

Having worked on open data frameworks for nearly 10 years now, I was surprised to see that the proposed new data release and sharing framework makes not specific mention of copyright and license selection.

I submit that the new legislation and related implementation guidance should defer to the ""Australian Government Intellectual Property Rules"" to determine the approach type of copyright protection. These rules were instituted based on a recommendation from the Gov 2.0 Taskforce and the AusGoal framework. In particular, the legislation should mandate that license selection is a mandatory step for the official release of PSI, and that this decision should be guided by the latest version of the Australian Government Copyright Rules (managed by the Department of Communications and the Arts).

For emphasis, I would like to draw your attention to the following statement about the intent of these rules:

""Entities are now required to make licensing decisions about whether to use Creative Commons licences (or other open content licences) when publicly releasing their PSI. Use of more restrictive licensing arrangements for new material should be reserved for special circumstance only.""

Having personally implemented these rules in an agency that deals with sensitive information at its core, the real benefit of mandating this step is that it ensure there is proper consideration of the risks related to the release of information takes place at the right stage of the process. These discussions often uncover concerns about data sensitivity and inappropriate use of data that can usually be addressed through the selection of an appropriate license. While the 5 safes approach is a good framework for uncovering risks related to open PSI, it does not help in determining the appropriate safeguards. Adding a license selection as formal step in the process based on whole of government policy will only improve the quality of decisions about how to apply appropriate safeguards to data that will maximize its availability and utility without compromising on quality and privacy.

The obvious example is data that has great potential to inform policy discussions, but it would not be appropriate for this data to be commercialised/monetised. In this case, a Creative Commons Non-Commercial license would address this particular concern.

The key point is that this is a decision that needs to be made on a case by case basis for individual data sets, and this needs to be a risk based decision that takes into account the potential uses and consequences of the data set being openly shared. The Australian Government Intellectual Property Rules and the Creative Commons framework that it incorporates, addresses this well by providing a spectrum of licensing options that can address a range of different risk and data sharing scenarios.

If the statutory framework does not mandate this step, then the risk of inappropriate release will be higher as well the risk of insufficient detail being recorded about what risks and other considerations were taken into account.

My second submission is that there should be a mandated minimum metadata set that accompanies all public sector data that is released. This metadata should include basic details such as:

data owner

data source

data sensitivity/classification

data quality

date of release

license/copyright

(legal) basis for release

who approved the release

This metadata will create a ""digital watermark"" that should always travel with the data, but it is important the legislation not specify the method for creating and attaching a digital watermark to PSI data - only the minimum content and the requirement to release it with the data. It should also be mandatory to record this metadata in the releasing agency's data release register.

I look forward to seeing these important issues incorporate into the final legislation and its implementation, and would be happy to meet with you to discuss any aspects of my submission in more detail.