Open data values: Calculating and monitoring the benefits of public sector information re-use

van Eechoud, M.

Published in:
Informationen der öffentlichen Hand: Zugang und Nutzung

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Open data values: Calculating and monitoring the benefits of public sector information re-use.

*Mireille van Eechoud*

Inhaltsübersicht

I. Re-use policy as economic policy 3
   1. From Guidelines to Public Sector Information Directive 4
   2. Key features Public Sector Information Directive 6

II. Economic studies 8
   1. Commercial exploitation of Europe’s public sector information (PIRA, 2000) 8
   2. Measuring European public sector information resources (MEPSIR, 2006) 9
   3. Assessment of the re-use of public sector information study (MICUS, 2008) 11
   4. Pricing of public sector information (POPSIS, 2011) 14
   5. National and local studies 17
   6. Value of economic studies 19

III. Open government studies 22
   1. Transparency studies 24
   2. Motivation studies 26
   3. Societal impacts and rankings 28
   4. Value of open government studies 30

IV. Twenty-five years on, where is re-use policy heading? 31
   1. Persistent lack of hard data 33
   2. From generic stand alone to a coordinated approach 35

The promotion of policies for opening up government information for broader use by industry and citizens goes hand in hand with promising estimates of its economic value. Such estimates vary widely but always appear impressive, from the whopping three trillion dollar per year McKin-
Mireille van Eechoud

sey¹ predicted the value of open data to be, to a more modest ten billion euro suggested in the influential ‘MEPSIR’ study.² Many more studies have been done at the European and national level. More recent are studies on the benefits at the local level, now that municipalities also embrace ‘open data’ policies to stimulate the local economy and improve public service delivery. The economic benefits for the private sector and the economy as a whole were central to the early re-use policy of the European Commission. Today these go hand in hand with a plethora of other objectives. They might be loosely grouped in two kinds. One are democratic-political objectives, aimed at fostering political accountability and citizen participation. The other comprises objectives aimed at improving the efficiency and quality of public service delivery. As the objectives of re-use policies become more varied, so does the range of research methods and tools to substantiate and evaluate policies.

In this contribution the focus is mainly on what the economic studies tell us about the estimated value of public sector information. It shows how challenging it is to get a clear sense of that value. The development of methods to actually measure the wider impacts of opening up government information is work in progress. A recent addition are instruments to rank public sector effort in opening up. What do these new strands of monitoring add to the economic figures? How do economic benefits relate to the many professed advantages – be it social, political or economic – of openness? As we start to unpack this, we can also look forward, to the long-term implications for the regulatory framework in the EU. Special focus will be on the Directive on re-use of public sector information a.k.a. the Public Sector Information Directive.³ It epitomizes EU re-use policy and

2 10 billion Euros was the lower range estimate of the EU market value of PSI (2006). With the benefit of hindsight, one of the MEPSIR authors later suggested even this might be too optimistic (see below). Makx Dekkers, Femke Polman, Robbin te Velde, Marc de Vries (Helm/Zenc), Measuring European Public Sector Information Resources (MEPSIR), Final Report of Study on Exploitation of public sector information. European Commission: Brussels 2006.
is generic in its reach across different domains of government-produced information.

I. Re-use policy as economic policy

In the 1980s, the US database industry grew while European companies lagged behind. This prompted the then European Community to elaborate the so-called ‘Guidelines for improving the synergy between the public and private sectors in the information market’ of 1989.¹ Aim was to encourage public sector bodies in Member States to grant companies access to public sector ‘raw’ data for commercial exploitation. The Guidelines also encouraged public sector bodies to refrain from providing commercial information services and products themselves.² This should lead to a bigger stake of European companies in the growing information markets. The Guidelines were not successful however, for one reason because they were not widely known. The Commission then started to entertain the idea of regulatory intervention. Initially, no European wide studies appear to have been done on the value of releasing government information. As to policy objectives, informed by private sector complaints the initial focus of the Commission seems to have been as much on curbing (anti)competitive activities of public sector bodies in information markets, for example in the area of company data services and mapping, as on stimulating the release of data for commercial exploitation.

The competition problem has not gone away since then. This is evident from the ongoing debate in the UK on the position of the national mapping organization Ordnance Survey³ and from complaints made against public

---

² Interestingly, PIRA concluded that public sector investment in the production of information in the US was significantly higher than in the EU, see PIRA, Commercial exploitation of Europe’s public sector information. Final Report. (European Commission DG Information Society) 2000, p. 16. [http://www.epsiplatform.eu/sites/default/files/media_672%20full%20report.pdf]
sector bodies in a number of countries for restrictive access and charging policies. The Public Sector Information Directive however does not apply to commercial activities of public sector bodies. Nor does it tell Member States what information production and dissemination activities should be part of public tasks or left to the private sector. It does create an obligation for Member States to allow re-use at non discriminatory terms, of information collected or produced in the course of public tasks. The information must be public and not subject to third party intellectual property rights. The focus of the PSI Directive therefore is firmly on the ‘spin off’ value of information created in the course of public tasks. It does not concern itself with the public sector share in information markets. A number of the economic studies we shall discuss do however address both aspects. The first studies that the EC commissioned in the 1990s were the so-called ‘PUBLAW’ studies. These dealt mainly with the existing legal frameworks for access to government information and data protection in Member States; not with economic value of government held data. Only in 2000 was the first report published which attempted to put a number on the value of PSI in the EU. Funded by the European Commission’s Directorate General for the Information Society under the eContent programme, it became known as the ‘PIRA study’. Its findings are discussed further below.

1. From Guidelines to Public Sector Information Directive

Despite sound supporting economic data, the EC 1991 work programme included action to draw up proposals for harmonizing rules on the “mar-

vey, Companies House and Land Registry) at [https://www.gov.uk/government/groups/public-data-group].

7 For an overview of recent cases in the Netherlands, Sweden and Austria, see: B. Lundqvist, Y. Forsberg, M. de Vries, M. Maggiolino, LAPSI 2.0 competition law issues position paper (2014), available at [http://www.lapsi-project.eu/lapsi-20].

8 On file with the author. Unfortunately, the reports do not seem to be available online anymore. A report of a workshop where the PUBLAW studies were presented is available at [http://cordis.europa.eu/news/rcn/1489_en.html]. The EC website where the content was originally hosted is no longer accessible [http://www2.echo.lu/legal/en/access/publaw/publaw.html].
keting of data files held by public and quasi-public bodies’. The European Information Industry Association had proposed a commercial right of access to public sector databases, to be regulated at EU level. They asked for an instrument that would oblige public sector bodies to allow commercial exploitation of their resources (if no legitimate interests argued against it) and stay clear of providing (electronic) services that the private sector could offer. The Association also suggested that public sector information should be excluded from copyright.

It appears industry was willing to ‘trade’ a separate instrument for a right of access incorporated into the Database directive. The directive was on the drawing board at the time. Special provisions for (access to) government databases did not make it into the directive. It makes no distinction between databases produced by the public and private sector. On the contrary, the copyright section of the Database directive confirms that collections of data created by public sector bodies enjoy copyright protection if the selection and structure of the dataset constitutes an original work (‘own intellectual creation’). What is more, the directive introduced a full-blown intellectual property right for the producer of any public sector database that has involved a substantial investment in either the obtaining, verification or presentation of the contents. Copyright and database rights can be an important means for public sector bodies to control the (re)use of their information. The recent trend however is for governments not to assert intellectual property rights but instead to make public sector data available as ‘open data’, that is, free from copyright and database right constraints and in a technical re-usable way.

