

Submission to Inquiry into Improving Access to Victorian Public Sector Information and Data

**by Brianna Laugher (bllaugher@wikimedia.org.au)
Secretary, Wikimedia Australia
September 5, 2008**

I would like to start by thanking the Parliament of Victoria's Economic Development and Infrastructure Committee for conducting this inquiry. The Discussion Paper is very progressive and for Victoria to adopt a free content licensing approach to much or all of its public sector information would make it a leader among governments around the world, in enabling access to government information and encouraging engagement with the public.

I am coming to the discussion with the perspective of the Wikimedia community, of which Wikipedia is a part. Wikipedia is an encyclopedia project, now a top ten website, available in hundreds of languages. It is available for anyone to use and adapt freely with minimal limitations. It is created and maintained by a large community, and exists as an educational resource in a strange, new space, without backing from any academic or educational institution, and yet clearly it is useful, and striving for ever-greater coverage, depth and accuracy. While most educational resources are rooted in politically defined area-specific curricula, Wikipedia and the other Wikimedia projects have no such remit.

In the sections of the Discussion Paper regarding benefits from the release of public sector data to "the wider community" or "social benefits", I find little description that would match what the Wikimedia movement is doing in creating resources such as Wikipedia. I will describe them below, so that they may be considered as potential beneficiaries of the release of such data.

Background on the Wikimedia projects

Wikimedia Australia is a not-for-profit association incorporated in Victoria and recognised by the US-based Wikimedia Foundation as a 'local chapter'. The Wikimedia Foundation is the parent organisation of various free-content projects, currently comprising the following websites (collectively, the Wikimedia projects):

- Wikipedia (<http://www.wikipedia.org/>): encyclopedia project
- Wiktionary (<http://www.wiktionary.org/>): a dictionary project including thesauri, rhymes, translations, audio pronunciations and etymologies
- Wikibooks (<http://www.wikibooks.org/>): textbooks and manuals
- Wikisource (<http://www.wikisource.org/>): collection of "source texts" (typically older works whose copyright has expired) and their annotations and translations
- Wikinews (<http://www.wikinews.org/>): News reporting with a "neutral point of view" policy
- Wikimedia Commons (<http://commons.wikimedia.org/>): a central repository for photographs, diagrams, maps, videos, animations, music, sounds, spoken texts, and other free media
- Wikiversity (<http://www.wikiversity.org/>): learning materials (curricula) and learning communities, as well as research
- Wikiquote (<http://www.wikiquote.org/>): a repository of famous or significant quotations, proverbs, mnemonics or slogans
- Wikispecies (<http://species.wikimedia.org/>): central species database for taxonomy
- MediaWiki (<http://www.mediawiki.org/>): open source software that runs all of the above websites

All of these projects have the following characteristics:

- involve wide-scale internet-based collaboration, with low barriers to participation
- multilingual (either separate subdomains for each language, or a single all-languages-in-one domain)
- produce or improve works available under free content licenses or in the public domain (termed "Free Cultural Works").

The Definition of Free Cultural Works (<http://freedomdefined.org/Definition>), which is the reference for the Wikimedia Foundation and hence its projects, considers the following four freedoms to be essential for a work to be considered a Free Cultural Work:

1. the freedom to use the work and enjoy the benefits of using it
2. the freedom to study the work and to apply knowledge acquired from it
3. the freedom to make and redistribute copies, in whole or in part, of the information or expression
4. the freedom to make changes and improvements, and to distribute derivative works.

Licensing for Free Cultural Works

From the above information I wish to draw attention to a fine distinction. Although Wikimedia Australia and the Wikimedia Foundation are not-for-profit organisations, and the Wikimedia projects are thus non-commercial endeavours, they nonetheless require that works must be allowed to be used commercially. Hence for data and documents to be maximally useful to Wikipedia et al, they must be available under terms that allow commercial use.

The Definition of Free Cultural Works makes no distinction between "commercial" and "non-commercial" use of a work. I mention it explicitly here because several of the Creative Commons licenses include a "NC" (non-commercial) clause which prohibits using material under that license for commercial purposes.

