



CIP-ICT PSP-2012-6

Grant agreement No 325171

LAPSI 2.0 Thematic Network

D2.2 – Position paper access to data



Deliverable number/Name: D2.2: Position paper on Access to Data

Dissemination level: Public

Delivery date: 12 December 2014

Status: Final

Author(s):

Coordinating author:	Mireille VAN EECHOU (IViR/University of Amsterdam, NL)
Contributors	Sam VAN VELZE (IViR/University of Amsterdam, NL) Marco CASPERS (IViR/University of Amsterdam, NL) Sarah Johanna ESKENS (IViR/University of Amsterdam, NL) Paddy LEERSEN (IViR/University of Amsterdam, NL) Linda AUSTERE (Centre for Public Policy PROVIDUS, LV) Heather BROOMFIELD (DIFI Agency for Public Management and eGovernment, NO) Cécile DE TERWANGHE (CRIDS, University of Namur, BE) Jo ELLIS (National Archives, UK) Georg HITTMAR (PSI Alliance) Federico Morando (Nexa Center, Turin, IT) Maja LUBARDA (Information Commissioner, Slovenia) Bastiaan VAN LOENEN (Technical University Delft, NL) Cristiana SAPPÀ (ICRI-iMinds, University of Leuven, BE) Marc DE VRIES (Citadel Consulting, NL)

Executive summary

Recommendations

Promoting re-use friendly access rights in MS

Although statutory instruments that grant public rights of access to government information are a backbone of re-use frameworks, such freedom of information laws are not especially tailored to helping public sector information to be discoverable and usable. The EC could promote re-use friendly reform of national/regional access laws, but there are a number of alternative routes that hold promise for faster results. Leading by example through the progress made by the EU, France, Italy and Germany in the context of the G8 Open Data Charter is one way; encouraging EU Member States that take part in the Open Government Partnership to more consistently include re-use promoting initiatives in the national action plans another.

Re-use as a standard concern in the EIF / e-government programmes

The EU develops policy and regulates domain on a broad number of domains with impact on access to and re-use of public sector information, e.g. in the field of statistics, transport, public spending, spatial information (notably through INSPIRE¹), basic registries and legal information. The removal of legal barriers to re-use should ideally be a standard element of such programmes, and especially of the legal interoperability agendas pursued as part of the European Interconnection Strategy and the EI Framework (EIF).

Develop deeper understanding of the impact of sector specific national laws

Apart from access laws, Member States have a host of other laws which regulate specific types of public sector information, notably registries in the areas of e.g. companies, land, electoral rolls as well as activities of shared facilities e.g. statistics, meteorological data and spatial information. The network has not studied these as it would have required research resources substantially beyond LAPSI's resources which do not cover research activities. However, from informal discussions held in the network it appears that the current state of regulation has profound (in some countries/domains enabling, in other disabling) impact on PSBs ability to open up data for re-use. It will be valuable if at least for such datasets as have already been identified as 'high value' for re-use, possible improvements of the legal framework for access and sharing are identified. The development of a comprehensive vision on the role of legal norms in realizing the economic and social potential of wider use of PSI will require a better understanding of current domain specific laws and their interactions, both in member states and at EU level.

Develop and adopt forward looking data protection safeguards

The PSI Directive recognizes that data protection rules and generally law protecting privacy must be adhered to, but how this can be done is an increasingly difficult question to answer. If public sector bodies adopt the precautionary principle, that is: when in doubt about

¹ Directive 2007/2/EC establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)

downstream data protection implications, do not release data for re-use at all, then many types of data might not be released. This is especially true in cases when (downstream) linking of data, which taken in isolation are not personal data as such, causes individuals to be identifiable and the combination of the data thus represents personal data.

