeliness but there is no universal standard because it varies due to the nature and value of data. It is noteworthy that three bus companies released their real-time arrival data through APIs in 2019, [16],[17] as well as four rail lines of the MTR.[18] This is a breakthrough after years of advocacy from civil society and business start-ups, as well as the coordination by the government.
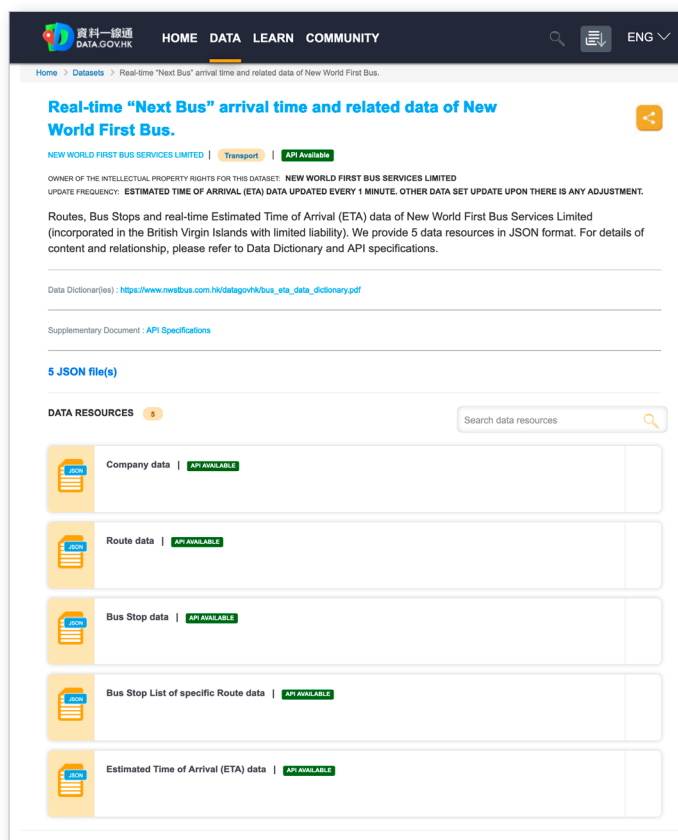
Source: Real-time "Next Bus" arrival time and related data of New World First Bus, Hong Kong Public Sector Information portal DATA.GOV.HK

Primary

Real-time data also fulfils the criteria set out by the indicator of "primary", which advocates that data should be published in its original and unmodified form without aggregation. Under the circumstances that aggregation is inevitable—e.g., for privacy and confidentiality concerns—the data should be disaggregated at the finest level of granularity when possible.

Hong Kong scores at 7.52 out of 10 on the primary indicator, performing moderately. It means a majority of datasets are released with raw data or at a proper granularity level.

For statistics, a challenge is to make the disaggregation meaningful for average readers. For example, the Consumer Price Index is released on the open data portal with a breakdown of its sub-indices. However, average readers would not know immediately what the sub-index of (A), (B) or (C) stands for, though there are footnotes underneath the table (footnotes may not be the best arrangement for machine-readable files such as CSV). A possible solution is to introduce the "linked data" approach in which every data entity (granule) would have an "ID" that allows users to link to the same granule in other datasets to make the granularity meaningful for cross-checking and analysis. Further research and deliberation on this are needed.

	A	B	C	D	E	F	G	H	I	J
1	Table052: Consumer Price Indices (October 2014 - September 2015 = 100)									
2										
3			Composite Consumer Price Index		Consumer Price Index (A)		Consumer Price Index (B)		Consumer Price Index (C)	
4	Year	Month	Index	Year-on-year % change	Index	Year-on-year % change	Index	Year-on-year % change	Index	Year-on-year % change
36	2006		74.7	+2.0	73.7	+1.7	74.8	+2.1	75.8	+2.2
37	2007		76.2	+2.0	74.7	+1.3	76.4	+2.2	77.8	+2.7
38	2008		79.5	+4.3	77.3	+3.6	79.9	+4.6	81.5	+4.7
39	2009		80.0	+0.5	77.8	+0.4	80.3	+0.5	82.0	+0.6
40	2010		81.8	+2.4	79.7	+2.7	82.2	+2.3	83.7	+2.1
41	2011		86.1	+5.3	84.2	+5.6	86.5	+5.2	88.0	+5.1
42	2012		89.6	+4.1	87.2	+3.6	90.2	+4.3	91.6	+4.1
43	2013		93.5	+4.3	91.7	+5.1	93.9	+4.1	95.1	+3.8
44	2014		97.7	+4.4	96.8	+5.6	97.8	+4.2	98.4	+3.5
45	2015		100.6	+3.0	100.6	+4.0	100.6	+2.9	100.5	+2.1
46	2016		103.0	+2.4	103.5	+2.8	102.9	+2.3	102.6	+2.1
47	2017		104.5	+1.5	105.1	+1.5	104.3	+1.4	104.2	+1.5
48	2018		107.0	+2.4	107.9	+2.7	106.7	+2.3	106.5	+2.2
49	1974	Jul	n.a.	n.a.	14.0	n.a.	14.4	n.a.	13.0	n.a.
50		Aug	n.a.	n.a.	13.8	n.a.	14.3	n.a.	13.0	n.a.
51		Sep	n.a.	n.a.	13.9	n.a.	14.3	n.a.	13.0	n.a.
52		Oct	n.a.	n.a.	14.2	n.a.	14.6	n.a.	13.1	n.a.
53		Nov	n.a.	n.a.	14.2	n.a.	14.6	n.a.	13.1	n.a.

Source: Consumer Price Indices (October 2014 – September 2015 = 100), Hong Kong Public Sector Information portal DATA.GOV.HK

DATASETS HIGHLIGHTS

DATASETS OVERVIEW

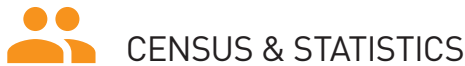|	OVERALL	DATA.GOV.HK	OTHER PUBLIC WEBSITES

A few government departments/organisations and data platforms in Hong Kong have set good examples of how open data should be managed and presented.
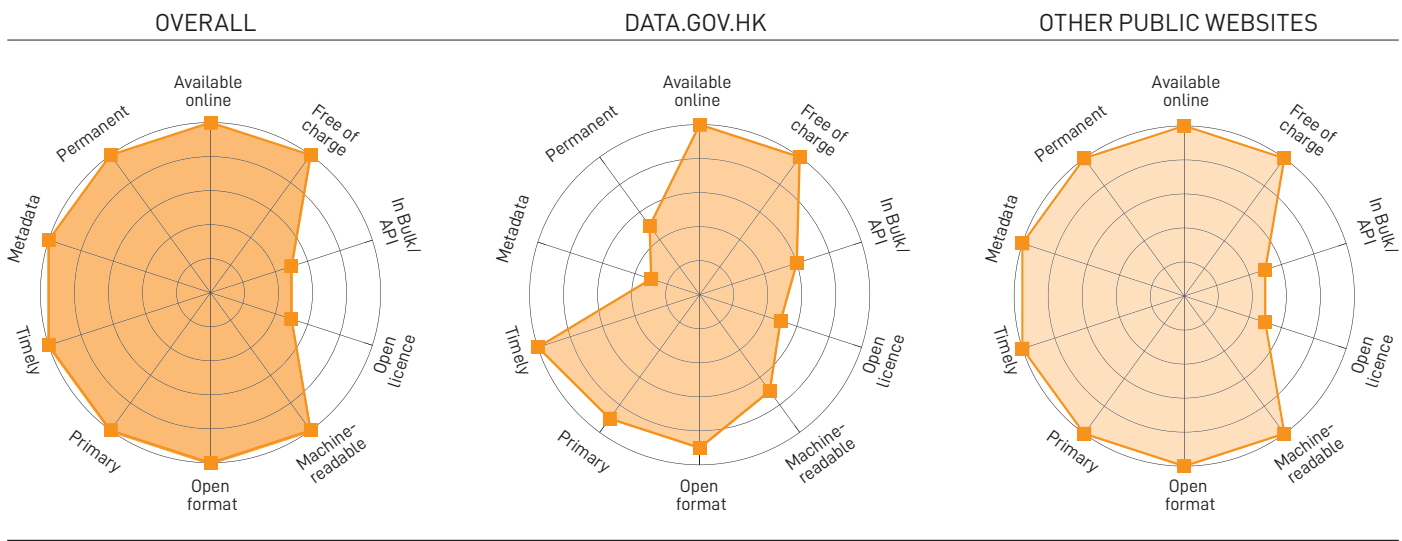
Websites of the Census and Statistics Department and Legislative Council are the front runners in the data opening process. The Hong Kong GeoData Store—a spatial-enabled data platform managed by the Lands Department—has showcased how essential features for an open data portal could be properly designed and arranged.

The Index also reveals the negative side: data in the five categories of land, business registration, justice & safety, housing and government operations are significantly less open than other datasets.

CENSUS & STATISTICS

Ranked 1st		Overall	data.gov.hk	Other official websites
A	Census & Statistics	90	73	90
	1. Census (e.g. population, location)	90	75	90
	2. Local income & balance of payments	90	73	90
	3. Trade performance	90	73	90
	4. Economic performance (e.g. CPI, PPI)	90	73	90

OVERALL   DATA.GOV.HK   OTHER PUBLIC WEBSITES

The data category of census & statistics outperforms others, because the Census and Statistics Department maintains its website in a way that mostly meets open data criteria (scored 90 out of 100), even better than the open data portal (scored 73) for the same types of datasets. The department's website receives significant higher scores than the portal on the indicator of metadata (10 vs. 3.13) and historical records (10 vs. 5).

Compared to other platforms and departments in Hong Kong, the only shortfall in the website of Census and Statistics Department is a lack of API (scored 0).

Note that the outstanding open data performance for the census & statistics datasets was a common phenomenon across the globe, evidenced in the previous rankings of Global Open Data Index and Open Data Barometer for the period 2013 to 2017. The implication of this might be that government statistics departments could contribute more to improve the open data process.

Source: Hong Kong Annual Digest of Statistics, Census and Statistics Department

LEGISLATURE & ADVISORY BODIES

Ranked 2nd		Overall	data.gov.hk	Other official websites
C	Legislature & Advisory Bodies	82	47	70
	1. Laws & Regulations	73	73	48
	2. Bills & Legislation	88	0	88
	3. Election results	88	88	70
	4. Legislative Council (legco) meetings	88	0	88
	5. District Council meetings	73	73	55

OVERALL   DATA.GOV.HK   OTHER PUBLIC WEBSITES
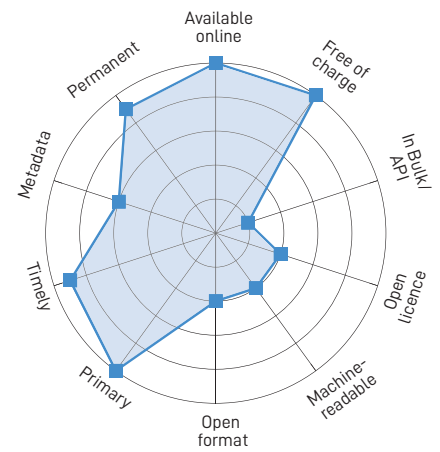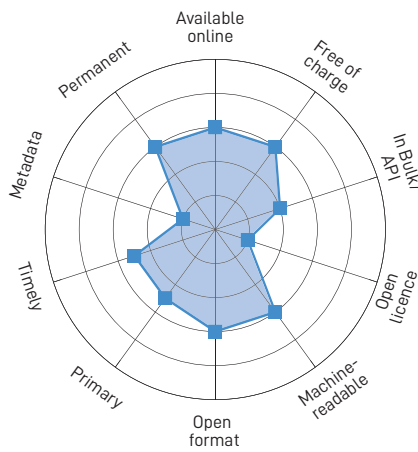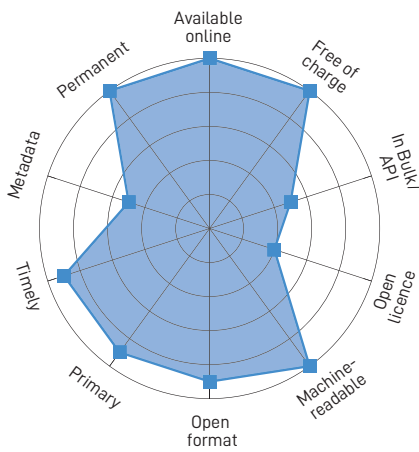
Data in the category of legislature & advisory bodies come from different sources, among which the website of the Legislative Council—covering the datasets "C.2 Bills & Legislation" and "C.4 Legislative Council (LegCo) meetings"—stands out as both types of datasets are scored 88.

