



Samuelson-Glushko Canadian Internet Policy and Public Interest Clinic
Clinique d'intérêt public et de politique d'internet du Canada
Samuelson-Glushko

ANALYSIS OF SHARE-ALIKE OBLIGATIONS IN MUNICIPAL OPEN DATA LICENSES

October 12, 2011

Report by:

David Fewer, Director
Kent Mewhort, Law Foundation of Ontario Public Interest Articling Fellow

Université d'Ottawa Faculté de droit 57 Louis-Pasteur, Ottawa (Ontario) K1N 6N5 Canada	University of Ottawa Faculty of Law Tel.: 613-562-5800 # 2553 cippic@uottawa.ca www.cippic.ca
--	---

I. Introduction

Municipalities create and own a valuable set of data. This data ranges from bike lane maps to trash collection schedules to city financial statements. Over the past few years, many Canadian municipalities have setup “open data portals” to release this data to the public. This increases the effective use of the data by allowing citizens to retrieve, view, and re-use city information.

All city open data portals require users to agree to the terms of a license that governs their use of the openly-released data. Unfortunately, although municipal governments intend for this data to be “open” and fully reusable, most licenses presently in use have some disconcerting restrictions that reduce the ways in which citizens can make use of the data.

This report focuses on one problematic restriction: the “share-alike” obligation. Although the share-alike obligation can serve a useful purpose in some contexts, it does not fit well with the objectives of municipal open data portals. Share-alike is designed for a different kind of project with different objectives.

II. Background to Share-Alike Conditions in Open Licenses

The share-alike concept originates from free and open source software licenses, where the popular General Public License (GPL) includes this type of provision. With share-alike, any user making an improvement to a work must share the changes back to the public, with no further restrictions. The foremost purposes of share-alike are to prevent others from turning open source software projects into closed-source proprietary ones, and to help build communities surrounding projects. For example, share-alike requirements enable open source software projects to generate the critical mass of fixes, modules, and improvements necessary to guarantee the viability of the project.

Free Software Foundation, "GNU General Public License Version 3" (2007),
<<http://data.vancouver.ca/termsOfUse.htm>>.

Share-alike provisions have found their way into the open data licenses for several Canadian municipalities, including Vancouver, Edmonton, Ottawa, and Toronto. The intention of this inclusion is found in the preamble of the original Vancouver license, which other municipalities have copied. The preamble sets out the standard goal of a share-alike term to “ensure that the recipients of these datasets give back to the community the benefits they derive from these datasets.”

City of Vancouver, “Open Data Catalogue Beta v2 – Terms of Use” (2010),
<<http://data.vancouver.ca/termsOfUse.htm>>.

However, municipal data and existing data portals are not conducive to such giving “back to the community”. A share-alike provision makes no sense in the city context for several reasons:

- **Multiple communities can thrive.** The diverse datasets found on city open data portals better lend themselves to numerous different communities, not the type of single monolithic community that share-alike attempts to establish.
- **No way to share-back.** Existing city data portals do not provide mechanisms for users to share-back their data. Share-alike is meant for use in collaborative repositories, not one-way data portals.

- **No competition for the data provider.** City governments have no need to protect their open data portals against appropriation by other data providers or other businesses, as nearly all possible uses of the data have the potential to provide benefits to the public.
- **Open data can help entrepreneurs.** To maximize taxpayer value, a municipal open data license should aim to promote use of the data by everyone, including entrepreneurs and businesses. There is no reason to limit the scope of potential users.

In short, the inclusion of share-alike clauses in city licenses was arguably a mistake. This is further evidenced by the fact that most, if not all, other open data portals around the world include no such provision. Share-alike is nowhere to be found in federal open data portals such as those found in Canada, the U.S., or Australia.

III. Pros and Cons of Share-alike

a) Municipal Data Benefits Many Communities, Not One

In the original software context, share-alike aims to create a cohesive and strong community of developers. As anyone making improvements must share-back their changes to the original community, most often these downstream contributors will simply join the original creators in their endeavours. This creates a stable community of contributors for a single software project.

Creative Commons licenses also offer a share-alike option for non-software works. The most prevalent aim of these licenses is similarly to build a single community of collaborators. For example, the Arduino project is an open hardware design upon which a community of users shares add-on designs to build devices ranging from 3D printers to robots (Power of Open, p. 37). Ficlet, a project started by America Online, uses a share-alike license to foster a community of "collaborative storytelling" (Power of Open, p. 28).

Creative Commons Corporation, *The Power of Open* (2011), <<http://thepowerofopen.org>>.

However, a city open data portal operates much differently than the software context and the typical Creative Commons context. A city data license should try to foster many different communities and many different uses of the data – not aim for one single use. The datasets that a city possesses are diverse, with a diverse range of possible users and uses. If a set of users wishes to spawn off their own project, rather than share-back to an existing one, the city should encourage this additional use. This will maximize the overall potential of the data. A license without a share-alike restriction will lead to a much wider set of possible uses by different individuals, communities and businesses – a fundamental objective of municipalities launching open data portals.

