

# Open Data Open Possibilities

**Opening Up Data:  
Facilitating Reuse of Public Information while Minimising  
Negative Consequences**

**Professor John Bacon-Shone**

Director of Social Sciences Research Centre, Former Chair of Human  
Research Ethics Committee, Lead on Research Data Management,  
Lead on Knowledge Exchange, HKU

Former Chair, Subcommittee on Privacy, Law Reform Commission

Member, Audit Team for Hospital Authority, PCPD

Leader, ODI-HKU

# Open Data in Hong Kong (context)

1. Two relevant Law Reform Commission reports on Government Records & Access to Information, they will stimulate expectation in the community that:
  - All useful government information is archived
  - It is made accessible, unless there is a good reason not to
2. Many people in Hong Kong (including civil servants!) want better access to government data for commercial, social and academic purposes
3. It is essential that making that data accessible must not infringe on the data protection (privacy) law or the census ordinance.
4. HKSARG role is crucial, because they
  - a) own/collect so much data
  - b) fund so much data (through contracts and grants)
  - c) control other data through licensing (e.g. public transport)

# Hong Kong Ranking on Open Data?

## Key criteria:

- openly licensed (all legal use allowed)
- machine readable (best is XML, worst is images)
- downloadable (no need to visit offices)
- up-to-date (for data that changes)
- publicly available (for everyone)
- free of charge

## Global Open Data Index:

Taiwan (#1)

UK (#3)

USA (#11)

Singapore (#17)

HK (#24)

# HK Weaknesses and Opportunities?

## **Weaknesses:**

Administrative Boundaries, Company Registry, Locations, Government Spending and Land Ownership get 0% score - datasets do not meet the open license, free of charge and other open data criteria.

## **Opportunities:**

- Implementing best practices such as XML format (already achieved by OGCI0 in some domains)
- Convert administrative boundaries into data sets that meet Open Data standards
- Change administrative practices, such as stopping income from sales of some Government Departments, which needs coordination at a senior government level, such as: Lands Registry (e.g. building and land ownership); Company Registry (e.g. company ownership); Lands Department (e.g. maps).

- Data sharing/reuse is good in principle, because of enabling better policy decisions and better decisions about individuals
- Major constraints: privacy, ownership and security
- HK people are well aware of very different approach in Mainland China when it comes to personal data (public good always trumps private good, e.g. social credit system)

Standard response to risks of sharing information: don't share it unless precedent, even if not personal data (e.g. police data on crimes at district level) as no government records law or Freedom of Information law, only very weak Code of access.

Selected Code of access exemptions:

- Disclosure would harm or prejudice competitive or financial position or the property interests of the Government.
- Information could only be made available by unreasonable diversion of a department's resources
- Disclosure could be misleading or deprive the department or any other person of priority of publication or commercial value

- C&SD solution for census data - destroy all questionnaires after input, make statistical tables available for free and allow secure access to datasets via enclave
- This makes future historical research on individual identifiable census data records impossible, but does allow analysis of individual records without identifiers, subject to check on privacy risks of resultant analysis

- Data protection: access and correction prior to proposed adverse administrative or private decision and matching approval only if competing social need overrides privacy interests. Use limitation exemption only applies for the original purpose, but this purpose is often not explicit (e.g. Judiciary)



- HK is special because public hospitals are nearly free, so are widely used and they keep very complete electronic records for those hospitals and public health clinics including sexual hygiene clinics
- Consequences in terms of eHealth Records (opt in is all or nothing) and Big Data analysis
- Big Data enthusiasts want access to the data (assume anonymity removes risks, but ignore high re-identification risk)

# Research Data Access?

- HKU has taken the lead in HK by requiring archiving of research data for new research projects
- Sole objective is to protect research integrity and there is still no obligation to ever share data, regardless of whether there is any privacy risk or benefit from reanalysis
- Academics often naïve about personal data, e.g. assume anonymised DNA raises no privacy concerns

- Access to data which was personal data, after anonymisation, requires careful assessment of the risk of re-identification as the consequences are often irreversible
- Income is rarely a good reason to charge for access to government data as the full benefits far outweigh the relatively small income and the costs have already been sunk, but this requires willingness to change long standing processes
- LRC consultations provide an opportunity to revisit old policies and increase access and reuse of government data (whether open or necessarily restricted)
- Need to re-examine whether restricting commercial use of government data is necessary or wise as commercial use may yield great public value and FB scandal indicates academic access is not always safer

# Links to Further Information

Open Data Initiative – Hong Kong (ODI-HK)

[http://www.ssrc.hku.hk/open\\_data\\_main.php](http://www.ssrc.hku.hk/open_data_main.php)

THANK YOU

