



Brussels, XXX
[...] (2014) XXX draft

Commission Notice

Guidelines on recommended standard licences, datasets and charging for the re-use of documents

1. PURPOSE OF THE NOTICE

Opening up public sector information (PSI) for re-use brings major socio-economic benefits. Data generated by the public sector can be used as raw material for innovative value-added services and products which boost the economy by creating new jobs and encouraging investment in data-driven sectors. They also play a role in increasing government accountability and transparency. These benefits have recently been recognised by the G8 leaders and enshrined in an Open Data Charter¹.

Yet, studies conducted on behalf of the European Commission show that industry and citizens still face difficulties in finding and re-using it. In response, in December 2011 the Commission adopted a package of measures² to overcome barriers to re-use and reduce the fragmentation of data markets. The key element was the recently adopted Directive 2013/37/EU amending Directive 2003/98/EC on the re-use of public sector information.

The amended Directive calls on the Commission to help the Member States implement the new rules by issuing guidelines on recommended standard licences, datasets and charging for the re-use of documents. The guidelines are an important element of the Commission's efforts to help the EU's economy to generate more value from data (including scientific data and "big data" from other sources than the public sector). They will also facilitate the roll-out of open data infrastructures under the Connecting Europe Facility (CEF).

In August 2013, the Commission launched an online consultation followed by a public hearing and a meeting of a Member States expert group on PSI. The aim was to gather the views of all interested parties on the scope and content of the future Commission guidelines.

The feedback received³ shows an increasing trend towards a more open and interoperable licensing system in Europe and agreement on the need for the speedy release of several high-value datasets. With regard to charging, it is clear that a wide range of approaches are in operation, but the newly introduced pricing principles were not called into question by the majority of respondents. This suggests that the PSI re-use market in Europe is still under development and that guidance on the key elements of the recently revised Directive is urgently needed if full advantage is to be taken of the commercial and non-commercial opportunities offered by the re-use of public data.

The purpose of this Commission Notice is to provide non-binding guidance on the best practices within the three subject areas of particular relevance for the re-use of public sector information in Europe.

2. GUIDELINES ON RECOMMENDED STANDARD LICENCES

Article 8(1) of the revised Directive provides that public sector bodies may allow for re-use of documents without conditions or may impose conditions, where appropriate through a licence. These conditions shall not unnecessarily restrict possibilities for

¹ <http://www.gov.uk/government/publications/open-data-charter>.

² http://europa.eu/rapid/press-release_MEMO-11-891_en.htm?locale=en

³ Final report summarising the outcome of the consultation:

<http://ec.europa.eu/digital-agenda/en/news/results-online-survey-recommended-standard-licensing-datasets-and-charging-re-use-public-sector>.

re-use and shall not be used to restrict competition. Recital 26 of Directive 2013/37/EU lists two such acceptable conditions by way of illustration: acknowledgment of source and acknowledgment of any modifications to the document. It also stipulates that licences, whenever used, should in any event place as few restrictions on re-use as possible, e.g. limiting them to an indication of source.

The revised Directive also encourages the use of standard licences, which must be available in digital format and be processed electronically (Article 8(2)). Recital 26 of the amending Directive encourages the use of open licences, which should eventually become common practice across the Union.

Thus, by stressing the need to avoid ‘unnecessarily restricting re-use’ and supporting the adoption of ‘common practice across the Union’, the Directive urges Member States’ in their licensing policies to deliver openness and interoperability.

It should be borne in mind that the Directive does not apply to documents for which third parties hold intellectual property rights. Such documents are not concerned by the present notice.

2.1. Notice or licence

Although public authorities often prefer to draft fully-fledged licences in order to retain control over their wording and updates, the Directive does not mandate the use of formal licences, but notes merely that they should be applied ‘where appropriate’. MS should consider whether in any individual instances and depending on a document, a notice could be used instead (in the form of a text, pop-up window or a hyperlink to an external website).

