

Is there such a thing as free government data?

The new European PSI Directive makes “marginal cost” the default charge for government data. How to implement this principle? A consultation is ongoing

Keywords: public sector information, open data, charging/fees, European PSI directive

Authors: Federico Morando (<http://nexa.polito.it/people/fmorando>), Raimondo Iemma (<http://nexa.polito.it/people/riemma>), Simone Basso (<http://nexa.polito.it/people/sbasso>)

Contacts: {federico.morando; raimondo.iemma; simone.basso} @polito.it

The recently-amended European public sector information (PSI) directive ([Directive 2013/37/EU](#), PDF, hereinafter “the directive”) rests on the assumption that “[d]ocuments produced by public sector bodies of the Member States constitute a vast, diverse and valuable pool of resources that can benefit the knowledge economy” (recital 1).

More specifically, European policy-makers submit that “[o]pen data policies which encourage the wide availability and re-use of public sector information for private or commercial purposes, with minimal or no legal, technical or financial constraints [...] can play an important role in kick-starting the development of new services [...], stimulate economic growth and promote social engagement” (recital 3).

Therefore, to keep financial constraints on re-use as low as possible, the directive provides that, “where charges are made by public sector bodies for the re-use of documents, those charges should in principle be limited to the marginal costs”. In practice, this should imply that most (natively digital) government data are free to re-use for any (lawful) purpose.

This article provides a brief review of the of the public-sector-information pricing issues. It then discusses the terms under which the ongoing consultation on the implementation guidelines of the PSI directive addresses pricing. In particular, this article discusses the calculation criteria for marginal costs.

The simple economics of PSI charging

Digital goods hold well known features: their creation entails high fixed costs, while reproducing them is almost costless. As a consequence, charging for them is typically tricky. This issue has been thoroughly debated by economists, who perhaps got inspired by the so-called ‘marginal cost controversy’, dating 1946, which involved Ronald Coase and Harold Hotelling.

Compared with a generic digital good, PSI has two extra features. First of all, it has great potential for re-use: governments, in fact, collect and manage tremendous amounts of information, which is assumed to be complete and accurate, and which, in many cases, is the only possible source of the data that one might want to embed in a digital service (Pollock, 2009). Secondly, the PSI production has been already funded through taxation, especially when data stems as a by-product of the public task ([LAPSI position paper nr. 1, 2011](#), PDF).

Because of the two features mentioned above, and because PSI entails both supply- and demand-side economies of scale, many authors support that public agencies should make available their data with charges that allow to recoup, at most, the cost of reproducing and disseminating the PSI itself. In this

vein, when transaction costs (i.e., the costs of getting paid) overcome the marginal cost-based price, the public administration should make the PSI available free of charge.

Yet, there are also reasonable arguments against charging to recoup transaction costs only. For instance, if a public agency is not the unique holder of a specific dataset, by giving the data away for free, the agency implements predatory pricing. Moreover, the free-of-charge strategy is typically coupled with a best effort level of service, but re-users generally expect a public administration to hold and release high-quality data.

What happens in reality is that two charging approaches co-exist and are applied to different segments. Public administrations currently make available small, previously undisclosed datasets at no charge. At the same time, national agencies extract profits from licensing access to databases of high interest, e.g., firm registries or geodata.

PSI charging: brief history of a long debate

Many authors agree that the [green paper on public sector information in the information society](#) (PDF), published in 1999 by the European Commission, is the first step towards the definition of a European PSI policy. The Green Paper contained a review of the main issues at stake, including pricing. Also, the Green Paper argued that the optimal charging principles should strike a balance between allowing affordable access to everyone, fostering the exploitation potential of PSI, and ensuring fair competition.

Five years later the European Commission issued the first piece of legislation addressing PSI, the [European directive 2003/98](#) (PDF). In a nutshell, with respect to charging, the 2003/98 directive stated that PSI holders could recoup the average costs of production and dissemination, provided that their return on investment was kept at “reasonable” levels.

In 2010 the European Commission promoted a public consultation in view of the revision of the PSI directive: the vast majority of the respondents signalled that PSI reuse had not achieved its full potential in Europe; around 40% of the respondents agreed on the marginal cost principle (reproduction and dissemination) for PSI pricing, while 36% disagreed; in any case, 54% of the participants were in favour of tightening and/or making more clear charging rules.

In June 2013, the amended version of the European directive on the re-use of public sector information was issued, containing, amongst other changes, updated prescriptions on charging.

Charges in the amended EU PSI Directive

Article 6 of the amended PSI directive discusses the principles governing charging, which we summarise below.

The new default rule is that charges for the re-use of the PSI have an upper bound in the “marginal costs incurred for [the] reproduction, provision and dissemination” of government data (§1 of art. 6).