Because the Wikimedia projects are predicated upon collaboration, it is more immediately obvious that they cannot use works that have a restriction on modification. This means material under Creative Commons licenses with a "ND" (no derivatives) clause cannot be used directly by the Wikimedia projects.

Consequently, there are two Creative Commons licenses which do make material available and useful to the Wikimedia projects:

- Creative Commons Attribution (CC-BY).
<http://creativecommons.org/licenses/by/3.0/>
- Creative Commons Attribution ShareAlike (CC-BY-SA).
<http://creativecommons.org/licenses/by-sa/3.0/>

A third acceptable option is releasing data into the public domain (which does not require Creative Commons).

Discussion paper questions

Regarding Discussion Paper **question 12**, the Wikimedia projects would benefit from material released under any license, old or new, that meets the requirements for a Free Culture License (as defined at <http://freedomdefined.org/Definition>). All Free Culture Licenses must offer licensees the following freedoms:

- The freedom to use and perform the work: The licensee must be allowed to make any use, private or public, of the work. For kinds of works where it is relevant, this freedom should include all derived uses ("related rights") such as performing or interpreting the work. There must be no exception regarding, for example, political or religious considerations.
- The freedom to study the work and apply the information: The licensee must be allowed to examine the work and to use the knowledge gained from the work in any way. The license may not, for example, restrict "reverse engineering".
- The freedom to redistribute copies: Copies may be sold, swapped or given away for free, as part of a larger work, a collection, or independently. There must be no limit on the amount of information that can be copied. There must also not be any limit on who can copy the information or on where the information can be copied.
- The freedom to distribute derivative works: In order to give everyone the ability to improve upon a work, the license must not limit the freedom to distribute a modified version (or, for physical works, a work somehow derived from the original), regardless of the intent and purpose of such modifications. However, some restrictions may be applied to protect these essential freedoms or the attribution of authors.

Permissible restrictions include:

- Attribution of authors: Attribution protects the integrity of an original work, and provides credit and recognition for authors. A license may therefore require attribution of the author or authors, provided such attribution does not impede normal use of the work. For example, it would not be acceptable for the license to require a significantly more cumbersome method of attribution when a modified version of the licensed text is distributed.
- Transmission of freedoms: The license may include a clause, often called copyleft or share-alike, which ensures that derivative works themselves remain free works. To this effect, it can for example require that all derivative works are made available under the same free license as the original.
- Protection of freedoms: The license may include clauses that strive to further

ensure that the work is a free work: for example, access to source code, or prohibition of technical measures restricting essential freedoms.

A number of licenses which are known to meet these requirements are listed at <http://freedomdefined.org/Licenses> .

Regarding **question 13**, in my opinion it makes no sense to use a license which restricts use by geographical area. Doing so will greatly limit the potential of such information to have an effect, by greatly limiting potential re-use. Such data is likely to be released in a digital format, and digital data crippled by such a restriction simply has no place in the networked world we live in today.

Terms of reference (b)

I would like to respond directly to point (b) of the terms of reference, *consider whether the use of open source and open content licensing models, including Creative Commons, would enhance the discovery, access and use of Government information*, by relating my experience as a Wikimedia editor in relation to US federal government data.

Wikimedia contributors are keen hunters of repositories of freely-licensed works. It is well known in the community that works produced by the US federal government are by default ineligible for copyright and hence in the public domain. (See http://en.wikipedia.org/wiki/Copyright_status_of_work_by_the_U.S._government)

US government agencies such as the military, navy, NASA, and the agriculture department have extensive image archives which are well-mined by Wikimedia editors -- to the extent that their photographs appear so often among our "featured" material (judged by the community as being of the highest quality), that the editors are occasionally accused of pro-US bias for displaying such content so often!

Wikimedia Commons, the media archive, currently has 1,483 "Featured Pictures", of which 126 (8.5%) are sourced from US government agency websites. By comparison just 44 (3.0%) are sourced from Flickr (<http://www.flickr.com/>), a well-known photography website that lets users indicate Creative Commons licensing. (The vast majority of other "Featured Pictures" are user-contributed.) This is a great example of how high-quality freely licensed government works will "rise to the top" via sites such as Wikipedia.