CHARGING POLICY: A CONCEPTUAL FRAMEWORK FOR EU GUIDANCE TO THE MEMBER STATES

MARCO RICOLFI
marco.ricolfi@studiotosetto.it

IP Law; Competition Law; Public Sector Information (PSI); PSI Re-Use; Principles Governing Charges.

Abstract
This memo intends to identify and link together a number of crucial topics concerning the charging for the re-use of Public Sector Information under the Directive 2003/98/EC on the re-use of public sector information (the PSI Directive) and the draft Proposal for a Directive amending the Directive unveiled December 2011. The purpose of the document is threefold: first, to assist in exercises concerning PSI charging which might be launched in the near future, such as the preparation of a multi-level conference on the issue, aimed at bringing together scholars, stakeholders, policy makers and the general public; second, to provide the conceptual terms of reference for the setting up of research projects, thematic networks or other cooperative research initiatives in the area; and third to prepare the ground for the discussion of possible components of guidance by the Commission to the Member States in specific connection with “charging and calculation of costs” as provided by Recital 18 of the draft Proposal.

1. Methodology
This memo, while building on the state of the art of the current debate, including the LAPSI contribution to it (Discussion Paper No 1),\(^1\) intends to focus on unsettled issues and give priority to the difficulties and complexities which arise when the PSI charging policy is looked at from different perspectives.

While it is quite obvious – and commonly accepted – that PSI charging policy has both a legal and an economic dimension, it is submitted that for present purposes the discussion should deal not only with the components of the two perspectives but also on their mutual interaction; that additional perspectives should be incorporated into the debate; and finally that added value of primary importance in view of the above indicated purposes may be extracted by taking into account overarching methodological issues.

Of course, an approach such as the one advocated here has an obvious snag. It is so expansive it runs the risk of taking off the ground, that is of turning mundane, detailed questions of cost-computing and of institutional detail of charging policy into metaphysical questions. This is a possibility; but not a necessity. What is advocated here is a multi-level approach consisting of several steps:

a) the identification of a starting point consisting of precise sets of analytical questions, dealing with legal and economic perspectives (chapter 2 and 3);

b) the broadening of the inquiry in a way that each set of questions is linked both to other separate sets of questions (chapter 3) or even to broader rationales (economic theory: below

---

chapter 3, history of ideas, chapter 4) and empirical findings relevant to verify or falsify hypothetical solutions;

c) the “weighting” of the importance of the various set of questions, in accordance with the findings made while carrying out the exercise; and

d) the narrowing down (or “pruning”) the alternatives to the limited number of options which are compatible with (i) the current normative and institutional setting; and (ii) the realities of our current market and societal conditions.

e) This may be a tall order; but it is not an impossibility.

2. Legal Perspectives

The current menu of legal issues concerning PSI charging policy are well known. These include:

a) Cost ceilings; interplay of rules and exceptions (current and proposed text of Art. 6 of the PSI Directive);

b) Analysis of the (possibly special) position of specific categories of PSI Holding entities (PSIH) such as Libraries, Museums and Archives (LMAs);

c) The content of contract terms with re-users;\(^2\)

d) The institutional design within which charging decisions are taken and reviewed, including burden of proof for moving from rules to exceptions, governance levels for decision making (including consideration of the respective roles of internal audit and of external review), structure and time-frame of redress mechanism;

e) A cross-cutting issue concerns the appropriateness of one-size-fits-all rules, which in this instance could be expressed asking for any and each of all the above listed items what is the extent of flexibility and modularity envisaged.

Further insights might be gained by a closer look at several other less debated topics. These should include:

a) taking into fuller account the differences between registration and non-registration based PSI,\(^2\) including the implications of the existence of two-sided markets; and

b) The issue which may described as “back-firing”: the question whether charging policies, which may imply unrecoverable costs for PSIHs, may have a negative impact on the decision by the same PSIHs whether to make PSI accessible in the first place deserves closer scrutiny. One obvious angle is to ask whether and to which extent the proposed change to the wording of Art. 3 takes care of the problem.\(^4\) Additional side issues should however be addressed in this very connection, such as the impact of governance levels referred to above in 4.d. Also the inter-temporal framework of charging policies would appear to have traditionally been neglected; but it turns out to be a crucial factor in investment decisions by

\(^2\) See §§ 7.3.1.2 to 7.3.1.6 of LAPSI, ‘The “principles governing charging” for re-use of Public Sector Information’, quoted at note n. 1.