Moreover, most city data portals do not even provide a mechanism for users to share-back data. If the city government wishes to integrate enhancements from users into its own data repositories, then a share-alike term could serve a purpose. However, this approach would be treading new tracks for public sector information. Most often, a government's sharing of information functions through a one-way portal to the public, not a multilateral collaboration platform. If a city plans to create a single large community of data-sharing, it should plan to dedicate significant resources to tracking, managing, and

maintaining a community repository. Otherwise, it should allow as many external communities as possible to take the data and thrive.

b) Competition Helps the City's Aims

Some projects use a share-alike license to keep the work open and try to prevent proprietary businesses from appropriating the work. For example, a share-alike clause can help ensure that a proprietary software vendor will not use open source software and, at the same time, keep improvements closed-off from the original community. The proprietary developer also cannot directly compete with the original creators using a modified, closed source version because any modifications are equally available to everyone.

However, in the context of municipal data, there is absolutely no need for the government to try to limit competition with itself. Given that the city gives away the data for free, it is highly unlikely that anyone will try to “compete” with the city by distributing data through their own data portal. Moreover, to the extent that anyone offers enhanced data to customers, a city government should promote this type of activity. When a person makes improvements to a dataset and sells it, this will benefit the person creating this business and the purchaser buying the added-value product – all at zero detrimental impact to the city. Furthermore, any type of distribution will only help the city in its public interest goal of getting the data out to the public.

c) City Data Should Be Open to Entrepreneurs and Businesses

Even if the City faces no competition for its data offering, the same cannot be said of entrepreneurs who choose to integrate city data into their own business models. Share-alike requires all users to freely share their own work on the applicable data to everyone, including competitors. This removes the competitive advantage that a business would normally enjoy. It acts as a strong disincentive for an entrepreneur or business to try to create a business model based on city data.

A share-alike clause will also persist through all downstream distributions of the data. Where businesses integrate data into their own databases, any resale of the databases will affect the competitiveness of all downstream businesses that use it.

IV. Common Misconceptions about Share-alike

Misconception #1: Without a share-alike clause, users will be able to appropriate the work without any attribution

This conception is false. Many licenses, such as the Creative Commons Attribution license, require all downstream users of licensed data to abide by terms such as attribution and waivers of liability – and all without resorting to a share-alike clause.

Open licenses, including those without any share-alike provisions, do not usually permit users to remove any obligations that the license imposes. If there is a clause requiring attribution to the original author, everyone else has an obligation under contract and/or copyright to attribute the original author. This operates irrespective of any share-alike obligation.

There are two key differences between a general obligation in an open license and a share-alike

obligation:

- A share-alike clause prevents a user from adding *further* restrictions onto a work, whereas a general obligation only requires a user to maintain *existing* obligations. With a general obligation license, a user may still introduce further restrictions, thereby permitting interoperability with many other open licenses.
- A share-alike clause often extends to all collections and datasets that include or integrate a licensed work, placing *everything else* in the collection under the same terms. A general obligation only applies to the *single original work* within a collection and to derivative works directly based off of the original.

For example, consider a cartographer who creates a set of interactive hiking trail maps spanning across Canada. If this cartographer includes an share-alike overlay of public washrooms from one single municipality, he or she must release the entirety of his or her across-Canada trail maps under the city data license.

Moreover, a share-alike term would also bar the cartographer from including any overlays from other municipalities. Most municipalities release their data under slightly different license terms, and share-alike clauses prohibit the inclusion of any other such terms.

A general attribution requirement, on the other hand, would only require the cartographer to attribute the municipality for his or her use of their component of the data. The public washroom data included in the maps would remain under the terms of the original city license and the cartographer could license his or her own work as desired.

Misconception #2: Share-alike is necessary for governments to fulfil information stewardship obligations

Municipal governments, as public bodies maintaining extensive document repositories, are guardians of public data. They are responsible for properly managing and maintaining their repositories. In both civil law and common law jurisdictions, provincial legislation often details these information stewardship obligations. For example, the *Loi sur les archives, LRQ, c. A-21.1* mandates management and administration policies for public body archives in Quebec. The *Ontario Archives and Recordkeeping Act, 2006*, SO 2006, c. 34 likewise attempts to ensure that public records are properly managed, kept and preserved.

These obligations generally apply to data that a municipality releases on an open data portal. In fact, the obligations play an important role ensuring that open data is reliable and sufficiently accurate.

However, these obligations should not extend to the users of data, nor require a share-alike clause. The many different end-uses of open data have entirely different and varying requirements for accuracy, currency, and availability. Imposing stewardship obligation to end users makes little sense.

Moreover, a simple attribution requirement -- which should require inclusion of the web address for the municipal data portal -- allows all downstream users of a dataset to find the original public source of the data. Even if a person re-releases city data under a commercial or highly restrictive license, the attribution requirement will always point end users to the city data portal, where they can always freely and openly download the original data.

V. Conclusion

The share-alike concept makes sense for some open source software projects that are built around a single community of developers. Although licensing restrictions such as share-alike reduce the ways in which people may use the work, this trade-off can be worthwhile to ensure a healthy community surrounding the original project. This goal likely inspired the inclusion of share-alike in existing city licenses. However, for public sector information, the situation is entirely different. Limiting the ways in which people can use data is antithetical to the aim of trying to maximizing overall taxpayer value.

When it comes to municipal data, share-alike is simply the wrong tool for the job. If a municipality wishes to foster communities around its open data, it should do so through the promotion of contests, hackathons, and advertising. The legal tool of share-alike will only cut-off potential members of open data communities, not promote their development.