The challenges are not only posed by technological developments (de-anonymisation techniques, linking of data from multiple sources). There are some fundamental tensions between core principles of data protection law and the notion of re-use, especially re-use as open data. More effort seems needed at the short term practical level, notably raising data protection awareness and developing mechanisms that help public sector bodies comply with data protection rules in the pursuit of implementing re-use policy. Of note, the move in recent years towards promoting open data as the re-use model to aspire to does not necessarily help to ease the tension between access/re-use on the one hand and data protection concerns on the other. Enabling anyone to use PSI for their own purposes – a central element of open data policy – implies that no terms/restrictions can be set in the interest of data protection. Such restrictions however are mandated by the PSI Directive and can very well comply with its non-discrimination principle. To the extent that open data licensing is promoted as the ideal to aspire to for PSBs, it might cause substantial numbers of datasets to be excluded from re-use. It would seem therefore that more moderate policies, or possibly enriching open licences with data protection terms can help PSBs overcome data protection concerns and more readily release data for re-use.

The underlying fundamental tensions need addressing too, through legal research backed by expert stakeholder input. Mapping the competing interests involved (e.g. public interest in economic growth, better service delivery, and accountable government) and establishing suitable balancing mechanisms that can be operationalized in instruments to aid public sector bodies in elaborating and implementing re-use policies. The upcoming Data Protection Regulation will of course be of pivotal interest.

1 Rights to information as re-use drivers

1.1 The Public Sector Information Directive at a glance

The Re-use of Public Sector Information Directive of 2013² (the PSI Directive) sets out a general framework for the conditions governing the right to re-use information resources held by public sector bodies, which includes provisions on non-discrimination, transparent licensing and the like. It is the main EU legal instrument for stimulating the creation of value added information products and services (tools, apps, content) that take public sector information or data as a (or the) source. The objective of the PSI Directive is to achieve through minimum harmonization of national rules and practices a more level playing field across the EU/EFTA, which allows European companies (and citizens) to exploit the full potential of re-using data produced by or for the public sector. As such, the PSI Directive is a key instrument for open data policies at all levels of government in Member States. Open government data are a feature of most, if not all, open government agendas, as is evident from the action plans submitted by the EU and Member States in the context of the Open Government Partnership.³ *Open* public sector data involves making machine-readable data (sets) available for any re-use without restrictions or licensing fees.⁴ As such it is a subset of the broader categories of re-usable information that the PSI Directive regulates.

In a nutshell, the PSI Directive requires that documents held by public sector bodies are made available for commercial and non-commercial uses, but only when there are no third party intellectual property rights involved and when the information is public under local access regimes. The PSI Directive requires that permission to re-use is granted through non-discriminatory and transparent terms. Any fees must be cost-based and pre-established; the default principle is that re-use is allowed against at most the cost of dissemination. The accompanying Guidelines on licensing favour the use of open, liberal licences,⁵ in keeping with the licensing guidelines drafted by the LAPSI network.⁶

1.2 Role of FOIA in regulatory framework

Re-use of government information naturally requires access to information. However, the PSI Directive itself does not oblige Member States to provide access. This can be explained by the fact that the legislative competences of the EU to regulate access to public sector

² Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information; also see Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information. A consolidated version of the PSI Directive as amended is available via <http://eur-lex.europa.eu/lexuriserv/lexuriserv.do?uri=consleg:2003l0098:20130717:en:pdf>

³ Updated list of participating countries and action plans, see: <http://www.opengovernment.org>

⁴ See the OKFNs open definition at <http://www.opendefinition.org>

⁵ EC Guidelines on recommended standard licences, datasets and charging for the re-use of documents (2014/C 240/01).

⁶ <http://www.lapsi-project.eu/sites/lapsi-project.eu/files/D5.2LicensingGuidelinesPO%20%281%29.pdf>.

information held by authorities in MS are limited.⁷ As noted, the PSI Directive therefore applies to ‘documents’ that are *already* publicly accessible under national rules for access to documents (art. 1(3) PSI Directive). It does not add to them or change them. What it does do is *oblige* Member States to allow re-use of such documents.⁸ In these terms, the PSI Directive of 2013 is a big step forward compared to the PSI Directive of 2003.

