

# Open Data use and re-use: Widening the approach towards “free flow of data”

Martina Barbero, PSI project coordinator, Public Sector consultant at Deloitte

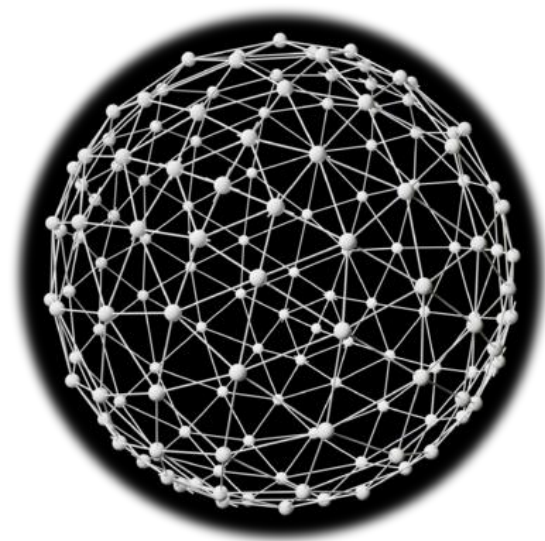
Annette Hillebrand, Managing Consultant Digital Security at WIK-Consult

Digital Executive Club of Idate and WIK  
4 October 2018, Brussels

# Common European data space

Towards a seamless digital area in the Digital Single Market

- Aim: Data should be available for re-use as much as possible, as a key source of innovation and growth and a precondition for development of Artificial Intelligence (AI)
- Measures to cover different type of data, e.g.
  - a proposal for a review of the Directive on the re-use of public sector information (PSI Directive)
  - political agreement on a Regulation on Free Flow of non-personal data (June 2018)
  - guidance on sharing private sector data among companies and with public sector bodies for public interest purposes (Data Support Centre)



- “PSI-Study”: Study to support the review of Directive 2003/98/EC on the re-use of public sector information (consortium of Deloitte (Belgium), WIK-Consult, Open Evidence, The Lisbon Council, Timelex, Sparks) (2016-2017)
- “Study on emerging barriers to data economy”: Study on emerging issues of data ownership, interoperability, (re-) usability and access to data, and liability (consortium composed of Deloitte (Belgium), WIK-Consult, Open Evidence, OpenForum Europe, Timelex) (2015-2017)
  - both studies published April 2018, in the context of the data package

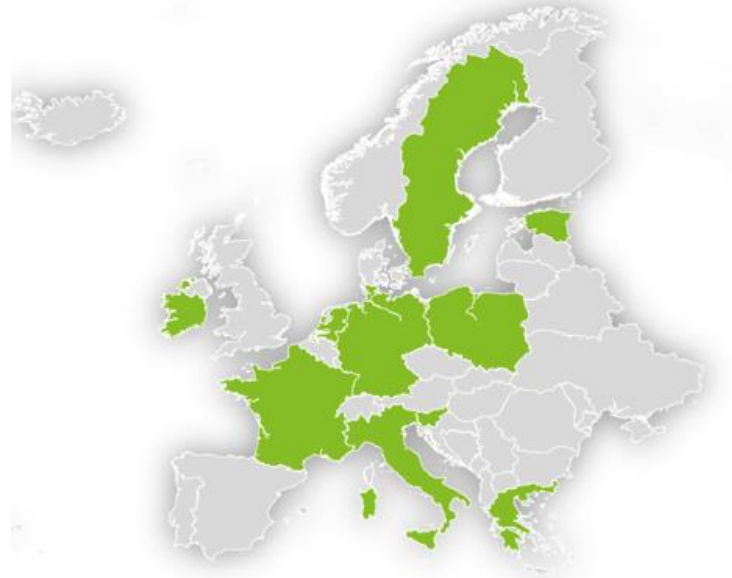
### Objective of the study

- **Evaluate** the functioning of the PSI Directive (with respect to the 2013 additions) based on the state of play regarding the re-use of public sector information in a representative number of EU Member States
- **Assess** the expected impacts of a limited number of options for policy interventions, as well as compare the different options (or combinations thereof)

	2003	2013
<b>Rationale</b>	Economic (traditional users)	Economic (extended users)
<b>Scope</b>	Public sector bodies	Addition: Museums, libraries, and archives
<b>Rights</b>	Access	Re-use
<b>Charging</b>	Recovery	Marginal with exceptions
<b>Licensing</b>	Without conditions or standard licences	Without conditions or standard licences
<b>Format / quality</b>	n/a	Open and machine-readable format

### Key facts and figures on the data collection and analysis

- 3 surveys carried out
  - *On research data*
  - *On cultural data*
  - *For re-users of PSI*
- Fieldwork in 10 Member States
- Legal data collection in 10 Member States
- 3 workshops
  - *Workshop on Open Research Data*
  - *High level round table*
  - *Final workshop*