10 This historic overview is derived from the document ‘Legal Advisory Board (LAB) File 1993/1’, Publaw 2 workshop 4 March 1993, Brussels: EC DGXIII, 1993.
12 That the sui generis right is available to the public sector can be deduced from ECJ 5 March 2009, Apis-Hristovich/Lakorda AD, Case C-545/07 (ECLI:EU:C:2009:132), ECR 2009 I-01627; ECJ 12 July 2012, Compass-Datenbank/Austria, Case C-138/11 (ECLI:EU:C:2012:449).
13 Key features of open data are: no intellectual property rights are asserted, the use of data is free (no use charges), data are machine-readable and available in open
Although ineffective, the Synergy Guidelines prepared the ground so to speak for hard regulation. As the realisation dawned that this soft law was not helping markets develop, the EC initiated regulation. A number of economic studies informed the decision-making process, from the 1998 Green Paper on Public Sector Information in the Information Society \(^{14}\) to the 2003 Directive on the re-use of public sector information (and its subsequent revision in 2013). These are discussed below. It is important to note that the focus of re-use policy has slowly been broadened from an economic one aimed at releasing the economic value of public sector information, to one that, on paper at least, includes social and political objectives associated with what is increasingly termed ‘open government’. The nature of recent studies designed to gauge the impact of re-use policies reflects this, as we shall elaborate later.

2. Key features Public Sector Information Directive

To be able to evaluate the studies, it is important to remind ourselves of the key features of the Public Sector Information directive:

- It applies to a wide range of public sector bodies;
- It applies to a wide range of information resources that are public under domestic access regimes (except when information is subject to third party owned intellectual property rights);
- With some exceptions it creates an obligation to allow re-use against as few conditions as possible (esp. public sector bodies that have to work under a cost recovery regime, e.g. many company registers, can continue charging and licensing);
- Terms and conditions must be non-discriminatory (similar for similar user groups, including commercial arms of public bodies)

Benefits of Public Sector Information re-use

- The default principle for charging is (at most) marginal cost of dissemination, no royalties or other use charges.

The marginal cost pricing principle is of crucial importance to the idea of re-use. Particularly the fact that information has properties of a (semi) public good notably that it is non-rival, makes it economically sound practice to allow re-use.\textsuperscript{15} Welfare maximization is said to result if information that is produced for public task purposes and tax-funded, is subsequently released for alternative uses, at no more than marginal costs of supply. Of note, this line of reasoning is particularly compelling for information that will be produced primarily for use within government. If information is collected and held to serve a variety of parties both in and outside the public sector, the economic case is more complicated. Examples are company registers and land registers. Part of the public task in such cases is already to provide information services to the private sector.

The economic rationale for making available PSI at no or low cost is theorized in the ‘Cambridge study’ of 2008 by Pollock et al.\textsuperscript{16} Showing that marginal cost pricing actually produces direct and indirect economic benefits in practice remains difficult because of a lack of data. The economic studies discussed in the next section mostly are based on three presumed direct advantages of making public sector information available for re-use free or at marginal cost of dissemination. The first are positive externalities. The wider dissemination beyond the public task purposes for which the information is produced leads to additional usage, which results in growth of information markets and hence more jobs and national income. Second, the resulting economic growth leads to a rise in tax returns.

\textsuperscript{15} The other public good aspect is excludability (once produced the good benefits multiple actors because access to it is difficult to control, in the case of information this may be done by keeping it confidential, or regulating access and use through contracts or technical measures). Intellectual property rights are an important instrument to artificially create excludability. I shall not go into the rationales for (not) having copyright and database rights in government information here. See for a discussion of the Dutch situation, M. van Eechoud, ‘Government Works’. In: P.B. Hugenholtz, A.A. Quaedvlieg & D.J.G. Visser (Eds.), A Century of Dutch Copyright Law. Auteurswet 1912-2012, Amsterdam: Delex 2012.

Third, dissemination free of charge yields savings by decreasing transaction and opportunity costs for both the public information provider and (private sector) user. For example, the public sector body will no longer have to commit resources to determining prices, handle creditor administration and police licensing terms.

II. Economic studies

In the next paragraphs, a number of influential studies are discussed in more detail. This will show how disparate the methods are which are used to estimate benefits, and how little hard data is available to feed those estimates. What is more, often no (clear) distinction is made between the different types of benefits, whether these are increased usage, growth in information markets, public sector savings or other public sector efficiency gains.

1. Commercial exploitation of Europe’s public sector information (PIRA, 2000)

Commissioned by the European Commission’s Directorate General for the Information Society, consulting firm PIRA set to quantify the value of public sector information for a number of domains in Europe. These were grouped as economic and business information (economic statistics, company data), environmental information, agriculture and fisheries information, social information (e.g. census, health), cultural information (museums, archives), legal system information (laws, cases, crime statistics, etc.) and political information (e.g. policy materials, public information services).

Data was gathered through desk research and interviews with organizations involved in the PSI ‘value chain’, mainly public sector bodies, companies active in the production of information for the public sector, companies creating information services and products based on government data. PIRA estimated values based on a combination of private sector spending on PSI and public sector investment value of PSI, i.e. govern-

17 I say influential because they are often referenced in other studies and policy documents on the benefits of PSI re-use.
Benefits of Public Sector Information re-use

The researchers made estimates for five EU countries and extrapolated from these for the entire EU. At the time, there were 15 Member States. They arrived at a figure of government investment in PSI of €9.5 billion per year. Demand side data was sparse. To arrive at an estimate of the economic value, the researchers looked to the value added by PSI users, and as a proxy, to the income public sector bodies derive from direct supply to final users. No private sector company would share its figures for value added, forcing the researchers to rely on (limited) statistics data from the US and on company annual reports.

The economic value --in PIRA terms this means the part of the national income attributable to industries and activities built on the exploitation of PSI-- ranged between €28 billion and €134 billion per year. In order of public sector investment size, geographic information ranked first, followed by cultural information and intellectual property data (patent registers). In terms of economic value, the study concluded that geographic information takes by far the greatest share, followed by industries based on social and economic data, and on company data.

PIRA further concluded that economic benefits are to be had from releasing public sector information at marginal cost of dissemination (or less) instead of maintaining models whereby public sector bodies attempt (full) cost recovery.

2. Measuring European public sector information resources (MEPSIR, 2006)

Run from 2004 to 2006, the MEPSIR study was one of two studies done for the European Commission, designed to feed into the PSI Directive statutorily imposed review of 2008. It set out a methodology and baseline assessment of five ‘framework conditions’, which would allow progress to be measured in later years once the Directive had been implemented by Member States. Among the framework conditions were actual availability of information for re-use, its accessibility (e.g. how user friendly data are supplied in terms of channels, response times) and transparency of conditions for re-use (e.g. different types of licences, online licences). The study also tried to estimate economic performance, understood as the size of demand and public sector resources spending on information production and dissemination. Values were derived mainly from data supplied by public sector information holders and re-users.
Like the PIRA study, MEPSIR focussed on what were thought to be the most promising re-use sectors. The categories were somewhat different this time. Cultural and scientific information were left out, as these were not within the Directive’s scope, although some types of cultural information (held by public archives, museums and libraries) are so now. Sectors covered included business information (company registers, patent registers), legal information (laws and court decisions), meteorological information, social data (e.g. economics, employment statistics, health, and census) and transport information (traffic information, vehicle data). Using desk research, a web survey and questionnaires, data were gathered for all EU Member States and Norway. Most data came from self-reporting.

