Deloitte.

open data institute

the Lisbon council
think tank for the 21st century

THE GREEN LAND

Final event

Impact assessment study on the list of High Value Datasets to be made available by the Member States under the PSI Directive

4 September 2020

Agenda

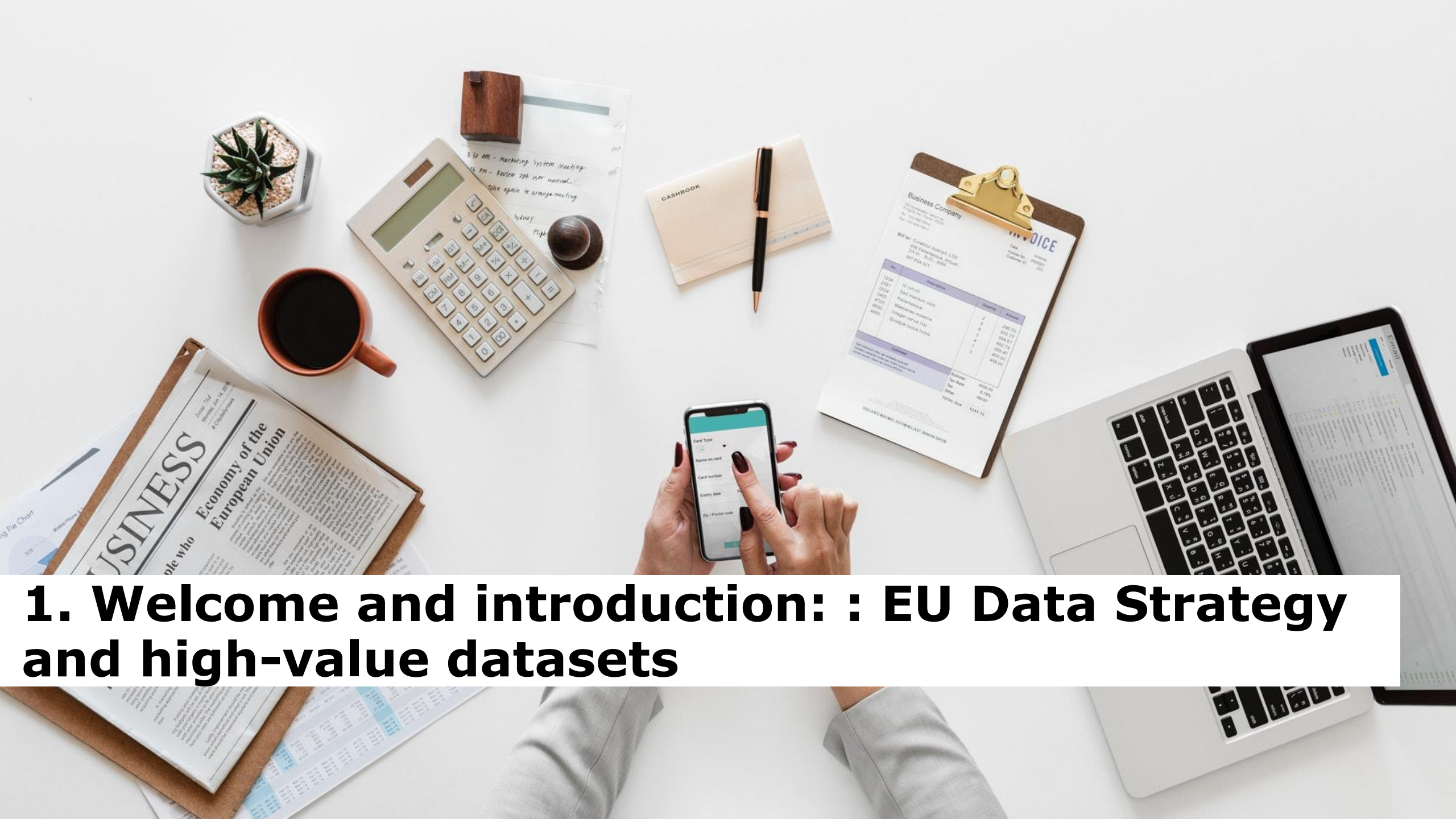
10h00	**10h20**	**Welcome and introduction: EU Data Strategy and high-value datasets**
10h20	**11h30**	**Presentation of the study results – Part I**
11h30	**12h00**	**Opinions, questions and answers – Part I**
12h00	**13h00**	**Lunch break**
13h00	**13h30**	**Presentation of the study results – Part II**
13h30	**14h45**	**Opinions, questions and answers – Part II**
14h45	**15h00**	**Closing**

House rules

The netiquette to follow for this online meeting

A **few simple rules** apply to this online focus group:

- **The moderator is introducing the topic/the question and then opening the floor for the discussions.** The moderator's goal is to generate a maximum number of different ideas and opinions from as many different people in the time allotted so he/she might ask you directly your opinion in case you have not had the chance to intervene on a certain topic yet.
- You can ask for the floor by **requesting it in the chatbox on GoToWebinar** (see control panel on the right of your screen). The moderator will give you the floor as soon as soon as possible.
- **Keep your intervention short and to the point.** This will allow us to stick to the time allocated for each part of the agenda.
- **Mute yourself when you are not speaking** to avoid echoes and background noises.
- **Enjoy the discussion**!