The Legislative Council is one of the earliest organisations in Hong Kong that took the initiative to practise open data. It published the voting results in machine-readable XML and JSON formats from the 2012-2013 session, and ever since it has publicised an extended database including the verbatim meeting records (Hansard) and bills with their APIs. Note that the earliest digital meeting records of the LegCo dates back to 8 December 1858, though PDF format is not ideally machine-readable.

Source: Open Data, Legislative Council of the Hong Kong SAR

Source: Current and Past Meeting Records, Legislative Council of Hong Kong SAR

In comparison, the websites of the district councils are not as well designed for open data as the LegCo website, since documents are inconsistent and barely in machine-readable formats.

Source: Central & Western District Council

RECREATION & CULTURE

Ranked 3rd		Overall	data.gov.hk	Other official websites
M	**Recreation & Culture**	80	70	56
	1. Parks, zoos & gardens	88	88	58
	2. Sports facilities	88	88	58
	3. Museums & other culture facilities	88	88	53
	4. Libraries	88	88	60
	5. Hiking trails	50	0	50

OVERALL   DATA.GOV.HK   OTHER PUBLIC WEBSITES
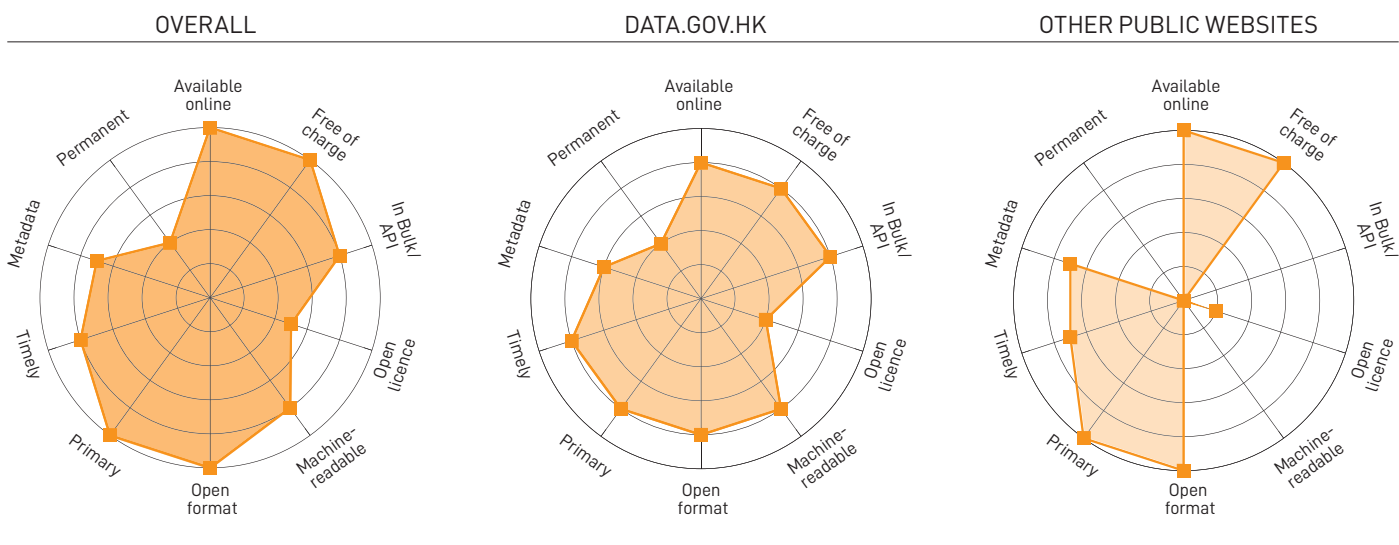
The recreation & culture category receives the third highest score at 80, mainly thanks to the Hong Kong GeoData Store that was launched in December 2018.[19] It is a geospatial information portal developed by the Lands Department to amass spatial-enabled datasets provided by various government departments for end users. The data includes the community and social services facilities, cultural, leisure and sports facilities, postal services facilities and job centres, many of which fall into the scope of the recreation & culture category.

When we searched the culture and recreation datasets in the main open data portal data.gov.hk, a number of results directed us to the GeoData Store instead. Therefore, our assessments on the datasets of parks, sports facilities, museums, and libraries were in fact based on the GeoData Store, and their scores turned out to be better than many other platforms and websites in Hong Kong.

Source: Hong Kong GeoData Store

A major feature of the GeoData Store is that it presents the data both in a reader-friendly manner with an interactive map and in at least five types of machine-readable formats, as well as with APIs. It is a well-designed and user-friendly data platform for both professionals and the general public.

LAND

Ranked 15th		Overall	data.gov.hk	Other official websites
G	**Land**	47	42	25
	1. Mapping	70	70	0
	2. Boundaries	33	33	0
	3. Land ownership	0	0	0
	4. Land utilisation	68	68	58
	5. Natural features or resources (e.g. trees)	45	45	30
	6. Urban planning (zoning, enforcement)	65	40	65

OVERALL   DATA.GOV.HK   OTHER PUBLIC WEBSITES
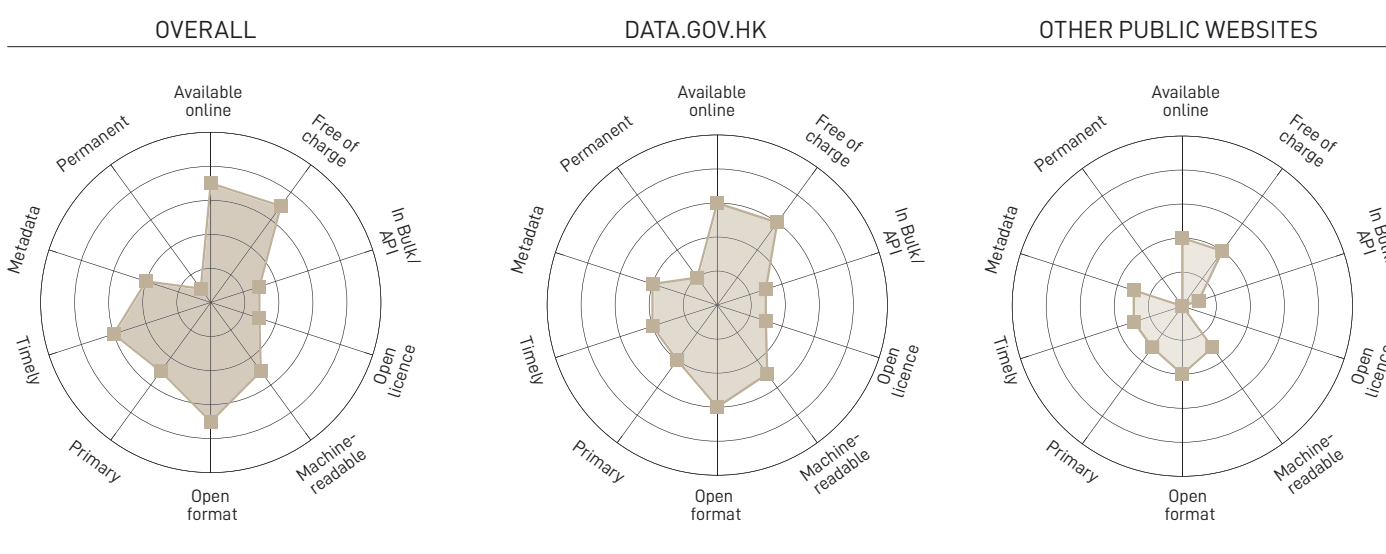
The category of land is scored the lowest among all 16 categories, mainly because the land ownership data is not freely available in Hong Kong.

Land ownership is the type of dataset examined by various open data assessment initiatives in the world including the Global Data Index and Open Data Barometer, and a number of jurisdictions won credits in this category. In Hong Kong, however, there are no free digests of land or house titles in the public sphere online. Users have to make inquiries

through the Internet Search Services in the Integrated Registration Information System and pay for records of the results.[20]

The datasets of boundaries receive a low score of 33, as only the district boundaries and boundaries for census are available in geo-referenced data format GML via the open data portal.[21,22] The detailed land boundary maps are for sale by the Lands Department.[23]

BUSINESS REGISTRATION

Ranked 15th		Overall	data.gov.hk	Other official websites
F	Business Registration	47	47	28
	1. Company register	48	48	33
	2. Licenses	93	93	50
	3. Beneficial ownership	0	0	0

OVERALL	DATA.GOV.HK	OTHER PUBLIC WEBSITES
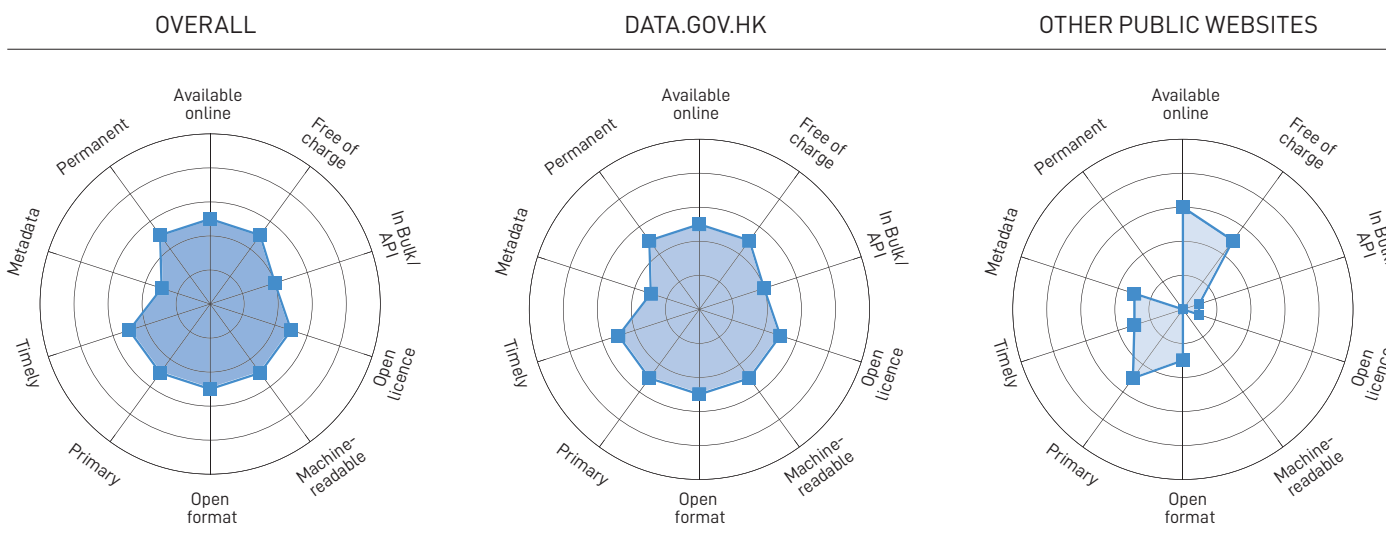
Business registration data is scored 47, tying with Land as the least open data category in Hong Kong, because the beneficial ownership data is not available in public in Hong Kong.

The dataset of beneficial ownership was initially required by the OECD Open Government Data initiative which has been developed since 2013., The OECD is a key advocate of transparency in corporate governance.[24]

In Hong Kong, the Companies Ordinance (Cap. 32) was amended to introduce the significant controllers registers from 1 March 2018, "to enhance transparency of corporate beneficial ownership in order to fulfil Hong Kong's international obligations."[25] However, the new registers are maintained by companies individually without a centralised database, and therefore there is no way to transform them into open data.

Members of the public can still search for the registered holders of a company via the Cyber Search Centre of the Integrated Companies Registry Information System (ICRIS), though they may not be the real beneficial owners. Like the Land Registry, the ICRIS charges users a fee, and so it is not open data.