Despite working five years after PIRA, the MEPSIR researchers too were hindered by a lack of hard data. Pre-existing data was sparse so the researchers relied heavily on data from surveys and interviews among producers and users of different types of PSI. The average response rate of PSI holders for the web survey was not very high, at 31%. Estimates given by different stakeholders about the size of their (national) PSI market varied considerably. Especially public sector bodies themselves had more modest perceptions of PSI value in their different domains. The upper range of aggregate estimates by (private sector) re-users was €47 billion for the EU plus Norway, with €26 billion at the lower limit. The combined public sector information holders ‘estimate however was much lower and ranged from €6 at the lower end to €26 billion as the upper limit.19 Given these large discrepancies, the researchers also constructed the estimate market size for PSI based on financial data supplied by stakeholders, e.g. data on turnover, data acquisition costs and staffing levels, with estimates of numbers of re-users and average turn-over from PSI per re-user. These data were considered more reliable, but still yielded a wide range: from 10 € billion to € 45 billion.

PIRA arrived at higher estimates in part because it included a wider range of domains, and based the economic value, defined as the share of PSI in national product, on less precise sources (national statistics). The MEPSIR team considered that the PIRA study does not allow for adequate identification of the added value that is directly attributable to public sector information. With hindsight, one of the principal authors of the MEPSIR study argued that although considerably lower than the PIRA estimates, the MEPSIR estimates had also proven to be too optimistic. He

---

19 MEPSIR Study, p. 33.
suggested the PIRA estimates themselves were perhaps a factor 15 to 20 too high.20

Bearing in mind that a primary objective of the PSI Directive is to stimulate growth of cross-border information goods and services, the study’s use of cross-border trade as a benchmark is of special interest. MEPSIR concluded that cross-border trade in all domains studied was low (average score 8 out of 100), with meteorological data (climate model data) and patent data ranking best with a score around 20 out of 100.21 There are two ways to read these results. They can support the argument that firmer re-use policy is needed for EU wide services to develop, so justifying EU regulatory intervention. But they can also be taken as an indication that there is in fact little justification for the EU to step in with stricter norms, since for most sectors there is hardly any cross-border trade. Whichever is the more compelling view depends of course on the reason why there are few public sector information based products and services that transcend Member State borders: because there are impediments to cross border re-use, or because there is no or little (latent) demand.

3. Assessment of the re-use of public sector information study (MICUS, 2008)

The EC commissioned the Study in 2007 as part of the preparation for the mandatory review of the effect of the PSI Directive. The research covers three broad areas that are considered as high potential for re-use: legal and administrative information, geographic information and meteorological data. Data was collected using online questionnaires, telephone interviews and written sources.

Here too, the researchers noted it was difficult to create a reliable picture that would be representative of the situation across EU Member States because of a lack of good pre-existing data and weak response rates. What also did not help is data was only collected for a number of countries and stakeholder types. For example, in the geographic information sector most respondents were from Germany, Spain, Italy and Sweden and the only

21 MEPSIR Study, p. 28.
public sector respondents were national mapping and cadastral bodies. For meteorological data, most private sector respondents came from Germany, but there was good public sector coverage across the EU. This makes sense because countries generally only have one public meteorological institute and these all cooperate in European associations. Therefore, they are straightforward to identify and approach. For legal information, nearly half the respondent re-users were from the UK and Germany, but the geographic spread of replies from PSI holders was broader.

The indicators studied to measure growth of PSI re-use were mainly income and use metrics, such as traffic to data portals/websites, volume of downloads and number of re-use/licensing requests. Web metrics such as server requests, number of unique visitors, number of downloads are only proxies for actual use of data. The 2000 PIRA study already flagged the unreliability of web statistics in this respect. For legal information, that is statutes, regulations and the like as well as court decisions the paucity of data was biggest. Only a small minority of PSI holders surveyed in the legal and administrative domain supplied data on use and income. On the re-user side, revenue for legal information publishers and service providers on average doubled in a five-year period, but this figure was based on only a handful of survey respondents.

The income from public sector bodies in the geographic information sector was estimated at 359 million EURO. It was not measured in detail, as it only covered turnover of ‘basic’ geographic datasets (not value added services) held by national organizations, and seems to have included ‘sales’ to other (local) public sector bodies who rely on their national agency for certain data they need to exercise public tasks. It also excluded supply by regional/local administrations to the private sector, although in many countries regional/local administrations hold essential data. Private sector income could not be measured accurately. At any rate, it included outsourced work, that is, private sector turnover from mapping activities done for public authorities or from the development of geographic information tools or systems for public authorities.

---

22 PIRA Study, p. 44-45.
23 The analysis in this section is based on the MICUS report, esp. p. 15-35.
24 The estimate for the Dutch market is that companies active in geographic information sector derive about half of their revenue from public sector bodies. W. T. Castelein, A. K. Bregt, Y. Pluijmers, ‘The economic value of the Dutch geo-
For legal and administrative information, there was no adequate data on volumes and the number of (re)users. However, for legal information respondents included large legal publishers that traditionally have a large share of the market. The authors did conclude that online access to legal materials in recent years led to ‘extremely high growth rates’—of what and what the causal link is between the availability of legal information re-use and market growth is not detailed unfortunately.

Interestingly, the user survey also showed that re-users were not familiar with the PSI Directive: more than half of responding re-users in the legal and administrative domain did not know the PSI Directive. As we have seen above, one important reason why the Synergy Guidelines of 1989 were deemed ineffective was that few knew of their existence. If the findings of MICUS are anything to go by, the PSI directive suffers from a similar problem, at least did so in the first years after its inception.

Overall, the MICUS study produced rather vague results. It concluded that ‘growth rates have been reported in every market segment under consideration. The improved access to information from the public sector in Europe contributes to this success.’ Because the indicators used in the MICUS study were different from the benchmarks developed in the MEP-SIR study, and the indicators used in PIRA, and the markets studied also differed, it is not possible to track development in re-use markets by comparing the outcomes of the studies.

The studies discussed above all acknowledge the impact of pricing strategies on re-use levels and described some case studies involving disputes over pricing. However, only in 2011 did the European Commission publish an in-depth report on models of supply and different pricing strategies and their impact on re-use levels. By 2009 the review of the PSI directive had led the EC to set up a working group on the ‘Economic case for marginal cost’, which supplied recommendations for the pricing study to be commissioned. This was to be the Pricing of Public Sector Information Study or ‘POPSIS’ of 2011.

---

25 MICUS study, p. 93.
26 MICUS study, p. 79.
27 MICUS study, p. 98.
4. Pricing of public sector information (POPSIS 2011)

The voluminous study29 (some 400 pages) analyses re-use markets and direct and indirect economic benefits based on twenty-one case studies from nine European countries. The choice of domains was informed by what previous studies (MEPSIR, PIRA) considered to have the highest economic potential: business registers (company data), meteorological and other geographic information (map data, building addresses, etc.) as well as legal information. Two-thirds of the cases concern geographic information and meteorology. Of note, national meteorological institutes cooperating in EUMETNET have committed to opening up essential data for re-use in the so-called Oslo declaration of 2009.