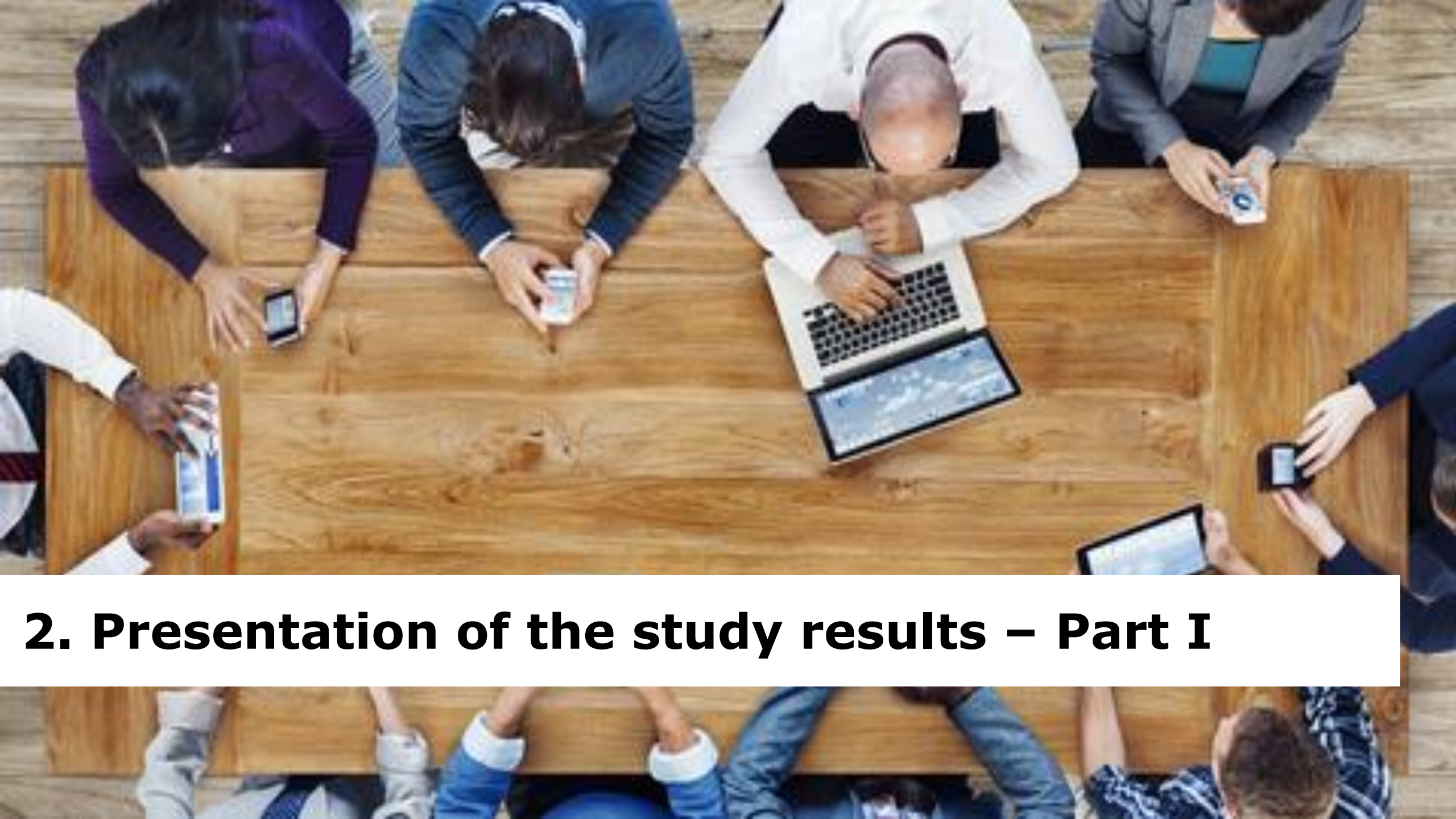


1. Welcome and introduction: : EU Data Strategy and high-value datasets

DG CNECT G.1

Yvo Volman

Head of Unit – Data Policy & Innovation



2. Presentation of the study results – Part I

About the study

Sebastiaan van der Peijl
Deloitte

About the study (1/3)
Objectives and scope

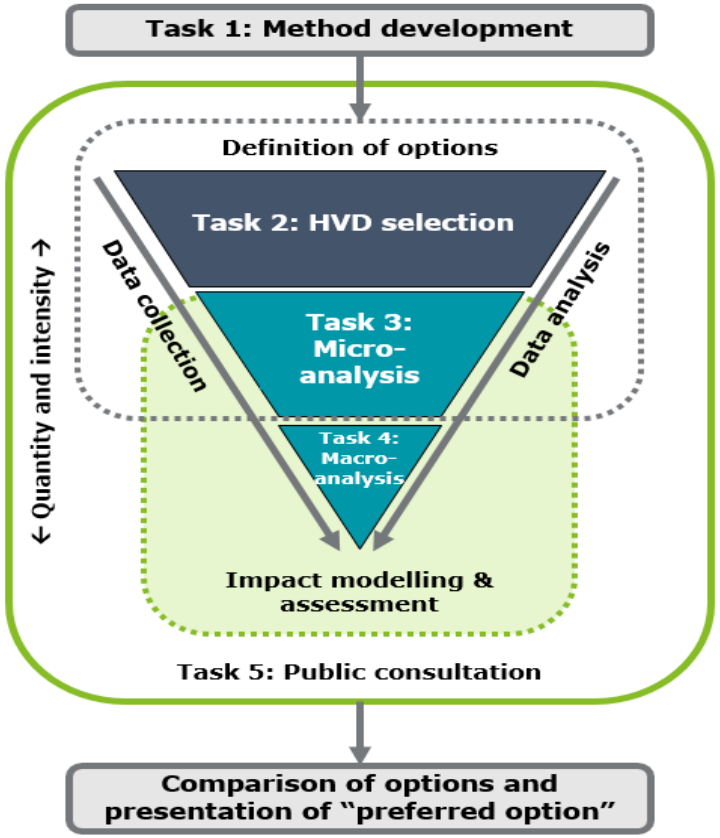
The overall objective of this assignment is to « **define concrete High Value Datasets that fall under the thematic categories included in the Annex of the revised Directive, based on an iterative process involving a number of cycles, by which an initial broad range of datasets will be narrowed down.** »

