UNLOCKING THE VALUE OF OPEN DATA

OPEN DATA CONFERENCE:
SATURDAY 9 JUNE 2028
HALL I & II, LEE SHAU KEE LECTURE CENTRE
CENTENNIAL CAMPUS, HKU

ONE-DAY CONFERENCE HOSTED BY THE SOCIAL SCIENCES RESEARCH CENTRE OF THE UNIVERSITY OF HONG KONG

A conference designed to raise awareness of the social and economic value of Open Data amongst government departments, data providers, data users and other stakeholders.

Sponsored by:
General Support Fund
Innovation and Technology Commission
HKSAR Government
Knowledge Exchange Office
The University of Hong Kong

Organised by:
Social Sciences Research Centre of HKU
Open Data Hong Kong
**OPEN DATA CONFERENCE:**
Unlocking the Value of Open Data
Saturday 9 June 2028
Hall II, Lee Shau Kee Lecture Centre, Centennial Campus, HKU

**Introduction:**
The objective of the one-day conference was to bring together open data practitioners to share best practices and quick-wins regarding providing and using open data and to identify the economic and social benefits of open data for diagnosing problems and raising competitiveness.

The event was a mix of presentations, open discussions, and unconference-style topic discussions. The conference hosted by HKU was designed to launch the ODI-HK and to sensitize the communities of around 250 people that include:

- Public and private sector data owners;
- Social and commercial data users;
- Academic data owners and users.

There were 216 registrations to the one-day conference. 113 people participated in the conference (including 11 “walk-ins” who registered on-site). Amongst the participants, 36 were from Government (including 7 from semi-government organisations or statutory bodies). Others from universities and private sectors companies. More than 50 of them considered themselves experienced and familiar with open data.

**Keynote Speech:**
**The Hon. Nicholas Yang, Secretary for Innovation and Technology**
The Secretary pointed out that innovation usually comes from a small group of passionate people. “The Open Data Conference hosted by HKU is timely”, he said. “The HKSAR Government fully intends to transform Hong Kong into a world class Smart City based on “The Smart City Blueprint for Hong Kong” and Open Data is very much a key enabler for Smart City”.

[http://gia.info.gov.hk/general/201712/15/P2017121400746_274170_1_1513302218815.pdf](http://gia.info.gov.hk/general/201712/15/P2017121400746_274170_1_1513302218815.pdf)
“The Steering Committee on Innovation and Technology chaired by the Chief Executive of Hong Kong was set up to tackle cross-bureau issues. The Committee includes the Chief Secretary and Financial Secretary and it is made up of 12 Secretaries and 10 Policy Bureaus”. The Secretary said the HKSAR Government fully encourages more new and existing government data to be release to the public as Open Data through the public-sector information portal DATA.GOV.HK. Through this portal, the Government releases dataset information, such as statistics, air quality and examination results in machine readable format. The datasets have been read 1.9 billion times, he said.


The Secretary concluded that the Government is very much supporting and driving “Open Data” as a policy initiative, and would like to work with the community towards turning Hong Kong into a smarter and more liveable city.

Policy Panel:
1. Ms Raina YEUNG, Assistant Privacy Commissioner Personal Data (Legal, Policy & Research)
2. Prof. John Bacon-Shone, Director Social Sciences Research Centre, HKU
3. Prof. Andy CHIU, Tony Yen Professor of Law, Beijing Normal University
4. Mrs. Stacy Belcher Lee, University Archivist, HKU
• The Assistant Privacy Commissioner pointed out that not all data are personal data. However, personal data that are already in the public domain – such as data posted by hackers in that open are still protected by law.

• Professor Andy Chiu suggested that enacting a Freedom to Information Act – sometimes called The Government in the Sunshine Law in Hong Kong, should in principle help the Open Data movement in Hong Kong. However, Hong Kong as a Common Law jurisdiction may not follow the practice in countries and territories such as Mainland China and Taiwan.