# Evaluation of 2013 changes

## Extension to cultural data

**The extension of the PSI Directive to cultural data has not created revolutionary changes yet.**

There are several explanations for this finding:

- It is too early to assess the impact of the Directive on museums, libraries and archives as in many Member States transposition happened only recently and these bodies need some time to adapt.
- There are a number of barriers limiting the effectiveness of the PSI Directive:
  - **Lack of awareness** on the PSI Directive
  - Issues related to **intellectual property rights**
  - **Costs of digitalisation**
  - **Data format**
  - **Charging provisions**

### Examples of costs related to the digitisation of cultural material:

- In the Netherlands, to clear IPR right of 8 pictures, it is necessary to staff 1 FTE for an entire day.
- The establishment of the open data portal of the Italian Central Institute for Cataloging and Documentation (ICCD) was of around 140.000euro out of which 110.000 concerned the digitization of the data and the development of data modelling.
- A Greek library digitized 905.000 pages of printed material from the University's libraries for about 260.000 euros.
- In France, the National Library (BNF) invests more than 12 million a year in digitisation projects.

# Evaluation of 2013 changes

## Changes to the charging provisions

It can be argued that the establishment of a marginal cost rule contributed to the achievement of the PSI objectives. However, there are a number of elements to consider:

- Not all EU countries have already established the *objective, transparent and verifiable criteria* which are needed for the definition of charges and therefore it is early for a full-fledged analysis of impact.
- Before 2013 already, *only a minority of public sector bodies was charging higher costs than the marginal costs*.
- Even in the cases in which public sector bodies was charging higher costs for data, *this revenue constituted only an extremely small component of the overall budget of the authority* (less than 1% of the budget in half of the cases according to POPSIS Study).

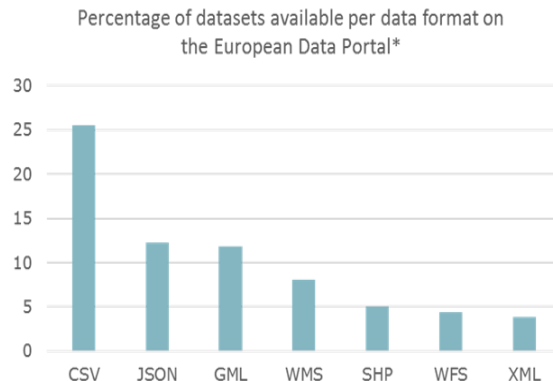
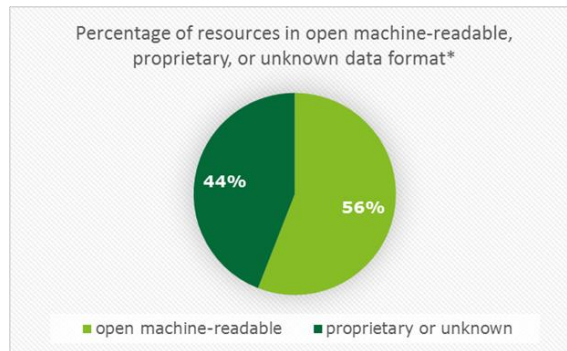
Furthermore, if these rules contributed to reduce the price of data and the number of bodies charging overall, **considerable differences still exist in terms of charging practices for the same datasets** across Member States.

# Evaluation of 2013 changes

## Changes in data format

*Public sector information should be published "(...) in open and machine-readable format together with their metadata. " AND "Both (...) comply with formal open standards." – PSI Directive, Article 5*

- Lack of adequate evaluation and monitoring tools;
- Open machine-readable format is gradually becoming the norm;
- Most of datasets are published together with their basic metadata, thanks to data portals;
- The main issue is on data quality and the lack of open data standards across EU;
- Need for adequate tools, infrastructures and clear guidance, as well as trainings and cultural shift (data literacy);
- Strong demand for public data APIs and dynamic data delivered in real-time





The Impact Assessment part of the study focused on the *four most relevant problem areas and future needs* identified in the evaluation:

- The pressure to open up other PSI held by other public bodies → the problems related to (1) the **inaccessibility of research data** and (2) **inaccessibility of data held by para-public bodies**.
- The need to continue reducing price of data → the problems related to (3) **charging of data**.
- The need continue working on making data (and especially dynamic data) attainable → the problems related to (4) **absence of APIs**.
- The need to maintain a level-playing field → the problems related to new types of (5) **exclusive agreements** emerging.

Each of the problem areas was analysed individually and in-depth following a number of steps:

- **Step 1:** definition of the causes, problems and effects
- **Step 2:** identification of possible policy sub-options to address these problems including
  - *No action*
  - *Non-legislative intervention (e.g. developing guidelines, raising awareness, disseminating best practices)*
  - *Legislative intervention (modifying the text of the PSI Directive)*
- **Step 3:** preliminary analysis of the sub-options according to 6 criteria: *effectiveness, efficiency, proportionality, legal feasibility and coherence, practical and technological feasibility and political feasibility.*
- **Step 4:** shortlisting of sub-options for further assessment

# Policy packages

## Possibilities for policy intervention

The team shortlisted a number of sub-options for each problem area and aggregated them in policy packages.