The focus of the Good practices analysis done by LAPSI⁹ was on access to (official) information laws at Member States level, that is, on ‘freedom of information laws’ or FOIA (increasingly referred to as ‘right to information’ laws). Freedoms of information laws are the most generic instruments in most national jurisdictions and in the past decade or so they have spread across the EU.¹⁰ This is why the network considered it important to gain insight in what particular attributes of freedom of information laws contribute to facilitating re-use.

2 Key findings good practices survey

Freedom of information laws were assessed along three dimensions¹¹, namely their contribution to: making PSI discoverable, because re-users will need to know what information or data public sector bodies hold and what its characteristics are; making PSI **available**, i.e. the inclusiveness of different information types and institutions as well as the conditions for access; and the contribution of FOIAs to ensuring data is **usable**, that is, providing norms that address e.g. format, machine-readability, granularity, timeliness, completeness, accuracy, or delivery modes.

The Good practices inventory and analysis took the perspective of re-users, and in particular worked from the assumption that ‘good’ access laws:

1. Ensure the widest possible access to resources
2. Limit the restrictions to what can be done with the information;
3. Provide legal certainty on what uses can be made;
4. Respect user preference with respect to e.g. format of supply;
5. Help to reduce search costs;
6. Give re-users a voice in decision-making on what data are actively released and how.

⁷ LAPSI 1.0 Policy Recommendation N. 6, Rights of Access to Public Sector Information (hereafter: Policy Recommendation 6), available via <http://www.lapsi-project.eu/outputs>.

⁸ The implementation date for this rule is Jan. 2015

⁹ [Http://www.lapsi-project.eu/sites/lapsi-project.eu/files/LAPSI_D2.1_GoodPracticesAccess%28final%29.pdf](http://www.lapsi-project.eu/sites/lapsi-project.eu/files/LAPSI_D2.1_GoodPracticesAccess%28final%29.pdf)

¹⁰ Especially the Member States in Central and Eastern Europe now score well on right to information ratings, see <http://www.rti-rating.org/files/docs/Report.13.09.Overview%20of%20RTI%20Rating.pdf>

¹¹ These three dimensions were already among those recognized by the Dutch government when it first started to develop an integral policy for opening up public sector information in the late 1990s. See Records of Parliament (Tweede Kamer) 1996-1997, 20644, nr. 30; and in A. Beers, *Informatica Publica*. Den Haag: Rathenau Instituut 1996; see also M. Backx, *Gebouwgegevens redden levens*, TU Delft, at http://www.metameba.nl/images/Gebouwgegevens_redden_levens_web.pdf.

Points 1 and 2 are central aspects of FOIA. For points 2-3 licences and (public domain) notices are key instruments. The LAPSI network has produced licensing guidelines for Public Sector Bodies which identify essential building blocks for any licensing policy and outlines steps a public sector body can take to establish if and how PSI can be licensed.¹² Points 4 through 6 are aspects of discoverability and usability, and as we shall see not a particular strong part of access laws.

To summarize the Good Practices findings, freedom of information laws are concerned primarily with availability, and much less with discoverability. The use of web directories, portals and in some countries FOIA specific “electronic journals” aids discoverability, but overall firm statutory obligations to pro-actively disclose information about what PSI is held, or pro-actively disclose PSI holdings itself for that matter, are rare¹³. Where formal obligations exist or are being considered they are mainly geared towards making transparent what a given public sector bodies’ tasks, policies and organizational structure are. It is safe to assume that such information does not have much (economic) re-use potential although from the public sector perspective and in the interest of transparency it is ‘high value’ information. Of note, the absence of formal obligations appears to have limited impact on the roll out of web directories and portals (with repository function) in Member States.¹⁴

Many laws have some provisions that affect usability, but a guiding principle in most FOIAs is that data/documents are provided ‘as is’. At most a party seeking access may have a choice between available ‘formats’ (print, digital, particular electronic format). The PSI Directive itself encourages PSBs to make documents available in machine-readable format with meta-data, using open standards and by doing so goes well beyond what the average FOIA prescribes. Note that elements of availability that pertain to the procedure for getting access and re-use permission have primarily been addressed in LAPSI Deliverables D4.1 and D4.2 and are therefore not discussed here.