As an exception, the directive allows the public sector bodies (PSBs) to charge higher fees in cases in which they are “required [by the law or by administrative practices] to generate sufficient revenue to cover a substantial part of the costs relating to their collection, production, reproduction and dissemination” (§2). If they charge higher fees, the PSBs must set charges according to objective, transparent and verifiable criteria; moreover the total income from supplying and allowing re-use of documents must “not exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment” (§3).

Another exception to the default rule is that libraries, archives and museums (LAMs) are generally allowed to charge above marginal costs. Moreover, LAMs charges can also take into account the costs of “preservation and rights clearance” (§ 4).

The ongoing consultation: guidelines on charging

Member States are free to apply lower charges and, in particular, no charges at all. This freedom is consistent with the directive, which primarily aims at maximising the PSI re-use and its economic benefits, as well as with the principle of minimum harmonisation.

Moreover, the principle of subsidiarity imposes that the criteria for charging above marginal costs are essentially left to Member States (recital 25).

However (and as stated in recital 36), the Commission shall help the Member States to implement the directive “in a consistent way by issuing guidelines, particularly on recommended standard licences, datasets and charging for the re-use of documents, after consulting interested parties.”

In the following paragraphs we focus on the ongoing consultation envisioned by the directive and, in particular, on its fourth section, which deals with the practical implementation of charges for the re-use of the PSI. By taking advantage of the questions as spelled in the consultation, we proceed to analyse the key open issues.

The calculation of the marginal cost of PSI

To implement the directive, an operational rule to calculate the marginal cost of “reproduction, provision and dissemination” is needed.

To this end, the consultation asks the responders whether the following cost items should contribute to the calculation of marginal costs: telecommunications costs, customer service, duplication, software licensing, database modification(s) for dissemination, hardware enhancements for dissemination (capacity, ports), value-added (activities) for dissemination (software enhancements, advertising), database development(s), hardware, data creation/collection, data maintenance, and archiving.

The consultation allows the responders to choose between the four following answers: “always”, “until amortized”, “never” and “no opinion”.

The standard definition of marginal cost as the change in total costs that arises when the quantity produced is increased by one unit (i.e., the cost of producing one more unit of a good) suggests the following answers:

1. duplication costs always contribute to the marginal costs. In practice, however, in a digital environment, the duplication cost is zero (except when the original data is in analog format and must be digitised);
2. telecommunications and customer service costs could or could not be marginal costs; in principle, one should answer with the no opinion option and should use the open answer option to provide the following explanation.

First and foremost: some marginal “telecommunications costs” do exist. For instance, the cost of adding network capacity to satisfy a certain request is a marginal cost; to recoup this cost, one can, for example, implement a “capacity charge” that captures the amount of capacity consumed by a user. Similarly, the customer-service costs generated by a user can be charged to her/him, e.g., by using a premium-rate telephone number (i.e., the 900- or 199- numbers, depending on the country).

However, not all telecommunication services allow their owners to charge their users in a simple way. A premium-rate telephone number, in fact, allows its owner to bill the user that makes the phone call, but the internet lacks a money-routing protocol and lacks a per-flow charging mechanism.

Therefore, when the costs of an internet-based service are significant, a reasonable answer is *until amortized*. However, this approach does not actually contribute to the definition of the “marginal cost” of re-use, but to what can be described as the “average incremental cost” of allowing the re-use of a certain dataset;

3. the sub-set of cost items “for dissemination” (i.e., database modifications, hardware enhancements such as capacity and ports, value-added activities such as software enhancements and advertising) contributes to the average incremental cost of allowing re-use as well.

Therefore, in principle, only customised modifications, enhancements and other value-added activities requested by the N^{th} user should be charged and just to him/her.

However, *until amortized* is again a reasonable, although theoretically incorrect, answer. In practice, instead of charging on the N^{th} re-user the costs he/she generates, all re-users are charged $1/N^{\text{th}}$ of the cost of allowing re-use (or a better/easier re-use);

4. software licensing could be comprised in the previous point, however, as a matter of policy, this could and possibly should be avoided: the rationale of not allowing to recoup licensing costs is that every needed activity in this domain can be performed with open source software at no charge (and at least some Member States may want to encourage this approach);
5. database development(s), hardware, data creation/collection, data maintenance, and archiving should *never* be considered, as they are typical examples of fixed costs, which are sunk at the moment of making PSI accessible and re-usable. An additional reason not to charge these costs on PSI re-users is that doing otherwise would create an incentive for the PSBs to charge on the re-users costs, which are actually related with the overall ICT management of the PSBs themselves.

In conclusion, there are several cases in which the marginal cost approach, if strictly implemented, would imply that just the N^{th} user (or, in certain cases, the first user) would pay a high fee, with users from $N+1$ onward receiving the improved service for free, if no further marginal costs are generated (e.g., user N requires some data to be published in a new format: a conversion tool is developed and paid by him/her, while the rest of the users can get the new format for free)¹. It may, however, be reasonable to treat these cases differently, guessing the total expected number of users (and/or shifting the costs on the next fiscal year) and charging on each of them *pro rata*: this is our understanding of the *until amortized* option offered in the consultation.