### Baseline policy package: no action for the problem areas identified

#### Policy Package 1 (PO1) – Low intensity legislative intervention

- Extend the scope of the PSI Directive to research establishments to cover administrative data and research results, focusing on the re-usability of already accessible data.
- Extend the scope of the PSI Directive to cover para-public bodies and private entities carrying out public tasks under the procurement Directive 2014/25/EU requiring the application of the provisions of the Directive to the data that these entities decide to make available for re-use
- Amending article 6 to limit recourse to exceptions to marginal cost charging: deleting the exception described in article 6.2(b) while specifying the eligibility of costs
- Defining a limited set of high-value datasets to be released as open data
- Entice public sector bodies to make dynamic data available for re-use immediately after collection via an API
- Introduce a procedural safeguard preventing the conclusion of agreements with a high risk of 'de facto' exclusivity.

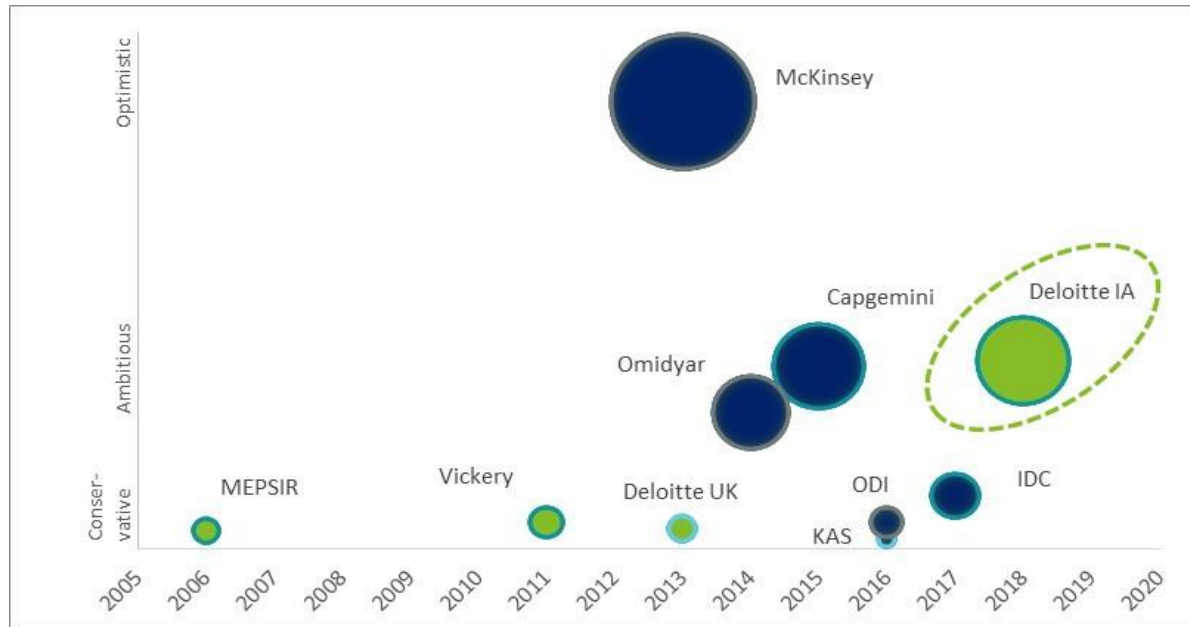
#### Policy Package 2 (PO2) – High intensity legislative intervention

- Extending the scope of the PSI Directive to research establishments so as to cover administrative data and research results and to mandate research establishments and research funding bodies to implement open access policies.
- Extending the scope of the PSI Directive to cover para-public bodies and private entities carrying out public tasks under the procurement Directive 2014/25/EU or on the basis of a concession bodies and entities under the default rules of the PSI Directive
- Introducing free of charge re-use as the only rule for all documents covered by the Directive.
- create an obligation to make dynamic data available for re-use immediately after collection via an API.
- Prohibit the conclusion of agreements between the public sector bodies and private companies that may lead to 'de facto' exclusivity.

# Policy packages

## Economic model

The economic model underpinning the analysis was based on primary data as well as data coming from similar economic analysis concerning the data economy. Many different studies tried to measure the benefits of open data for the European Union with very different results.



Based on economic analysis and modelling, Policy Package 1 – Low intensity intervention was identified as the preferred policy package.