• In any case, there will be no guarantee of free and unlimited release of Government information. There need to be procedural requirements and exemptions and they will probably be discussed when the LRC (Law Reform Commission) releases the public consultation paper on the subject. “In practice, it is not realistic or practical to set up a separate court or administrative unit to handle the procedural matters, such as: right to access information, exemptions, who owns the data and the Right to be Forgotten”, says Professor Chiu.

• Mrs. Stacy Belcher Lee, HKU Archivist pointed out that there’s nothing to access to if nothing has been archived. The archivist profession has been compared to that of a museum curator as it is still new to Hong Kong. Archivists look at records to determine if they are worth holding. Decisions to hold data or not are based on: evidence value, legal value, historical value, emotional value and intrinsic (i.e. monetary) value. “At HKU, less than 15% or records are kept after review. The records could be a letter, a signature or a document.
And archivists keep value in mind when making the record keeping decisions”, says Mrs. Belcher Lee.

**Community Presentations/Lightening Talks/Project Demos:**

1. **Victor Hung, Adjunct Associate Professor of The University of Hong Kong**
   *(Topic: Tap on the Census Data to Search for Factors Affecting Employment Opportunities Across Districts)*

   Victor Hung outlined how he used TPU (Tertiary Planning Units) data from the Census and Statistics Department to count and predict employment opportunities from various sectors through the Hong Kong Island, Kowloon and New Territories.

2. **Lyndsey Dixon, Editorial Director, APAC, Taylor & Francis**
   *(The Future of Data Sharing: A Publisher’s Perspective)*

   Ms Dixon introduced the Data Spectrum – from internal access, to named access, group-based access, public access by anyone. As publisher, Taylor & Francis encourages the author make open the data associated the paper if it does not violate human rights protection or privacy laws. And in many situations, data availability statement and data citations are highly encouraged or mandatory. However, it is the author’s choice to apply licence to data sets relating to the paper.

3. **Dr. Scott Edmunds, ODHK & GigaDB**
   *(Topic: An Open Data Users Perspective of The Lack of Open Data in Hong Kong”)*

   Scott Edmund, a data scientist by profession, outlined the eight types of Open Data as classified by Open Knowledge International (OKI) and discussed the continuous improvements led by the HKSAR Government through Open Data Portal coordinated by OGCIO. He also pointed out the difficulties faced by data users at a practical level and that more data sets and more data formats are not necessarily a good thing. He believes data owners in the private sector still do not understand Open Data and Hong Kong is still going through a transformation process. However, Hong Kong will eventual reach the goal since there is active cooperation between government units and other stakeholders.
4. **Sammy Fung, President, Open Source Hong Kong**  
*(Topic: From Developer’s Perspective: where are we in open data?)*  
As the president of Open Source Hong Kong, Mr Sammy Fung is a strong advocate of more Open Data and Open Data as value assets of Hong Kong as a Smart City. He pointed out that while some HKSAR Government Department are doing well in leading the Open Data initiative in Hong Kong, data sets such as transport data fall far behind neighbouring cities in Mainland China and Taipei. Data users and web developers still face an uphill battle in dealing with private sector data owners and some government departments are not trying help enough to help the data users to get access.

5. **Gene Soo, Founder of HK Startups and Creator of Citymapper**  
*(Topic: Citymapper as a Transport Data App)*  
Mr. Gene Soo shared his experience of how he worked with Hong Kong’s tram operator in improving their real-time data offer to the public. The Tram company eventually stopped operating their open Apps and now only releases their real-time data as Open Data to the public and to third party Apps creators. However, Gene says that not all transport operators in Hong Kong follow the best practice in releasing real-time data and fare concessions on transfers between different public transports.