\(^1\) The distinction between registration and non-registration based PSI is current in the literature (see e.g. NEWBERRY, D., BENTLY, L. & POLLOCK, R 2007, Models of Public Sector information provision via Trading Funds, Study commissioned jointly by the Department for Business, Enterprise and Regulatory Reform (BERR) and HM Treasury, 53); PSI which is collected as a result of a statutory obligation by individuals or entities to supply certain information (e.g. on real estate owned; on vehicles owned; on financial data) is usually described as registration based. It is often noted that, even though the same information might theoretically be collected also by private businesses, it would not have the legal effects which usually are linked to inclusion in public registrars; so that also the notion of “sole source” information is resorted to in this same connection.

\(^4\) In this connection, one should not only consider the issue whether a Member State may deny access and re-use to already digitized documents, but also the issue whether Member States are under a duty to digitize.
would-be re-users.\textsuperscript{5} In the same vein, a list of “additional topics” should be compiled and their impact on the “current menu” should be further explored.

The PSI charging issues just referred to, which we may here for the sake of convenience lump together as “classical legal issues”, to do not come into a vacuum. To look at them we should adopt a multi-level approach, considering that their appreciation involves at a minimum taking into account competition law, tax law and the principle of subsidiarity.

It is well known that \textit{competition law} is relevant to PSI charging in several ways. The issue of competition of private re-users with vertically integrated PSI IHs has been well explored, particularly in connection with downstream markets. The notions of “upstream” and “downstream” markets are familiar to antitrust lawyers, in particular in connection with questions of abuse of dominant position; these positions might require being to certain extent adapted to PSI, where, for registration-base PSI, it may be questioned whether the information generated, collected and disseminated by the PSI IH within its public task may be understood as belonging to a market.\textsuperscript{6} The prohibition of discrimination among categories of re-users and of cross-subsidies has been thoroughly researched; the same applies to the debate on the advisability of providing for separate bookkeeping and accounting of different branches of vertically integrated PSI IHs.

Further inquiry may however prove appropriate in connection with the impact of charging options on upstream markets, e.g. depending whether we are dealing with registration (or sole source) data sets or not. It might for instance be questioned whether charges below cost for non registration based PSI, where private competition may exist or come into existence in the future, may amount to predatory prices.\textsuperscript{7}

Also the institutional design of competition scrutiny of PSI charging policy would seem to deserve more focus. While it is well accepted that PSI re-use tends to have a cross-border dimension, not necessarily this reality has sufficiently been taken into account in imagining antitrust intervention in the area. One possible instance is cooperation between national antitrust authorities, which might deserve some kind of “easy access route”.

The proposal here is twofold: to complete the list of relevant competition law questions (which include e.g. the treatment of contractual exclusivity and the assessment of the potential of public task and general interest to defuse antitrust scrutiny) on the one side and to systematically explore how they link back both to “classical legal” PSI charging issues and to the other groups of topics discussed below on the other side.

The same conceptual sequence should be adopted in the next areas. Consider \textit{tax law}. What are the tax law assumptions of the various PSI charging scenarios? In particular, what is the role given in each of them to general taxation and, if adopted, to special purpose taxes and levies? Does the tax law angle change depending whether we consider registration or non-registration based data? Is the law of State aids and of public subsidies relevant?

This barrage of questions leads to the next level: \textit{subsidiarity}. One can imagine here the starting point: what is the legal basis for EU action in this area and its limit? What is the level of sovereignty retained by Member States in the area of PSI charging policy? What are the implications of the principle of proportionality? However, the main point here is again to engage in a two-steps exercise: what are the relevant issues, especially the ones which are left out of the mainstream debate in connection with PSI charging and how do they interact with the rest of the matrix (both above and below)?

\textsuperscript{5} See below, chapter 3.

\textsuperscript{6} See the Opinion of Advocate General Jääskinen 26 April 2012, in Case C-138/11, case Compass – Datenbank GmbH.