### Effectiveness of Policy Package 1:

- ✓ This policy package would contribute addressing all problems without nonetheless solving all of them completely

### Efficiency of Policy Package 1:

- ✓ The benefits of this policy package outweigh its costs

### Coherence of Policy Package 1:

- ✓ This policy package is coherent with other EU legislative measures
- ✓ This policy package is coherent with the approach and history of the PSI Directive

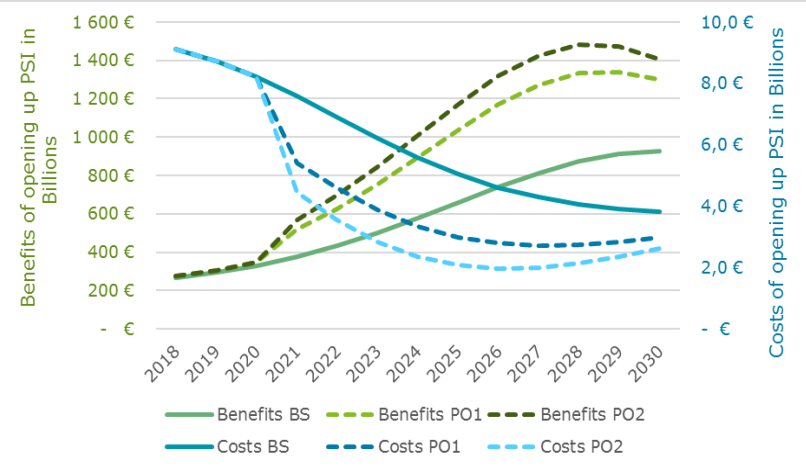
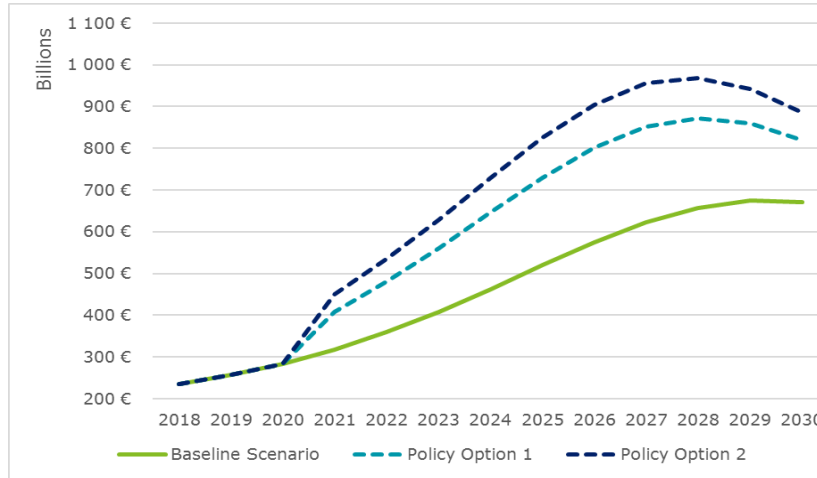
### Proportionality of Policy Package 1:

- ✓ This policy package is the most proportionate

# Policy packages

## Preferred policy package 2/2

Policy Package 2 – High Intensity intervention would be more effective than PO1 in addressing the barriers identified and increasing the value of the PSI economy in Europe. However, PO1 is more efficient, coherent and politically feasible.



- The PSI Directive has led to many positive achievements in the domain of public sector information and it has greatly benefit all stakeholders (public sector bodies, re-users and citizens).
- Through a number of limited modifications to the legislation (low intensity intervention – thus limiting burden for stakeholders) further progresses can be made with respect to:
  - The [availability of additional datasets](#) through gradual extension of the Directive in new (relevant and promising) domains
  - The [increase in the take up of APIs](#) by public sector bodies
  - The [reduction of barriers for accessing data](#) including those linked to *costs of datasets* and *exclusive agreements*



The European Commission followed only partially the recommendations of the study:

- Reduce market entry barriers, in particular for SMEs, by **limiting the exceptions** that allow public bodies to charge for the re-use of their data more than the marginal costs of dissemination;
- **Increase the availability of data by bringing new types of public and publicly funded data into the scope of the Directive**, such as data held by public undertakings in the utilities and transport sectors and research data resulting from public funding;
- Minimise the risk of excessive first-mover advantage, which benefits large companies and thereby limits the number of potential re-users of the data in question, by **requiring a more transparent process for the establishment of public-private data arrangements**;
- Increase business opportunities by **encouraging the dissemination of dynamic data via application programming interfaces (APIs)**.





# Para-public bodies

## Problems, causes and effects 1/2

*Data of para-public bodies is of strategic public value and highly important for fostering innovations. Some reluctance to allow data re-use seems unlikely to change in near future*

### Problems

- Data of para-public bodies is still locked in most cases
- Dormant opportunities particularly in the energy and transport sector, e.g.
  - Energy consumption analysis and management
  - Door-to-door travel services
- Fragmented landscape, depending on
  - business strategies and
  - open data strategy in MS
- Existing regulation of data appears insufficient (ITS, Third Energy Package)

### Causes

- Major utilities are not among open data first movers, influenced by traditional monopoly-based business culture
- Many para-public bodies act in very competitive markets and use data for own benefit only (effectiveness, attractive customer service)
- Companies as well as their PSB shareholders avoid additional costs
- Operators fear to violate data protection and privacy provisions if they allow access to data