Below is the table that summarizes characteristic provisions in freedom of information acts that play a role in stimulating re-use. Items in light type are uncommon provisions in Freedom of Information Acts.

¹² LAPSI D5.2. Licensing Guidelines (Feb 2014). See also the July 2014 EC Guidelines on recommended standard licences, datasets and charging for the re-use of documents (2014/C 240/01), which equally advice the use of open licences.

¹³ A notable exception is in the data themes of INSPIRE (see section 4.2 of this paper).

¹⁴ M. de Vries et al., The Pricing Of Public Sector Information Study, Open Data Portals (Oct. 2011), noted a steady increase of data portal/registries during the 2009-2011 period; we assume this trend has continued quite separately from freedom of information / right to information legislation reform.

Table: Freedom of information laws characteristics conducive to promoting re-use of PSI (LAPSI 2.0 Good practices, June 2014)

Helps make PSI: Good FOIA practice:	Discoverable (known what PSI is where)	Available (what PSI public, how to get it, terms& pricing)	Usable (fit for purpose re-user)
Broad scope of information & bodies covered		√	
Few and narrowly described limitations		√	
Access for all (non-discrimination)		√	
Access to bulk & dynamic data			√
Quick response times			√
Pro-active disclosure duties	√	√	
<ul style="list-style-type: none"> • Online 	√	√	
<ul style="list-style-type: none"> • Open format 			√
<ul style="list-style-type: none"> • User led 			√
Search support	√		
Referral system	√		
<ul style="list-style-type: none"> • Information registers / portals 	√		
User preferred form			√
<ul style="list-style-type: none"> • Open formats 			√
<ul style="list-style-type: none"> • Machine readable 			√
No or low costs of access		√	
Efficient review access procedure		√	√

3 Alternatives to FOIA as instrument for re-use policy

From the above it is clear that although access is the at the basis of re-use regimes, access to information laws play a limited role in stimulating discoverability and usability of PSI for re-

use purposes. Discoverability is a recognized problem for cultural heritage institutions as well,¹⁵ although flagship projects like Europeana which make metadata on collections available for re-use under the Creative Commons public domain dedication are important exceptions.

As was noted in the Good Practices, current freedom of information laws serve primarily as instruments for re-active disclosure of existing documents on a case-by-case basis, in aid of making governments accountable. Pro-active disclosure obligations come second at best.¹⁶ Since the wheels of FOIA reform turn slowly¹⁷ and are driven by Member States themselves, it is therefore a legitimate question how much productive it would be for the EU to focus on FOIA as instruments of re-use or even of open data policy.

3.1 Knowledge of sector specific instruments

A point for future research are other types of legal instruments that regulate either active disclosure of certain PSI or rights to access certain PSI. It was well beyond the scope of this network (which is not a research network) to analyse what type of information obligations and access rights are 'out there' beyond freedom of information acts, to assess which ones are most pertinent because they apply to content that stakeholders identify as attractive for exploitation, and how they contribute to the creation of cross-border information products and services, whether commercial or not-for profit. The Best Practices examples on meteorology and companies registers only scratch the surface so to speak. EU and national policy makers cannot hope to promote efficient re-use policies if at the source it is unclear what data has (legal) status as publicly accessible. Of particular interest are 'key' or 'base' registries and datasets that are considered as core components of national information infrastructures (notably on persons, vehicles, land ownership, companies, procurement as well as spatial data under the INSPIRE directive). The ISA¹⁸ programme already addresses some 'legal' interoperability issues of base registers; although semantic and technical interoperability issues for cross-border exchange of register data dominate.¹⁹ Its focus however is on facilitating cross-border cooperation between public sector bodies and ultimately allowing better public service delivery, not on re-use. The impact of the PSI Directive does however show that ISA Work Program Access to Base Registries final report²⁰ advises PSBs operating base registers to opt for re-use enhancing business