HOUSING

Ranked 12th		Overall	data.gov.hk	Other official websites
H	Housing	58	48	45
	1. Stock of flats	68	68	58
	2. Building information	75	75	45
	3. Rental & sale transactions	48	48	38
	4. Public housing (e.g. waiting time)	40	0	40

OVERALL	DATA.GOV.HK	OTHER PUBLIC WEBSITES
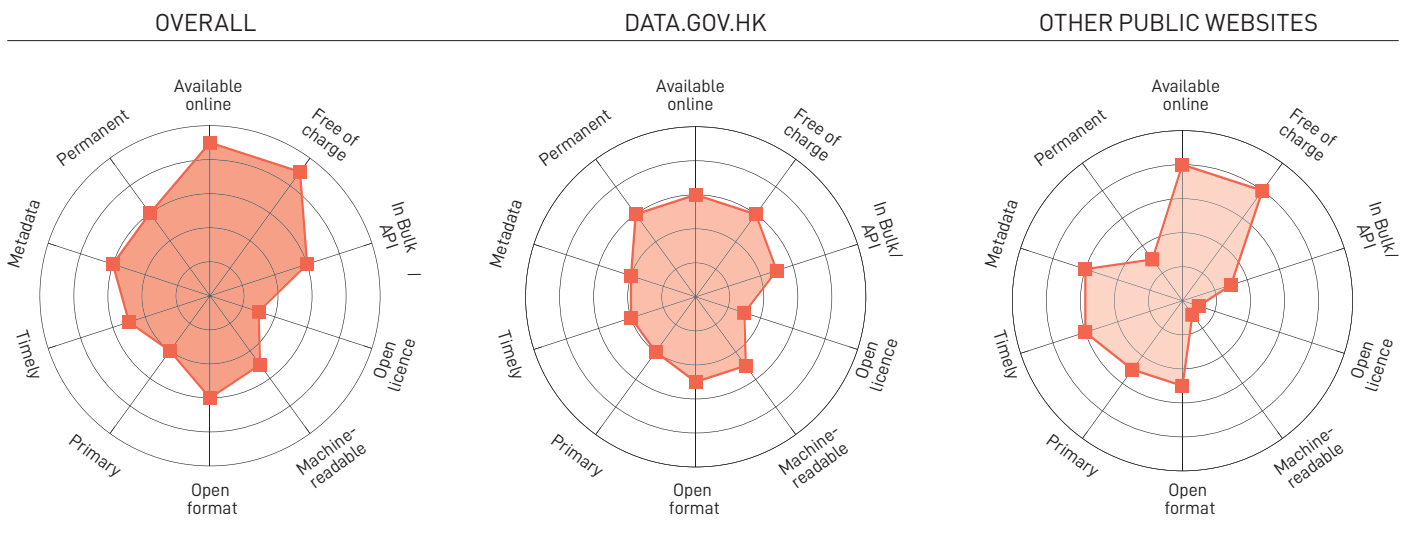
The openness of housing data is scored 58 as one of the lowest five categories in our assessment. The major shortfall in it is that the datasets are usually statistics without proper disaggregation or they should have been released as raw data with deidentification, due to the significance of the housing issue in Hong Kong.

For example, the allocation and waiting time for public housing is frequently discussed in Hong Kong, but such data is hidden in a nine-page PDF document,[26] seemingly updated only once a year and without historical records. The data is aggregated and the numbers are rounded to the nearest hundred.

Table 3
AWT as at end-June in recent years

	As at end-June 2013	As at end-June 2014	As at end-June 2015	As at end-June 2016	As at end-June 2017	As at end-June 2018	As at end-June 2019
AWT for general applicants	2.7 years	3.0 years	3.4 years	4.1 years	4.7 years	5.3 years	5.4 years
AWT for elderly one-person applicants	1.5 years	1.7 years	1.9 years	2.4 years	2.6 years	2.9 years	2.9 years

Table 4
Distribution of waiting time of general applicants who were housed to PRH between July 2018 and June 2019

District housed	Waiting Time	Household size					Total
		1-P Note 5	2-P	3-P	4-P	5-P+	
Urban	≤3 years	620	190	190	50	40	1 100
	>3 - ≤ 4 years	1 500	50	60	30	20	1 700
	>4 - ≤ 5 years	110	220	110	80	40	560
	>5 - ≤ 6 years	60	860	830	70	90	1 900
	>6 years	100	1 000	1 700	2 200	700	5 700
	Subtotal	2 400	2 300	2 900	2 400	890	10 900

Source: Special Analysis of the Housing Situation of General Applicants for Public Rental Housing as at end-June 2019, Subsidised Housing Committee of the Hong Kong Housing Authority

JUSTICE & SECURITY

Ranked 14th		Overall	data.gov.hk	Other official websites
E	**Justice & Security**	56	31	45
	1. Judiciary & judgement	35	0	35
	2. Crimes	53	53	38
	3. Emergency services (e.g. temporary shelters	73	73	45
	4. Travel alerts	63	0	63

OVERALL | DATA.GOV.HK | OTHER PUBLIC WEBSITES
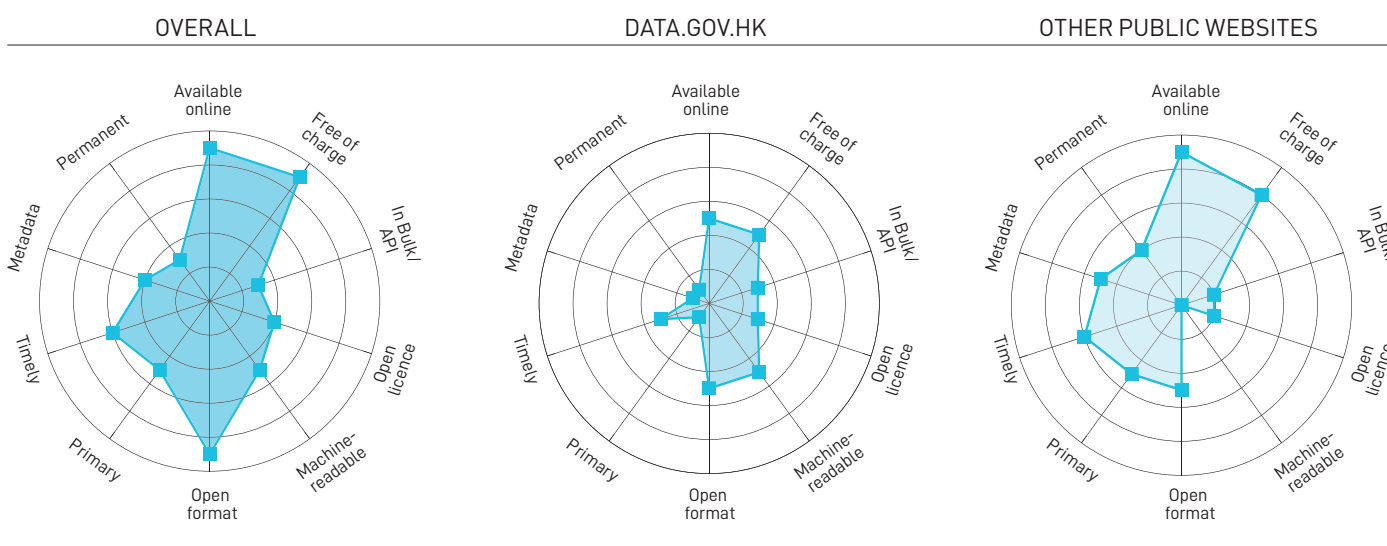
The data openness of the Justice & Security category is ranked the 3rd lowest with a score of 56, even though the datasets are all available online. The major problem is that they fail to meet key open data criteria such as being machine-readable and primary.

Within this category, the judiciary datasets achieve the lowest score at 35, because the court judgements are not completely available online and not in ways that facilitate bulk access and computer processing. Nevertheless, the judgements are uploaded online immediately after the rulings have been made.

The crime statistics are scored at 53, as they are not disaggregated by age, gender, or ethnicity, which are the common elements for insights into the crime situation of a city, in addition to breakdowns such as type of crime.

Travel alerts is another dataset that should be more open, since Hong Kong is an international business and transports hub with many frequent travellers. The automated release

of travel restriction information around the world with API and in machine-readable data form is valuable for many individuals and companies.

Source: Legal Reference System, retrieved on 25 April 2020

🏛 GOVERNMENT OPERATIONS

Ranked 12th		Overall	data.gov.hk	Other official websites
D	**Government Operations**	63	31	56
	1. Government contact points	65	0	65
	2. Government structure & personnel	50	35	50
	3. Declarations of interest	48	0	48
	4. Salaries (pay scale)	88	88	63

OVERALL   DATA.GOV.HK   OTHER PUBLIC WEBSITES
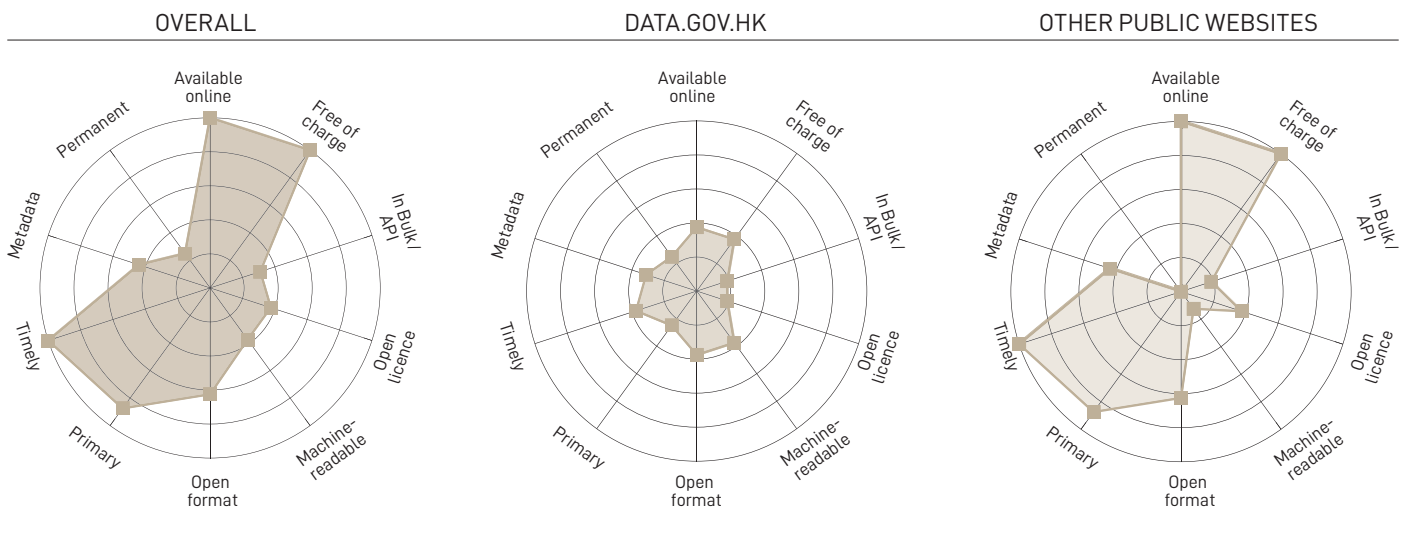
The data regarding government operations in Hong Kong is scored at 63, the 5th lowest. These datasets are usually available on respective departments' websites but not on the open data portal, for example, the declarations of interests of major officials.

In Hong Kong, members of the Executive Council and other politically appointed officials are required to declare their investments and interests and make them available to the public.[27] There is no one-stop online platform to publish these declarations, but they are released on the website of the Executive Council and websites of policy bureaux individually.[28]

However, the declarations of interests are published as non-machine-readable images in PDF format, created from scanned paper files, posing a crucial challenge for cross-checking and analysis by the public.

Source: Register of Members' Interests, Executive Council of the Government of Hong Kong SAR

Source: Website Chief Secretary for Administration, The Government of Hong Kong SAR

Serial No.	Particulars of Interests
1.	One residential flat and one parking space in Zhongshan, Mainland China (solely owned by spouse and for self-use)

Date _____ 1 July 2019 _____   Signature : _____

Source: Registration of Financial and Other Interests, website of the Chief Executive of Hong Kong SAR

METHODOLOGY

The Hong Kong Open Data Index is an assessment tool to test the openness of public data.

Assessment period

October 2019—March 2020

Methodology development

The Index is drafted based on a review of 19 established open data standards at international or regional levels, including 10 sets of standard-setting principles and nine assessment tools. A local stakeholder consultation was conducted from December 2018 to January 2019.

10 Standard-setting principles

Open Definition	2005-
Eight Open Government Data Principles	2007
Ten Principles for Opening Up Government Information	2010
Five-Star Open Data	2010
Open Data Policy Guidelines	2012/13/14
Open Government Data: The Book	2012/14
Open Data Handbook	2012/15
G8 Open Data Charter	2013
International Open Data Charter	2015
Internet Universality	2018

9 Assessment tools

Global Open Data Index	2013-2017
Open Data Barometer	2013-2018
ePSI Platform Scoreboard	2013
G20 Anti-corruption Open Data Principles Assessment	2015
Open Data Inventory	2015-
OECD OURdata Index on Open Government Data	2017
Open Standards Directory	2017
Taiwan Open Government Report	2017
China Open Data Index	2017-

Based on the review and consultation, the structure of Index has been firmed up: it consists of 12 indicators (including two non-scoring indicators) to evaluate the datasets in the Hong Kong across 16 categories covering 69 types of datasets.

Indicators

The Index consists of 12 indicators (see I. Indicators below), including 10 primary indicators, and two secondary indicators for reference only. The scores of the secondary indicators are not incorporated into the final overall scores, because they are less clearly defined and recognised by the international open data community. Please refer to Appendix III: Indicator description for the rationale of each indicator.

Datasets

The public data that the Index examines is the data produced by the government, organisations funded by public budgets, and the data owned by private businesses but of public interest. For example, bus services are operated by private companies in Hong Kong, but their operational data such as estimated arrival times should be regarded as public data and therefore fall into the scope of the Index assessment.