# Para-public bodies

## Problems, causes and effects 2/2

*Data of para-public bodies is of strategic public value and highly important for fostering innovations. Some reluctance to allow data re-use seems unlikely to change in near future*

### Effects

- Activities today are limited to some MS and to few companies in the EU
- Voluntary activities have only small impact on overall European Data Economy
- Severe Limitations on
  - Intermodal transport services
  - Cross-border services / single market generally
- Large data collectors might bypass slow voluntary efforts and supersede them

### Evidence from country cases

- Legislation in France (2016), relevant regulation in Finland and Slovenia
- Transport companies invest in hackathons, making available selected data sets for third parties and development of company specific apps, based on corporate (not social) economic logic
- Energy companies engage in innovation competitions, MS promote hubs for data access

# Para-public bodies

## Possible policy options 1/2

Baseline scenario	Ineffective to enhance the development of innovative applications
Non-regulatory options	Important flanking measures to enhance voluntary efforts. Sufficient to extend and increase availability of data substantially? More innovation around para-public bodies needed, but can hardly be imposed by legislation. Commission and MS could do more to promote mutual benefits of sharing data, communicate good practices, and help overcome (perceived) risks and negative effects of data sharing.
Regulatory options <i>Extend the scope of the PSI Directive to cover para-public bodies and private entities carrying out public tasks under the procurement Directive 2014/25/EU ...</i>	<ol style="list-style-type: none"><li>1. Recommendation with high intensity of legislative intervention: ...</li><li>2. Recommendation with low intensity of legislative intervention: ...</li></ol>

# Para-public bodies

## Possible policy options 2/2

### Regulatory options

*Extend the scope of the PSI Directive to cover para-public bodies and private entities carrying out public tasks under the procurement Directive 2014/25/EU ...*

1. Recommendation with high intensity of legislative intervention:  
*... or on the basis of a concession under the default rules of the PSI Directive (marginal cost charging, transparency, data formats, processing of requests, etc.)*
  - Disproportionate
  - Diminishes voluntary action
  - Imposes high costs without considering which data is attractive for re-use
2. Recommendation with low intensity of legislative intervention:  
*... but require the application of the rules of the PSI Directive only to data made available for re-use voluntarily*
  - + Enables Member States to extend their open data strategies to additional fields of activity
  - + Ensures a higher level of data access and re-use for third parties
  - + Allows Companies to follow their open data strategies in full control of costs

### European Commission's choice

- Increase the availability of data held by public undertakings in the utilities and transport sectors
- Transparency requirements for public-private agreements to avoid exclusive arrangements
- Delegated act: listing a set of high-value datasets that are to be provided in a timely manner and free of charge

# Example non-regulatory option: mFUND initiative

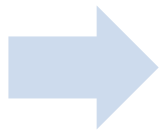
Open Data non-legislative measures in Germany

- Widening Open Data Strategies beyond the „national level“ – to the local level

- Access to raw data
- Access without charges
- APIs
- Dynamic data
- Open Data and public undertakings