As its title suggests, POPSIS focussed particularly on the effects of different pricing strategies, including on public sector bodies operating under (full) cost recovery models and public sector bodies who charge at zero or marginal cost of dissemination. In this respect, it tries to assess whether the economic advantages of marginal cost pricing as theorized in the 2008 Cambridge study actually materialize. The Cambridge study had set out the economic theory behind potential welfare effects of marginal cost pricing of major public sector information holdings. It looked to geographic and company data in the UK in the form of large datasets held by the Ordnance Survey, Met Office, Land Registry, UK Hydrology Office (water ways), Companies House (company register data) and vehicle data from the NVDL register.30 The study emphasized the importance of adequate governance structures and regulation to prevent negative incentives, inefficiency and poor performance of public sector information holders, regardless of pricing strategy adopted.31

The cases compiled by POPSIS show it is difficult for public sector bodies to assess cost-based price of (a unit of) information or data. This can be used as an argument against pricing of course and POPSIS does

---

just that. One might also argue that in effect, the difficulty of it is characteristic of digital information products and services as such, precisely because fixed costs of production are high and marginal cost of an extra unit supplied are close to zero. Private sector suppliers of information services face the same accounting challenge, but they are normally free to set their prices as they wish. The PSI Directive however prescribes that any charges be based on real costs, and the charges need to be specified up front.

The cost recovery through sales to private sector re-users as part of the entire budget of public sector bodies was found to be 'low' (e.g. 10-30% in the higher ranges)\(^{32}\), which according to the authors is a strong argument against the efficiency of charging for re-use. Of note, when public sector bodies operate under a cost recovery scheme and supply (mostly) public sector customers, pricing and the terms of use underpin the entire funding model of the public sector body. From that perspective, the fact that little revenue comes from commercial re-users is of limited significance for funding model choices.\(^{33}\) Reducing charges and use restrictions for private sector customers may jeopardize the long term sustainability of a cost recovery scheme vis-à-vis public sector customers,\(^{34}\) as the private sector starts to offer competing products to the public sector customers. Ultimately, ‘user pays’ funding models might have to be abandoned for models where public sector information is funded through general-purpose taxes. Unless of course, costs can be shifted onto captive users on the input side. This might be the case if public or private parties are under an obligation to supply data and can be charged for doing so, e.g. registration at land registries, vehicle registries or business registers. Charging more

---

\(^{32}\) POPYSIS Study, p. 33.


than the actual costs of registration however might fall foul of EU law. Captive users in the public sector are for example bodies that have to use data from a certain public source (e.g. a designated key registry), or have to publish notices in official publications.

Taken together, the POPSIS cases show that a more liberal re-use policies including the lowering of charges fairly consistently results in growth in number of re-users and volume of data delivered. Many of the case studies concerned access over the internet of public sector information free of charge. The growth of use was typically represented in web metrics, e.g. as percentage increase in website views or downloads. Particularly National statistics offices that have started to make data viewable and downloadable on the web have seen user numbers rise significantly over the years.

At the level of individual cases, POPSIS provides insights on a variety of indicators: use metrics such as growth in number of licenses, change in re-use related income for PSB, number of re-users, effects on re-user turnover and size of staff. The available indicators differ per case study and often are estimates. Consequently, the study could not provide firm numbers on actual overall economic effects. Instead, it focusses on the consumer surplus that might result from releasing information at no or marginal costs. Income from selling information products and services by public or private actors is but one representation of value.

Free information also has value, even if this does not show in direct returns. The authors of POPSIS argue that consumer willingness to pay is an important value to take into account when estimating positive economic effects. Public sector bodies that make raw data and added value information available for free could thereby cause a reduction of market size in some areas (driving out of the market companies that provide similar services), but ‘market size by itself is not a sign of economic efficiency and societal welfare.’ They posit that ‘consumer surplus and indirect economic impacts are substantial, and are often far higher than the direct revenues and jobs created.’ This might be true, but if it is already the case that it is difficult to estimate the effects on commercial re-users of increased avail-

35 POPSIS, p. 59, pointing out that for business registries under the relevant EU directives, charges for registration may not exceed the actual cost of such registration.
36 POPSIS, p. 36 (geographic information), p. 49 (meteorological cases), p. 61 (legal information).
Benefits of Public Sector Information re-use

ability of free public sector data, this is of course even more true for quantifying consumer surplus and (other) indirect economic effects.

As to wider societal benefits, for example in the political domain, here too the study posits rather than substantiates positive effects, e.g. by stating that increased access to free PSI on the web leads to higher levels of citizen participation. More broadly, the many positive effects hailed by POPSIS remain anecdotal and speculative. In this respect, it suffers from the same paucity of hard data that also affected studies that came before.

5. National and local studies

The studies detailed above are special in that they concern the European market and were designed to inform EU policy making. There is also a growing body of research that focusses at a specific national or local territory. These studies too suffer from a lack of reliable data. This is illustrated by the 2013 Market Assessment study commissioned by the UK government’s department for Business and Skills. It estimated the value of (existing) commercial use of a wide range of public sector data. Of note, in the United Kingdom relatively a lot of research has been done on the value of public sector information. This is mainly due to persistent controversy over how large public sector information holders operate and charge users. Pricing strategies, licensing terms and providing value added services in competition with private companies all affect private sector markets. In the UK ‘trading funds’ and especially the Ordnance Survey and Companies House have long been under fire for allegedly aggressively pursuing revenue maximizing strategies to the detriment of broader access to their products and the development of private sector information services. Both are separate legal entities that have to operate under a cost recovery model, so need to cover their costs by charging public and pri-


38 Newspaper The Guardian has both fuelled and tracked debate; see the ‘Free our Data’ section on the newspaper’s website [www.guardian.co.uk]. However, debate over the dangers of (full) cost recovery charging for public sector geographic information was already raging in the 1980s. See M. Blakemore & G. Singh. Cost recovery charging for Government Information: A False Economy? London 1992.
vate sector customers. Indeed, some argue that the wider open government agenda in the UK has come to serve ‘neo-liberal’ policy objectives, aimed at privatizing and scaling back government functions. According to Bates\(^{39}\): “In the case of the UK’s OGD [open government data, author] initiative, for example, the overall ends of the coalition government can be understood as the continuation of a neoliberal capitalist form of state through a period of crisis.”

The authors of the Market Assessment study, using similar methods as had been used in a study for the Office of Fair Trading\(^{40}\) seven years previously, noted that there still was a paucity of data around public sector information. It therefore remained difficult to ‘generate accurate figures on the size, value and potential of public information’. \(^{41}\) Their careful estimate is that ‘narrow value’ to consumers, businesses and the public sector itself ranges between £1.2 billion and £2.2 billion [2011/12 fiscal year], and that the overall broader economic and societal value of public sector information for the UK is between £6.2 billion and £7.2 billion. \(^{42}\)

An example of attempts to substantiate the value of PSI at local levels\(^{43}\) is the 2014 ‘Digitales Gold’ study on the economic potential of open data for the city of Berlin. \(^{44}\) It extrapolates from earlier studies, including the PIRA and MEPSIR studies. For the Berlin area, PSI market size would be

---


\(^{40}\) Office of Fair Trading, The commercial use of public information, London: OFT 2006. The aim of the study was to assess how UKs (large) public sector information holders can improve supply of raw data to private industry. OFT estimated that with improvements the net economic value contributed by the industry could grow to 1 billion GBP annually. Because OFT had access to better data from relevant public sector bodies than PIRA, they considered that their more conservative estimate of 590 million in net economic value (net consumer surplus plus net producer surplus) were more accurate (p. 114-115).