¹⁵ This is at least suggested by the Curtis+Cartwright Consulting final report on *PSI re-use in the cultural sector* (CC462D011-1.1), May 2011, commissioned by the EC. [http://ec.europa.eu/information_society/newsroom/cf//document.cfm?doc_id=1148].

¹⁶ On the role of pro-active disclosure obligations, see Darbishire, H. (2011) *Proactive Transparency: The future of the right to information? A review of standards, challenges, and opportunities*, World Bank Institute, Washington.D.C. <http://siteresources.worldbank.org/EXTGOVACC/Resources/DarbishireProactiveTransparency.pdf>

¹⁷ Some countries do take drastic steps however: a new Greek law (number 4305/31.10.2014) for example obliges public sector bodies to register any datasets/resources that they hold and indicate for each resource whether it is exempt from the new default rule that all government information made public is freely re-usable [information supplied through LAPSI network, personal email of Greek Ministry of Administrative Reform and E-Government, Department of Permanent Law Codification.

¹⁸ Interoperability Solutions for European Public Administrations.

¹⁹ http://ec.europa.eu/isa/actions/documents/isa_1.2_d3.1_final_report.pdf

²⁰ ISA Work program *Access to base registries*, final Report, Feb. 2014 [http://ec.europa.eu/isa/documents/final-report_en.pdf].

models. The EU has also legislated to improve public access to cross-border company information²¹; but re-use appears not to be an aspect taken wholly on board yet.

3.2 High value datasets

The importance of assessing the “re-use friendliness” of legislative frameworks for access to public information resources outside the scope of FOIA is also obvious if one considers the ‘high value’ datasets identified for priority release in the EC Guidelines and in the G8 Open Data Charter. The latter²² identifies a number of areas with high value data sets that the G8 countries (EU members: UK, France, Italy) undertake to release for re-use purposes under the ‘open by default’ principle. The lists include company data, statistics, mapping, postcodes and transport data such as timetables of public transport: typically resources that are subject to specific laws and regulations, and agreements in public private partnerships. The EU has endorsed the Charter. EU funded research by Price Waterhouse Coopers has identified over 250 data sets at EU institutions that are considered ‘high value’ and could be prioritized for active release. The research took in account the results of a number of (national) public consultations, surveys, and reports by other consultants.²³ In the G8 EU Action plan the EU has committed not only to releasing high value datasets held by its own institutions, but also to encourage MS to do the same.²⁴ Although the EC PSI Guidelines (2014/C 240/01) recommend Member States release high value datasets and give guidance on how to do this to make them attractive for re-use (e.g. machine-readable form, online, with meta-data), the question whether this can be achieved within Member States’ existing regulatory frameworks is left undiscussed.

Promoting re-use as a tool toward innovative information services may require sector specific regulatory approaches. For example, the provision of multi-modal transport data and (other) traffic information services takes place in a complex and highly dynamic field of public and private actors, where (unfair) competition issues are a particular concern. This is why the authors of the final report on EU-wide real-time traffic information services (for DG MOVE) suggest that formalising rules for access and re-use of data sets might be desirable.²⁵ Further specification of rules in the PSI Directive is not a realistic option as it is an instrument designed to operate at a more abstract level across broad ranges of public sector information types.

²¹ Directive 2012/17/EU on interconnection of business registers, OJ 2012, L 156/1.

²² <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>. Technical annex at 6.2.