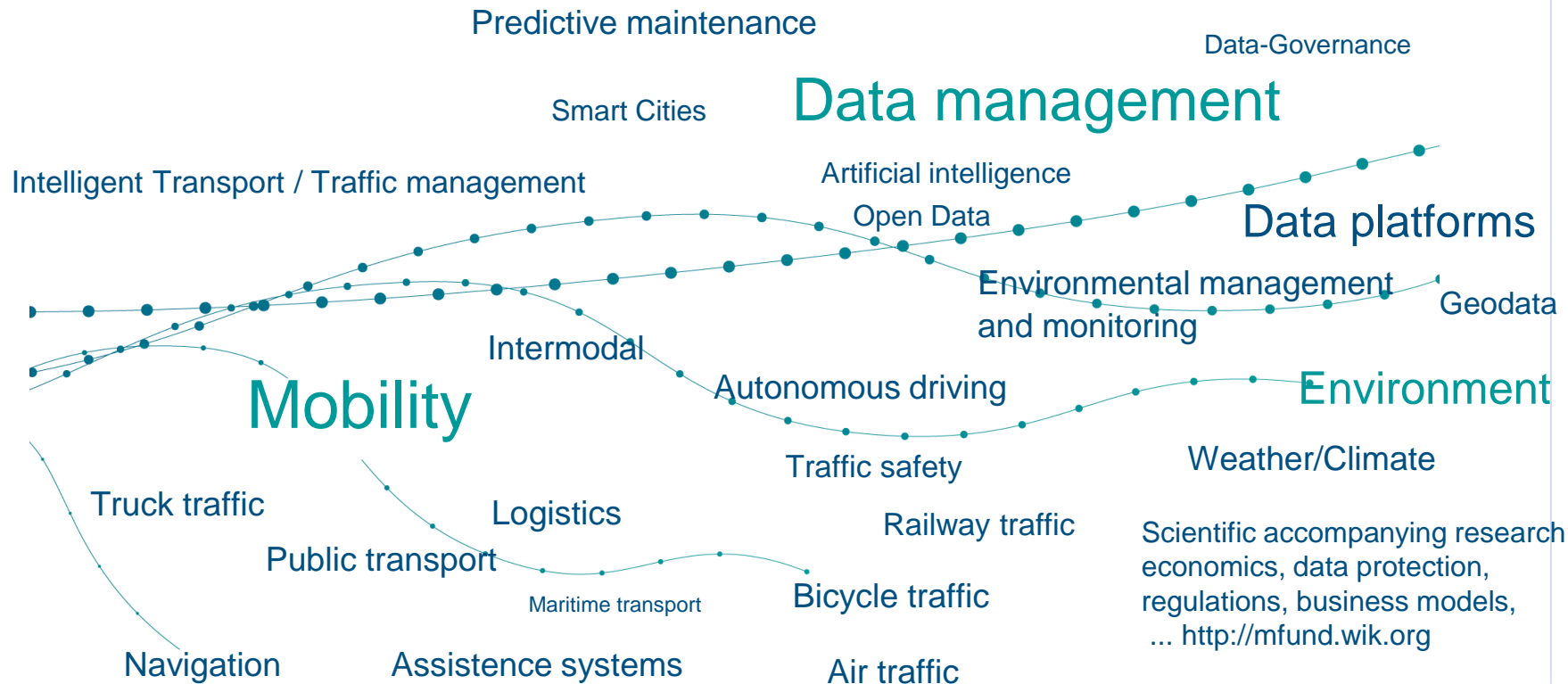
- Development and implementation of strategies too slowly in local administrations, but also on national level and in businesses



Supporting projects, products and new business opportunities and publish best practices

# Funding areas & projects of mFUND

Wide variety of solutions for mobility 4.0



# mFUND: Supporting Mobility 4.0

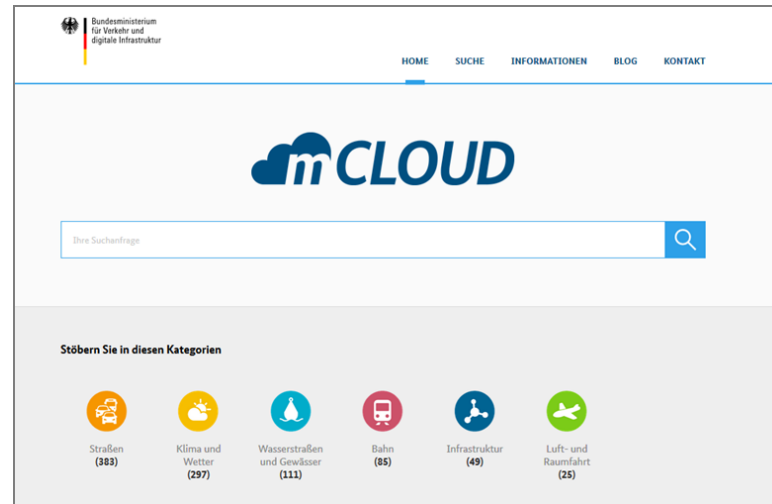
Ministry for Transport and Digital Infrastructure

Aim	Development of digital, data-based business models based on mobility, geospatial and weather data bases
Participants	Founders, start-ups, universities, local administrations, administrations on national level (Bund, Land), SMEs and large companies, research insitutions (public and private)
Funding	150 million Euro (Mai 2018), projects from 100,000 Euro to 3 million Euro funding; in general no full funding of projects
Examples	Indoor navigation at railway stations, wildlife accident prediction and management, heavy gods transport navigation, public transport optimization in rural areas, city event management control

# mCloud – Data source for projects

And also the public

- Research platform for open data from all kind of sources of the Ministry of Transport and Digital Infrastructure
- Data from
  - Roads
  - Railway
  - Waterways
  - Air traffic
  - Weather
  - Climate
  - Water bodies



[www.mcloud.de](http://www.mcloud.de)

- More than 80 project consortia are currently using mCLOUD data or other open data
- Number of projects still increasing



# Study on the emerging barriers to the data economy

## Objective and context of the study

**Objective of the study:** To identify the most important barriers to the development of the data economy and the use of IoT, robots and autonomous systems: **liability, (re-)usability of and access to (third party) data, and interoperability.** → *geographical/localisation restrictions fell outside the scope of the study.*