Main content scope

- The categories of high value datasets identified in Annex I of the Directive (*geospatial, earth observation and environment, meteorological, statistics, companies and company ownership, transport*);
- The **other areas of legislation** (at the European level) which cover the provision of certain datasets by public sector authorities and must be taken into account for ensuring that the overall legislative framework on public sector information is coherent and effective.

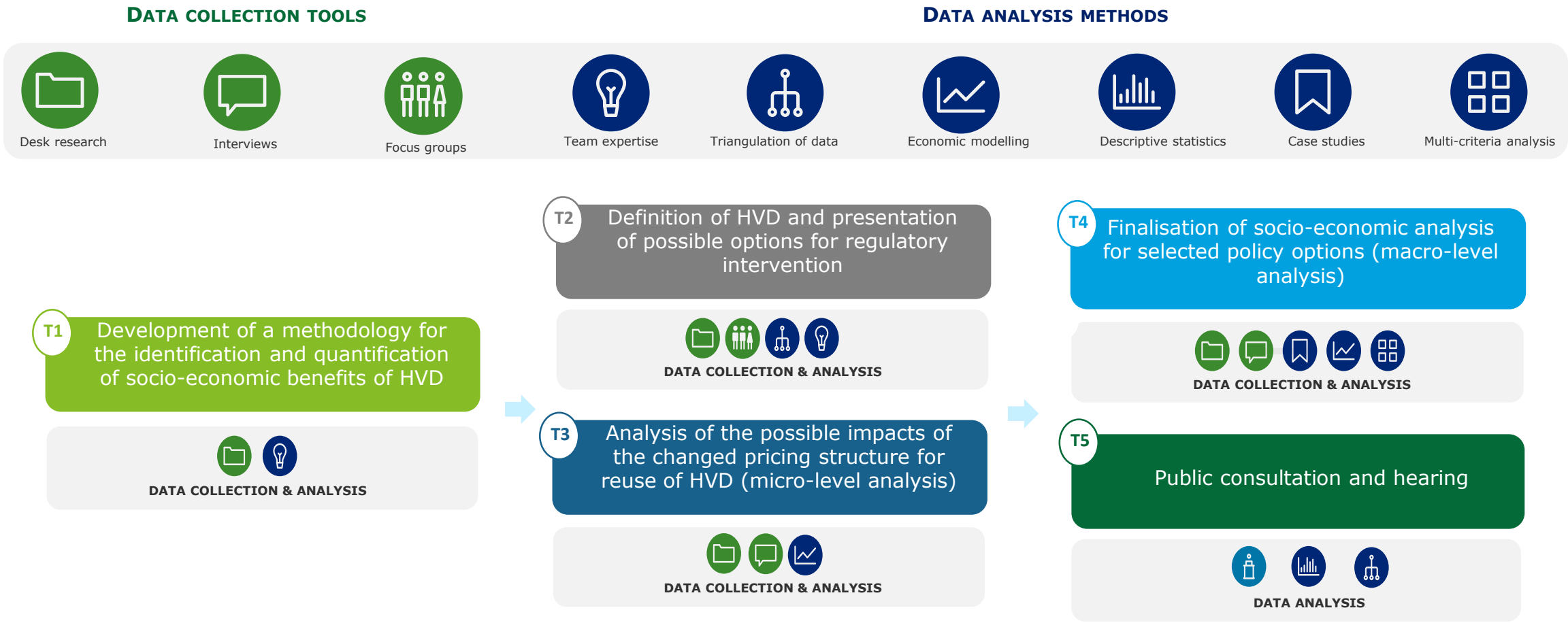
Additional aspects to take into account

- **Business models** of bodies who are currently charging for data and how these will be affected by the Directive;
- **Competition issues** for public undertakings opening up high value datasets whereas their private competitors do not have the same obligation;
- Importance of high value datasets for **SMEs and for building AI applications** and services.