\textsuperscript{7} On the question whether registration-based (or sole source) PSI may amount to an essential facility under EU antitrust law see the Opinion of Advocate General Jääskinen 26 April 2012, in Case C-138/11, case Compass – Datenbank GmbH.
This line of reasoning leads us to the next point. The inquiry should move on to asking what other legal domains are relevant to PSI charging, directly or indirectly. Some replies are already pretty clear. Data protection rules may be relevant, albeit only indirectly, in particular in the calculation of costs (see the issue of anonymization). IPRs rules, as well, may affecting the cost of generating and re-producing some kinds of PSI, and condition some of the exceptions (one of them has been already envisaged in the proposed new Art. 6) to the default charging rule. Subsidiarity and constitutional rights are interlinked and the constitutional underpinnings of Freedom of Information (FOIA) and access legislation also are pretty well settled; their connection with cost issues is not necessarily so clear, though. The link to Fundamental Human Rights (including the EU Charter) would, at least at first glance, seem more remote. However, other areas would appear likely to yield greater insights on charging. It would appear, for instance, that the debate on re-use may suffer, if it is mainly (or even only) linked to its economic and market oriented, cross-border based dimension. The impact on charging policies of the imperatives of accountability, transparency, open government and e-democracy would appear to have been explored mainly by US scholars such as Uhlir.\(^8\) Maybe the time has come for us to follow in their footsteps.

The next question, then, is how to bring these additional domains into the rest of the matrix, along the lines proposed earlier.

Of course, this approach needs an antidote. It may easily slip into the compilation of wish lists or expressions of interest; so some kind of “pruning off” is required and must be carried out.

### 3. Economic Perspectives

From an economic viewpoint, PSI charging policies may be looked at in a number of ways. Four would appear to be relevant in this context: costs and markets, theory and empirical evidence.

Cost calculation. If we look upstream, at PSIHs which collect, generate, organize, re-package, store and disseminate PSI, we may say that several of the conceptual issues already have been addressed.\(^9\) However, it might help to look further into several questions:

a) Metadata.\(^10\) When we look at costs, do we include metadata (which are very relevant for MLAs)? When we discuss metadata, do we have clear whether we refer to “legacy” metadata, the ones the PSIH cannot by definition work without; or to “revamped” or “upgraded” metadata, e.g. metadata which are generated in the process of digitization and are either standardized, or interoperable, or both? Is this distinction better captured by contrasting metadata generated for internal use by the PSIH to metadata created also (or only) in view of re-use? Have we asked ourselves how this impacts (not on cost, but) on value? What is the impact on this assessment of the overarching rationale for making available PSI?\(^11\)

b) How does this link to the engineering principle of separating data from interaction?\(^12\) Does

---

\(^8\) New research on “The New Ambiguity of ‘Open Government” has been just released by Yu & Robinson.

\(^9\) Including by LAPSI, “The ‘principles governing charging’ for re-use of Public Sector Information”, quoted above, chapter n. 1, § 3. In particular see the discussion of the notion of “extra-costs” linked to requests from specific re-users; of the notion of “upfront” costs which are a prerequisite to dissemination; of transaction costs, etc.

\(^10\) Also here the inquiry should start with a clarification concerning the notion of metadata. The very distinction between data and metadata is for from obvious. Under some approaches, such as the discussions conducted in connection with the semantic web, metadata are “just” data as any other datum (see W3C NOTE 1997: “RDF encourages the view of "metadata being data"”; [<http://www.w3.org/TR/NOTE-rdf-simple-intro>]). See also [http://en.wikipedia.org/wiki/Metadata].

\(^11\) The distinction between costs which are an unavoidable incident of the main, “public task” activity of the PSIH and costs which are separate and additional to them is specially relevant here and also has an obvious link with the issue of the public good character of PSI, see below, chapter n. 3.

the popular mantra “raw data now” apply also to metadata?

c) Is cost calculation relating to PSI affected by the existence of two-sided markets, where taxpayers are charged for the entry (registration) of data into land, vehicle and similar registries to begin with and whatever charge re-user may end up paying is on top of that first payment? How does the standard analysis (concerning the demand elasticity of the two groups engaged in the two sides of the market; on the cross-externalities between the two groups; on the positive externalities which may be created by them; on the welfare effects of the above) apply to the specific case of registration-based PSI?

d) How is cost calculation affected by the fact that the data sets include one off data (e.g. metadata), bulk data and continuous data streams (e.g. weather, environmental data)?

e) What are the cost calculation methodologies adopted? Do they fit the different categories of data, including the ones mentioned above?

f) What is the empirical evidence and what is it telling us on the above? How do case studies (e.g. for Europeana) contribute to the clarification of the above?

If we look downstream, at markets, of course costs become prices. What do we know on:

a) The impact of price on quality of data (which may be seen as a facet of the “backfiring” issue);

b) The impact of price on the quality (and nature: e.g. in terms of interoperability) of metadata (which again may be seen as a facet of the “backfiring” issue);

c) What is the impact of the “extra-price components” of charges, e.g. the time horizon for which the price is set, on re-use and investment in view of re-use?

d) What is the impact of price (e.g. for non single source data sets) on competition? See also the other price-related competition issues above.

e) Again, what is the empirical evidence and what is it telling us on the above? How do case studies contribute to the clarification of the above?