### Context:

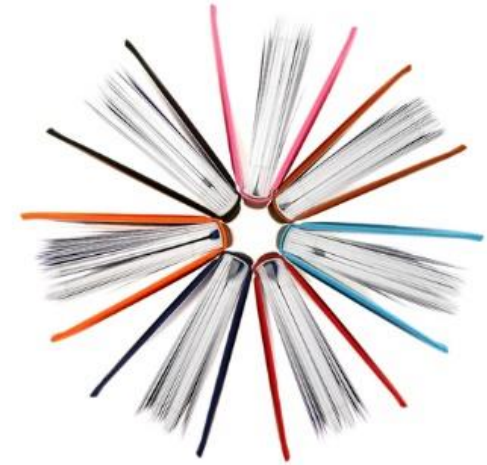
- Markets that are still in their infancy ('emergence phase')
- Less than an estimated 10% EU companies are intensive data users (6.3% in 2016, EU study)
- Analysis of more than 100 business models within the study show lack of data use
- The few companies already engaged in the data economy face uncertainties and barriers

# Study on the emerging barriers to the data economy

## Methodology for the study

The study involved different data collection and analysis tools

- *Expert interviews and literature review*
- *One major survey* with companies sharing and reusing data (250 respondents) in multiple sectors and across 6 EU countries
- *Qualitative research in industry sectors* (health, finances, machinery, agricultural and food, energy, chemistry, aerospace, automotive, transport, telecom, fast moving consumer goods (FMCG) industry)



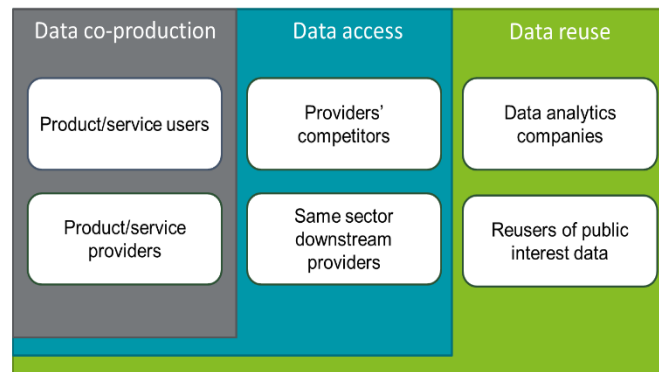
# Study on the emerging barriers to the data economy

## Main finding of the study

There are **three preconditions** that can help determine the types of barrier a company is likely to face in the data market:

- its position in the value chain;
- its size (SMEs versus larger companies); and
- the sector it is in.

In terms of position in the value chain, companies can fall in the data production, data access or data (re-)use category.



# Study on the emerging barriers to the data economy

Main finding of the study

The study highlighted the different types of barriers each category of companies is likely to face based on its position in the data value chain and identified primary and secondary barriers.

Stakeholder	Legal uncertainties	Data Ownership	Access to data	Data (re-)use	Liability	Data portability	Interoperability	Skills	Valuing data	Unequal bargaining power	Cost of data
Product/service users											
Product/service provider											
Providers' competitors											
Same-sector downstream provider											
Data analytics companies											
(re-)users of public interest data											

# Study on the emerging barriers to the data economy

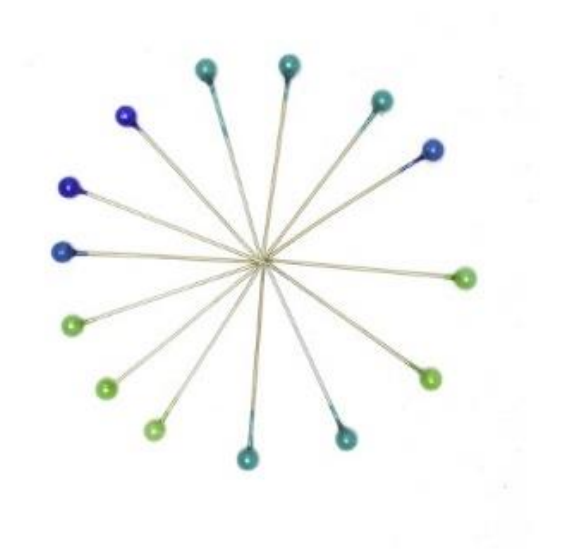
## Primary barriers

- **Access and to (re )use of data:** Companies cannot access the data they need or would like, and they face strict (contractual) limitations when wanting to (re )use data
- **Data liability:** Existing liability laws are based on the concept of tangible products. Companies cannot be sure whether they can have recourse to this legislation for data-based products, so prefer to fall back on contractual liability on a case-by-case basis
- **Data interoperability:** Different standards and specifications are used for the same data and for different datasets
- **Unequal bargaining power:** Smaller companies (SMEs) and companies in a weaker position in the value chain do not have the bargaining power to obtain access to certain data, whether for free or at a cost
- **Skills:** There are not enough people now with the right skills, and the problem is likely to get worse in future

# Study on the emerging barriers to the data economy

## Secondary barriers

- **'Data ownership'**: The concept of 'data ownership' is far less controversial for companies than thought when the study was launched; access to and (re )use of data are much more important
- **Data portability**: This is not a bar to companies sharing, accessing and re-using data, except in very particular cases
- **Intellectual property rights (IPR)**: There is not felt to be a need to have recourse to the exclusive protection conferred by IPR when sharing, accessing and re-using data as this tool seems inadequate in most cases
- **Valuing data**: The cost of data is an obstacle for data (re-) users, but if a company is interested in sharing data, it will find a means of valuing it