For research purposes, we divide the public data in Hong Kong into 16 categories that cover 69 types of datasets (see II. Datasets below). The categorisation is also based on the 19 international standards, as well as a "potential priority list" in the consultancy study for the smart city blueprint.[29] Each type of dataset is examined from two types of sources separately: the open data portal data.gov.hk and individual websites of the organisations that produce the data.

Scoring

Each type of dataset is scored by 10 primary indicators respectively in increments of 0/5/10, based on one of the possible answers: "No", "Partial", or "Yes".

The score of each dataset type is calculated as the sum of the scores on 10 indicators, and the final score for the HKODI is the average score of all 16 dataset types. The scoring for each dataset type is also determined based on the two types of sources separately — the website data.gov.hk and other public websites. We compare the scores of each dataset type on the two sources, from which we choose the higher one to be aggregated into the calculation of the overall score for Hong Kong.

I. INDICATORS

Indicator	Questions for evaluation	Score	Chaining logic
1 Available online to anyone	1.1 Can users access the dataset online without submitting requests?	5	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial THEN 2.5 ELSE IF "1.1" = Yes THEN 5
	1.2 Can users access the dataset without being required to register or identify themselves?	5	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "1.2" = Partial OR "1.2" = Yes) THEN 2.5 ELSE IF "1.1" = Yes AND (IF "1.2" = No THEN 0 ELSE IF "1.2" = Partial THEN 2.5 ELSE IF "1.2" = Yes THEN 5)
2 Free of charge	Is the dataset available free of charge?	10	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "2" = Partial OR "2" = Yes) THEN 5 ELSE IF "1.1" = Yes AND (IF "2" = No THEN 0 ELSE IF "2" = Partial THEN 5 ELSE IF "2" = Yes THEN 10)
3 In Bulk/API	3.1 Is the dataset downloadable in bulk?	5	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "3.1" = Partial OR "3.1" = Yes) THEN 2.5 ELSE IF "1.1" = Yes AND (IF "3.1" = No THEN 0 ELSE IF "3.1" = Partial THEN 2.5 ELSE IF "3.1" = Yes THEN 5)
	3.2 Is the dataset provided with an Application Programming Interface (API) when applicable?	5	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "3.2" = Partial OR "3.2" = Yes) THEN 2.5 ELSE IF "1.1" = Yes AND (IF "3.2" = No THEN 0 ELSE IF "3.2" = Partial THEN 2.5 ELSE IF "3.2" = Yes THEN 5)
4 Open license	4.1 Is the dataset released under an open license?	5	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "4.1" = Partial OR "4.1" = Yes) THEN 2.5 ELSE IF "1.1" = Yes AND (IF "4.1" = No THEN 0 ELSE IF "4.1" = Partial THEN 2.5 ELSE IF "4.1" = Yes THEN 5)
	4.2 Is the open license presented in an explicit manner?	5	IF "4.1" = No THEN 0 ELSE IF "4.1" = Partial AND (IF "4.2" = Partial OR "4.2" = Yes) THEN 2.5 ELSE IF "4.1" = Yes AND (IF "4.2" = No THEN 0 ELSE IF "4.2" = Partial THEN 2.5 ELSE IF "4.2" = Yes THEN 5)
5 Machine-readable	5.1 Is the dataset provided in machine-readable formats?	5	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "5.1" = Partial OR "5.1" = Yes) THEN 2.5 ELSE IF "1.1" = Yes AND (IF "5.1" = No THEN 0 ELSE IF "5.1" = Partial THEN 2.5 ELSE IF "5.1" = Yes THEN 5)
	5.2 Is the data organised in a structured or standardised manner?	5	IF "5.1" = No THEN 0 ELSE IF "5.1" = Partial AND (IF "5.2" = Partial OR "5.2" = Yes) THEN 2.5 ELSE IF "5.1" = Yes AND (IF "5.2" = No THEN 0 ELSE IF "5.2" = Partial THEN 2.5 ELSE IF "5.2" = Yes THEN 5)

Indicator		Questions for evaluation	Score	Chaining logic
6	Open format	Is the dataset provided in an open format, which can be processed with at least one non-proprietary application?	10	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "6" = Partial OR "6" = Yes) THEN 5 ELSE IF "1.1" = Yes AND (IF "6" = No THEN 0 ELSE IF "6" = Partial THEN 5 ELSE IF "6" = Yes THEN 10)
7	Primary	Is the dataset released at the finest possible level of granularity available, not in aggregate or modified forms?	10	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "7" = Partial OR "7" = Yes) THEN 5 ELSE IF "1.1" = Yes AND (IF "7" = No THEN 0 ELSE IF "7" = Partial THEN 5 ELSE IF "7" = Yes THEN 10)
8	Timely	Is the dataset up to date?	10	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "8" = Partial OR "8" = Yes) THEN 5 ELSE IF "1.1" = Yes AND (IF "8" = No THEN 0 ELSE IF "8" = Partial THEN 5 ELSE IF "8" = Yes THEN 10)
9	Metadata	9.1 Is the dataset provided with core metadata ?	5	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "9.1" = Partial OR "9.1" = Yes) THEN 2.5 ELSE IF "1.1" = Yes AND (IF "9.1" = No THEN 0 ELSE IF "9.1" = Partial THEN 2.5 ELSE IF "9.1" = Yes THEN 5)
		9.2 Is the dataset provided with accompanying documentation describing the context?	5	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "9.2" = Partial OR "9.2" = Yes) THEN 2.5 ELSE IF "1.1" = Yes AND (IF "9.2" = No THEN 0 ELSE IF "9.2" = Partial THEN 2.5 ELSE IF "9.2" = Yes THEN 5)
10	Permanent	Are historical copies of the dataset kept accessible at a stable online location?	10	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "10" = Partial OR "10" = Yes) THEN 5 ELSE IF "1.1" = Yes AND (IF "10" = No THEN 0 ELSE IF "10" = Partial THEN 5 ELSE IF "10" = Yes THEN 10)
(11)	Identifier	Is the dataset provided with Uniform Resources Identifiers (URIs) to denote its key elements.	(10)	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "11" = Partial OR "11" = Yes) THEN 5 ELSE IF "1.1" = Yes AND (IF "11" = No THEN 0 ELSE IF "11" = Partial THEN 5 ELSE IF "11" = Yes THEN 10)
(12)	Human-readable	Is the data written in plain and clear language that can be understood by the general public?	(10)	IF "1.1" = No THEN 0 ELSE IF "1.1" = Partial AND (IF "11" = Partial OR "11" = Yes) THEN 5 ELSE IF "1.1" = Yes AND (IF "11" = No THEN 0 ELSE IF "11" = Partial THEN 5 ELSE IF "11" = Yes THEN 10)

II. DATASETS

Indicator		Types of Datasets	
A	Census & Statistics	1	Census (including language, population, location)
		2	Local income & Balance of payments
		3	Trade performance
		4	Economic performance (e.g., CPI, PPI)
B	Public Finance	1	Government budget
		2	Government accounts
		3	Calls for tender
		4	Contracts let
C	Legislature & Advisory Bodies	1	Laws & Regulations
		2	Bills & Legislation
		3	Election results
		4	Legislative council (LegCo) meetings
		5	District councils meetings
D	Government Operations	1	Government contact points
		2	Government structure & personnel
		3	Declarations of interest
		4	Salaries (pay scales)
E	Justice & Safety	1	Judiciary & Judgement
		2	Crimes
		3	Emergency services (e.g., police stations, fire stations, temporary shelters)
		4	Travel alerts
F	Business Registration	1	Company register (name, unique identifier, address)
		2	Licenses
		3	Beneficial ownership
G	Land	1	Mapping
		2	Boundaries
		3	Land ownership
		4	Land utilisation
		5	Natural features or resources (e.g., trees, rivers, streams, mountains)
		6	Urban planning (e.g, zoning, enforcement cases)
H	Housing	1	Stock of flats
		2	Building information
		3	Rental & sale transactions
		4	Public housing (including waiting time and allocation status)

Indicator		Types of Datasets	
I	Transport	1	Public transport timetables (real-time)
		2	Public transportation statistics
		3	Parking lots & Charge stations
		4	Traffic flow (road, highways, tunnels, etc.)
J.	Welfare	1	Unemployment benefits
		2	Elderly benefits
		3	Services for the elderly
		4	Services for the needy
K.	Education	1	List of educational institutions
		2	Performance of schools (public exams, etc.)
		3	Performance of higher education institutions
		4	Budgets of educational institutions
L.	Health	1	Healthcare facilities
		2	Healthcare service performance
		3	Hygiene inspection & Food safety
		4	Healthcare practitioners (including register and declaration)
		5	Diseases
M.	Recreation & Culture	1	Parks, zoos & gardens
		2	Sports facilities
		3	Museums and other cultural facilities
		4	Libraries
		5	Hiking trails
N	Environment & Energy	1	Air quality
		2	Water quality
		3	Pollutant emissions
		4	Energy consumption
O	Climate & Weather	1	Temperature
		2	Wind
		3	Rainfall distribution
		4	Lightning locations
P	Science & Technology	1	Research & development supports (funds, tax cut, etc.)
		2	Wifi spots
		3	Other telecommunications (broadband penetration, mobile usage, etc)
		4	Patents
		5	Granted projects and research reports (including UGC and PICO)

ENDNOTES

1. OECD (2018), Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact, OECD

2. Open Knowledge Foundation, Open Definition, https://opendefinition.org/

3. Open Government Working Group (2007), Open Government Data Principles, https://public.resource.org/8_principles.html. The eight principles that define "open" government data were: complete, primary, timely, accessible, machine-processible, non-discriminatory, non-proprietary, and license-free. Many of the 30 advocates were influential in promoting an open culture of the internet, e.g, Tim O'Reilly popularised the terms "open source" and "Web 2.0"; Lawrence Lessig founded the open licensing initiative Creative Commons.

4. Lau, S. (2016),Open Data, Legislative Council Secretariat, https://www.legco.gov.hk/research-publications/english/essentials-1516ise18-open-data.htm#endnote3

5. HKSAR Government (2019), Smart City Blueprint: Development Plans, https://www.smartcity.gov.hk/develop_plans/government/

6. HKSAE Government (2019), "Over 650 datasets to be released in first year of government open data plans", Press Release, https://www.info.gov.hk/gia/general/201901/03/P2019010300255.htm

7. Open Data Charter (n.d.) 'Principles', esp. Principle 4 – Comparable and Interoperable, https://opendatacharter.net/principles/

8. HKSAR Government (2019), 'Over 650 datasets to be released in first year of government open data plans', press release, https://www.info.gov.hk/gia/general/201901/03/P2019010300255.htm

9. Legislative Council Panel on Information Technology and Broadcasting (2018) 'Open Data Policy', LC Paper No. CB(4)283/18-19(03), https://www.legco.gov.hk/yr18-19/english/panels/itb/papers/itb20181210cb4-283-3-e.pdf

10. Taipei City Government Department of Information Technology (2017) 臺北市資料大平台，https://data.taipei/#/members/wantmore

11. Shanghai City Government (2020) 上海市公共數據開放平台，https://data.sh.gov.cn/view/interactive-community/index.html

12. A.Brandusescu, C. Iglesias, D.Lämmerhirt and S. Verhulst (2019), 'Open data governance and open governance: interplay or disconnect?', World Wide Web Foundation, https://webfoundation.org/2019/02/open-data-governance-and-open-governance-interplay-or-disconnect/

13. Office for the Government Chief Information Officer (2017) 'Opening remarks by S for IT at press conference on Smart City Blueprint for Hong Kong', press release, https://www.ogcio.gov.hk/en/news/press_releases/2017/12/pr_20171215_1.htm

14. Open Knowledge Foundation (n.d.) 'Conformant Licenses', https://opendefinition.org/licenses/

15. Open Knowledge Foundation (n.d.) 'License Approval Process', https://opendefinition.org/licenses/process/

16. HKSAR Government (2019) 'Transport Department welcomes opening up of real-time arrival data by New World First Bus Services Limited and Citybus Limited', press releases, https://www.info.gov.hk/gia/general/201903/21/P2019032100598.htm?fontSize=1

17. HKSAR Government (2019) 'Transport Department welcomes New Lantao Bus Company (1973) Limited's opening up of real-time arrival data', press releases, https://www.info.gov.hk/gia/general/201904/17/P2019041700260.htm?fontSize=1

18. HKSAR Government (2019) 'More public transport operators participate in opening up real-time arrival data', press releases, https://www.info.gov.hk/gia/general/201905/31/P2019053100305.htm?fontSize=1

19. HKSAR Government (2018) 'Revamped GeoInfo Map and new Hong Kong GeoData Store facilitate smart city development and geospatial data applications', press release, https://www.info.gov.hk/gia/general/201812/10/P2018121000341.htm

20. The Land Registry, HKSAR (2020) 'Public Services: (B) Search of Land Records', https://www.landreg.gov.hk/en/services/services_b.htm