Special cases: full cost recovery scenarios

Article 6 of the PSI directive provides that, where such full cost recovery is allowed, the total income from supplying and allowing re-use of PSI “shall not exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment”. Accordingly, the consultation asks which of the following costs may be included in the calculation of fees for re-use:

¹ If this approach is chosen, notice that at least one exception should apply: no charging should be made in case the customisation request actually consists in a bug fixing, because this signalling activity should be subsidised, as a matter of policy, since it generates a public good (for all re-users and possibly for the PSB itself).

overheads, non-incremental database development costs, non-incremental hardware costs, data maintenance.

Considering the generic language of the directive and its permissive rationale, all these costs could *possibly* be considered (the other available answers being *always*, *never* and *no opinion*). That said, the consultation may arguably be criticised for its lack of precision in defining costs such as "non-incremental hardware costs"².

A related question concerns a definition of a "reasonable return on investment". The consultation investigates what percentage above the fixed interest rate on the main refinancing operations set by the ECB (currently 0.5%) should be considered "reasonable".

The question is prone to any kind of guessing and the authors would put their bets on the 2-5% range. That said, from the economic point of view, one should arguably assess the "reasonableness" of the return on an investment comparing it with the average profits made on similar markets by players who are not in a monopoly position. Depending on how risky these markets are, even the highest option offered by the consultation (i.e., less than 10%) may be "reasonable", however, and considering the PSBs do not typically suffer all the economic consequences of the risks they take, comparing their rate of return with the ones of risky markets may not be appropriate. This is why, intuitively, we submit that a moderate 2-5% premium over the main refinancing rate of the ECB could be provided as a reference point.

Finally, another special case concerns libraries, museums and archives. Not only they are always free to charge more than marginal costs, but they can also recoup additional cost elements, i.e., the cost of preservation and rights clearance. The consultation asks how these costs should be calculated, but the questions appears to be partly tautological: any cost of preservation and rights clearance could arguably be recouped, possibly including the cost of digitization itself and the cost of copyright searches (e.g., to assess whether a works is in the public domain). In this regard, the most delicate point does not concern charges, but the fact that public domain material should remain in the public domain, i.e., as soon as one has a copy of a public domain piece of content, he or she should be free to use, re-use and share it as he or she sees fit. It is therefore very difficult to imagine on which basis it could be possible to charge any costs on the re-users, unless LAMs are allowed to contractually void the public domain status of these works of most of its meaning.

Final remarks and policy implications

Broadly speaking, the new European prescriptions concerning PSI charging may seem the result of an act of balance between the need of a wider and easier circulation of information, and the current budget constraints of public agencies. In practice, the radical option of marginal - and *de facto* zero - cost, attractive as it may be on paper, might tend to be nothing more than a formal default option.

In several cases, a public administration might decide not to charge at all. This is what has started to happen for all datasets made available through open government data portals, even before the issuing of the amended directive. However, in other cases, what would be actually charged according to the approach that can be inferred by the aforementioned consultation is likely the average incremental cost to allow data re-use, which is conceptually and practically different from the marginal cost of reproduction and dissemination. As discussed in the previous sections, this might result as inefficient. Finally, under the current rules, it seems quite easy to take advantage of the allowed exceptions, especially for public agencies who so far relied on income deriving from PSI dissemination. Not by

² A Google search for "non-incremental hardware costs" just returns the text of the consultation, confirming the impression that this concept is far from being a commonly understood one.

chance, databases with higher potential for commercial re-use are arguably held by those agencies, and feed mature re-use markets, usually holding strong barriers to entry.

References

- Coase, R. (1946). The Marginal Cost Controversy. *Economica, New Series*, 13(51), 169-182.
- Davies, R., (2008). Recommendation of the ePSIplus network to the EC review of the Directive on PSI re-use. *Paper presented at the first Communia conference*. Retrieved from URL <http://www.communia-project.eu/node/112>.
- Pollock, R. (2009). The Economics of Public Sector Information. Retrieved from URL http://rufuspollock.org/papers/economics_of_psi.pdf
- Ricolfi, M., Drexler, J., van Eechoud, M., Janssen, K., Maggolino, M.T., Morando, F., Sappa, C., Torremans, P., Uhler, P., Iemma, R., de Vries, M. (2011). The 'principles governing charging' for re-use of public sector information. *Informatica e Diritto*, 1-2, 105-128.
- Uhler, P. (2009). *The Socioeconomic Effects of Public Sector Information on Digital Networks: An Analysis of Different Access and Reuse Policies: Workshop Summary*. Washington, D.C.: National Academies Press.