\(^{41}\) Deloitte, Market assessment of public sector information, Study for UK Department for Business, Innovation and Skills, 2013. p. 43.

\(^{42}\) Deloitte, Market assessment of public sector information, Study for UK Department for Business, Innovation and Skills, 2013, p. 38.


nearly € 300 million a year. 45 Noticing that market size or economy share estimates are very inexact because they rely on assumptions with little hard data for support, a much lower estimate of the direct and indirect benefits was produced in a bottom-up approach. Applying a number of multipliers 46 on the estimate costs of running the Berlin open data portal the author arrives at an estimate economic value of 22 to 54 million euros per year. 47 However, these figures in turn rely on a various assumptions and extrapolations (e.g. of web metrics) which make the spread between conservative and progressive estimates large.

Outside Europe, a recent US study estimated the value of public sector information by taking annual revenues of commercial firms active in ‘government data-intensive business activities’, comprised of traditional information services providers (news, media, broadcasting), internet companies, management consultants, real estate sector and marketing/polling companies. It arrived at a crude estimate in the range of 24 to 221 billion USD per year. The report refrains from making a broader estimate of benefits, but sets out how the various federal statistics activities contribute to better decision-making in the public and private sector alike thus helping to optimize the use of resources. Using net consumer surplus type methodologies, estimates are given of the value of certain information types that are freely provided to the public, such as weather forecasts. 48

6. Value of economic studies

What then, may we conclude about the value of economic studies for public sector information policy? Taken together, the studies do show that once produced for the public task, making available public sector information for re-use has economic benefits that overall outweigh costs. However, it is wise to take the numbers with a large pinch of salt. First, be-

45  Digitales Gold Study, p. 34.
46  E.g., the multiplier for direct economic effects takes account of estimated efficiency savings, additional tax income generated because of increased private sector earnings and loss of direct earnings for public sector bodies who switch from charging fees to free supply.
47  Digitales Gold Study, p. 43.
cause there is a lack of hard data and specific data on both the costs side and benefit side. This means all studies depend on assumptions and extrapolations, leading to numbers that are not very firm. It also causes the estimates to range greatly. We have seen for that PSI market size estimates range from €6 to €26 billion per year (MEPSIR), even though the authors of this particular study managed to collect a wide range of data. For OECD, Vickery summarized existing studies and concluded for the EU the direct narrow value might amount to €32 billion (for 2010) and the broader direct and indirect benefits would be in the order of 140 billion. This brings us to a second reason for caution. It is not possible to meaningfully compare the estimates of value in the different studies. The methodologies, geographic scope and public sector information types studied differ. This also has implications for the extent to which results from one study can be used to estimate values for e.g. the national or local level.

There was some overlap in types of information covered because the economic studies tended to focus on large public sector information holdings like national statistics, meteorological data, mapping, cadastral data, business registers and legislation. To what extent the results apply to all types of data cannot be ascertained. This is a third reason why the results must be viewed with caution. Yet the large studies inform re-use policy across all levels of government, at the EU and in Member-States. Indeed, the PSI Directive is generic in that it applies to information of all types held across governments, with only some exceptions like for public broadcasters and education and research institutions.

The actual numbers then do not provide a very solid basis for policymaking. Yet as the meme of government information or data as the new gold or oil takes hold, the impressive estimates of older studies take on a life of their own and may get lost in translation as they travel outside their proper context. For example, the Berlin study presents the Dutch royal meteorological office KNMI as a success story. It states that KNMI’s move towards marginal cost pricing led to an 80% decrease in income and the creation of 300% more jobs. In effect, the price for meteo-data sets affected by the policy changes sank by 80%, not the income. The total turnover dropped by 60% but this is not attributable solely to the change in charging policy. The estimate was that some 100 FTE jobs were created in the private sector over a period of eleven years (1999-2010), with a 15

49 G. Vickery, Review of recent studies on PSI re-use and related market developments, Information Economics, Paris 2011.
Benefits of Public Sector Information re-use

million EURO growth in turnover. As part of the policy change KNMI abandoned certain activities such as providing detailed weather forecasts for the public. It left the provision of these services to private sector companies eager to expand their markets. In the same period 84 FTE were lost at KNMI, and so was 40 million of turnover. Before it was disbanded, the commercial arm had provided 25 jobs. In terms of net jobs created then, a figure of ‘300% job growth’ obviously sounds more impressive than what really took place, namely the net creation of between 75 and 16 jobs in more than a decade.

Further, what studies share is a focus on the potential benefits of opening up information. The evidence that there is actual (substantial) uptake by the private sector and that the presumed economic benefits materialize is thin and anecdotal. Here also, the KNMI example is instructive. It is a flagship example that features in MEPSIR, POPSIS, the Berlin ‘Digitales Gold’ study and Vickery’s OECD report. Although it is safe to say the policy change of KNMI had positive welfare effects, the size of these – beyond a modest amount of job creation- remains unclear. The head of the KNMI itself publicly expressed disappointment with the uptake of open meteorological data in a talk with the telling title ‘Where are the innovative weather products?’ He argues the policies to open up raw data are too much technology and supply-side driven. More than a decade after the meteorological institute opened up its data for broader re-use, the development of sustainable business models by the private sector remains a challenge.

What is important to remember about the large EC commissioned studies and a number of others, is that they do not give estimates for the value of re-use within the meaning of the PSI Directive. Rather, as we have seen, the studies primarily look at the size of markets or share in the economy of industries that provide information services or goods that are (in part) based on government data. Crucially, this includes public sector

50 POPSIS study p. 268-270.
body spending on information that is necessary for the exercise of public tasks (including spending for outsourced production and services).

Why do studies consider the overall market size and broader economic benefits of information production? It can be the case that government agencies involved in large-scale production of information feel it is in their interest to highlight the particular contribution that their activities make to the economy and public policy. In the value of public sector information lies a justification of the public tasks associated with its collection and dissemination; in times of austerity and political ambition to reduce the size of government, extolling the virtues of public sector information ‘markets’ can be a survival strategy for government agencies.\footnote{Defending its policies, even very existence was obviously a motive behind the UK’s Ordnance Survey report Oxford Economic Research Associates, ‘The economic contribution of Ordnance Survey GB’, London 1999. The recent study initiated by the US statistics departments also may well have been informed by the need to publicly justify government spending on statistics. See U.S. Department of Commerce, Fostering Innovation, Creating Jobs, Driving Better Decisions: The Value of Government Data. Washington DC 2014.}

### III. Open government studies

The purported benefits of opening up government information are manifold. The influential studies discussed above mainly target the economic value of public sector information for the economy as a whole and the private sector in particular. They are not only steady suppliers of component parts of the public information infrastructure, but are also expected to develop new (value added) services and products, with public sector data as ‘raw material’. In this section, we highlight a number of other rationales for re-use of public sector information and the outcomes of research in corresponding domains.

