

# Consultation on guidelines on recommended standard licences, datasets and charging for the re-use of public sector information

The revision of the PSI Directive is one of the key actions of the Digital Agenda for Europe. Governments can stimulate content markets by making public sector information available on transparent, effective and non-discriminatory terms. As demonstrated by a number of studies analysing the sector (e.g. the 2008 Cambridge Study or the 2011 review by G. Vickery), Public data is an important source of growth of innovative online services.

The re-use of public information resources has been partly harmonised by the PSI Directive adopted in 2003 and revised in June 2013 to address the remaining barriers to re-use.

Difficulties and uncertainties surrounding compliance with the licensing and charging provisions were identified as one of the main obstacles to PSI re-use as they curb both the release of data by public bodies and the actual re-use. The revised PSI Directive therefore calls on the European Commission to assist the Member States in implementing the Directive in a consistent way by issuing guidelines on recommended standard licenses, datasets and charging for the re-use of documents. The intention is to bring about a more uniform set of practices across the EU, which are necessary for the development of cross-border information products and services based on public sector information.

The objective of the consultation is therefore to seek the views of stakeholders on specific issues to be addressed in the 3 sets of guidelines. Where appropriate, the guidelines will take into account and build on the existing experiences and good practices in the Member States where they may help to achieve the crossborder dimension of PSI re-use.

## 1. Background Information

1.1. Please specify, which category of stakeholders you belong to.

\*  
(compulsory)

- Citizen
- Public authority
- Commercial re-user
- Non-commercial re-user
- Re-use expert
- Other

1.3. Please provide your name, and where relevant the name of your organisation

\*  
(compulsory)

TechAmerica Europe; Christian Wagner

1.4. Please provide your e-mail address

\*  
(compulsory)

christian.wagner@techamerica.org

1.5. Please provide your country of residence / establishment

\*  
(compulsory)

Other

## 2. Recommended datasets

In its Communication on Open Data published in December 2011 the European Commission called on the EU Member States to formulate and implement open data policies, which includes releasing a wide range of datasets. Among the huge amounts of public data waiting to be released, some are of higher immediate value to re-users than others. The value accrues when the data is re-used in a crossborder context.

In this respect, the G8 has made a commitment to adhere to an Open Data Charter, which comes into effect on 18 June 2013 and identifies (1) datasets to be made available and discoverable for anyone to use in June 2013 (National Statistics, National Maps, National Elections and National Budgets), (2) datasets in at least two of the agreed categories to be published by the end of 2015 ( Legislation, Public Company register/registry, Environmental information and Government contracts and spending).

### 2.1. Please rank the desired consequences of characterising a dataset as a core one (1 being the most important).

	1	2	3	4	5
2.1.1. Release with highest priority (optional)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
2.1.2. Release in a machine-readable format (optional)	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.1.3. Release in an interoperable format (optional)	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.1.4. Guarantee quality of datasets (regular updates, data accuracy) (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.1.5. Other (please refer to the question below) (optional)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

2.2. Please specify another consequence of characterising a dataset as a core one. (optional) (maximum 150 characters; count: 0)

It is essential to guarantee the source / provenance of core data sets, for example by ensuring that a digital signature is attached. If it is important enough to make publication obligatory, then users must be able to verify that the data really came from the group that is publishing it.

2.3. Please rank the characteristics that should be used to identify core datasets to be released with highest priority (1 being the most important).

	1	2	3
2.3.1. High value for commercial re-use (optional)	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.3.2. High value for non-commercial re-use (optional)	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.3.3. Other (please specify in the open comment box below) (optional)	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

2.4. Please specify other characteristics of a core dataset (optional) (maximum 75 characters; count: 0)

It is often not possible to differentiate between commercial and non-commercial re-use. Therefore more emphasis should be given to the “high” “value” of a dataset meaning the tangible probabilities for re-use that would produce high and measurable long or short term ROI in the commercial or non-commercial sphere. At the same time, it should also be borne in mind that large scale data analytics are fertile ground for innovation, and it may be difficult to determine prima facie the - commercial or non-commercial - value that can be derived from a particular dataset. The lack of tangible probabilities for re-use in the immediate future does not mean that a dataset won’t yield substantial value in the longer term. Therefore the release of such datasets, even if not of the highest priority, is still very important. With these view all datasets below can be core.

2.5. Please indicate core datasets from the list below.

For the purpose of this question core datasets are to be understood as those that should be released with highest priority, in machine-readable and interoperable formats because of their particular value for improving democracy and/or encouraging innovation.

(optional)

- Companies: e.g. company/business register
- Crime and Justice, e.g. crime statistics, safety
- Earth observation, e.g. meteorological/weather, agriculture, forestry, fishing
- Environment: e.g. pollution levels, energy consumption
- Geospatial: e.g. topography, postcodes, national maps, local maps
- Education, e.g. list of schools; performance of schools
- Finance and contracts: e.g. calls for tender, future tenders, local budget, national budget

(planned and spent)

- Government Accountability: e.g. government contact points, election results, legislation and statutes, salaries (pay scales), hospitality/gifts
- Global Development: e.g. aid, food security, extractives, land
- Health: e.g. prescription data, performance data
- Statistics: e.g. national Statistics, Census, infrastructure, broadband penetration, wealth, skills
- Transportation: e.g. public transport timetables, access points
- Other datasets to be considered as core datasets
- No opinion

### 3. Licensing

In line with Art. 8 of the PSI Directive, several Member States have developed national licences for re-use of public sector data. In parallel, public sector bodies at all levels sometime resort to homegrown licensing conditions. Some crossborder solutions are offered through licenses with a supranational vocation, such as Creative Commons.