21. HKSAR Home Affairs Department (2020) 'District boundary',on DATA.GOV.HK, https://data.gov.hk/en-data/dataset/hk-had-json1-hong-kong-administrative-boundaries

22. HKSAR Planning Department (n.d.) '2016 Population By-census (Boundaries of New Towns)', on DATALGOV.HK, https://data.gov.hk/en-data/dataset/hk-pland-pland1-boundaries-of-new-towns

23. HKSAR Lands Department () 'Maps and Services', https://www.landsd.gov.hk/mapping/en/digital_map/sampledata.htm

24. OECD (2003) 'Maritime Security - Ownership And Control of Ships: Options to Improve Transparency', discussion paper, DSTI/DOT/MTC(2003)61/REV1, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/DOT/MTC(2003)61/REV1&docLanguage=En

25. HKSAR Companies Registry (2018) 'Overview', https://www.cr.gov.hk/en/legislation/scr/overview.htm

26. HKSAR Housing Authority (2019)'Special analysis of the housing situation of general applicants for PRH as at end-June 2019', https://www.housingauthority.gov.hk/en/common/pdf/about-us/housing-authority/ha-paper-library/SHC62-19EN.pdf

27. HKSAR Constitutional and Mainland Affairs Bureau (2012) 'Code for Officials Under the Political Appointment System', https://www.cmab.gov.hk/en/issues/PAO_Code_1.7.2012.pdf

28. HKSAR Governemnt (2019) 'Annual updating of declarations of interests by CE, politically appointed officials and ExCo Members', press release, https://www.info.gov.hk/gia/general/201907/30/P2019073000389.htm

29. PWC (2017) Report of Consultancy Study on Smart City Blueprint for Hong Kong, retrieved from https://www.smartcity.gov.hk/report/

APPENDICES

I.   ASSESSMENT RESULTS BY TYPES OF DATASET

Data category Type of datasets		data.gov.hk	Other websites	Overall
A	**Census & Statistics**	**73.13**	**90**	**90**
1	Census (including language, population, location)	75	90	90
2	Local income & Balance of payments	72.5	90	90
3	Trade performance	72.5	90	90
4	Economic performance (e.g., CPI, PPI)	72.5	90	90
B	**Public Finance**	49.38	71.25	71.25
1	Government budget	**77.5**	**87.5**	**87.5**
2	Government accounts	80	87.5	87.5
3	Calls for tender	0	55	55
4	Contracts let	40	55	55
C	**Legislature & Advisory Bodies**	**46.5**	**67.5**	**79.5**
1	Laws & Regulations	72.5	47.5	72.5
2	Bills & Legislation	0	87.5	87.5
3	Election results	87.5	70	87.5
4	Legislative Council (LegCo) meetings	0	77.5	77.5
5	District Council meetings	72.5	55	72.5
D	**Government Operations**	**8.75**	**56.25**	**56.25**
1	Government contact points	0	65	65
2	Government structure & personnel	35	50	50
3	Declarations of interest	0	47.5	47.5
4	Salaries (pay scales)	0	62.5	62.5
E	**Justice & Safety**	**31.25**	**45**	**55.63**
1	Judiciary & Judgements	0	35	35
2	Crimes	52.5	37.5	52.5
3	Emergency services (e.g., police stations, fire stations,temporary shelters)	72.5	45	72.5
4	Travel alerts	0	62.5	62.5
F	**Business Registration**	**46.67**	**27.5**	**46.67**
1	Company register (name, unique identifier, address)	47.5	32.5	47.5
2	Licenses	92.5	50	92.5
3	Beneficial ownership	0	0	0
G	**Land**	**42.5**	**25.42**	**46.67**
1	Mapping	70	0	70
2	Boundaries	32.5	0	32.5
3	Land ownership	0	0	0
4	Land utilisation	67.5	57.5	67.5
5	Natural features or resources (e.g., trees, rivers, streams, mountains)	45	30	45
6	Urban planning (e.g, zoning, enforcement cases)	40	65	65
H	**Housing**	**47.5**	**45**	**57.5**
1	Stock of flats	67.5	57.5	67.5
2	Building information	75	45	75
3	Rental & sale transactions	47.5	37.5	47.5
4	Public housing (including waiting time and allocation status)	0	40	40

Data category Type of datasets		data.gov.hk	Other websites	Overall
I	Transport	76.88	42.5	76.88
1	Public transport timetables (real-time)	42.5	25	42.5
2	Public transportation statistics	87.5	60	87.5
3	Parking lots & Charge stations	90	45	90
4	Traffic flow (roads, highways, tunnels, etc.)	87.5	40	87.5
J	Welfare	77.5	55.63	77.5
1	Unemployment benefits	70	65	70
2	Elderly benefits	85	42.5	85
3	Services for the elderly	77.5	57.5	77.5
4	Services for the needy	77.5	57.5	77.5
K	Education	71.88	56.88	75
1	List of educational institutions	87.5	57.5	87.5
2	Performance of schools (public exams, etc.)	47.5	60	60
3	Performance of higher education institutions	80	60	80
4	Budgets of educational institutions	72.5	50	72.5
L	Health	76	60	76
1	Healthcare facilities	70	50	70
2	Healthcare service performance	77.5	55	77.5
3	Hygiene inspection & Food safety	72.5	60	72.5
4	Healthcare practitioners (including register and declaration)	75	57.5	75
5	Diseases	85	77.5	85
M	Recreation & Culture	70	55.5	80
1	Parks, zoos & gardens	87.5	57.5	87.5
2	Sports facilities	87.5	57.5	87.5
3	Museums and other cultural facilities	87.5	52.5	87.5
4	Libraries	87.5	60	87.5
5	Hiking trails	0	50	50
N	Environment & Energy	76.25	61.88	76.25
1	Air quality	80	75	80
2	Water quality	85	77.5	85
3	Pollutant emissions	65	42.5	65
4	Energy consumption	75	52.5	75
O	Climate & Weather	36.88	65.63	70
1	Temperature	65	67.5	67.5
2	Wind	0	55	55
3	Rainfall distribution	0	75	75
4	Lightning locations	82.5	65	82.5
P	Science & Technology	69	59.5	75
1	Research & development supports (funds, tax cut, etc.)	67.5	65	67.5
2	Wifi spots	87.5	47.5	87.5
3	Other telecommunications (broadband penetration, mobile usage, etc)	72.5	52.5	72.5
4	Patents	35	65	65
5	Granted projects and research reports (including UGC and PICO)	82.5	67.5	82.5

II.   ASSESSMENT RESULTS BY INDICATORS

			Overall	data.gov.hk	Other websites
1	Available online to anyone	1.1 Can users access the dataset online without submitting requests?	3.60	4.47	4.51
		1.2 Can users access the dataset without being required to register or identify themselves?	3.71	4.50	4.62
		Total	**7.27**	**9.02**	**9.13**
2	Free of charge	**Is the dataset available free of charge?**	**7.35**	**8.93**	**9.24**
3	In Bulk/API	3.1 Is the dataset downloadable in bulk?	3.67	2.03	3.74
		3.2 Is the dataset provided with an Application Programming Interface (API) when applicable?	1.15	0.40	1.29
		Total	**4.83**	**2.43**	**5.03**
4	Open license	4.1 Is the dataset released under an open license?	3.65	2.38	4.07
		4.2 Is the open license presented in an explicit manner?	0.00	0.00	0.00
		Total	**3.65**	**2.38**	**4.07**
5	Machine-readable	5.1 Is the dataset provided in machine-readable formats?	3.63	1.05	3.68
		5.2 Is the data organised in a structured or standardised manner?	3.43	0.93	3.53
		Total	**6.75**	**1.98**	**6.90**
6	Open format	**Is the dataset provided in an open format, which can be processed with at least one non-proprietary application?**	**7.13**	**6.23**	**8.51**
7	Primary	**Is the dataset released at the finest possible level of granularity available, not in aggregate or modified form?**	**5.87**	**7.38**	**7.52**
8	Timely	**Is the dataset up-to-date?**	**6.07**	**7.73**	**8.04**
9	Metadata	9.1 Is the dataset provided with core metadata?	2.12	2.24	2.57
		9.2 Is the dataset provided with accompanying documentation describing the context?	1.56	3.27	2.66
		Total	**3.52**	**5.51**	**5.08**
10	Permanent	**Are historical copies of the dataset kept accessible at a stable online location?**	**5.18**	**3.88**	**6.38**
(11)	Identifier	**Is the dataset provided with Uniform Resources Identifiers (URIs) to denote its key elements?**	**3.01**	**3.74**	**4.15**
(12)	Human-readable	**Is the data written in plain and clear language that can be understood by the general public?**	**4.07**	**9.02**	**6.53**
		Total/Overall	**57.62**	**55.46**	**69.90**

III.   INDICATOR DESCRIPTIONS

1. Available online to anyone

Question(s)

1. Can users access the dataset online without submitting requests?

2   Can users access the dataset without being required to register or identify themselves?

Description

Data is not meaningfully open if it is not available on the internet to the widest range of users, and therefore this indicator is the premise for others. Typical barriers to availability are: 1) requirements to submit requests, and 2) mandatory registration.

Information requests by citizens may be standard procedures under the Freedom of Information legislation in most jurisdictions—though Hong Kong has yet to have this—but many open data standards and assessment tools indicate that data should be at least readily accessible online without requests. For example, the Ten Principles For Opening Up Government Information (2010) developed by the Sunlight Foundation elaborated the barrier of requests as "requirements to visit a particular office in person", or "making data accessible only via submitted forms."

Another barrier is the requirement of registration, in particular when the authorities ask potential data users to identify themselves. Such requirements "may dissuade some individuals from using open data for fear that their activity could be monitored or they may be subject to reprisals," (International Open Data Charter, 2015). Accessing data without registration is also called the "non-discriminatory" principle, because it ensures anyone can use the data without discrimination based on their ethnicity, nationality, profession, etc.

Guidance

To receive credit on this indicator, the answers to both questions must be YES, because both are cornerstones of open data. Also, given that this indicator is a precondition for others, no credit on this means other indicators will receive zero credit by default.

If the dataset is readily accessible at an online location operated or authorised by the government, the answer to Question 1.1 is YES.
If the dataset is only found on a website built by or for non-governmental actors (e.g., volunteers, NGOs, private businesses), the answer to Questions 1.1 is NO.

If the dataset is downloadable without asking users to register any information, the answer to Question 1.2 is YES. (If users have to accept the use of cookies in order to view or download the data, the answer to Question 1.2 is NO.)

2. Free of charge

Question(s)

Is the dataset available free of charge?

Description

Access to government information sometimes comes with a charge, which deters the public from using it. Governments justify the fees with a number of reasons, for example as "cost of collection, production, reproduction and dissemination" in a 2003 European directive on the re-use of public sector information (PSI). However, the Ten Principles For Opening Up Government Information (2010) argues that "the existence of user fees has little to no effect on whether the government gathers the data in the first place." A study in 2017 commissioned by the EU to review the directive also suggested "the trend towards zero charges should be strengthened. "

Most of the open data standard initiatives uphold the 'free of charge' principle, though their bases vary and some subordinate the principle to others. For example, some regard fees as a restriction to access along with requirements of requests or registration, some interpret 'free of charge' as a built-in element of the open license principle, whereas some others single it out as an independent criterion. Given that the 'free of charge' principle has been recognised and adopted extensively, we choose to let the indicator stand alone for the assessment purpose.

Guidance
If users have to pay a fee to access the dataset, even if it is *de minimus*, then the answer to Question 2 is NO.

3. Downloadable in Bulk/API

Question(s)
1. Is the dataset downloadable in bulk?

2. Is the dataset provided with an Application Programming Interface (API) when applicable?

Description
Providing bulk access and an API are two common methods to open up datasets for the public, though their advantages and disadvantages depend on the circumstances.

Bulk access refers to putting all of the data into a file or a set of files, so that all of the data can be acquired with a few simple downloads (Joshua Tauberer, 2014). Compared to searching the database by query languages like SQL or accessing via API, bulk access is easier to use with fewer restrictions. People do not have to learn database programming. Open Data Handbook (2015) listed three advantages of bulk access: 1) data is still available when there is a restructure of data provider; 2) anyone can obtain a copy and redistribute it; 3) others can develop their own services using the data beyond its original purposes.

API is a technical contract that commits the system to working in a certain way. For open data, an API allows users to get some slice of the data by placing specific queries, and makes it possible for programmers to automate the data access process (Tauberer, 2014). When data changes in real time, like weather temperatures and traffic, the automation potential of API makes it a better way to release data than in bulk.

Both the book Open Government Data (2014) by Joshua Tauberer and Open Data Handbook (2015) agreed that publishing data in bulk should be the primary concern for all open data initiatives, because API must do everything bulk access does first. Tauberer argued that API might violate the 'non-discriminatory' principle because they often require registration first.