About the study (2/2)

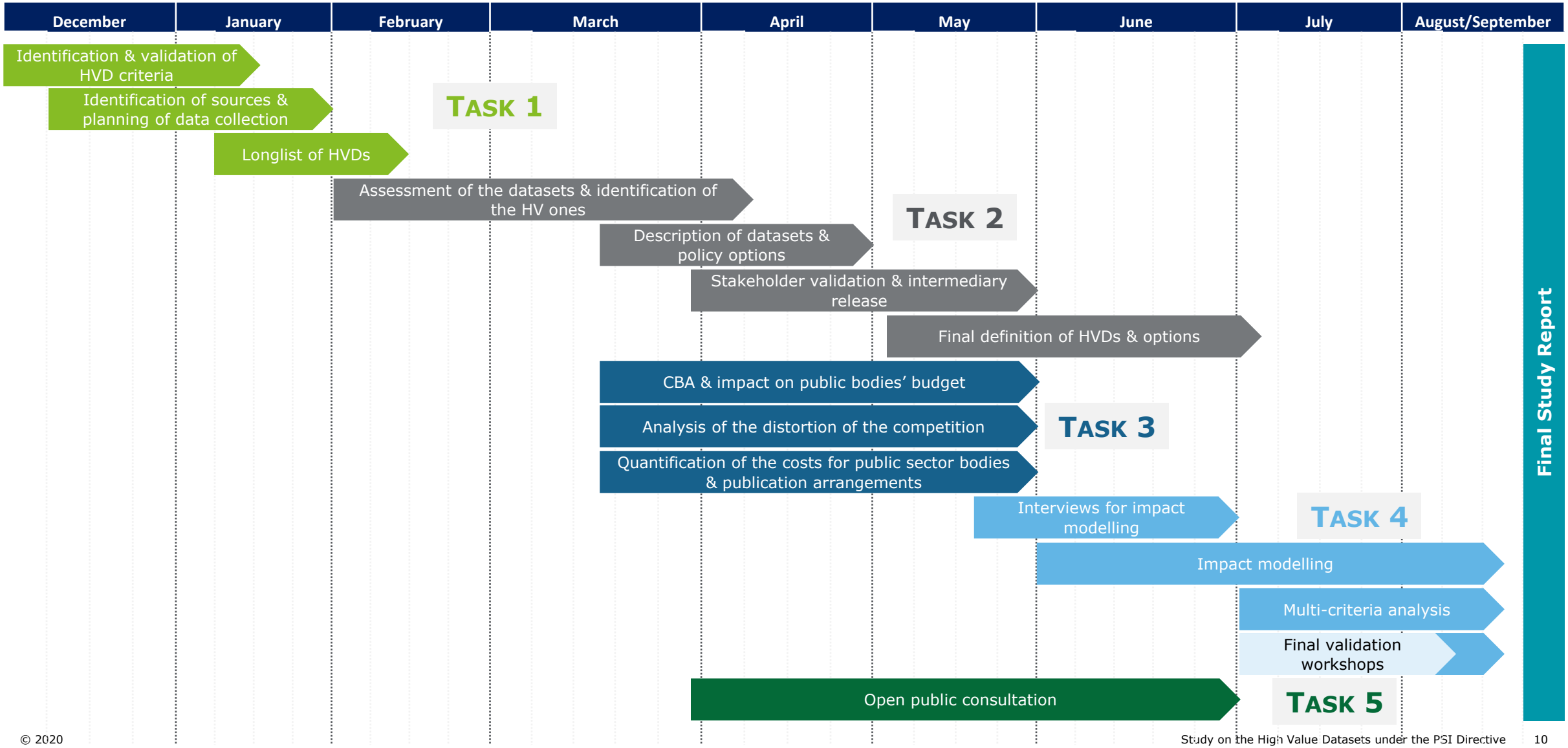
Overall methodological data collection and analysis tools



Main data collection activities across tasks: interviews, online focus groups and desk research.

About the study (3/3)

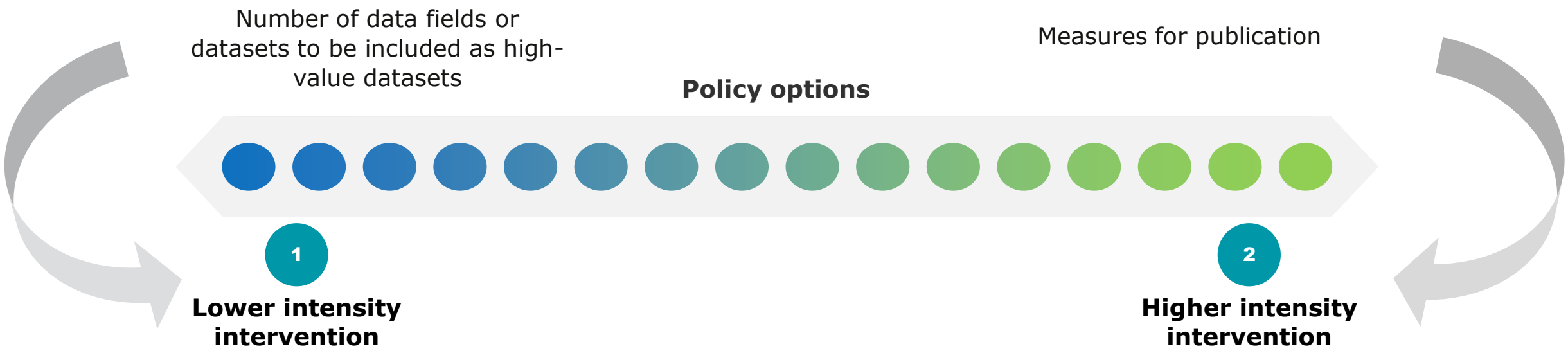
Timeline



About this session
Contents

- Value framework for HVDs
- Datasets in scope of each thematic area
- Policy intervention options per thematic area

Parameters examined for policy intervention options



A value framework for HVDs

Josh D’Addario
The ODI

Finalised value framework for HVDs (1/12)
Methodology

A literature review was conducted to collate categories of value related to the **six macro characteristics of potential value**:

Climate Change	**Economic**	**Innovation and AI**
Public Services and Administration	**Reuse**	**Social**

The review carried out by the study team generated **32 categories of value, supported by 126 quantitative and qualitative indicators**, within the six characteristics.