A simple notice (e.g. the Creative Commons public domain mark)⁴ clearly indicating legal status is specifically recommended for documents in the public domain (e.g. where IPR protection has expired or in jurisdictions where official documents are exempt from copyright protection by law).

In any case, a reference to the conditions under which re-use is allowed should appear prominently at the point of display of, or accompanying, the information.

2.2. Open licences

Several licences that comply with the principles of ‘openness’⁵ described by the Open Knowledge Foundation to promote unrestricted re-use of online content, are available on the web. They have been translated into many languages, centrally updated and already used extensively worldwide. Open standard licences, for example the most recent Creative Commons (CC) licences⁶ (version 4.0), could allow the re-use of PSI without the need to develop and update custom-made licences at national or sub-national level. Of these, the CC0 public domain dedication⁷ is of particular interest. As a legal tool that allows waiving copyright and database rights on PSI, it ensures full flexibility for re-users and reduces the complications associated with handling numerous licences, with possibly conflicting provisions. If the CC0 public domain dedication cannot be used, public sector bodies are encouraged to use open standard licences appropriate to a member state's own national intellectual property and contract law and that comply with the

⁴ <http://creativecommons.org/publicdomain/mark/1.0/>.

⁵ <http://opendefinition.org/>.

⁶ <http://creativecommons.org/licenses/>.

⁷ <http://creativecommons.org/publicdomain/zero/1.0/>.

recommended licensing provisions set out below. In the light of the said recommendations, consideration should also be given to the possibility of developing a suitable national open licence.

2.3. Recommended licensing provisions

2.3.1. Scope

This provision should define the temporal and geographical scope of the rights covered by the licensing agreement, the types of rights granted and the range of re-use allowed. In order to facilitate the creation of products and services that re-use combined content held by different public sector bodies and licenced under different open licences (issue often referred to as ‘interoperability of the licence’⁸) a generic formulation rather than detailed lists of use cases and rights would be preferable.

In order to proactively promote the re-use of the licenced material, it is advisable that the licensor grants worldwide (to the extent allowed under national law), perpetual, royalty-free, irrevocable (to the extent allowed under national law) and non-exclusive rights to use the information covered by the licence.

It is advisable that rights not covered by the licence be set out explicitly and the types of right granted (copyright, database right, and related rights) be defined broadly.

Finally, the broadest possible wording could be used to refer to what can be done with the data covered by the licence (terms such as: use, re-use, share can be further described by an indicative list of examples).

2.3.2. Attribution

Where licences are required by law and cannot be replaced by simple notices, it is advisable that they cover attribution requirements only, as any other obligations may limit licensees’ creativity or economic activity, thereby affecting the re-use potential of the documents in question.

The aim of attribution requirements is to oblige the re-user to acknowledge the source of the documents in a manner specified by the licensor (public sector body). It is recommended that (depending on the law applicable) the obligations be kept to a minimum, requiring at most:

- a) a statement identifying the source of the documents; and
- b) a link to relevant licensing information (where practicable).

2.3.3. Exemptions

Where re-usable datasets are being made available in conjunction with non-re-usable datasets (e.g. as different parts of the same document or table) it is advisable to explicitly indicate which datasets are not covered by the licence.

This provision is designed to ensure greater legal certainty for re-users and the public sector body and could be accompanied by feedback arrangements whereby users can report cases in which datasets appear to have been distributed under the licence in error or in which datasets appear to have been erroneously excluded. A disclaimer would be appropriate in such cases.

⁸ LAPSI 2.0 Licence Interoperability Report,
http://lapsi-project.eu/sites/lapsi-project.eu/files/D5_1_Licence_interoperability_Report_final.pdf

2.3.4. *Definitions*

It is advisable that the main terms of the licence (licensor, use, information, licensee, etc.) are defined concisely and as far as possible in layman's language and in line with those of the Directive and national transposing legislation.

