(GIGA)^n SCIENCE

RE: Submission to Legislative Council Panel on Information Technology and Broadcasting

5th December 2018

Dear Chairperson Dr Hon Elizabeth Quat,

I would like to make a submission to the Panel on Information Technology and Broadcasting and the panel chairperson Dr Hon Elizabeth Quat. This is in response to the agenda items and request for public feedback and advice on the proposed new open data policy discussed on the 10th December 2018.

Specifically I would like to ask some questions and provide some feedback on the Data Usability part (Section 12) on the requirements.

For requirement (b): updated timely, particularly for real-time data like traffic and weather (this also relates to Guideline 3 on frequency at the end of the report)

Can I query does this mean that you will overwrite previous data, as this point should be clarified? While Hong Kong still does not have an archiving law there is nothing to stop this happening, but time series data is essential for users to look at trends and put data in context. If data.gov.hk is preserving dynamic datasets they also need to confirm that they should make a commitment to preservation of historic data, even if within stated time limits.

For requirement (c): annotated with proper data description

Is this trying to say "metadata? The international community is moving towards making data FAIR (Findable, Accessible, Interoperable, Reusable), but at a minimum has OGCIO looked at including schema.org metadata? With the launch of the new Google Dataset Search data portal, simple schema.org metadata is required to maximize its discoverability by google[1].

For requirement (d): removing unnecessary terms and conditions (e.g. restrictions on commercial use)

OGCIO should be applauded for simplifying their licensing terms and conditions, but having a custom license means still it is legally ambiguous and non-interoperable with all the other open data in the world. Has any legal advice been sought on harmonization and standardising the licensing to make it compatible with other Open Government Licenses and the Open Definition[2]? Creative Commons is the de facto global standard for open data, and other governments either use this or ensure their license is compatible under the Open Definition. All the terms and conditions they already have are likely to be compatible, and slightly changing the wording to be compatible is the easiest and cheapest win in moving Hong Kong towards true open data. See the license approval procedure from the Open Definition Advisory Council[3].

Sincerely,

Scott Edmunds, PhD
Open Data Hong Kong/GigaScience

Further Reading
1. https://www.blog.google/products/search/making-it-easier-discover-datasets/
2. https://opendefinition.org/
3. https://discuss.okfn.org/c/projects/OpenDefinition