6. **Edmon Chung, CEO of Dot.Asia**  
*(Topic: Open Standards, Open Data & Universal Acceptance: Open Works at DotAsia)*  
Mr. Edmon Chung links Open Data to Smart Cities and Open Cities. He points out Open Data is part of the soft infrastructure of modern cities and why Hong Kong should be at the forefront of this movement practiced by the leading cities around the world. He also describes the work of “Dot.Asia”, a domain name he founded and how Dot.Asia (a non-profit organisation) is involved in promoting Open Data in Asia.
7. **Guy Freeman of [https://dataguru.hk](https://dataguru.hk)**
   Guy Freeman highlighted his website which gathers some data and information not found on OGCIO data portal. Mr. Freeman points out that his data source is not necessarily a comprehensive data source.

8. **Jesse Xiao, Data Architect, GigaDB**
   Mr. Jesse Xiao, highlighted the work of work and data architecture of GigaGB. GigaScience and GigiDB is an online, open-access journal and database published in collaboration between BGI and Oxford University Press. GigaDB is an attempt to meet the needs of a new generation of biological and biomedical researches as the science of genomics enters the era of “big-data”

9. **Just Tang, Open Data Initiative, SSRC-HKU**
   Ms. Just Tang shared that a private discussion among a group of Chief Information Officers (CIO) of big companies in Hong Kong revealed that their knowledge of Open Data is very low. Their current perception is there is no money in Open Data so they do not spend time on it. They further commented that the HK government has not put the necessary policies in place so as to create a sufficient environment for Open Data to start. On the other hand, a Corporate Development Manager of one of the property developers commented that under the following three conditions, his employer would consider sharing data with other companies:
   a) If other competitors also share data, we will share data so as to reduce cost on collecting competitive intelligence;
   b) If Open Data can empower our colleagues to innovate and achieve return on investment;
   c) If donating some of our data will yield Corporate Social Responsibility outcomes to do good to the society.
Afternoon Discussion – Beginner’s Track:

1. **Just Tang, Open Data Initiative, SSRC-HKU**
   *(Topic: What type of Open Data would you like to have?)*
   Ms. Just Tang asked the general audience what type of Open Data they would like to have and found that the awareness of the subject is low. The audience is generally unaware of the use and benefits of releasing government-generated data as Open Data. Ms. Tang went on to describe the work of Open Data Initiative of the Social Sciences Research Centre of The University of Hong Kong and the objectives of the project.

2. **Bastien Douglas, Open Data Hong Kong & Open Data Initiative, SSRC-HKU**
   *(Topic: What is Open Data)*
   Mr. Bastien Douglas is the chair of Open Data Hong Kong and an active promoter of Open Data in Hong Kong. He pointed out that Open Data is the foundation and platform for development and open dialogue between government and the civil society. He stated that free and open access to Open Data is about: “Searchable/Findable”, “Machine-usuable/app-makeable/Excel Usable”, and use “Without Restriction/Limitations”. He highlighted the initiative led by former Financial Secretary John Tsang and current Chief Executive Carrie Lam in helping to drive the “Open Data” initiative and highlight best practices from around the world, including examples from the Canada and Taiwan.

3. **Professor John Bacon-Shone, Open Data Initiative, SSRC-HKU**
   *(Topic: Managing Research Data)*
Professor John Bacon-Shone, statistician and a data user at professional level, highlighted his experience in trying to release openly the research report and dataset funded by public source. Professor Bacon-Shone is a strong believer in Open Data and believes that public funded researches should normally be released to the public as Open Data by default as this will bring more benefits to the community rather than being locked up in government files.

4. Dr. Scott Edmunds, ODHK & GigaDB
*(Topic: FAIR data for research and beyond)*
Dr. Scott Edmunds explains what is “FAIR” in Open Data – “Findable”, “Accessible”, “Interoperable” and “Re-usable”. To some NGOs (such as Justice Centre), government’s response to request for information takes and long time and information is not released for unclear reasons. And some government units are not covered by Code on Access to Information.