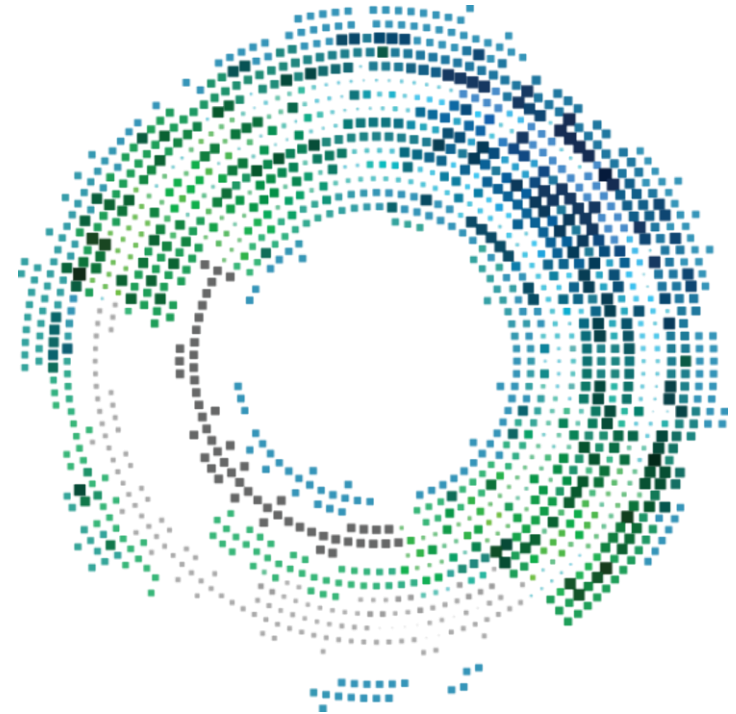


# Study on the emerging barriers to the data economy

Non-legislative measures preferred

## Recommendations of the study

- *Non-legislative measures* (e.g. exchanges of best practices, examples of contract clauses, trainings) are to be preferred at this stage of the development of the markets
- *Continuous monitoring of barriers*, liability issues and cases, and business models - to be ready to regulate when and if it is necessary



# Study on the emerging barriers to the data economy

Non-legislative measures preferred

## Political agreement (19 June 2018)

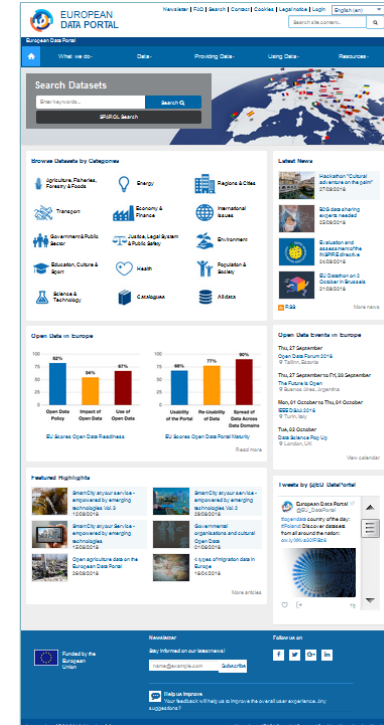
- Ensure the *free flow of data across borders*: Data storing and processing across the EU, prohibiting data localisation restrictions. Member States have to communicate any regulations on public sector data processing. GDPR is complementary to this provision.
- *Ensure data availability for regulatory control*: Public authorities will be able to access data everywhere in the EU.
- Encourage *creation of codes of conduct* for cloud services to facilitate switching between cloud service providers



# Setting up a Support Centre for data sharing

Towards a seamless digital area in the Digital Single Market

- Funded by the European Commission (1.5 million Euro)
- Based on Service Platform within European Data Portal
- Objective: facilitate data transactions and data analytics
- Method: collect, develop and disseminate tools and technical expertise and provide practical support
- Public sector data or data held by private companies can be made available to other organisations (public sector bodies or private companies) for use and re-use
- Provide public authorities with information on data sharing



# Points for discussion

- Data flows as preconditions for the development of an European AI economy?
- Catching up U.S. and China in terms of data availability?
- Limits to non-legislative measures? Need for action?
- Sustainable business models for (re-)use of data?
- Lack of availability of specific data?
- Specific support for SMEs needed?





Annette Hillebrand

WIK-Consult GmbH

Postfach 2000

53588 Bad Honnef

Deutschland

Tel.: +49 2224-9225-0/-53

Fax: +49 2224-9225-63

eMail: [a.hillebrand@wik-consult.com](mailto:a.hillebrand@wik-consult.com)

[www.wik-consult.com](http://www.wik-consult.com)