In line with the considerations in point 2.3.1 above and in order not to undermine interoperability, it is advised that 'use' or 're-use' is defined using an indicative rather than exhaustive list of rights.

2.3.5. *Disclaimer of liability*

This provision should be used (to the extent allowed under the applicable law) to draw attention to the fact that the licensor provides the information 'as is' and assumes no responsibility for its correctness or completeness.

Where the public sector body is not in a position to guarantee the sustained supply of, and access to, the information in question, this should also be clearly stated in the licence.

2.3.6. *Consequences of non-compliance*

The consequences of non-compliance with the terms of the licence could be spelled out, in particular if they include automatic and immediate revocation of the re-user's rights.

2.3.7. *Information on licence compatibility and versioning*

This provision could be used to indicate other licences with which the licence is compatible, i.e. the information derived from different sources under different compatible licences can be re-used together as long as any of the licences is complied with.

Finally, it is important to maintain and refer to a clear licence versioning and date scheme so as to indicate updates.

2.4. **Personal data**

Useful guidance and best practice in the area of the re-use of personal data are set out in Opinion 06/2013 (on open data and public sector information) of the Article 29 Data Protection Working Party⁹ and in related documents of the European Data Protection Supervisor (EDPS)¹⁰.

Opinion 06/2013 advises strongly that, where re-usable information includes personal data, re-users be made aware of the rules on the processing of such data from the outset. This could be done by including an appropriate provision in the licence and thus, making personal data protection a contractual obligation, which could also be used to prevent the re-identification of anonymised datasets. Another option is to adopt a provision excluding personal data from the scope of open licensing altogether. Other solutions, e.g. 'smart notices'¹¹ also exist, when the public

⁹ http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp207_en.pdf.

¹⁰ EDPS Opinion of 18 April 2012 on the 'Open Data Package' of the European Commission and EDPS Comments of 22 November 2013 in response to the public consultation on the planned guidelines on recommended standard licences, datasets and charging for the reuse of public sector information, <http://edps.europa.eu>

¹¹ See footnote 8: LAPSI 2.0 Licence Interoperability Report, Recommendation No 5, p. 17.

sector body decides to allow the licencing of personal data. Such notices could be separate from the licence, stored in a permanent online location, indicate the original purpose of personal data collection and processing and serve as a reminder of the obligations with regard to EU rules on personal data protection and national law transposing these rules. Not being a part of the licence itself, the notices would not discourage the mixing of public sector information covered by different licences.

3. GUIDELINES ON DATASETS

Public sector data in certain thematic fields constitute a valuable asset for the economy and society at large. International initiatives on opening up government information (e.g. the G8 Open Data Charter¹² and the Open Government Partnership¹³) recognise this by putting the emphasis on strategic datasets identified via feedback from the public or with the help of experts.

Access to and the re-use of such datasets not only speeds up the emergence of value-added information products and services, but also encourages participatory democracy. In addition, their wider use across the administration itself leads to tangible efficiency gains in the execution of public tasks.

3.1. Categories of data – priorities for release

Inspired by the abovementioned international initiatives and guided by the preferences expressed in the open consultation, the following five thematic dataset categories can be said to be those in highest demand from re-users across the EU and could thus be given priority for being made available for re-use:

Category	Examples of datasets
1. Geospatial data	Postcodes, national and local maps (cadastral, topographic, marine, administrative boundaries, etc.)
2. Earth observation and Environment	Space and in situ data (monitoring of weather, land and water quality, energy consumption, emission levels, etc.)
3. Transport data	Public transport timetables (all modes of transport) at national, regional and local levels, road works, traffic information, etc ¹⁴ .
4. Statistics	National, regional and local statistical data with main demographic and economic indicators (GDP, age, health, unemployment, income, education, etc.)
5. Companies	Company and business registers (lists of registered companies, ownership and management data, registration identifiers, balance sheets, etc.)