Finalised value framework for HVDs (2/12)
Methodology

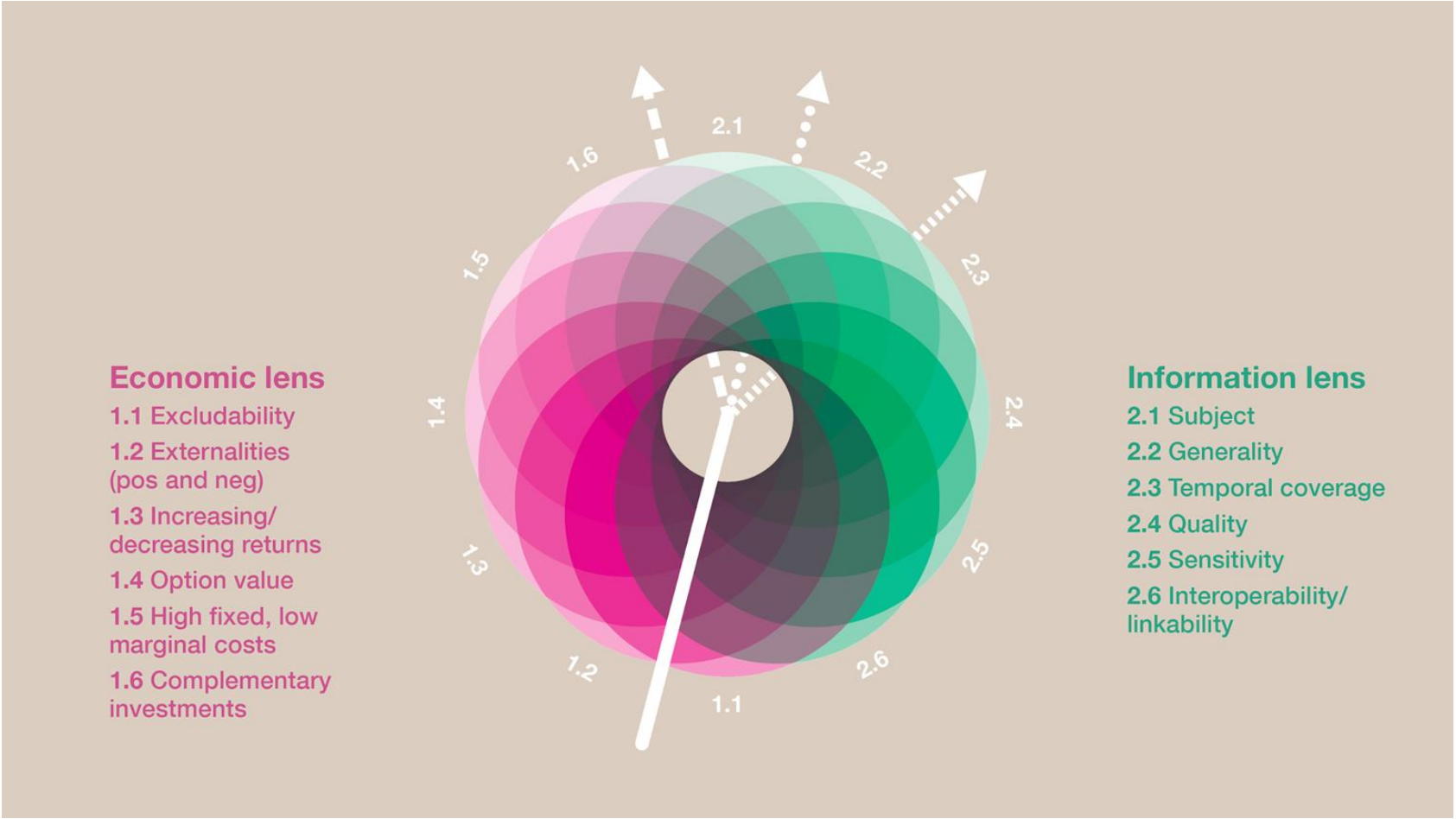
The realisation of value with open data can be considered through **two lenses**