PSI Markets deserve being closely looked at in their own right. The main divide to be taken into account is between upstream and downstream markets on the supply side, and between intermediaries (business re-users), on the one side, and end (re-)users, on the other, on the demand side. Additional relevant distinctions would appear to emerge between the following:

a) Final Users Markets;

i. Paying;

1. High value added/business client: e.g. weather for oil rigs; data sets used by professionals for environmental assessment;

2. Low price/individual consumer: apps for cyclers wishing to avoid showers; scalability of apps;

ii. Non paying (including advertising driven);

iii. Hybrids between i. and ii. (again including advertising driven);

b) Intermediaries (e.g. data markets);

c) Communities, institutions, including cities; other PSIH.

The idea here is that taxonomies breaking down markets should and might still be improved. The

13 For non-registration-based PSI only, however: see above, chapter n. 2.
attempt above is probably very elementary, but it still might be more detailed than the one employed in standard empirical research, which tends to simplify its job by confining itself to the most basic distinctions. For instance, we are not sure that the role of communities, which are neither businesses nor individual consumers, on the demand side has been fully explored yet. We are not aware of any research visualizing PSI as intermediaries between an upstream PSI and citizens or communities either.

We mention this point here because one of the most characteristic features of PSI, the unpredictability of ultimate re-uses, can best be visualized in the perspective of the multiple re-use and end-use markets which may demand PSI.

After visualizing cost and markets, the easy part of the job is to apply to them the approach sketched above: first we should complete the list of issues and second we should systematically explore how they interact with the other components of the matrix (beginning with the legal issues above).

The more difficult part of the job is to incorporate into the analysis of PSI charging policy two additional dimensions, one grounded in economic theory, the other on a more massive resort to empirical evidence.

On the first level, theoretical, our query should be: has PSI features which make it a public good? Given that re-use of data sets is not rival;\(^\text{14}\) and that the externalities flowing from re-use are magnified by digital networks, how is this (arguably public) good provided? Which of Paul David’s three Ps, Patronage, Property or Procurement,\(^\text{15}\) should apply? What adjustments of public goods theory are suggested or required when the resource is digital and may be made available over digital networks?\(^\text{16}\) How does this theoretical framework impact on charging policy, considering that Property (actually: Intellectual Property) would support profit maximization, Procurement might go the opposite way, enabling free or subsidized provision of the good, except that here we have a (probably unusual) case of Procurement of complementary goods, or, more to the point, Procurement of a main public good (let us say: land title registrations) to which an ancillary public good (cadastral data) may be incidental as a sort of “spin off”. The current framework gives a clear inducement to address an important complication in the model: also in this connection we should take into account the difference between PSI which is re-usable in its original form and just waits to be made accessible and re-usable on the one side and PSI which requires that the PSIH invests additional resources to enable access and re-use to the PSI. Here the choice of P may specifically involve the “re-usable master copy”, as the same may not be available in the first place.

It is often said that one of the features of PSI is that the re-uses to which it may be put depends on a very large number of decentralized decisions by re-users, by trial and error or via unexpected needs and intuitions. It is arguable that PSI may be viewed as an intermediate (non final) experience good. The question here is: how does this character affect the analysis above? Can we say that this feature speaks in favour of zero (or close to zero) charges, as no willingness to pay may be expected \textit{ex ante}? As H. Ullrich recently said, also technologies are intermediate experience goods, in that both

\(^{14}\) Even though excludable, under current EU arrangements which build on IP protection of data sets. To be sure, in public goods theory, goods which are non rival in consumption but may be made excludable are considered impure public goods or also “club goods”: see Frischmann, B.M. 2005, ‘An Economic Theory of Infrastructure and Commons Management’, Minn. L.R., p. 943, and originally Buchanan, J.M. 1965, ‘An Economic Theory of Clubs’, in Economica, p. 1


licensors and prospective licensees must anticipate what final price the market will bear in connection with goods incorporating them. In which way is the analysis for PSI so different that we may consider zero cost rather than profit maximizing price as ideal yardstick?

That these theoretical issues have an impact on PSI policy is quite obvious; what exactly the impact may be is probably more arduous to say. What is likely is that the impact should be quite large in specific connection with charging policy.