Other categories may be considered ‘core’ or ‘high-value’ data, depending on the circumstances (relevance to strategic goals, market developments, social tendencies,

¹² See footnote 1.

¹³ <http://www.opengovpartnership.org/>.

¹⁴ Sector-specific rules (e.g. EU railway law) make take precedence

etc.¹⁵). It is therefore recommended that the responsible public authorities assess in advance, preferably with feedback from the relevant stakeholders, which data sets should be released as a priority. Primarily, this should involve gauging the expected impact in the three areas referred to above: innovation and business creation, government transparency and accountability, and improved administrative efficiency.

3.2. Other recommendations

In order to maximise the intended benefits of these ‘high-demand’ datasets, particular attention should be paid to ensuring their availability, quality, usability and interoperability.

However, both the supply and demand side of data re-use are subject to technical constraints which play a key role in reducing or maximising the potential value of public sector data for society and the economy.

To facilitate the use of data in the public sector while significantly increasing the value of datasets for subsequent re-use, it is recommended that datasets be:

- a) published online in their original, unmodified form to ensure timely release;
- b) published and updated at the highest possible level of granularity to ensure completeness;
- c) published and maintained at a stable location, preferably on the highest organisational level within the administration, to ensure easy access and long-term availability;
- d) published in machine-readable¹⁶ and open formats¹⁷ (CSV, JSON, XML, RDF, etc.) to enhance accessibility;
- e) described in rich metadata formats and classified according to standard vocabularies (DCAT, EUROVOC, ADMS, etc.) to facilitate searching and interoperability;
- f) accessible as data dumps (massive outputs of data) as well as through application programming interfaces (APIs) to facilitate automatic processing;
- g) accompanied by explanatory documents on the metadata and controlled vocabularies used, to promote the interoperability of databases; and
- h) subject to regular feedback from re-users (public consultations, comments box, blogs, automated reporting, etc.) to maintain quality over time and promote public involvement.

4. GUIDELINES ON CHARGING

This section refers to situations in which documents held by public sector bodies are made available for re-use against payment, provided the activities in question are covered by the Directive, i.e. where the documents were produced for a public task, taking into account the scope of the Directive as set out in Art. 1, and will be used

¹⁵ ISA report on high-value datasets can be taken as reference: http://ec.europa.eu/isa/actions/01-trusted-information-exchange/1-1action_en.htm

¹⁶ See recital 21 of Directive 2013/37/EU for a definition of ‘machine-readable format’.

¹⁷ See Art. 2(7) of the Directive.

outside the public task remit by an external re-user or by the public sector body itself¹⁸.

The policy of lowering charges has been supported by research¹⁹ and by the outcome of public consultations conducted by the Commission²⁰.

4.1. Marginal cost method

The revised Directive (Article 6(1)) lays down a principle applying to all charging for public sector data re-use in the EU, except the situations specified in Article 6(2): public sector bodies may charge no more than the marginal²¹ cost of reproducing, providing and disseminating the documents.

4.1.1. Cost items

Practice has shown that in the context of PSI re-use, the three main cost categories relate to:

- a) data production (including collection and maintenance);
- b) data distribution; and
- c) sales and marketing or the provision of value-added services.

When these categories are compared with what could be considered as marginal costs according to the Directive, it is clear that (a) and (c) go beyond reproduction, provision and dissemination. Instead, the principle of marginal cost charging fits best within the broad category of 'data distribution', which in the context of data re-use could be defined as costs directly relating to, and necessitated by, the reproduction of an additional copy of a document and making it available to the re-users.

The level of charges may vary further according to the dissemination method used (offline/online) or the format of the data (digital/non-digital).