INFORMATIONAL

Associated with the relation and conveyance of information

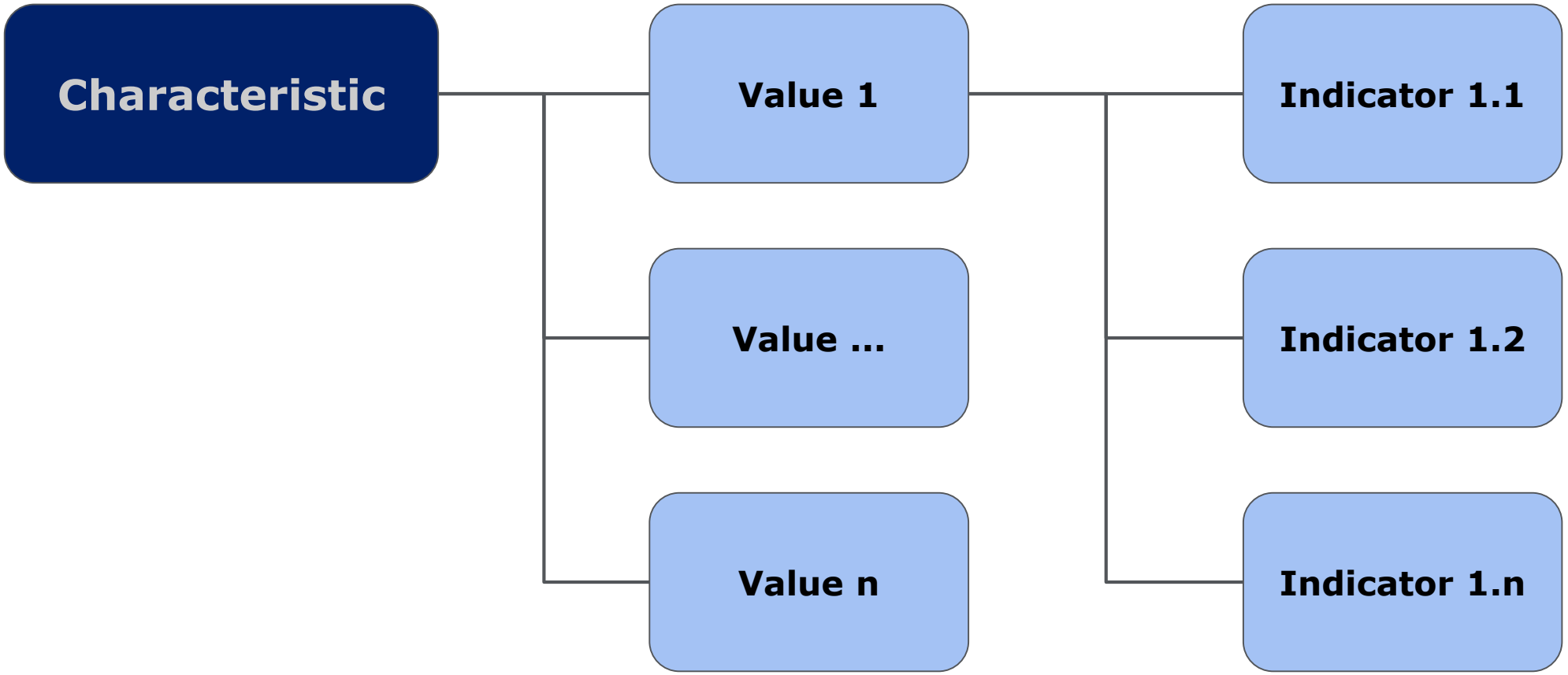
ECONOMIC

Associated with the production, distribution and consumption of data



Open Data Institute & Nuffield Foundation (2020), "Valuing data: foundations for data policy",

Finalised value framework for HVDs (3/12)
Framework



Finalised value framework for HVDs (4/12)

Framework

Climate Change	Open data exploited for understanding and improving environmental conditions and addressing climate change
Citizen engagement in addressing climate change	Citizen understanding of climate change issues; Number of citizen-led environmental initiatives
Climate change	CO2 emissions
Energy management and efficiency	Levels of energy use
Environment management	Air quality; Citizen use of waste disposal sites and schemes; Number of flood prevention schemes; Responsiveness to environmental emergencies

Finalised value framework for HVDs (5/12)

Framework

Economic 	**National income attributable to industries and activities that are based on the exploitation of open data - the value added of open data with respect to the economy as a whole and private sector expenditure on public sector information - and the creation of quality and decent jobs**
Competition	Exit of inefficient firms; Improvements in performance of old firms; Lower firm input costs; Number and level of entry of firms
Consumer benefits	Attractiveness of products to consumers; Purchasing power of consumers
Economic output	Economic output from digital sectors; Market size; Rate of sector growth; Value added
Economy monitoring	Use of imports and exports information; Use of inflation and GDP information
Employment	Number of countries in which new jobs have been created; Number of new high value jobs created; Number of new jobs created; Number of new jobs in high value sectors created; Number of sectors in which jobs created; Rate of increase in the creation of new jobs; …

Finalised value framework for HVDs (6/12)

Framework

Economic 	**National income attributable to industries and activities that are based on the exploitation of open data - the value added of open data with respect to the economy as a whole and private sector expenditure on public sector information - and the creation of quality and decent jobs**
International competitiveness	Level, range, and value of exports
Product market dynamism	Attractiveness of products to consumers; Diversity of products; Number of products bought by consumers; Prices optimization; Product improvement
Productivity and commercialisation	Citizen and firm ease of paying taxes; Efficiency of logistics; Level of business costs; Level of fraud; Level of market insight; Level of reporting demands on business; Level of sales by firms; Number and types of new business model; Number of new products and services; …
Public-private coordination	Quality of coordination between public and private organisations
Trust and transaction costs	Efficiency of market transactions; Reliability and transparency of transactions in the real property sector; Reliability of market transactions

Finalised value framework for HVDs (7/12)

Framework

Innovation & AI 	**New services and applications created through the use of open data in the private sector - in particular by small and medium-size enterprises - or the public sector**
Citizen innovation	Level of self-servicing by citizens
Entrepreneurialism and private sector innovation	Commercial use of public sector technology; Growth of firms using public data; Number of firms using public sector data; Number of new products and services
Public sector innovation	Number of registered developers of apps using public sector information
Public-private coordination	Number of hackathons and similar initiatives for private sector collaboration

Finalised value framework for HVDs (8/12)

Framework

Public Services and Administration	**Open data exploited for public service delivery through access, efficiency, and quality improvements**
Access to public services	Ease of access by citizens to healthcare facilities; Efficiency of school choices made by citizens; Quality of comparisons made by citizens between schools; Use of libraries
Public administration transparency, accountability & engagement	Access to voting locations; Public participation in political and social activities; Public understanding of legislative processes; Public understanding of political processes; Public understanding of public administration and spending
Public sector procurement	Level of competition in bidding