²³ PwC (I. Bargiotti et al.) , Value based prioritisation of Open Government Data Investments (EPSI platform topic report no. 2014/08) [<http://www.scribd.com/doc/238741083/Value-Based-Prioritisation-of-Open-Government-Data-Investments>] ; See also ISA report on http://ec.europa.eu/isa/actions/documents/isa_1.2_d3.1_final_report.pdf; http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information.

²⁴ EU Implementation of the G8 Open Data Charter (Oct. 2013) [<http://ec.europa.eu/digital-agenda/en/news/eu-implementation-g8-open-data-charter>].

²⁵ Van de Ven & Wedlock et al., *D5 – Final Report Action B - EU-wide real-time traffic information services*, Brussels July 2014, p.42-3. [<http://ec.europa.eu/transport/themes/its/studies/doc/2014-07-its-action-plan-d5-action-b.pdf>].

INSPIRE²⁶

The INSPIRE Directive (2007/2/EC) lays down general rules aimed at the establishment of the Infrastructure for Spatial Information in the European Community, for the purposes of Community environmental policies and policies or activities which may have an impact on the environment (art. 1 INSPIRE Directive). This requires measures that address exchange, sharing, access and use of interoperable spatial datasets and spatial data services across the various levels of public authority and across different sectors (recital 3 INSPIRE Directive). INSPIRE applies to a wide range of 34 content themes, from specific geographic reference data themes (e.g., transport networks, cadastral parcels, buildings, orthoimagery, elevation, statistical and administrative units) to environmental themes (e.g., geology, habitats and biotopes, human health and safety, meteorology, hydrology, oceanographic features).

All INSPIRE datasets and services are required to be pro-actively published through the INSPIRE geoportal at the European community level.²⁷ With the requirement to provide dataset and services metadata, the obligation to conform to INSPIRE data specifications, and the requirement to provide access through discovery, view, and download services, INSPIRE makes a very important contribution, in particular to the legal and physical attainability of public sector spatial information as well as to its usability.

4 Data protection

Possible the biggest challenge to re-use and open data policies will prove to be the need to develop instruments which ensure that data protection laws are complied with while maximizing re-use of PSI. From the work done in the first LAPSI network in 2011-2012 it was already apparent that current data protection law and re-use policy can be difficult to reconcile.²⁸ The 2013 Opinion of WP29²⁹ on open data and public sector information ('PSI') re-use proposed that part of the solution is for public sector bodies to engage in 'data protection by design and default' and perform data protection impact assessments so that data protection concerns are addressed prior to any release of PSI for re-use purposes. This is sound advice of course to help prevent (inadvertent) breaches of data protection laws, but does not deal with the some fundamental problems.

4.1 Key tensions under proposed Data Protection Regulation

²⁶ Based on Loenen, B, van, and M. Grothe, INSPIRE empowers re-use of public sector information (accepted for publication in International Journal of Spatial Data Infrastructures Research)

²⁷ The INSPIRE portal is available at <http://inspire-geoportal.ec.europa.eu/>

²⁸ See LAPSI Policy Recommendation 4: Privacy and Personal Data Protection (Cristina Dos Santos et al), 2012. The Recommendation sketches problems not only for the public sector body allowing re-use, but for the re-using parties as well. That analysis still holds true.

²⁹ Working Party on the Protection of Individuals with regard to the Processing of Personal Data (WP29), Opinion 06/2013 on open data and public sector information ('PSI') re-use, adopted June 2013. See also the earlier opinion 7/2003 on the re-use of public sector information and the protection of personal data - Striking the balance (WP 83).

For those public sector resources that contain personal data, making it available for re-use – that is for uses other than the public task it was collected for – is highly problematic. Using personal information for unforeseen purposes or in a different context easily breaches the purpose specification principle which is central to the Data Protection Directive (and the proposed General Data Protection Regulation – (G)DPR). Furthermore, disclosing personal data, being an act of processing personal data, requires either that the data subject gives consent or that the disclosure is lawful on one of the other enumerated grounds for processing.