In calculating charges, the following costs could be regarded as eligible:

- infrastructure: cost of development, software maintenance, hardware maintenance, connectivity, within the limits of what is necessary to make documents available for access and re-use;
- duplication: cost of additional copy of a DVD, USB key, SD card, etc.;
- handling: packaging material, preparation of the order;
- consultation: phone and e-mail exchanges with re-users, costs of client service;
- delivery: postage costs, including standard postage or express carriers; and
- special requests: costs of preparing and formatting data on request.

4.1.2. Calculation of charges

The Article 6(1) of the Directive does not preclude a zero-cost policy: it allows for documents to be made available for re-use free of charge. At the same time it limits any charges to the marginal costs incurred for the reproduction, provision and dissemination of documents.

¹⁸ The exact scope of the Directive is laid down in Article 1; the term 're-use' is defined in Article 2(4).

¹⁹ <https://ec.europa.eu/digital-agenda/en/news/economic-analysis-psi-impacts>.

²⁰ Commission staff working document SEC(2011) 1552 final; see footnote 3.

²¹ In economics terminology, 'marginal' refers to the difference made by one additional unit.

Where non-digital documents are disseminated physically, the charge may be calculated on the basis of all the above cost categories. In an online environment, however, total charges could be limited to the costs relating directly to the maintenance and functioning of the infrastructure (electronic database), subject to what is necessary for reproducing the documents and providing them to one more re-user. Given that average database running costs are low and falling, the figure is likely to be close to zero.

It is therefore recommended that public sector bodies regularly assess the potential costs and benefits of a zero-cost policy and a marginal cost policy, bearing in mind that charging itself comes at a cost (invoice management, monitoring and policing payments, etc.).

In conclusion, the marginal cost method may be applied to ensure recovery of expenditure relating to the additional reproduction and physical distribution of non-digital documents, but where digital documents (files) are disseminated electronically (downloaded) a zero-cost method could be recommended.

4.2. Cost recovery method

Article 6(2) sets out circumstances under which the principle of marginal cost charging will not apply to certain public sector bodies or certain categories of documents. In such cases, the Directive allows for the recovery of incurred costs ('cost recovery').

4.2.1. Cost items

The Directive stipulates that total income from supplying and allowing re-use cannot exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment.

Practice has shown that the following direct costs may be regarded as eligible:

- A) Costs relating to the creation of data
 - production: generation of data and metadata, quality-checking, encoding;
 - collection: gathering and sorting of data;
 - anonymisation: deletion, obfuscation, impoverishment of databases;
- B) Costs relating broadly to 'distribution'
 - infrastructure: development, software maintenance, hardware maintenance, media;
 - duplication: cost of additional copy of a DVD, USB key, SD card, etc.;
 - handling: packaging material, preparation of the order;
 - consultation: phone and e-mail exchanges with re-users, costs of client service;
 - delivery: postage costs, including standard postage or express carriers;
- C) Costs specific to libraries (including university libraries), museums and archives
 - preservation: data curation and storage costs;
 - rights clearance: time/effort spent identifying and obtaining permission from rights-holders.

Regarding the overhead costs, only those strictly related to the above categories may be eligible.

4.2.2. *Calculation of charges*

The Directive requires that the process of calculating charges be guided by a set of objective, transparent and verifiable criteria, but leaves the responsibility for defining and adopting these entirely with the Member States.

The first stage of cost calculation is adding up all relevant and eligible cost items. It is advisable that any income generated in the process of collecting or producing documents, e.g. from registration fees or taxes, be subtracted from the total costs incurred so as to establish the ‘net cost’ of collection, production, reproduction and dissemination²².

Fees may have to be set on the basis of the estimated potential demand for re-use over a given period (rather than an actual number of re-use requests received), as the charging limit relates to total income, which is not known at the time of the calculation.

While calculating costs per individual document or dataset would be burdensome, it is essential that a quantifiable output of public sector activities is used as a reference in order to ensure that fees are calculated on a correct and verifiable basis. This requirement is preferably met at database or catalogue level – it is recommended that such an aggregate be used as a reference in the calculation of charges.