Finalised value framework for HVDs (9/12)

Framework

Public Services and Administration	**Open data exploited for public service delivery through access, efficiency, and quality improvements**
Public sector revenue	Level of tax revenue
Public services management	Cost of maintaining information on companies; Cost of publishing open data relative to previous costs, and returns; Efficiency of public spending allocation; Level of citizen requests for information from public sources; Location accuracy of public services provision; …
Public services performance	Ability of public health services providers to predict risks to health; Improvements in school performance; Quality of healthcare services; Responsiveness of emergency services

Finalised value framework for HVDs (10/12)

Framework

Reuse 	**Open data exploited by intermediaries for a high number of users, with potential for being combined with other datasets, and potential beneficiaries of value-added services and applications based on these datasets**
Demand for information	Level of demand by country; Number of companies using information; Number of customer groups using public sector information portals; Number of established and returning users; Number of licences issued, delivered, sold; Number of new users; …
Trust and confidence in information	User confidence in the market for re-use of public sector information
Volume and range of information	Download volume; Number of types of public sector information used

Finalised value framework for HVDs (11/12)

Framework

Social 	**Open data exploited for improving transparency and accountability, and creating other important benefits for society**
Crime and justice	Compliance with the law; Crime monitoring; Legal system efficiency; Public knowledge of legislation
Disease prediction and prevention	Disease detection; Disease prevention
Mobility access	Number of public transport users; Number of public transport users from marginal groups; Use of foot and cycle paths; Use of healthy mobility schemes by citizens

Finalised value framework for HVDs (12/12)

Framework

Social 	**Open data exploited for improving transparency and accountability, and creating other important benefits for society**
Mobility efficiency	Access to apps by public transport users; Awareness of mobility delays; Hours saved when searching for parking; Savings made from moving from SMS alerts; Time saved from avoiding congestion and delays; Time spent on public transport; Transport cost savings
Mobility planning	Ease of citizen transport planning; Numbers of people walking or cycling; Transport user satisfaction; Use of travel information data
Mobility systems planning	Mobility planning; Transport system integration

Company & Company Ownership

Maruca De Ramón
Deloitte

Company and company ownership

Datasets in scope

Basic information		Companies' documents and accounts	Companies ownership		Companies insolvency status
Non personal	*Personal*		Non personal	Personal	
• Name of the company (full version, in different languages when applicable); • Company status (active, resolved, in liquidation, reconstruction, merger…) • Founding date; • Cessation date (if applicable); • Historical names; • Addresses (i.e. legal, visiting postal); • Legal form; • Identifiers (registration number / company identifier / tax identification number / phone number / e-mail address); • Member State where registered; • NACE code (of the predominant and secondary activities and the code's source); • Number of employees; • Turnover; • Capital; • Detailed information on branches • All changes (to individual companies and list of companies dissolved), and date of the last update.	• (Name(s) of company legal representative(s); • Name of company's directors ;) • The appointment, termination of office and particulars of the persons who either as a body constituted pursuant to law or as members of any such body: ✓ are authorised to represent the company in dealings with third parties and in legal proceedings; it shall be apparent from the disclosure whether the persons authorised to represent the company may do so alone or are required to act jointly; ✓ take part in the administration, supervision or control of the company • All changes (to individual companies and list of companies dissolved), and date of the last update.	• Legal entities; • Detailed data on branches; • Accounting documents, incl. financial statements, non-financial statements, management reports, and other reports (e.g. financial reports, audit reports, corporate governance reports); • Intra-group transactions • Date of the last update; • Other companies documents which are provided to the authority (i.e. companies' meeting minutes).	• Share (percentage) of ownership, and nature and extent of Beneficial Interest held (in shareholdin g and/or voting rights) as well as legal ownership; • Capital links between companies; • All changes, and date of the last update.	• Name of the owner; • Month and Year of birth; • Nationality; • Owner identifier; • Names of shareholder s; • Country of residence of shareholder s / owners.	• Type of insolvency proceeding; • Time limit for lodging claims; • Date of closing main insolvency proceedings; • The court before which the decision opening insolvency proceedings is to be lodged; • All changes (i.e. to individual companies and list of companies dissolved), and date of the last update

Company and company ownership
Expected costs & benefits

Expected costs	
Infrastructural costs	• Costs related to infrastructural investments such as portals, APIs, Servers (could), etc.
Data transformation	• Costs related to data processing including data cleaning, preparation of metadata, aggregation, anonymisation, etc.
Operational costs	• Costs related to data updates, replies to user requests, corrections of errors in the datasets, etc.
Other costs	• Any other costs such as legal advice on GDPR, training costs, etc.