Therefore, for the purpose of assessment, the bulk access requirement will only apply to all datasets whereas API will apply to a few.

Guidance
The answer of all datasets must be YES to Questions 3.1 to receive credit.
(Datasets that are required to answer to Question 3.2 are: Transport, Climate & Weather.)

Regarding Question 3.1, note that publishing data is an ongoing process, so downloads "might be organised by month or year for time-series data, or broken down into sub-files for very large datasets" (Open Data Barometer, 2017). As long as the sub datasets are logically linked and they can be easily retrieved to become a whole set (it depends, but putting together dozens or hundreds of files is not 'easy'), the answer should be YES.

If a list of quarterly statistics are published on one single web page, the answer to Question 3.1 is YES.
If you have to visit school websites individually to put together a dataset of school contacts, you should answer NO to Question 3.1.

4. Open license

Question(s)
1. Is the dataset released under an open license?

2. Is the open license presented in an explicit manner?

Description
A published work is subject to intellectual property rights, for example, you must seek the author's permission to republish his/her articles, and in some cases you need to pay. An open license is one which grants permission to access, reuse and redistribute a work with few or no restrictions (Open Definition, 2005). Almost every established standard agrees to waive some of the major copyright restrictions to open data. In addition, many require that the open license must be 'explicit' to users so that they can easily find and understand the conditions.

Open Definition provides a number of conditions that an open license should satisfy and lists conformant licenses for data publishers to choose from. One of the most popular licensing systems in the world is Creative Commons (CC), founded in 2001 in the US and introduced to Hong Kong in 2008. CC features a set of visual and constantly updated labels that help users quickly understand their rights to using the licensed works. Many influential organisations including Wikipedia adopt one of several CC licenses when they share their works online and plan to inform the public explicitly.

Guidance
To receive credit on this indicator, the answers to both Question 4.1 and 4.2 must be YES.

If there are words on the webpage of the dataset clearly stating that it is released under an open license (e.g., Creative Commons), the answers to both Question 4.1 and 4.2 is YES.

If the data publisher's permission to reuse is only found in its 'terms and conditions', and only in legal language instead of plain language, the answer to Question 4.1 is YES, to Question 4.2 is NO.

5. Machine-readable

Question(s)
1. Is the dataset provided in machine-readable formats?

2. Is the data organised in a structured or standardised manner?

Description
People nowadays cannot talk about open data without mentioning machine-readability, a term referring to formats readily processable by a computer (Open Definition, 2005). The principle is also called 'machine-processable'. Joshua Tauberer, an American open government activist who founded GovTrack in 2005—the first website making US congressional documents available in open and structured formats—elaborated the importance of this principle, "as the sizes of data sets grow, the most interesting, informative, or innovative applications of government data require the use of a computer to search, sort, or transform it into a new form. Most broadly, this is analysis." (Open Government Data: The Book, 2014)

A machine-readable document must fulfill two criteria. First of all, the document format must be one that is 'readable' by a computer. For example, image formats such as jpg and gif, or scanned copies in pdf format do not meet this criterion. Secondly, data in the document must be structured or standardised. Xml is a typical format for machine-readable documents, but simply transforming a press release from pdf to xml format does not necessarily facilitate an analysis by computer. Instead, a matrix of numbers with clearly defined column and row titles are more meaningful to a machine. A standardised data format is a series of guidelines that define the way in which data should be collected or recorded, supporting compatibility and interoperability between datasets (Open Data Barometer, 2017).

Guidance
To receive credit on this indicator, the answers to both Question 5.1 and 5.2 must be YES.

If the dataset is provided in xml, csv or json format, the answer to Questions 5.1 is YES.

6. Open format

Question(s)
Is the dataset provided in an open format, which can be processed with at least one non-proprietary application?

Description
According to the Open Definition, open format is "one which places no restrictions, monetary or otherwise, upon its use and can be fully processed with at least one free/libre/open-source software tool." At the core of the definition is that no one should own exclusive right to the form, and so it is also called 'non-proprietary'. Proprietary formats are troublesome for open data because data is not open if it is not open to all (Tauberer, 2014).

Typical proprietary formats are those developed for commercial softwares and you properly have to pay for them, for example, xls and doc by Microsoft, and pdf by Adobe. To conform to the open format principle, you can substitute xls with csv, doc/pdf with odf/xml.

Guidance
Common open/non-proprietary formats are: csv, xml, html, png, odf.
Common proprietary formats are: doc, xls, pdf.

7. Primary

Question(s)
Is the dataset released at the finest possible level of granularity available, not in aggregate or modified forms?

Description
This indicator advocates that data should be collected from primary source and published in its original, unmodified form without aggregation. A critical value of open data is to increase government transparency and hold them accountable, so it must allow the public to carry out analysis based on raw data instead of second-hand information processed or screened by governments. Unless for privacy or security concerns, data should not be aggregated.

In the circumstances that aggregation is inevitable (e.g., census), the data should be disaggregated to the lowest level possible. The disaggregation can be by gender, age, income and other categories (International Open Data Charter, 2015).

Guidance
For the purpose of assessment, datasets that do not contain private information should not be aggregated. Non-private datasets are: public finance, legislature, business registration, land, housing, transport, education (schools), recreation & culture, climate & weather, science & technology.

For datasets that may concern private information, researchers should use their best judgement to decide if they are 'at the finest possible level of granularity'.

8. Timely

Question(s)
Is the dataset up-to-date?

Description
The principle of timeliness is two-layered: 1) data should be released as quickly as it is gathered and collected; 2) the dataset should be regularly updated. The purpose is to preserve the value of data, as elaborated in the International Open Data Charter (2015), "Effective and timely access to data helps individuals and organizations develop new insights and innovative ideas that can generate social and economic benefits, improving the lives of people around the world."

There is no universal standard to measure timeliness because it depends on the nature of the data: public transport data should be ideally real-time, economic performance is announced monthly, whereas the census is conducted every few years. Open Data Barometer provided a list of update frequencies to review 16 categories of datasets, but it is more a reference and the researchers were encouraged to use their best judgement to provide rationale and justifications. So it is for this assessment.

Guidance

1	Census & Statistics	Yearly
2	Public Finance	Yearly
3	Legislature & Advisory Bodies	After each new election/ legislation
4	Government Operations	Yearly
5	Justice & Safety	Yearly
6	Business Registration	Monthly
7	Land	Monthly
8	Housing	Monthly
9	Transport	Real-time
10	Welfare	Yearly
11	Education	Yearly
12	Health	Yearly
14	Environment & Energy	Monthly/Real-time
15	Climate & Weather	Real-time
16	Science & Technology	Monthly

9. Metadata

Question(s)
1. Is the dataset provided with core metadata ?

2. Is the dataset provided with accompanying documentation describing the context?

Description
Metadata was originally used in the catalogues of libraries to enable users to find books. For the purpose of open data, metadata provides several kinds of information that defines and explains a dataset, and so users can easily find a specific category of data by searching on the internet or within the data portal.

A typical kind of metadata is core metadata, which provides important, fundamental information about a dataset. According to the International Open Data Charter (2015), core metadata elements may include the dataset title, source, publication date, and format, among others that support discoverability.

A few open data standards also advocate to accompany datasets with documentation to provide context, so that users will understand their strengths, weaknesses, analytical limitations, and security requirements, as well as how to process the data. G8 Open Data Charter (2013) recommended a few best practices of such documentation: 1) to provide explanations about the data fields used; 2) data dictionaries to link different data; and 3) a user's guide that describes the purpose of the collection, the target audience, the characteristics of the sample, and the method of data collection.

Guidance
(A YES to either Question 9.1 or 9.2 will ensure half credit on this indicator. If both answers are YES, full credit will be granted.)

You should indicate what information is included in the core metadata (e.g., date of first publication, last updated, coverage, licence, published by).

You may also indicate whether guidance documentation is accompanying the dataset to explain the source, strengths, weaknesses, and analytical limitations of the data.

10. Permanent

Question(s)
Are historical copies of the dataset kept accessible at a stable online location?

Description
This indicator concerns archiving published datasets on the internet (not digitising the archives, which should be encouraged but is another topic). In the context of open data, 'permanent' means 1) retaining copies of all published datasets available online; 2) stable formats with version-tracking; and 3) stable online locations.

Permanent availability and stable formats ensure comparative analysis over time. Permanent web addresses help the public share documents with others by pointing directly to the source rather than providing instructions on how to find it (Tauberer, 2014).

Guidance
If all datasets of a category are listed on the same webpage by time, the answer to Question 10 is YES.
If a webpage only maintains the current version of dataset but provides valid links to its historical versions, the answer to Question 10 is YES.

(11) Identifier

Question(s)
Is the dataset provided with Uniform Resources Identifier (URI) to denote its key elements?

Description
Tim Berners-Lee, founder of the World Wide Web, coined the concept of 'linked data' in 2006, advocating the use of URIs to identify things and to link to each other for information sharing in an automatic manner. This principle was adopted in his later work Five Star Open Data (2010), and became an assessment criterion of Open Data Barometer in 2013.

A Uniform Resource Identifier is a string of characters identifying a resource. The most common form of URI is the Uniform Resource Locator (URL), colloquially referred to as web address. For the purpose of this assessment, you can treat URI as the same as regular web addresses (e.g., http://data.gov.hk)

Guidance
If there are unique web pages with specific URIs (web addresses) provided for the dataset maintained by the authorities, the answer to Questions 11 is YES.

(12) Human-readable

Question(s)
Is the data written in plain and clear language that can be understood by the general public?

Description
Human-readability is newer principle in the open data community than others—it had not been strengthened until the International Open Data Charter (2015) placed equal emphasis on it as machine-readability. It is aligned with the 'non-discriminatory' principle (1. Available to anyone) that ensures anyone can access and use open data, regardless of whether they have a certain type of programming skill. The Open Data Standards Directory (2017), recognised it as one of the assessment metrics, among human-readable, open license, up-to-date and others.

To ensure that the data can be understood by people, both the data itself and the dataset's accompanying documentation should be written in plain and clear language.

Guidance
To examine human-readability, two elements should be considered: 1) format and 2) language. The data should be released in user-friendly formats, such as doc, xls, txt and html (They are very likely opposite to the machine-readable requirement). Plain and clear language is required and the language has to be one that is used by the majority of local people. It means in Hong Kong, it should be Chinese.

IV. Indicator and Dataset references

Indicator references

1. Available online to anyone

Questions for evaluation	References
1. Can users access the dataset online without submitting requests? 2. Can users access the dataset without being required to register or identify themselves?	**ONLINE WITHOUT REQUESTS** **Open Definition (2005-)** 1.2 Access - The work...should be downloadable via the Internet without charge **Eight Open Government Data Principles (2007)** 4. Accessible - Data is available to the widest range of users for the widest range of purposes. **Ten Principles Opening up Government Information (2010)** 4. Ease of Physical and Electronic Access - Barriers to physical access include requirements to visit a particular office in person or requirements to comply with particular procedures (such as completing forms or submitting FOIA requests). **5-Star Open Data (2010)** * available on the Web (whatever format) under an open license **Open Government Data: The Book (2012/2014)** 1. on the Internet for free - If information can be obtained only by request in person, the information is essentially unavailable to the vast majority of citizens. **G20 Anti-corruption Open Data Principles Assessment (2017)** 10. D-Datasets - Are key national anti-corruption-related datasets available online in any form? **Global Open Data Index (2013-2017)** Is the data available online without the need to register or request access to the data? **Open Data Barometer (2013-2018)** b. Is it available online from government in any form? **Internet Universality** E.2 Do government departments and local government agencies have websites which are available in all official languages and through all major browsers? **NO REGISTRATION** **Eight Open Government Data Principles (2007)** 6. Non-discriminatory - Data is available to anyone, with no requirement of registration **Ten Principles for Opening up Government Information (2010)** 6. Non-discrimination - Barriers to use of data can include registration or membership requirements. **International Open Data Charter** Accessible and Useable - without registration **Open Data Policy Guidelines (2012/2013/2014)** 10. Remove Restrictions For Accessing Information - no technical restrictions such as registration requirements **Open Government Data: The Book (2012/2014)** 6. Non-discriminatory - Data are available to anyone, with no requirement of registration. **G8 Open Data Charter (2013)** 3. Usable by All - should be done without bureaucratic or administrative barriers, such as registration requirements, which can deter people from accessing the data. **International Open Data Charter (2015)** 3. Accessible and Useable - Release data without mandatory registration, allowing users to choose to download data without being required to identify themselves. **G20 Anti-corruption Open Data Principles Assessment** 19. D-Accessibility - Does the government release the datasets without mandatory registration? **Global Open Data Index (2013-2017)** Is the data available online without the need to register or request access to the data? **OECD Survey on Open Government Data 3.0 (2016)** Q.66.b. Do visitors need to go through a registration process?