Re-use policy has become a key part of broader ‘open government’ agendas. In the US, the open government agenda means ‘working to ensure the public trust and establish a system of transparency, public participation and collaboration’.\footnote{https://www.whitehouse.gov/open.} The European Commission takes it to mean ‘opening up public data and services and facilitating collaboration for the design, production and delivery of public service’, as well as ‘making government processes and decisions open, in order to foster citizen partic-
Benefits of Public Sector Information re-use

Of note, in the current open government discourse the term ‘open data’ is more common than the term re-use. Open data is data that is easily re-useable, because it is made available without technical or legal constraints (i.e. intellectual property rights, terms and conditions). In effect, it is a particularly liberal way of opening up government data for re-use purposes. The model has the support of the European Commission, as is clear from the EC licensing guidelines of July 2014.

Not surprisingly maybe, given its internal market background, the PSI Directive itself only makes cursory reference to political or wider societal benefits of access and re-use. The G8 Open data charter of 2013 however, as endorsed by the EU, lists a multitude of benefits. Open data it says, leads to transparency and awareness, thus ‘promotes accountability and good governance, enhances public debate, and helps to combat corruption.’ It can empower individuals and other actors ‘to fuel better outcomes in public services such as health, education, public safety, environmental protection, and governance.’ It serves as ‘catalyst for innovation in the private sector, supporting the creation of new markets, businesses, and jobs’. Furthermore, releasing data as open would ‘help to improve the flow of information within and between countries’.

Better and more efficient public services, more transparency, more participation of citizens in political and social life and development of new innovative applications also feature in EC policy documents and speeches.

56 Recital 16 of the PSI Directive states: ‘Making public all generally available documents held by the public sector - concerning not only the political process but also the legal and administrative process - is a fundamental instrument for extending the right to knowledge, which is a basic principle of democracy. This objective is applicable to institutions at every level, be it local, national or international.’
One purported benefit of governments actively disclosing information is that it makes them and the institutions they fund work more efficiently because it helps performance monitoring and compliance. The ever-watchful public would take on the role of armchair auditor. On the political side, access to information and particularly active disclosure of it, is also considered to be of crucial importance for upholding the legitimacy of public decision-making. Legitimacy requires trust and citizen participation in democratic processes.

To measure efficiency gains in public service delivery, increased citizen participation and better accountability, different methodologies are required. The following sections sample various strands of research, to give an idea of how multifaceted the work is that is being done in the larger ‘open government’ field. The research is grouped in transparency studies (e.g. on effects of transparency), motivation studies (e.g. on what drives people re-using government information or participating in e-initiatives) and projects which focus on ranking countries or regions on ‘open government’ performance.

1. Transparency studies

Academics increasingly try to map transparency effects. So much so that ‘transparency studies’ is fast becoming an independent multidisciplinary field of research. Active dissemination of government information is regarded as a means to increase (or restore) trust in government which is...
Benefits of Public Sector Information re-use

important to ensure legitimacy. This idea is both defended and challenged. Some opponents argue inter alia that more disclosure may actually lead to less trust in government.61 In addition, the capacity of citizens or citizen groups to harness information and understand its meaning(s) is an important consideration in assessing effects. Open data in its purest form—raw, lowest level of granularity, complete, machine-readable—takes work and interpretation to transform into useful information. This means its immediate usefulness is limited.62 Another problem is that the data available (e.g. on government spending or budgets) is not organized in such a way as to answer questions the public has: the ‘my data can’t tell you that’-problem.63

In public administration research, empirical studies are still not commonplace. There is however a growing body of experimental research that tests causal relationships between different forms of transparency, the quality of decision-making processes, citizen trust and participation. By way of example, one study is recounted here. Inspired by theories that argue for and against a causal relationship between government transparency and citizen trust, the author set out to test assumptions made in transparency theories through empirical research.64

A series of experiments were run to test the effects on citizen trust of the provision of online information on a municipal website about environmental policy. Aspects of trust in government were identified based on


existing literature. They are competence, benevolence, and honesty. The hypothesis was that if citizens have more specific knowledge, this has a positive effect on the trustworthiness of the body as perceived by citizens. Various dimensions of information were part of the design: completeness of information (levels ranging from none to full), usability of information (comprehensible, timely) and the framing (positive, balanced or negative policy messages). The outcome in this setting was that the overall effect of transparency on trust is very limited. The act of transparency itself positively affected how honest people felt government to be. Whether people believe government to be benevolent (i.e. acting in the interest of the people) hardly depends on levels of transparency. That is, unless bad policy results are presented, as this adversely affects perceived benevolence as well as competence.

No generalizations can be made from this particular experiment of course, and the researcher is meticulous in discussing the possible weaknesses in the experiment design and caveats of generalization. However, it does show that there is no simple causal relationship between active disclosure of government information and citizens’ trust. Re-use policy is increasingly based on numerous assumptions about the benefits that opening up government information brings in the political domain. Results like those of the Utrecht experiment raise important questions: on the robustness of assumptions driving open government agendas, on whether assumptions lack context-specificity, on what the quality of assumptions tells us about the effectiveness of policies.

2. Motivation studies

One strand of research directly targets the question what motivates citizens/organizations to participate in 'open government' projects and what the success factors are. The presumption of policymakers (and open government advocates) is that opening up information and asking for response will increase participation in public-decision making and lead to improved public service delivery.

---

65 Competence meaning: how do citizens perceive competence of a particular public sector body/actor? Benevolence is the term used as shorthand for whether citizens perceive of a public authority as working in the public’s interest.
Davies (2010)\textsuperscript{66} analysed the use made of government data released through data.gov.uk, the UKs central data portal. He particularly looked to motivations of users and to what type of democratic engagement and public sector reform is actually supported by open government data. Technology developers appeared to be primary users. In Davies’ sample, users had various motives for engaging with open data; but many shared a technology oriented focus towards public sector reform. Developing proof-of-concept innovations and using government data to deliver services to public authorities were common drivers, but using data to engage directly in public planning was not. How robust and representative these outcomes are is difficult to establish. By academic standards the research design and its limitations are not set out clearly.

In a different setting, researchers did a small survey (161 responses) on citizens’ willingness to participate in German ‘open government’ projects. Three projects were picked to represent different types of involvement: collaborative democracy (asking citizen input for funding decisions), crowdsourcing (citizens identify and report road problems like ‘potholes’) and citizen assisted innovation (asking the public for input to solve administration problems). The finding was that respondents did not seem particularly technology oriented. They were also less willing to participate if they thought the project was too complex or felt they lacked to knowledge to meaningfully contribute. Fun seemed to be an important motivational factor.\textsuperscript{67}

The findings of these examples tie in with observations made in public management literature, which take issue with the ‘linear’ mode of thinking that characterises much policy documents.\textsuperscript{68} How citizens respond, what their motivations are, whether the parameters for success are in place is much more context specific than grand political statements on the effects of ‘open government’ suggest. If much depends on context, one would ex-


pect that policies too should be tailored to local circumstances. Yet the growing body of instruments for ranking ‘openness’ suggests otherwise, as we shall see below.

