

PSI directive: Input from the digital industry

Brussels, 31 October 2018

KEY MESSAGES

- DIGITALEUROPE's Members value the importance of the PSI directive and ask for improvements to make it work better in practice.
- Harmonisation of PSI data access rules between Member States should be sought.
- The ICT industry acknowledges the costs of making data available for re-use and believes that charging marginal costs may be an acceptable solution in certain cases.
- Data should be accessible in easily readable formats and documents explaining how to use the data should be provided.
- Real-time data should be made available. When such dynamic data is not available, updated data should be accessible soon after being collected.
- Several sectors such as energy or transport are key for the development of data-based services and should be covered by the Directive.
- Public undertakings operating in a competitive environment may be exempted.
- Open access to research data should be voluntary and on a case-by-case basis.

SOURCE OF THIS INPUT

DIGITALEUROPE represents over 60 ICT companies and about 40 National Trade Associations (NTAs), representing the European digital sector, notably thousands of SMEs. We estimate that we represent in total over 35,000 ICT companies. *For more details, see the last page of this document.*

In order to compile this input, we have discussed the PSI directive recast with our Members and received information and comments from several of them, including both NTAs and corporate Members.

Our Members are interested in the PSI directive for two main reasons:

- **Direct interest:** for both SMEs as well as larger industry, use of the public data to create/expand/improve their services.
- **Indirect interest:** offer of data-related services (cloud, data management) to other companies as well as public institutions.

All our Members express strong support to the principle of facilitating the re-use of public sector data. However, they believe that the functioning in practice could be improved. Find below our comments.

INPUT AND SPECIFIC EXAMPLES

1. Uneven implementation

The implementation of the PSI directive is uneven at European level. The fact that **access regimes differ from one Member State to another** slows down the emergence of EU-wide services and products based on public sector information.

Generally, there is a vast difference in to which extent data is available in practice. Some countries have extensively developed open access policies on part of the public data but did not for the rest, which may still be locked under restrictive and differing licensing systems. Thus, access to data may be facilitated in certain cases yet rather complicated in others.

This is the case in Denmark, where a lot of effort has been put into the Government's Basic-data Initiative since 2012¹. This 'basic data' includes notably information about individuals, businesses, real properties, buildings and addresses. Accessing other types of data, either government data not considered as basic data or data from decentralised-local governments, is much more difficult. There is for instance no single entry for an overview of what data is available and in which conditions/formats, etc.

2. Charging costs

DIGITALEUROPE's Members consider that the **costs for licences to re-use public sector information is usually expensive**. Large firms may hesitate to contract such licences and SMEs will not be able to. This means that these companies will not be able to offer innovative services that would have been based on this data or to improve their core business functioning.

For instance, meteorological institutes are now opening their weather data models to the private sector. This data can be used for environmental research, business with logistical planning needing to foresee weather conditions, etc. Yet the most important institutes are charging excessive costs for the use of their data, which prevents most companies from using such data. Some of our member companies have experienced cases when research projects have been halted or discontinued because researchers did not have enough funding to pay disproportionate charging costs.

We understand that **making available data has a cost for public institutions**, particularly when it comes to data delivery. If the costs of making available data are too important for the public sector, it may have negative side effects for the proper functioning of services based on the re-use of this data. This is the case with the European Space Agency, which has a policy of open and free access to their data². However, it has been noted that the Agency cannot allocate sufficient resources for their data repository infrastructures, which entails that access to and delivery of data is extremely slow.

¹ Agency for Digitisation, "Basic data", <https://en.digst.dk/data-and-it-architecture/basic-data/>

² European Space Agency, "How to access data", https://www.esa.int/Our_Activities/Observing_the_Earth/How_to_access_data

DIGITALEUROPE believes that **charging marginal costs is appropriate**, particularly when this includes data delivery costs. Marginal costs should help public institutions to cover their costs related to the making available of data and should ensure that access to data is provided in good conditions.