Public sector bodies are advised to regularly conduct cost and demand assessments and adjust charges accordingly. The ‘appropriate accounting period’ referred to in the Directive can in most cases be assumed to be one year.

The calculation of total income could therefore be based on costs:

- a) falling under one of the categories in the list above (see point 4.2.1)
- b) relating to a quantifiable set of documents (e.g. database);
- c) adjusted for the amount of revenue generated during production or collection;
- d) assessed and adjusted on an annual basis; and
- e) augmented by a sum equivalent to a reasonable return on investment.

4.2.3. *Special case of libraries (including university libraries), museums and archives*

The abovementioned institutions are exempt from the obligation to apply the marginal cost method. For these institutions, the steps described under 4.2.2 remain relevant, with three important exceptions:

- (a) these institutions are not required to take into account the ‘objective, transparent and verifiable criteria’ to be laid down by the Member States; and
- (b) the calculation of total income may include two additional items: data preservation and rights clearance costs. This reflects the special role of the cultural sector, which includes a responsibility to preserve cultural heritage. The direct and indirect costs of data maintenance and storage and the costs of identifying third party rights-holders, excluding the actual cost of licensing permits, should be considered eligible.

²² For additional guidance, see the Judgment of the EFTA Court of 16 December 2013 in Case E-7/13 *Creditinfo Lánstraust hf. v Þjóðskrá Íslands og íslenska*.

- (c) when calculating a reasonable return on investment, these institutions may consider prices charged by the private sector for the re-use of identical or similar documents.

4.2.4. *Reasonable return on investment*

While the Directive does not specify what constitutes a ‘reasonable return on investment’, its main features could be outlined by reference to the reason for departing from the principle of marginal costs, which is to safeguard the normal running of public sector bodies that may face additional budgetary constraints.

The ‘return on investment’ can therefore be understood as a percentage, in addition to eligible costs, allowing for:

- a) recovery of the cost of capital; and
- b) inclusion of a real rate of return (profit)

In the case of commercial players in a comparable market, the rate of return would take account of the level of business risk. However, it is not appropriate to refer to business risk in relation to PSI re-use since the production of PSI is part of public sectors' bodies remit. The Directive requires the rate of return to be ‘reasonable’, and that could be slightly above the current cost of capital but well below the average rate of return for commercial players, which is likely to be much higher due to the higher level of risk incurred.

As the cost of capital is closely linked to credit institutions’ interest rates (themselves based on the ECB’s fixed interest rate on main refinancing operations), the ‘reasonable return on investment’ could not generally be expected to be more than 5% above the ECB’s fixed interest rate. This expectation was shared by the respondents to the Commission’s public consultation, with a mere one in ten replies indicating a rate above 5%²³. For non-eurozone Member States, the ‘reasonable return’ should be linked to the applicable fixed interest rate.

4.3. **Transparency**

The Directive (Article 7) requires that the following information be pre-established and published, online where possible and appropriate and so as to relate visually and functionally to the documents subject to re-use:

- a) applicable conditions, calculation basis and amounts of standard charges (i.e. charges that can be applied automatically to the pre-defined documents or sets of documents, not requiring a case-by-case assessment);
- b) factors to be taken into account in the calculation of non-standard charges; and
- c) requirements to generate sufficient revenue to cover a substantial part of the costs relating to the collection, production, reproduction and dissemination of documents for which charging above marginal cost is allowed under Article 6(2)(b).

In line with the results of the open consultation, public sector bodies are also encouraged to publish the amounts of revenue received through charging for the re-

²³ See p. 14 of the final report summarising the outcome of the consultation:
<http://ec.europa.eu/digital-agenda/en/news/results-online-survey-recommended-standard-licensing-datasets-and-charging-re-use-public-sector>.

use of the documents they hold. Such information should be compiled at an aggregate (database or whole institution) level and updated annually.