3. Societal impacts and rankings

The strand of research that seeks to develop models for measuring the impact of open data and that allow for comparison of geographic areas and societal domains takes place mainly largely outside academic settings. Some projects emerge from international governmental organizations like OECD and the World Bank. Others are driven by the IT sector (e.g. the World Wide Web foundation).69 Interestingly, international civil society groups, sometimes working alongside industry or governments, play a large role. Only a few are run by universities. Some initiatives focus on developing countries, others on transparent government more generally. For example, the Canadian funded ‘Emerging impacts’70 multi-year study comprises a host of national projects which monitor the effect of e.g. opening up of government budget data, or awareness and uptake in the judiciary setting or governance in education. The OECD also runs a project on improving impact assessments for countries so that effects of open government data initiatives on economic, social and good governance values can be measured. The World Bank developed transparency indicators for use in various rating instruments, e.g. the CPIA71 transparency, accountability, and corruption in the public sector rating or the Easy of doing business index.72 The Right to Information Rating run by Access-Info Europe and the Centre for Law and Democracy assesses and ranks the strength of the legal framework for access to government information in

---

69 My focus is on organizations/projects concerned predominantly with policy, not on organizations working on technical aspects of accessibility (for example standards for linked open data by the W3C, Opengeospatial consortium).
70 http://www.opendatasearch.org/emergingimpacts.
71 CIPA stands for ‘Country Policy and Institutional Assessment’.  
Benefits of Public Sector Information re-use

over a 100 countries. Its assessment is based on some 60 indicators; many relating to characteristics of domestic freedom of information acts. The project is concerned with access to information as a political and civil right, not the economic re-use dimension.

The Open Government Partnership, a ‘coalition’ of civil society and government administrations also measures how participating countries do on various aspects of open government. The partnership has grown rapidly in the space of four years of its inception in 2011, to a membership of over sixty countries. Its objective is to be ‘international platform for domestic reformers committed to making their governments more open, accountable, and responsive to citizens’. Member states submit action plans in which they make commitments, e.g. on datasets to be made available as open data, on citizen participation and improvement of public service delivery. They submit self-assessments regularly. Reviewers for the ‘Independent Reporting Mechanism’ (IRM) report annually on the fulfilment of commitments.

To track the adoption and actual usefulness of open data policies at national levels, the Open Knowledge Foundation uses a standardized tool to produce an annual ranking of countries, based on details about the availability of what it considers ten key datasets. A number of these will be produced by public authorities in any country (e.g. legislation, national statistics, election results, national budget), but others might be produced or managed by semi-public or private entities (e.g. postcodes, public transport timetables) or concern private sector performance (education, health). The ranking therefore has an (implicit) normative load about what information production and dissemination activities should be public sector tasks, and how these should be funded namely through general purpose taxation rather than charging (re)users. The World Wide Web foundation published its first ‘Open data barometer’ in 2013. It scores almost 90 countries on open data ‘readiness’ and implementation of open data policies, as well as impacts measured through the proxy of media and academic mentions of data use and impacts. The implementation score is based on the availability as open data of 15 different types of datasets across government domains, from education to health, transport, budget and public procurement. The open data barometer –like the OKFN index— takes a

75  https://okfn.org/about/our-impact/open-data-index/.
76  http://www.opendatabarometer.org/.
political stance, e.g. posing that company registries and land ownership data should be disclosed at no cost to the user.\footnote{See the 2015 report http://barometer.opendataresearch.org/report/analysis/implementation.html.}

Another example of an international comparative tool is the Open data 500 project. The GovLab at New York University initiated the Opendata 500 study.\footnote{http://www.opendata500.com/}. Its aim is to comprehensively list US companies that use open government data in their products and services to both the public and private sector. The type of companies included range from suppliers of software to public sector bodies (e.g. Adobe, SAP), legal publishers, management consultants (e.g. Deloitte), credit rating agencies, to companies offering healthcare advice, or providing location services which are in part based on government procured satellite imaging. GovLab set up the ‘OD500 Global Network’ to enable researchers from other countries to do a similar mapping exercise. According to the founders, this should make it possible to compare but not necessarily rank countries and arrive at a better understanding of the impact of open government data. Of note, the lists compiled are by no means limited to companies that somehow ‘re-use’ public sector information within the meaning of the PSI Directive. Many of them supply services to governments that the latter need as part of their normal operations (e.g. the human resource systems supplied by SAP, or document management software). The Opendata 500 mapping exercise is very illuminating as it provides some insight into the highly complex web of information relationships between private and public sector. It cannot however tell us anything about the actual effect of open government data or its less liberal re-use alternatives.

4. Value of open government studies

Obviously, the above overview does not do justice to the wide range of research activities that I have conveniently placed under the heading of ‘open government studies’. Their existence is important to note for a number of reasons. First, because of a shared characteristic: the focus lies not on the size or development of public sector information markets, but on the (presumed) benefits of transparency for the political-democratic process and the performance of public sector bodies. As we have seen, the
benefits of public sector information re-use

initial policies of the EU were firmly rooted in the (direct) economic effects of giving the private sector access to public sector information. However, over time fostering the commercial exploitation of government information has become but one objective of what today is termed ‘open government’. Consequently, the further development of re-use policies needs to be informed by a wider set of considerations, and open government studies are an essential source.

Second, the ‘open government studies’ shed light on the multifaceted and complex relationships that lurk under seemingly straightforward concepts like open data, transparency and open government. It can help lawyers and policymakers unpack the many countervailing interests and competing norms, in order to find ways to develop appropriate legal frameworks. Third, by querying and testing commonly held assumptions about societal benefits of open government, transparency studies especially provide insights for making successful information policy. Fourth and final, the ranking and indicator projects testify to a development of semi-formal assessment and review mechanisms for public sector information policies.

What is striking about this development is not just the important role that non-governmental bodies play or how diverse the interest groups involved are. That the initiatives are global in scope sets them apart from the dominant regulatory model for public information policy, which is national (especially access to information law) or European (right to re-use).

IV. Twenty-five years on, where is re-use policy heading?

The potential of public sector data re-use is celebrated as much today as it was twenty-five years ago. The European Commission’s 1989 Synergy Guidelines exemplified the attempt to stimulate information industry growth by getting public sector bodies to unlock data resources. Today, rather than speaking of commercial exploitation, or re-use of public sector information, ‘open government data’ is the common shorthand for policies aimed at improved access and use conditions for public sector information. In the past five years, accountability, participatory democracy and better public service delivery have joined private sector growth as objectives served by re-use policy. In recent communications, the European Commission presents re-use and open data as part of an even larger concept: big data. Riding the hype cycle wave, the numbers put forward to justify re-use policy seem to grow rather than shrink. In October 2014, shortly before she handed over office, the Vice-President of the European
Commission Neelie Kroes signed the ‘big data’ deal, a Big Data Value public private partnership involving companies and research institutions that aims to ‘build a thriving data community in the EU.’ The EC committed to investing alongside IT industry, which should “put Europe at the forefront of the global data race”, making European companies take 30% of the global market, and result in the creation of 100,000 data related jobs. The availability of high quality datasets from the public sector is seen as an important part of the EU big data agenda. Echoes of the 1989 Synergy guidelines abound.

The commercial exploitation of public sector information thus remains a factor in economic growth predictions. However, it might become a less visible policy objective due to the big data hype. The role of opening up data to help public sector make savings and deliver better services has gained lots of traction – as is evident from the ‘open data’ agendas of the G8, World Bank and the European Commission’s data value chain approach to public sector information. Political-democratic objectives – cementing citizen trust and participation – are as we have seen also new additions to re-use policy.