The guidelines will, where appropriate, build on the existing experience and good practices with the aim to promote interoperable conditions for crossborder re-use.

3.1. Which option do you consider as preferable for re-use of public data in default cases?

 (optional)

- Disclaimer explaining the rights of re-users
- A licence (click-wrap or negotiable)
- No conditions at all
- Other
- No opinion

3.3. The Commission favours the least restrictive re-use regime possible. Which of the following conditions would you consider as generally compliant with this feature?

 (optional)

- Obligation to acknowledge the source of data
- Obligation to acknowledge that original data has been altered
- Obligation not to distort the original meaning or message of the documents
- Liability waiver of the data source for any consequences stemming from re-use

- Every condition is restrictive
- Other
- No opinion

3.5. Are there any conditions that should be considered as 'black-listed' or 'hardcore' in **default (non-exceptional) cases** because particularly harmful to re-use?

(optional)

- Yes
- No, although some conditions should be used only in duly justified circumstances
- No, any condition should be allowed
- No opinion

The core value of re-used datasets is likely to come from data analysis often in combination with other data sources. The product of data analysis should not be restricted due to licensing limitations or other conditions.

Non-commercial licenses should be blacklisted in all circumstances in order to foster innovation.

3.8. Are there any conditions that should be considered as 'black-listed' or 'hardcore' in exceptional re-use cases because particularly harmful to re-use?

(optional)

- Yes
- No although some conditions should be used only in duly justified circumstances
- No, any condition should be allowed
- No opinion

The core value of re-used datasets is likely to come from data analysis often in combination with other data sources. The product of data analysis should not be restricted due to licensing limitations or other conditions.

Non-commercial licenses should be blacklisted in all circumstances in order to foster innovation.

3.11. Can any of the existing supranational (such as Creative Commons) or national licensing models be used as a basis for achieving interoperability at EU level?

(optional)

- Yes
- No
- No opinion

CC?

3.16. Is it justified to use different licences for commercial vs. non-commercial re-use?

(optional)

- Yes
- No
- No opinion

Comment: It is often impossible to differentiate between commercial and non-commercial re-use”

3.18. What are the necessary characteristics of an interoperable licence?

(optional)

- Machine-readability
- Common terminology
- A limited number of allowed conditions
- Common sets of conditions for specific datasets
- Other
- No opinion

3.21. Are there any additional characteristics/issues to be addressed to ensure interoperability of licensing conditions?

(optional)

## 4. Charging

Under the revised PSI Directive, charges for re-use shall in principle be limited to the marginal costs incurred for their reproduction, provision and dissemination. There are exceptions where full cost recovery is allowed. This section of the consultation deals with the practical implementation of these provisions.

### 4.1. Which cost elements can be considered when calculating the marginal cost of reproduction, provision and dissemination?

	always	until amor- tized	never	no opinion
4.1.1. Telecommunications costs (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.2. Customer service (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.3. Duplication (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.4. Software licensing (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.5. Database modification for dissemination (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.6. Hardware enhancements for dissemination (capacity, ports) (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.7. Value-added for dissemination (software enhancements, advertising) (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.8. Database development (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.9. Hardware (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.10. Data creation/collection (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.11. Data maintenance (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.12. Archiving (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.1.13. Other (please refer to the next question) (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

4.2. Please specify what other elements can be considered when calculating the marginal cost of reproduction, provision and dissemination? (optional) (maximum 100 characters; count: 0)

Marginal cost is the reproduction of an additional unit. In the digital era this cost is zero in all

cases. Some of the abovementioned items can be taken into account in the calculation of cost recovery charging.

## Open data

4.3. In its art. 6 the PSI Directive provides that: "Where charges are made, the total income from supplying and allowing re-use of documents shall not exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment". Accordingly, where such full cost recovery is allowed under the Directive, can any of the following costs be included in the calculation of fees for re-use?

	always	possibly	never	no opinion
4.3.1. Overhead costs (optional)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
4.3.2. Non-incremental database development (optional)	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.3.3. Non-incremental hardware costs (optional)	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.3.4. Data maintenance (optional)	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.3.5. Other (optional)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4.4. Where full cost recovery is allowed under the Directive, what other elements can be included in the calculation of fees for re-use? (optional) (maximum 100 characters; count: 0)

4.5. At what level should Public Sector Bodies calculate their charges for re-use? (optional)

- Individual item
- Database
- The whole organisation
- Other
- No opinion

4.7. In case of charges above marginal cost, what % above the fixed interest rate on the main refinancing operations set by the ECB (currently 0,5%) would you consider a reasonable return on investment?

(optional)

- <2%
- 2-5%

- 5-8%
- <10%
- No opinion
- Other

4.9. With regard to libraries, museums and archives, what cost elements should be considered when calculating the cost of preservation and rights clearance?

(optional) (maximum 150 characters; count: 0)

4.10. Should public bodies publish their revenues from re-use?

(optional)

- Yes
- No
- No opinion

## 5. Additional comments

5.1.

Would you have any other comments or input that you wish to give regarding the guidelines?

(optional) (maximum 200 characters; count: 0)