Crucially, we should ask ourselves what would change in the above range of questions if we assumed that PSI, while best described as a public good, should be subject to an access regime worked out on the basis of the notions of commons, of common pool resources or even infrastructure? This is probably the single most important issue to design a charging policy: therefore it should have priority in the analysis and it should be taken into full account while “weighting” the importance of the various components of this exercise ad indicated above in lett. c. of item 4.

Turning to empirical evidence, the starting point is that the substantial amount of evidence collected (including the Vickery Study, POPSIS, the study on re-use of cultural material, etc.) is likely to have an impact at all levels discussed above. I do not know that anybody has done this yet: but imagine just going through the 598 submissions made during the Consultations on the revision of the Directive having in mind a simple question: what kind of re-use is contemplated here? who is the re-user replying? Has she a business model? Or what re-user is considered in the reply? What impact on her/him/it may have the different charging options in discussion here? If this work has not yet been done, it should be definitely be carried out; and we would have a very large amount of information relevant for designing charging policy. This approach should be (selectively) replicated for many of the issues mentioned above. This interfacing job is not often carried out, because exactly as lawyers tend to keep to legal literature, economists even more so tend to keep to the dismal science. Except that PSI is at the intersection of so many that we must at some point change tack.

4. Other Domains

Actually, we should be in favour of further extending consideration to a number of other domains which would appear to be directly relevant to charging policy. Let me just mention two of them which would appear to have a clear priority in this connection.

a) The first one is history of ideas. What is the philosophy or ideology behind the effort to open up PSI? One might well look at it as a last strand of liberalization and privatization cycles, a latter day Thatcherism finally encompassing information collected by public bodies. This understanding might find some comfort in the parallel history of the much earlier American movement. However, it is also possible to see totally different narratives

---


behind it; these may range from radical ideas of empowering the citizens (and communities) advocated by open data constituencies to thoughtful ideas of restoring a (post-national) provision of public goods.\textsuperscript{22} The case may also be made that, in the current moment of financial turmoil and economic crisis, the EU is finding itself once again at the crossroads, as it was after the oil embargo of the early Seventies, when it was able to move out from the idea of full harmonization to the much more effective approach based on minimum harmonization plus mutual recognition;\textsuperscript{23} it may also be argued that fostering the creation of a common pool of data sets which are amenable both to the creation of cross-border, information-based services and of a digital network driven innovative environment is a similar chance before the EU right now.

b) It would appear however that there is not much consensus on any of these (very different) accounts. It is submitted that this lack of a general reference point may greatly weaken the understanding of the rationale of many of the policy options in the field of re-use charging.

c) The second has to do with the changing landscape of social actors, where new players (digital communities) come to the fore and possibly ancient players (such as local communities and cities) respond to new, network-enabled, demands. If this is the case, then the question arises as to the best way to incorporate into the process shaping charging policy the demands coming from these constituencies, which might be left out if governance of the relevant processes is left to the views and institutions prevailing before the change.\textsuperscript{24}

5. how does this feed into guidance?

It should be kept in mind that the options we are striving to work out should ideally come in a format which is fit for the purpose. Competition law guidelines, such as the various Communications prepared by the EU Commission to deal with vertical restrictions, or mergers, are the ideal blueprint here. It would seem that § 7 of the LAPSI Discussion paper No. 1 is an intermediate step into this direction. If this is so, it should be noted that what is proposed here is a continuation and a deepening of the approach on which the Discussion paper was originally based. The time has come to go beyond that initial effort, with a view of imagining a set of comprehensive guidelines to be issued by the Commission to Member States in the twin areas of “charging and calculation of costs”.

\textsuperscript{24} It is submitted that an intergenerational factor comes into play do shape the dynamics involving PSIHs and re-users and, particularly, re-user communities. Indeed, re-user communities tend to be formed by individuals belonging to the younger generation; PSI tends to be controlled by entities the top positions of which are dominated by not-so-young bureaucrats. If this intuition is confirmed by hard data, we would experience one more case in which Europe trails behind the U.S. because it fails to empower the younger generations. It may be questioned however whether this perspective has direct relevance for charging policies, as opposed to broader policies such as access regimes, of which charging is just one component. It may well be that also sociological dynamics may be relevant: just as it may be surprising how large is the group of librarians which have turned into activists in favor of access to knowledge, it may be worth its while to chart the existence of pro-reuse groups in the lower branches of several public administrations. It may questione whether this component, which certainly is remarkable as far as formats, interoperability and standards are concerne, plays any role in shaping charging policies.