The LAPSI 1.0 network Recommendation on data protection focussed on the situation under the current Data Protection Directive. The now advanced plans for a Data Protection Regulation suggest the challenge to reconcile the regulatory framework for re-use with data protection norms has, if anything, become bigger. This is not the place to analyse the implications of the proposal in detail, but by highlighting a few issues we can give an idea of the complexities involved. For one thing, public sector bodies might have less leeway than they do now, no longer being able to base processing on the ‘legitimate interest’ ground that is available under the Data Protection Directive.³⁰ Instead, disclosure for re-use purposes might only be possible if there is a legal obligation to do so, or if it is part of the performance of a task carried out in the public interest (which must be clear and precise).

Also the Data Protection Regulation will likely contain a fairly specific test for establishing whether a use is in conformity with the purpose of initial collection, taking into account reasonable expectations of the data subject and the consequences of the intended further processing for data subjects. Such case-by-case consideration of the effects of making data available for re-use is hardly compatible with the access for all; re-use for any purpose principles that are key to re-use and especially to open data policies.

The pervasive application of data protection rules to all types of PSI that until a few years ago would not be considered to count as personal data, also has major implications for re-use policy. In light of the ever increasing ease with which combining data leads to (near) identification of persons, more data held by the public sector will be considered personally identifiable information, which means disclosure (and downstream use) must be (G)DPR compliant. In addition, as the WP29 Opinion 05/2014 on Anonymisation Techniques shows, advances in de-anonymisation techniques make it increasingly difficult to robustly anonymise personal data.

The LAPSI network takes no position on whether the broad application of data protection rules is undesirable from the perspective of stimulating PSI re-use. The PSI Directive itself recognizes that data protection must be respected, and this is also made explicit in many implementing laws in Member States.³¹ It seems likely that the (G)DPR will contain a very explicit consideration on the primacy of data protection law over the right to re-use public

³⁰ According to the a legitimate interest must be an interest that is: lawful, sufficiently clear to allow balancing against privacy interests, reflect a real and present interest (not future or theoretical). See Article 29 Working Party, Opinion 6/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, adopted on 9 April 2014 (WP 217), pg. 25, available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf (last accessed 9 December 2014).

³¹ See the overview in the LAPSI 1.0 Policy Recommendation on Data protection.

information as enshrined in the PSI Directive.³² At the same time, the latest proposal discussed in the Council still provides that the disclosure of (personal) information by public sector bodies under access laws should be allowed, if it is provided for in law and on condition that access laws balance the public interest in access with privacy concerns. It is uncertain how this hierarchy of norms plays out where pro-active publication of data is concerned. If it is indeed the case that pro-active transparency stimulates re-use most, the lack of clear legal obligations to release data in the absence of a specific request might prove to be an Achilles heel.

The (G)DPR will likely leave manoeuvring room for Member States to maintain or introduce national provisions to further specify the application of the data protection rules of the regulation to public sector bodies. One could take the perspective that this space can serve as a ‘test bed’, possibly yielding different ways of reconciling national pro-active and re-active disclosure duties and corresponding re-use permission with data protection rules.

4.2 Re-use v open data policies

The current trend to frame re-use as “open data” policy has a dual effect. Open data means data that is among other things available for everyone to put to any use.³³ The mission to provide access to all with no conditions attached makes it particularly difficult for public sector bodies to comply with data protection rules. The net-effect of pushing open data as the model to aspire to may be that PSBs are driven to exclude all personal data and data relating to identifiable persons from their re-use policies. This is not what the current Data Protection Directive and the proposed (G)DPR require.³⁴ In effect, this seems to be happening in the Netherlands, where central government equates re-use policy with open data policy, but does not regard data that is subject to data protection regulation as ‘open data’. Admittedly, this approach has the advantage of being simple, which is important to bring data holders on board. However, it is only elegant at first sight, considering the broad application of data protection law, especially when one considers it applies to data that pertains to identifiable persons (whereby to determine whether this is the case account should be taken of ‘all the means likely reasonably to be used’ either by the PSB itself or any third parties) and that linking data makes identification easier. By constructing open data as not involving data on (identifiable) persons, it becomes a shrinking area.