5. Just Tang, Open Data Initiative, SSRC-HKU
*(Topic: How overseas government tackle legal & political issues of Open Data)*
Ms. Just Tang quoted Mr. Emil Chan on how New York City, Guiyang (Guizhou, China), Taiwan, London established their Open Data initiatives, which were all driven by the political will of the political leaders to either win votes or develop the economy of the cities quickly. The takeaway from the presentation is Hong Kong must catch up with the rest of the world. Ms. Tang also commented that the OGCIO team have been moving forward with Open Data in the HKSARG. However, it is a big challenge to transfer the knowledge and data know how from OGCIO to the person-in-charge of Open Data in various government units.

**Afternoon Discussion – Intermediate Track “Table Discussions” (Unconference)**
The unconference had 4 pre-established presentation topics scheduled.

The table discussions varied as much as the topics. Some had a lot of group participation, some were led by a coordinator. Some discussions were fiery, some were more guided. Some discussions were packed, some had a few around a table engaged in more focused discussion. Most were in English, but also some were in Cantonese. This was all intended. Participants knew to move around if the topic didn’t interest them or wasn’t to their liking. Some groups even wanted to keep at their session topic into the following session. Many connections seem to have been made, and the sessions lasted right until the closing time of the conference.

The schedule of topics were the following (each round was 45 minutes):

- Round 1:
  - Transport data
  - Success of Open Data abroad
  - Environment data
  - Privacy & Control vs. Open Data
- Round 2:
Challenges of Open data for research data
Web analytics and netizen behaviour
Developing with Open Data: Challenges and pitfalls

Round 3:
Challenges of Citizen access of Open Data
Access to Info and Archives Law
Working with lack of transit data

Some photos of the group discussions:
Participants would like to get hold of a wide range of data sets as Open Data. This list of wanted data sets gathered at the opening of the Intermediate session:

1) Transport data
   - bus data:
     - live bus location
     - routes
   - MTR data
   - Travel patterns
     - cross-border
     - maritime (cross-border)
1) Traffic Accidents data

2) Employment data
   o Job Well-being
   o Women in STEM
   o Salaries

3) Public Health data
   o Flu

4) Crime data
   o offence & conviction rate
   o geo-coded

5) Lands data

6) Government expenditure data
   o presentation
   o vs budget

7) Construction data
   o materials, delays, budgets, and costs

8) Opinion Polls data
   o civil liberties relation to information

9) 1823 Complaints data

10) Macroeconomics data
    o Forecasting and planning

11) Biodiversity data

12) Property Data
    o Features which determine a price
    o Rental rates
13) **Power Consumption**

14) **Company Registry data**  
   o **historical directorship**  
   o **reverse look-up**

15) **HKIA Vending Machines data**

16) **Weather - more detailed data**

17) **Supply Chain Data - how does Amazon plan delivery success**

18) **Financing of Mortgages**

**META WISHLIST**

19) **Data Security**

20) **Web Analytics (and how it’s predictive of behaviour)**

21) **HKO API Design**

22) **OD Data Usage - incl. from a privacy perspective**

23) **Traffic Accident Apps**

24) **Census Ordinance Review (keep the raw responses)**
APPENDIX II

Feedback from the participants
(Beginner’s Session)

There were some useful questions from the beginner’s session, for example:

1. What is the relationship of the following elements? Open data, big data, common spatial data infrastructure and smart cities?

Other questions from participants in the beginner’s session generally come from a personal viewpoint, for example participants would like to know:

2. Salaries levels, not just a general average value by gender/sector, but also by position (job title), by company, to help people find out if they are underpaid working hours, like above.

3. Company/organization’s income/revenue/no. of employees. A great amount of companies tell their employees there is not enough resources to hire new employee, and force the employees to work overtime without getting over-time pay. With reference to company revenue, average salaries for similar positions, people could find out if they get what they deserve for their hard work, without being exploited by employers.

4. Health information;
5. Medical information;
6. Travelling related data: special places to visit;
7. Distribution of restaurant categories; noise distribution;
8. Number of visitors around the target;
9. Hotel price rate;
10. Plants in HK: distribution of species; how seasonal changes is per species?