3. Format and documentation

DIGITALEUROPE's Members believe that **restrictive formats are one of the main barriers to the re-use of public sector information**. Some Members received datasets that were deemed almost unreadable because they could not be easily accessed, opened and analysed due to format restrictions. Extra operations to open and convert illegible formats reduce the effective use of data, especially when it comes to dynamic data. Institutions should support the release of machine-readable and 'user-friendly' datasets.

Different institutions or departments within the same institution may also be releasing data in different formats, which complicate even more its re-use. At EU level, disparities between Member States are also quite important. It would be opportune to harmonise all these discrepancies into easily readable formats.

Additional **documentation should be provided whenever possible**, especially when restrictive formats are used. This would help re-users to make the most of the data they have access to, by facilitating the operations needed to use such data. Ideally, documentation should include tutorials that users can follow to get hands-on experience with data.

In general, documentation to be provided should not only cover the explanation about the use of the data delivered, but also include basic information on how the data was collected. Data re-users should be able to understand when the data was created, the methodology used to create it, how to interpret the values contained in it, and if there are any licenses that may limit how the data can be used. Documentation should also include a method to contact someone who can answer questions about the data.

If data is not documented, its audience will be inherently limited. There are times when users will do the 'detective' work required to interpret poorly documented data, but a lack of documentation will usually frustrate users to the point that they will simply not trust the data.

4. Availability

When available, **real-time data** ('dynamic data') **should be made accessible for re-use**. This is particularly needed to develop services relying on data from specific sectors, such as transport, energy and environment. Member agree that the cost of giving access to real-time data may be important, particularly to send data to third parties wanting to re-use it. This data delivery cost may be integrated to marginal costs.

Our Members noted that **data available for re-use is often outdated**. Sometimes, it takes more than a year for sets of data to be updated, even if there are no manual changes to be made. We understand that removal of confidential information, personal information for privacy reasons, etc. may extend the processing time for data to be made available. However, when there is no justification, data should be made available for re-use as soon as possible after its collection.

We believe that it is also of high importance that **data made accessible should still be available in the future**. Ensuring that data will be obtainable on an ongoing basis provides certainty to businesses which can develop services without the risk of investing and then having to discontinue a service due to datasets not being available anymore. Developers will not make significant effort to create tools or applications based on data if they have no assurance that the data will still be available in the future.

Finally, we encourage public institutions to use cloud solutions, particularly when sharing large volumes of dynamic data. Cloud services provide great platforms for storage, low-latency access and transfer of data. Users can retrieve data directly from the source, which assures them that they can reliably access a trustworthy copy of the data.

5. Data of interest for business and society

Our Members highlighted several fields that are of strong interest for them: transportation, energy, health, Earth observations (including satellite imagery, maps, meteorological data), official statistics data (e.g. census data) and any official documentation of real economic activity (e.g. government spending data is more useful than government budget data).

Opening access to data in these fields would have major positive consequences for the economy and society in general. Public data and statistics may be used by citizens to propose new ways of interacting in their communities but also to scrutinise government activities and make public institutions accountable. For businesses, the possibilities are endless as public data can foster extremely diverse services towards the population, other companies and public institutions themselves.

Allowing the **re-use of public data leads to important economic impacts**. In the European Union, the total market value of Open Data is expected to reach €75.7 billion in 2020³. Public institutions would actually be the first beneficiary, as public sector activities linked to the re-use of public data should reach a value of €22 billion by the same year. In 2020, activities based on access to public data should support up to 100,000 direct jobs.

Regarding access to public undertakings' data, DIGITALEUROPE proposes a balanced approach that would take into consideration different market situations. In some instance, public companies may have a monopoly in their field of activities. In this case, it is only logical that such public undertakings are considered 'classic' public institutions falling under the scope of the PSI directive and give access to their data.

However, when public undertakings are facing competition, particularly from private companies, they should not be required to comply with the same open access obligations than other public institutions. This should be tailored to specific cases, for instance where a public undertaking is only partly operating in a competitive environment; in this situation, data stemming from the activities carried without competition should be made available for re-use.