Note that the adoption of ‘open data’ is not what the PSI Directive prescribes. In the PSI Directive what matters are transparent, accessible and non-discriminatory terms of use. If re-use is to be data protection compliant (as the PSI Directive recognizes it should be) the possibility of imposing use restrictions and conditions is arguably of paramount importance.

4.3 Tailoring re-use to data protection norms

³² See Presidency note to Council of 1 December 2014, Doc. 16140/14 on the General Data Protection Regulation,

³³ See opendefinition.org for a more elaborate explanation of the 8 (some would argue 9 if one includes costs) principles of open data. What is relevant to our purposes are the twin conditions that access is open to all, and with no conditions attached (or at most an obligation to recognize the source or share alike).

³⁴ For example, both contain provisions that explicitly allow re-use for scientific, statistic and historical purposes, on condition that certain additional safeguards are in place.

In its EC PSI Guidelines (2014/C 240/01) the Commission took up WP29 recommendation that PSBs actively inform re-users of the applicability of data protection rules. It might be worthwhile to explore whether it is useful to develop standard 'privacy notices' or 'privacy licensing terms' which give re-users more (formal) guidance on how to ensure processing is data protection compliant. Ensuring early development of standard notices might prevent proliferation of notices and licensing terms which in the context of intellectual property are considered as a barrier for the development of information services and products built on different data sources. It is conceivable that various privacy licensing terms are developed depending on the level of privacy risk involved, e.g. release of postal codes carries more risk (depending on the granularity of the local postcode system of course) than the release of say aggregate housing prices. The PSI Directive, in allowing public sector bodies to set different terms for different types of re-uses is no barrier to such differentiation. All it demands is that the terms and conditions are non-discriminatory against individual re-users.

A more tailored approach might ensure that the benefits of opening up PSI for re-use are not lost simply because data protection law applies. The proposed provisions in the (G)DPR on the processing of personal data for historical, statistical or scientific purposes and for archiving purposes, and the special status for persons and institutions using personal data for academic, artistic and journalistic purposes all point towards the possibility –desirability even– of distinguishing between different types of re-use.

At any rate, even if a specific provision would make it into the (G)DPR which explicitly allows disclosure of personal data contained in PSI for re-use purposes, there still is a requirement to balance the public interest in re-use against the concrete privacy interest at stake.³⁵ The prospects of one-size-fits all open data policies therefore seem limited.

The underlying fundamental tensions need addressing too. This requires in depth research into normative aspects: what are the competing interests involved (e.g. public interest in economic growth, better service delivery, and accountable government) and what are the balancing mechanisms suited for resolving them? How can these be operationalized in instruments that can aid public sector bodies in deciding which datasets can be released under what conditions and for which re-use purposes? What tools might be developed to aid citizens in exercising some control over the re-use of data that relates to their person? How can the potential for re-use be maximized in the space of the (soon to be) Data Protection Regulation? These are complex questions that would benefit from a thorough legal analysis flanked by adequate expert technological, public administration, business and civil society input.

³⁵ The latest proposal by the Presidency (16140/14 of 1 Dec. 2014), art. 88a reads: 'Personal data in in public sector information held by a public authority or a public body or a private body for the performance of a task carried out in the public interest may be disclosed by the authority or body in accordance with Union law or Member State law to which the public authority or body is subject in order to reconcile the re-use of such official documents and public sector information with the right to the protection of personal data pursuant to this Regulation'. The Commission has made a reservation on this draft provision, as it judges it incompatible with the PSI Directive (which leaves data protection law intact).