2.   Free of charge

Questions for evaluation	References
Is the dataset available free of charge?	**Open Definition (2005-)** 1.2 Access - The work...should be downloadable via the Internet without charge **Ten Principles for Opening up Government Information (2010)** 10. Usage Costs - Imposing fees for access skews the pool of who is willing (or able) to access information. **Open Data Handbook (2012/2015)** 3. Make the data available - Data should be priced at no more than a reasonable cost of reproduction, preferably as a free download from the Internet. **Open Data Policy Guidelines (2012/2013/2014)** 10 Remove Restrictions For Accessing Information - no technical restrictions such as...access fees... **Open Government Data: The Book (2012/2014)** 1. on the Internet for free **G8 Open Data Charter (2013)** 3: Usable by All - should be available free of charge in order to encourage their most widespread use. **International Open Data Charter (2015)** Accessible and Useable - Release data free of charge, under an open and unrestrictive license. **G20 Anti-corruption Open Data Principles Assessment** 18. D-Openness - Does the government release open data that is free of charge and under an open and unrestricted licence? **Global Open Data Index (2013-2017)** Is the data available free of charge? **Open Data Barometer (2013-2018)** Is the dataset available free of charge? **OECD Survey on Open Government Data 3.0 (2016)** Q.31. are there formal requirements for free of charge with an open license?

3.   In Bulk/API

Questions for evaluation	References
1. Is the dataset downloadable in bulk? 2. Is the dataset provided with an Application Programming Interface (API) when applicable?	**Open Definition (2005-)** 1.2 Access - must be provided as a whole **Ten Principles for Opening up Government Information (2010)** 4. Ease of Physical and Electronic Access - providing an interface for users to download all of the information stored in a database at once (known as "bulk" access) **Open Data Handbook** Make data available - In bulk **Open Government Data: The Book (2012/2014)** 4. Accessible - Data are available to the widest range of users for the widest range of purposes.- If the data is accessible through an interactive interface, it must also be possible to download the complete data set in raw form and in bulk through an automated process (i.e. a bulk data download) **Open Data Policy Guidelines (2012/2013/2014)** 19 Publish Bulk Data **Global Open Data Index (2013-2017)** Is the data downloadable at once? **Open Data Barometer (2013-2018)** d. Is the machine-readable data available in bulk?

4. Open license

Questions for evaluation	References
1. Is the dataset released under an open license? 2. Is the open license presented in an explicit manner?	**Open Definition (2005-)** 1.1 Open License or Status - The work must be in the public domain or provided under an open license. **Eight Open Government Data Principles (2007)** 8. License-free - Data is not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed. **Ten Principles for Opening up Government Information (2010)** 8. Licensing - Maximal openness includes clearly labeling public information as a work of the government and available without restrictions on use as part of the public domain. **5-Star Open Data (2010)** * available on the Web (whatever format) under an open license **Open Data Handbook** Apply an open license - We recommend...the licenses conformant with the Open Definition and marked as suitable for data. **Open Data Policy Guidelines (2012/2013/2014)** 11 Mandate Data Be Explicitly License-Free **Open Government Data: The Book (2012/2014)** 8. License-free - Dissemination of the data is not limited by intellectual property law such as copyright, patents, or trademarks, contractual terms, or other arbitrary restrictions. **International Open Data Charter** Accessible and Useable - Release data...under an open and unrestrictive license. **G20 Anti-corruption Open Data Principles Assessment** 5. D-Openness - Does the government release open data that is free of charge and under an open and unrestricted licence? **Open Data Standards Directory (2017-)** 1. Open License **Global Open Data Index** Is the data openly licensed/in public domain? **Open Data Barometer** f. Is the data openly licensed? **China Open Data Index** Data layer - Data standard - Open license **Internet Universality** E.4 Can individuals and organisations use and share data which have been made publicly available

5. Machine-readable

Questions for evaluation	References
1. Is the dataset provided in machine-readable formats? 2. Is the data organised in a structured or standardised manner?	**Open Definition** 1.3 Machine Readability - The work must be provided in a form readily processable by a computer and where the individual elements of the work can be easily accessed and modified. **Eight Open Government Data Principles (2007)** 5. Machine processable - Data is reasonably structured to allow automated processing. **Ten Principles for Opening up Government Information (2010)** 5. Machine readability - information should be stored in widely-used file formats that easily lend themselves to machine processing. **Open Data Handbook (2012/2015)** Make Data Available in an open, machine-readable format. **Open Data Policy Guidelines (2012/2013/2014)** 8 Mandate Data Formats For Maximal Technical Access - in machine-readable formats, that are structured (or machine-processable) appropriately. **Open Government Data: The Book (2012/2014)** 5. Analyzable - the most interesting, informative, or innovative applications of government data require the use of a computer to search, sort, or transform it into a new form. Most broadly, this is analysis. **International Open Data Charter** 3. Accessible and Useable - machine-readable and human-readable **G20 Anti-corruption Open Data Principles Assessment** 4. D-Formats - Does the government release datasets in machine-readable and reusable formats? **Open Data Standards Directory (2017-)** 7. Machine Readable **Global Open Data Index** Is the data in open and machine-readable file formats? **Open Data Barometer** c. Is the dataset provided in machine-readable formats? **OECD Open Government Data** Q.36. are there formal requirements to provide data In machine readable format/With their associated metadata **China Open Data Index** Data layer - Data standard - Open technology - a.Machine readable format **Internet Universality** E.3 Do government and other public stakeholders provide easy online access to anonymised publicly-held data sets, including machine-readable access to original data?

6. Open format

Questions for evaluation	References
Is the dataset provided in an open format, which can be processed with at least one non-proprietary application?	**Open Definition (2015)** 1.4 Open Format - places no restrictions, monetary or otherwise, upon its use and can be fully processed with at least one free/libre/open-source software tool. **Eight Open Government Data Principles (2007)** 7. Non-proprietary - Data is available in a format over which no entity has exclusive control. **Ten Principles for Opening up Government Information (2010)** 7. Commonly owned or open Standards - Freely available alternative formats often exist by which stored data can be accessed without the need for a software license. Removing this cost makes the data available to a wider pool of potential users. **5-Star Open Data (2010)** *** available in a non-proprietary open format (e.g., CSV instead of Excel) **Open Data Handbook (2012/2015)** 3. Make the data available - In an open, machine-readable format - Re-use of data held by the public sector should not be subject to patent restrictions. **Open Data Policy Guidelines (2012/2013/2014)** 8. Mandate Data Formats For Maximal Technical Access - "open formats" refer to a rolling set of "open standards," often defined by standards organizations, that store information in a way that can be accessed by proprietary or non-proprietary software means. **Open Government Data: The Book (2012/2014)** Non-proprietary - Data are available in a format over which no entity has exclusive control. **G8 Open Data Charter (2013)** 3. Usable by All - release data in open formats wherever possible, ensuring that the data are available to the widest range of users for the widest range of purposes. **International Open Data Charter (2015)** 3. Accessible and Usable - Release data in open formats to ensure that the data is available to the widest range of users to find, access, and use. "Open formats are those which do not require proprietary software in order to be used." **China Open Data Index** Data layer - Data standard - Open technology - b. Open format

7. Primary

Questions for evaluation	References
Is the dataset released at the finest possible level of granularity available, not in aggregate or modified forms?	**Eight Open Government Data Principles (2007)** 2. Primary - Data is as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms. **Ten Principles for Opening up Government Information (2010)** 2. Primacy - Datasets released by the government should be primary source data. This includes the original information collected by the government… **Open Government Data: The Book (2012/2014)** 2. Primary - data as collected at the source, with the finest possible level of granularity, not in aggregate or modified forms. **G8 Open Data Charter (2013)** 2: Quality and Quantity - to the extent possible, data will be in their original, unmodified form and at the finest level of granularity available. **International Open Data Charter (2015)** 2. Timely and Comprehensive - To the extent possible, release data in its original, unmodified form…release data that is disaggregated to the lowest levels of administration, including disaggregation by gender, age, income, and other categories. **OECD Survey on Open Government Data 3.0 (2016)** Q.5. publish data in a disaggregated way when applicable **G20 Anti-corruption Open Data Principles Assessment** 12. D-Granularity - Does the government release datasets at the finest level of granularity available without data aggregations?

8.   Timely

Questions for evaluation	References
Is the dataset up to date?	**Eight Open Government Data Principles (2007)** 3. Timely - Data is made available as quickly as necessary to preserve the value of the data. **Ten Principles for Opening up Government Information (2010)** 3. Timeliness - Whenever feasible, information collected by the government should be released as quickly as it is gathered and collected. **Open Data Policy Guidelines (2012/2013/2014)** 22 Mandate Ongoing Data Publication And Updates - specific requirements should be put in place to make sure government data is released as close as possible to the time that it is gathered and collected. **Open Government Data: The Book (2012/2014)** 3. Timely - Data are made available as quickly as necessary to preserve the value of the data. **G8 Open Data Charter (2013)** 2: Quality and Quantity - release data as early as possible **International Open Data Charter** 2. Timely and Comprehensive - Release high-quality open data in a timely manner, without undue delay. **G20 Anti-corruption Open Data Principles Assessment** 11. D-Timeliness - Are the available datasets timely and updated? **Open Data Standards Directory (2017-)** 8. Requires Up-To-Date Data **Global Open Data Index** Data should be updated every [Time Interval]: Is the data up-to-date? **Open Data Barometer** g. Is the dataset up to date? h. Is the publication of the dataset sustainable? (Is the dataset being kept regularly updated?) **OECD Open Government Data** Q.38. are there formal requirements to provide data in a timely manner? **China Open Data Index (2017-)** Data layer - Data sustainability - updates

8. Metadata

Questions for evaluation	References
1. Is the dataset provided with core metadata ? 2. Is the dataset provided with accompanying documentation describing the context?	**Ten Principles for Opening up Government Information (2010)** 1. Completeness - Metadata that defines and explains the raw data should be included as well, along with formulas and explanations for how derived data was calculated. **Open Data Policy Guidelines (2012/2013/2014)** 13 Require Publishing Metadata - Providing a common and fully described core metadata scheme (as well as other documentation) can be useful for the public and government alike. **Open Government Data: The Book (2012/2014)** 4. Accessible - Data must be provided with sufficient documentation so that the data user understands the structure of and abbreviations in the data. **G8 Open Data Charter (2013)** 1: Open Data by Default - Each portal will include a registry file that lists all the data and metadata used on the portal. **International Open Data Charter (2015)** 4. Comparable and Interoperable - Ensure that open datasets include consistent core metadata (Core metadata elements may include the dataset title, source, publication date, and format, as well as other relevant information that describes the dataset and supports discoverability) **G20 Anti-corruption Open Data Principles Assessment** 8. D-Metadata - Does the government ensure that the datasets include consistent core metadata including at least a descriptive title, data source, publication date and available formats? **Open Data Standards Directory (2017-)** 9. Takes into Account Associated Metadata for the Dataset - Each standard should readily make available the time and date of the data's creation, the author, location of the data on the computer networks, and information about any standard applied to the raw data. **Open Data Barometer** i. Was it easy to find information about this dataset? "You will be also asked to indicate whether core metadata (descriptive dataset metadata to facilitate discovery and identification) and/or any guidance documentation is accompanying the dataset to explaining the source, strengths, weaknesses, and analytical limitations of the data." **OECD Open Government Data** Q. 36. are there formal requirements to provide data In machine readable format/With their associated metadata Q. 75.b. what do the metadata usually cover? (Purpose of the collection/Characteristics of the sample/Method of data collection/Period covered/Terminology/Contextual elements/Expected date of registering updated data/Data limitations) **China Open Data Index** Basic metadata coverage

10. Permanent

Questions for evaluation	References
Are historical copies of the dataset kept accessible at a stable online location?	**Ten Principles for Opening up Government Information (2010)** 9. Permanence - For best use by the public, information made available online should remain online, with appropriate version-tracking and archiving over time. **Open Data Policy Guidelines (2012/2013/2014)** 23 Create Permanent, Lasting Access To Data - Once released, digitized government data should remain permanently available, "findable" at a stable online location or through archives in perpetuity. **Open Government Data: The Book (2012/2014)** 9. Permanent - Data should be made available at a stable Internet location indefinitely. When data changes over time, persistence means 1) retaining copies of all published versions of the data, and 2) maintaining stability of format from version to version. **International Open Data Charter (2015)** 2. Timely and Comprehensive - Apply consistent information lifecycle management practices, and ensure historical copies of datasets are preserved, archived, and kept accessible as long as they retain value. **OECD Open Government Data** Q. 15. Are there requirements for Central/federal ministries/agencies to systematically maintain access to historical copies of datasets as long as they remain valuable for users? **China Open Data Index (2017-)** Data layer - Data sustainability - archives

(11) Identifier

Questions for evaluation	References
Is the dataset provided with Uniform Resources Identifiers (URIs) to denote its key elements	**5-Star Open Data (2010)** **** URIs to denote things (also RDF) **Open Data Policy Guidelines (2012/2013/2014)** 15 Mandate The Use Of Unique Identifiers - The use of unique identifiers within and across data sets improves the quality and accuracy of data analysis. **Open Data Barometer (2013-2018)** j. Are data identifiers provided for key elements in the dataset? **OECD Survey on Open Government Data 3.0 (2016)** Q. 75. what proportion of the data are provided Use Uniform Resource Identifiers (e.g. RDF) to denote elements **China Open Data Index (2017-)** Data layer - Data standard - Open technology - c. Fixed links (RDF)(2%)

(12) Human-readable

Questions for evaluation	References
Is the data written in plain and clear language that can be understood by the general public?	**G8 Open Data Charter (2013)** 2: Quality and Quantity - Ensure that information in the data is written in plain, clear language, so that it can be understood by all. **International Open Data Charter (2015)** 3 Accessible and Usable - providing data in multiple, standardized formats, so that it can be processed by computers and used by people 4 Comparable and Interoperable - Ensure that open datasets include consistent core metadata and are made available in human- and machine-readable formats **Open Data Standards Directory (2017-)** 6. Human Readable - Human readable requires a medium of data or information that can easily be understood by people.