³ European Commission, "Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources", European Data Portal study, 2015, https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

In general, we believe that public undertakings have a lot to benefit from opening access to their data. Transport for London, the public body responsible for transportation in London, offers a great example of the advantages of allowing the re-use of its data for public undertakings:

Transport for London (TfL)'s data policy is an excellent example of the many benefits generated by the re-use of public data. The local government body responsible for the transport system in Greater London has opened for re-use most of its data since 2010.

A 2017 study⁴ showed that 13,000 developers are registered to have access to the over 80 data feeds made available for re-use by TfL. This led to the creation of more than 600 apps by professional or amateur developers, used by an estimated 42% of London's inhabitants.

In total, this policy is generating economic benefits and savings of up to £130 million (about €150 million) per year. Companies using this data have generated between €14 and 17 million with their data-powered services, which led to the support and/or creation of 500 direct jobs and 230 indirect ones.

This data policy benefits TfL as well. As hundreds of apps have been created using TfL's data, the company does not need to develop and support an extensive app ecosystem on its own, which leads to potentially important cost-savings. Through partnerships made with some of the companies exploiting its data, TfL is also able to retrieve valuable data it does not possess (for instance crowdsourced traffic data collected through third-party apps). Finally, as there is now an extensive ecosystem of apps surrounding TfL's services updating users and passengers, the Contact Centre of TfL can reduce its communication and outreach spending. TfL considers that the benefits of allowing the re-use of data considerably outweigh the costs of publishing such data (€1,15 million per year).

Overall, DIGITALEUROPE's Members expressed great interest in sensor data, in comparison to man-made data, which can contain mistakes during creation or copy to datasets. **Extremely reliable, machine-generated data is therefore valuable for businesses** to develop high-end services. Sensor or machine-generated data can also be made automatically available as 'dynamic data'. When technically possible and financially feasible, we thus advocate to avoid human interventions on datasets – or at least reduce them to the strict minimum.

6. Research data

While DIGITALEUROPE's Members support open access to scientific publications, they believe that **research partners should have the freedom to decide if they want to give access to research data**. Depending on the case, public and private research consortiums members should be able to decide on a voluntary and case-by-case basis whether access can be granted, and if so, to which data and to whom, according to the principle of access to research data being **"as open as possible, as closed as necessary"**. This principle is fundamental as research data may contain sensitive competitive information which, if released, could undermine years of

⁴ Deloitte, "Assessing the value of TfL's open data and digital partnerships", July 2017, <http://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf>

investments in innovation and reduce the competitiveness of companies. Thus, this approach respects the essential protection of intellectual property, security, confidential information and personal data as well as legitimate commercial interests.

Most of the research partially or fully EU-funded is carried through Horizon 2020, the largest research programme in Europe. Many different entities are joining consortiums to participate to research activities under Horizon 2020: public universities, research centres, but also the private sector, from SMEs to larger companies.

All these actors of the research landscape are benefitting from working together on research projects, by engaging in peer-to-peer collaboration to exchange new ideas for state-of-the-art research, by creating networks of skilled researchers and by supporting cross-border R&D&I pre-competitive cooperation between different stakeholders, even competitors. This helps building partnerships and synergies with critical mass, ensuring both the excellence and strengthened competitiveness of European research. On top of the above benefits, the industry also brings to research partners its own knowledge, networks and resources (access to market and to industry networks, business expertise, research infrastructures, targeted technical resources, etc.). These advantages of collaborative research apply to most public-private research projects, whether they are funded at EU or national level.

Despite the many benefits for the private sector of participating to publicly-funded research projects, companies would likely no longer join such projects if open access to research data was mandatory and applied by default. This would affect the whole research landscape as industry and public-sector researchers would no longer bring their complementary expertise to each other.