Not surprisingly, because so many ambitions, objectives and interests travel on the public sector re-use and open government bandwagon(s), confusion over precise meanings is rife and assessing whether policies actually produce the desired effects is a herculean task. Further, what the pursuit of multiple disparate objectives implies for the existing regulatory framework is unclear. It could well be that because of its generic one-size-fits all approach the Public Sector Information Directive has reached its limits. To conclude this contribution, a few observations are made on this triad of problems: multiplicity of objectives, a generic legal instrument, and uncertainty about if and to what extent benefits ascribed to re-use pol-

79 [http://www.bigdatavalue.eu].
Benefits of Public Sector Information re-use

Policy materialize. It will lead to two main conclusions. First, that policy makers and legislatures should be much more specific (and modest) about the objectives to be attained, and rely more on evidence of effectiveness and less on assumptions about effects. Second, that if re-use policy is to serve different masters, it must be tailored to and integrated in other domains that more directly serve those objectives.

1. Persistent lack of hard data

The economic studies overall show there are direct economic gains to be had from improving access and re-use possibilities. A lack of hard data on the actual size of effects persists --or on the size of public sector information markets for that matter. Grand claims about the enormous benefits in terms of new products and services, growth in data driven jobs and efficiency savings in the public sector should therefore be taken with proper amounts of salt.

In the economic studies, to focus on overall market size and large public sector information providers masks the need for suitable instruments for cost-benefit analysis at the level of individual public sector bodies or information holdings. It is unclear whether the findings for large information producers (like statistics, mapping and meteorology) can be generalized to all information types. An implicit assumption of the Public Sector Information Directive is that it can; hence its obligation to allow re-use and do so in principle at no or marginal cost and with minimal re-use restrictions.

Furthermore, if the studies as is argued above do not estimate the value of opening up public sector information for re-use within the meaning of the PSI Directive, the same can be said for ‘open government data’, which after all is the most liberal kind of permitted re-use. As we have seen the Commission Guidelines on licensing and charging champion the disclosure of government data as open data. Given this trend, it is important to note that the value of PSI markets as estimated in the economic studies is by no means the same as the economic benefits that result of opening up government data in electronic, structured form for any re-use purpose, with as no conditions imposed or fees charged. Yet EU, national and local

---

authorities alike tend to phrase re-use policy as open data policy, and continually justify open data policies by pointing to the economic value of public sector information markets.

The indicators and rankings so popular among international governmental organizations and non-governmental organizations driven by civil society and industry might well serve as a PR stick that spurs governments into action. However, they are no measure of effectiveness. In the end the question is however whether the (legal) instruments and actions achieve the objectives: do companies create value added services, do citizens take part more actively in public decision making, do inefficiencies in government lessen, does fraud decrease? We only have anecdotal evidence.\(^8\)

Above I already noted that the much applauded re-use policy of the Dutch royal meteorological institute has by its own admission not led to much innovative private sector products. The open data policy initiated by president Obama in the US is hailed as a success, but critics argue this claim is not supported by proper empirical data.\(^8\) We have seen also that in the nascent discipline of empirical transparency studies, experiments show that the positive relationship between transparency of government information and citizen trust is not nearly as straightforward as conventional wisdom would have it. Social impact studies are at an early stage and the claims of benefits of ‘open government’ on for example citizen participation in policy debate need further research. Nonetheless, within political sciences and public management studies a body of empirical work is evolving that should be taken into consideration in policy design and the (regulatory) implementation.

In light of the lack of hard data on the kind and size of potential benefits size and actual effects, it would seem sensible for policy objectives to be phrased more precisely and tailored to local possibilities and needs. As Meijer et al argue in the context of enhancing political-democratic val-

---

\(^{8}\) The Deloitte Market Assessment study (2013) observes that one of the key gaps in the evidence base is ‘knowledge about how public sector information datasets are being used and re-used by customers and others downstream’, p. 38.

Benefits of Public Sector Information re-use

2. From generic stand alone to a coordinated approach

The Public Sector Information Directive has reached full maturity. Because it sets down standards that apply across an incredibly large domain of public sector activity at all levels of government, it seems inescapable that the different modes of re-use regimes will be elaborated side by side and at different speeds as part of sector specific information policy.87 This is possible in many areas.

One of the better-known examples is the INSPIRE88 programme. It was set up initially to develop an EU wide infrastructure for spatial information in the service of EU environmental policies, but that has grown into a body of technical, organizational and legal standards which serves more diverse policies, although (commercial) re-use is not yet an integral part.89 This may be explained because the INSPIRE focus originally was squarely on the exchange of information in the context of public tasks. Re-use policy is already becoming part of a wider public sector innovation agenda. The EU e-government initiatives are an obvious area where further integration can take place, e.g. where it concerns so-called base or key registries, or transport data, public spending data and legal information.90


87 A similar point is made in the EC Legal Aspects of Public Sector Information network position paper D. 2.2: M. van Eechoud et al., LAPSI 2.0 Thematic Network D2.2 – Position paper access to data. Brussels 2014. Available at [http://www.lapsi-project.eu].


89 See [http://www.inspire.ec.eu].

There is growing agreement over the need to link base registries and core data sets (e.g. company information, cadastral data, statistics) of Member States for use in and outside the public sector, but overcoming the financial, legal, technical and organizational hurdles proves to be difficult.

In the area of company data – regarded in all studies, by governments and industries as of high re-use value – the EU has also legislated to improve public access to cross-border company information. However, the directive does not mandate public access for re-use purposes. As part of the EU's anti-money laundering policy, the European Parliament’s insisted that the Fourth Anti-Money Laundering Directive should oblige Member States to keep public registers on (beneficial) ownership of companies. In the end, these registers will not be fully open, since users have to show a legitimate interest to gain access. This would give for example investigative journalists access, but not members of the public.

A final instructive illustration concerns satellite data. This is one type of data that seems to consistently be much more widely used once access is opened and charges are removed. This was already apparent from the US ‘Landsat’ case where the outsourcing of the production and distribution of earth observation data adversely affected use levels. Even in areas like these, where it seems to be common opinion that public funding coupled with releasing data liberally is beneficial to the economy, it does not follow that the EU succeeds in getting this organized. When the newly appointed Commission presented its plan for deregulation and published a list of instruments under development that were candidates for repeal, one

---

91 See e.g. the E-government action plan [ec.europa.eu/information_society/activities/egovernment/action_plan_2011_2015/docs/action_plan_en_act_part1_v2.pdf].
of the pending proposals was the directive that should have made high-resolution satellite data (earth observations) available for commercial exploitation.\footnote{Proposal for a Directive of the European Parliament and of the Council on the dissemination of earth observation satellite data for commercial purposes. COM(2014)0344 final. See also ‘Report on Progress GEEO 2011-2013, Ministerial Summit Geneva 2014’. GEO (Group on Earth Observations), 2014.}

From an initial focus on direct economic impacts and a relatively stand-alone approach at EU level, re-use of government information has now become part of public sector innovation and transparency agendas in many domains. Such agendas are pursued not just at the EU level, but also at local, national and international levels of governance. The EU examples above then show that in the stages of development in the legal framework for re-use are likely to be decentralized and involve complex balancing acts. All the more reason for policymakers to advance well-argued, evidence supported re-use objectives, rather than mere grand political claims about the incredible benefits of ‘open government’. 

\textit{Benefits of Public Sector Information re-use}