Dataset references

A. Census & Statistics

Datasets for evaluation	References
1. Census (including language, population, location)	**Global Open Data Index**
	Open Data Barometer
2. Local income & Balance of payments	National Statistics
	G8 Open Data Charter
3. Trade performance	National Statistics
	Census
4. Economic performance (e.g., CPI, PPI)	Statistics-infrastructure
	Statistics-wealth
	Statistics-skills
	OECD Open Government Data
	National Statistics
	Census
	Statistics - Wealth
	Taiwan Government Open Data
	Report Census Data

B. Public Finance

Datasets for evaluation	References
1. Government budget	**Global Open Data Index**
	Budget
2. Government accounts	Spending
	Procurement
3. Call for tender	**Open Data Barometer**
4. Contracts let	Detailed budget data
	Public contracting data
	Government spend data
	G8 Open Data Charter
	OECD Open Government Data
	budget
	call for tender
	contracts let
	Transaction spend
	future tenders
	Global Development-Aid
	Taiwan Government Open Data
	Government Accountability - Budgets and Fiscal Balance
	Government Operations - Procurement
	China Open Data Index
	Taxation and finance
	Consultancy Study on Smart City Blueprint for Hong Kong
	Government budget and spending-related information (Budget and variations by B/Ds)
	Tenders and awards aggregated by office - Tenders (tender name, tender description, tender status), Awards (award title, award description, value of the award, suppliers name)
	Transactions - Individual record, Date, B/D, Description of the transaction, Name of vendor, Amount

C. Legislature & Advisory Bodies

Datasets for evaluation	References
1. Laws & Regulations	**Global Open Data Index**
	National law
2. Bills & Legislation	Draft legislation
	Election Results
3. Election results	
	Open Data Barometer
4. Legislative Council (LegCo) meetings	Legislation data
	National election results data
5. District Council meetings	**G8 Open Data Charter**
	OECD Open Government Data
	Government Accountability and Democracy-legislation and statutes
	Government Accountability and Democracy-election results
	Lobbying meetings
	Taiwan Open Data Report
	Law and Organization - Laws and Regulations
	Government Accountability - Election and Recall

D. Government Operations

Datasets for evaluation	References
1. Government contact points	**G8 Open Data Charter**
	Government Accountability and Democracy-Government contact points
2. Government structure & personnel	**OECD Open Government Data**
	Declarations of interest
3. Declarations of interest	**Taiwan Open Data Report**
	Government Structure and Personnel
4. Salaries (pay scales)	Civil Service Ethics and Accountability
	Government Operations - Basic Information of Incorporated Foundations
	National Defense and Diplomacy
	Public Inspection and Violation

E. Justice & Safety

Datasets for evaluation	References
1. Judiciary & Judgements	**G8 Open Data Charter**
	Crime and Justice-Crime statistics
2. Crimes	Crime and justice - Safety
	OECD Open Government Data
3. Emergency services (e.g., police stations, fire stations, temporary shelters)	Disaster relief data
	Risk management
	Taiwan Open Data Report
	Public Safety - Crime and Incident
4. Travel alerts	Law and Organization - Judiciary and Judgment
	China Open Data Index
	Public security
	Consultancy Study on Smart City Blueprint for Hong Kong
	Law and Security - Data related to law enforcement
	Law and Security - Crime map

F. Business Registration

Datasets for evaluation	References
1. Company register (name, unique identifier, address) 2. Licenses 3. Beneficial ownership	**Global Open Data Index** Company Register **Open Data Barometer** Company registration data **G8 Open Data Charter** Company/business register Company/business register **OECD Open Government Data** Beneficial ownership **Taiwan Open Data Report** Economic Activities - Production and Various Licenses **China Open Data Index** Organisations Business and commerce **Consultancy Study on Smart City Blueprint for Hong Kong** Company register-related information Name of company Unique identifier of the company Company address

G. Land

Datasets for evaluation	References
1. Mapping 2. Boundaries 3. Land ownership 4. Land utilisation 5. Natural features or resources (e.g., trees, rivers, streams, mountains) 6. Urban planning (e.g, zoning, enforcement cases)	**Global Open Data Index** National Maps Locations Administrative Boundaries Land Ownership **Open Data Barometer** Mapping data Land ownership data **G8 Open Data Charter** **OECD Open Government Data** Geospatial Topography - National & local maps Global Development - Land Geospatial-Topography Zipcodes/Postcodes Earth observation-agriculture Earth observation-forestry Earth observation-fishing Earth observation-hunting **China Open Data Index** Urban development and housing Agriculture and rural area **Consultancy Study on Smart City Blueprint for Hong Kong** Hong Kong map and land-related/geospatial information Markings of public/private buildings, roads, tunnels, railways, street lighting Land boundary records Stock of lands and related land status information Markings of natural features (e.g. streams, rivers, mountains) Base map of appropriate scale including administrative areas and zones (2D and 3D) Markings of trees

H. Housing

Datasets for evaluation	References
1. Stock of flats 2. Building information 3. Rental & sale transactions 4. Public housing (including waiting time and allocation status)	**G8 Open Data Charter** **OECD Open Government Data** Social mobility and welfare - Housing **Consultancy Study on Smart City Blueprint for Hong Kong** Housing price to income distribution Property-related information Stock of flats (by type, age, district) Building information (e.g. land lease expiry, use, inclusion of public open space) Usable floor areas of new buildings completed and of new building projects for which consent to commence work has been given Property rental and sale transaction including property address, transaction date, price, period of the lease, areas of the property

I. Transport

Datasets for evaluation	References
1. Public transport timetables (real-time) 2. Public transportation statistics 3. Parking lots & Charge stations 4. Traffic flow (roads, highways, tunnels, etc.)	**Global Open Data Index** Public transport timetable data **G8 Open Data Charter** Public transport timetables **OECD Open Government Data** Motor vehicle registration statistics **Taiwan Open Data Report** Public Services - Infrastructure Public Inspection and Violation **Consultancy Study on Smart City Blueprint for Hong Kong** Traffic flow (e.g. roads, highways, tunnels, real time bus services, real-time traffic status, cargo flow) Availability of parking lots and charge stations Usage of curbside loading/unloading bay Deployment status of emergency vehicles Data on public transportation (e.g. journey time, schedules and passenger capacity of buses and trains in operation) Travel related statistics in city (e.g. number of passenger entering / exiting in each MTR exit, using Octopus card for MTR travel, which can help predict mobility bottlenecks and travel time)

J. Welfare

Datasets for evaluation	References
1. Unemployment benefits 2. Elderly benefits 3. Services for the elderly 4. Services for the needy	**G8 Open Data Charter** **OECD Open Government Data** Social mobility and welfare-health insurance Social mobility and welfare-unemployment benefits Social mobility and welfare - Ageing society **China Open Data Index** Employment and social security **Consultancy Study on Smart City Blueprint for Hong Kong** Finance - Information on investment loss/gain of MPF Social Welfare - Usage and provision of selected social services

K. Education

Datasets for evaluation	References
1. List of educational institutions 2. Performance of schools (public exams, etc.) 3. Performance of higher education institutions 4. Budgets of educational institutions	**Open Data Barometer** Primary and secondary education performance data **G8 Open Data Charter** **OECD Open Government Data** Education-List of schools Education-performance of schools Education-digital skills **Taiwan Open Data Report** Public Services - Basic Education **Consultancy Study on Smart City Blueprint for Hong Kong** Individual primary/secondary/international school Individual primary/secondary/international school - Public examination results, Awards Number and statistics of students, classes, teachers (including academic qualifications and years of tenure) Subjects taught (and language used), Facilities, Extra-curricular activities Association with religious body Tertiary and higher-education institutions Tertiary and higher-education institutions (Number and statistics of students including international students, Tertiary and higher-education institutions - Details about teaching staff (including number, expertise/specialism, academic qualifications and years of tenure), administrative staff Tertiary and higher-education institutions - Number of intellectual properties (IPs) and academic publications Amount of research grants Tertiary and higher-education institutions - Statistics on the degree/diploma awarded (e.g. average GPA) Tertiary and higher-education institutions - Facilities Tertiary and higher-education institutions - Awards Tertiary and higher-education institutions - Association with religious body

L. Health

Datasets for evaluation	References
1. Healthcare facilities 2. Healthcare service performance 3. Hygiene inspection & Food safety 4. Healthcare practitioners (including register and declaration) 5. Diseases	**Open Data Barometer** Health sector performance data **G8 Open Data Charter** **OECD Open Government Data** Health-Performance data Health-Prescription data Global Development-food security **Taiwan Open Data Report** Public Services - Healthcare **China Open Data Index** Health **Consultancy Study on Smart City Blueprint for Hong Kong** Usage of public hospitals, clinics and other healthcare facilities Data on patients (e.g. prescribing rates, treatment options and mortality rates by disease group) Hygiene information after inspection of food premises Statistics and trends on disease incidence rate, mortality rate, and cost of treatment for different types of diseases Information on expenditure of the government on healthcare Supply of practitioners in public hospitals, clinics and other healthcare facilities

M. Recreation & Culture

Datasets for evaluation	References
1. Parks, zoos & gardens 2. Sports facilities 3. Museums and other cultural facilities 4. Libraries 5. Hiking trails	**China Open Data Index** Culture and Leisure Livelihood **Consultancy Study on Smart City Blueprint for Hong Kong** Sports and recreational facilities (e.g. availability, booking) Libraries (e.g. usage) Museums and performance venues (e.g. visits) Parks, zoos and gardens (e.g. locations)

N. Environment & Energy

Datasets for evaluation	References
1. Air quality 2. Water quality 3. Pollutant emissions 4. Energy consumption	**Global Open Data Index** Air quality Water Quality Pollutant Emissions **Open Data Barometer** National environmental statistic data **G8 Open Data Charter** **OECD Open Government Data** Global Development-extractives Energy and Environment-Pollution levels Energy and Environment-energy consumption **Taiwan Open Data Report** Public Safety - Environment **China Open Data Index** Natural resources and environment **Consultancy Study on Smart City Blueprint for Hong Kong** Air quality (e.g. particulate matter levels, sulphur oxides, nitrogen oxides, volatile organic compounds, carbon monoxide) Data on water and air quality in detail Volume of waste generated by categories (e.g. domestic vs non-domestic, food waste, construction waste) Statistics on energy consumption, e.g. percentage of renewables / non-renewable, peak usage and time Utilities consumption, e.g. water Environment - Noise level

O. Climate & Weather

Datasets for evaluation	References
1. Temperature 2. Wind 3. Rainfall distribution 4. Lightning locations	**G8 Open Data Charter** **OECD Open Government Data** Earth observation-Meteorological/weather **Consultancy Study on Smart City Blueprint for Hong Kong** Temperature and sea surface temperature Maximum and minimum temperatures (over a pre-defined period) Relative humidity Wind Visibility Lightning locations Rainfall distribution Times of sunrise, sun transit, sunset, moonrise, moon transit, moonset and tidal predictions Tropical cyclone-related information Ultra violet (UV) index

P. Science & Technology

Datasets for evaluation	References
1. Research & development supports (funds, tax cut, etc.) 2. Wifi spots 3. Other telecommunications (broadband penetration, mobile usage, etc) 4. Patents 5. Granted projects and research reports (including UGC and PICO)	**G8 Open Data Charter** **OECD Open Government Data** Science and Research-Genome data Science and Research-research and educational activity Science and Research-experiment results Transport and Infrastructure-access points broadband **China Open Data Index** Education, science and technology **Consultancy Study on Smart City Blueprint for Hong Kong** IT and Broadcasting - Information on number of mobile device in usage, bandwidth consumption