7. Annex – List of proposed high-value datasets

Our proposed list of high-value datasets:

CATEGORY	EXAMPLE OF DATASETS
Administrative Divisions	Administrative divisions, such as regions, districts, municipalities or other administrative units, delineated for the purpose of administration by the Member States
Companies	Company/business register
Crime and Justice	Crime statistics, safety
Earth observation	Meteorological/weather, agriculture, forestry, fishing, and hunting
Education	List of schools; performance of schools, digital skills
Energy	Energy consumption
Environment	Air quality (Pollution levels), water quality

Finance and contracts (incl. gov. budget)	Transaction spend, contracts let, call for tender, future tenders, local budget, national budget (planned and spent)
Geospatial (locations and maps)	Topography, postcodes, national maps, local maps
Global Development	Aid, food security, extractives, land
Government Accountability and Democracy	Government contact points, election results, legislation and statutes, salaries (pay scales), hospitality/gifts
Health	Prescription data, performance data
Land Ownership	Mapping data on boundaries of lots/parcels of land, including land registry data
National Law	Legislative, regulatory and administrative measures Draft measures, including procedural information related to their adoption Measures which have been amended, repealed or are no longer in force Accompanying documents, such as explanatory statements, impact assessments, opinions of advisory bodies and voting records Case law
Statistics	National Statistics, Census, infrastructure, wealth, skills
Social mobility and welfare	Housing, health insurance and unemployment benefits
Transport and Infrastructure	Public transport timetables, access points broadband penetration

--

For more information please contact:

Julien Chasserieu, DIGITALEUROPE Policy Officer

+32 4 92 27 13 32 | julien.chasserieu@digitaleurope.org

ABOUT DIGITALEUROPE

DIGITALEUROPE represents the digital technology industry in Europe. Our Members include some of the world's largest IT, telecoms and consumer electronics companies and national associations from every part of Europe. DIGITALEUROPE wants European businesses and citizens to benefit fully from digital technologies and for Europe to grow, attract and sustain the world's best digital technology companies. DIGITALEUROPE ensures industry participation in the development and implementation of EU policies.

DIGITALEUROPE's Members include in total over 35,000 ICT Companies in Europe represented by 63 Corporate Members and 39 National Trade Associations from across Europe. Our website provides further information on our recent news and activities: <http://www.digitaleurope.org>

DIGITALEUROPE MEMBERSHIP

Corporate Members

Adobe, Airbus, Amazon, AMD, Apple, Bosch, Bose, Brother, Canon, Cisco, Dell, Dropbox, Epson, Ericsson, Fujitsu, Google, Hewlett Packard Enterprise, Hitachi, HP Inc., Huawei, IBM, Intel, JVC Kenwood Group, Konica Minolta, Kyocera, Lenovo, Lexmark, LG Electronics, Loewe, Mastercard, Microsoft, Mitsubishi Electric Europe, Motorola Solutions, MSD Europe Inc., NEC, Nokia, Nvidia Ltd., Océ, Oki, Oracle, Panasonic Europe, Philips, Pioneer, Qualcomm, Ricoh Europe PLC, Samsung, SAP, SAS, Schneider Electric, Sharp Electronics, Siemens, Sony, Swatch Group, Tata Consultancy Services, Technicolor, Texas Instruments, Toshiba, TP Vision, VMware, Western Digital, Xerox, Zebra Technologies.

National Trade Associations

Austria: IOÖ	France: AFNUM, Syntec Numérique, Tech in France	Romania: ANIS, APDETIC
Belarus: INFOPARK	Germany: BITKOM, ZVEI	Slovakia: ITAS
Belgium: AGORIA	Greece: SEPE	Slovenia: GZS
Bulgaria: BAIT	Hungary: IVSZ	Spain: AMETIC
Croatia: Croatian Chamber of Economy	Ireland: TECHNOLOGY IRELAND	Sweden: IT&Telekomföretagen, Teknikföretagen
Cyprus: CITEA	Italy: Anitec-Assinform	Switzerland: SWICO
Denmark: DI Digital, IT-BRANCHEN	Lithuania: INFOBALT	Turkey: Digital Turkey Platform, ECID
Estonia: ITL	Netherlands: Nederland ICT, FIAR	Ukraine: IT UKRAINE
Finland: TIF	Poland: KIGEIT, PIIT, ZIPSEE	United Kingdom: techUK
	Portugal: AGEFE	