Expected benefits	
Increase in reuse benefits	• Increase overall number of users
Increase economic benefits	• Companies' data hold significant economic value for reusers and constitute a considerable share of the information sector
Increase in AI & innovation	• Citizens' innovation, private sector and public sector innovation would all be fostered by the greater possibilities of reuse of these datasets
Increased social benefits	• Crime fight, public engagement and understanding, and government accountability
Increased env. & climate change benefits	• Information contained in companies' accounts and reports can be used to analyse their environmental impact and sustainability outlook, giving consumers better information for choosing their products & services

Company and company ownership

Lower intensity intervention – Datasets in scope

Basic information		Company documents and accounts	Company ownership		Company insolvency status
Non personal	*Personal*		Non personal	Personal	
• Name of the company (full version, in different languages when applicable); • Company status (active, resolved, in liquidation, reconstruction, merger…) • Founding date; • Cessation date (if applicable); • Historical names; • Addresses (i.e. legal, visiting postal); • Legal form; • Identifiers (registration number / company identifier / tax identification number / phone number / e-mail address); • Member State where registered; • NACE code (of the predominant and secondary activities and the code's source); • *Number of employees;* • *Turnover* • *Capital;* • *Detailed branches of branches* • *All changes (to individual companies and list of companies dissolved), and date of the last update.*	• (Name(s) of company legal representative(s); • Name of company's directors ;) • *The appointment, termination of office and particulars of the persons who either as a body constituted pursuant to law or as members of any such body:* ✓ *are authorised to represent the company in dealings with third parties and in legal proceedings; it shall be apparent from the disclosure whether the persons authorised to represent the company may do so alone or are required to act jointly;* ✓ *take part in the administration, supervision or control of the company* • *All changes (to individual companies and list of companies dissolved), and date of the last update.*		• Share (percentage) of ownership, and nature and extent of Beneficial Interest held (in shareholding and/or voting rights) as well as legal ownership; • Capital links between companies; • *All changes, and date of the last update.*	• Name of the owner; • *Month and Year of birth;* • *Nationality;* • *Owner identifier* • *Names of shareholders;* • *Country of residence of the shareholders/owners*	

Company and company ownership

Lower intensity intervention – Recommended measures

Dimensions		Basic information (non-personal and personal)	Company documents and accounts	Company ownership
Openness-data specification	License (terms of use)	CC.BY 4.0 (or equivalent open license) Terms of use concerning personal data and registration No database right		
	Format	XML		
	Machine-readability	Mandatory		
	Availability of API, bulk download	Both API and bulk download		
Documentation	Metadata (dataset content description)	Complete (*.csv document available)		
	Data linking	No specific recommendation		
	Documentation (incl. structure and semantics)	Complete and web-available		
	Shared vocabularies	Not mandatory		
	Taxonomies	Not mandatory		
Completeness	Traceability	Not necessary		
	Update frequency and timeliness	When available (minimum weekly)		
	Granularity	Individual company level (plus identifier)		
	Key attributes	Company code for disambiguation		

Company and company ownership

Higher intensity intervention – Datasets in scope

Basic information		Company documents and accounts	Company ownership		Company insolvency status
Non personal	*Personal*		Non personal	Personal	
					

Company and company ownership

Higher intensity intervention – Recommended measures

Dimensions		Basic information (non-personal and personal)	Companies documents and accounts	Companies ownership	Companies status
Openness-data specification	License (terms of use)	CC.BY 4.0 No terms of use No database right			
	Format	XML - Json			
	Machine-readability	Mandatory			
	Availability of API, bulk download	Both API and bulk download			
Documentation	Metadata (dataset content description)	Complete (*.csv document available), DCAT-AP			
	Data linking	No specific recommendation			
	Documentation (incl. structure and semantics)	Complete and web-available			
	Shared vocabularies	ISA² Core Vocabularies			
	Taxonomies	Recommended			
Completeness	Traceability	Not necessary			
	Update frequency and timeliness	Real time (minimum daily for insolvency data)			
	Granularity	Individual company level (plus identifier)		Individual owner (plus identifier)	Individual company level (plus identifier)
	Key attributes	Company code for disambiguation		Beneficial owner code for disambiguation	Company code for disambiguation

Geospatial

Francesca Dechiara
FBK

Geospatial

List of HVDs and their value

Datasets	Short description	Use Cases
Administrative Units	Units of administration, dividing areas where Member States have and/or exercise jurisdictional rights, for local, regional and national governance, separated by administrative boundaries. Land Administrative Units and Maritime Units are the basic units. Land Administrative Units are covering mostly land surface, while Maritime Units are covering territorial waters.	Mapping or use as statistical units, manage emergency rescue, waste management plans, protect water ecosystems, find responsible party for policy implementation and administration, forest management, subsidies for farmers, forecast agricultural production, spatial planning, monitoring of regional and urban policy implementation using territorial typologies based on administrative units, maritime spatial planning, integrated coastal management
Place Names	Geographical names or place names (or toponyms) are the proper nouns applied to topographical features and settled (and used) places and spaces on the earth's surface. Toponyms represent an important reference system used by individuals and societies throughout the world.	Emergency response Economic, social and environmental analysis Cultural identity and heritage Mapping and navigation Providing a link / index function to other spatial and aspatial data
Addresses	Location of properties based on address identifiers, usually by road name, house number, postal code. The basic unit of addressing is a building; a permanent construction, intended or used for the shelter of people, having at least one entrance from publicly-accessible space.	Geocoding of statistical surveys, manage emergency rescue, locate where people are, accessibility studies, manage incidents; locate economic activities in ecosystem accounting
Buildings	Geographical location of buildings. Constructions above and/or underground, intended or used for the shelter of humans, animals, things, the production of economic goods or the delivery of services that refer to any structure permanently constructed or erected on its site [from INSPIRE Data Specifications on Buildings].	Buildings are 3D topographic objects and, as such, may influence the propagation of physical phenomena. These data are required for serving citizens (e.g. school, hospital), assessments for air and noise pollution or risk assessments to various kinds of risks (earthquake, fire, flood etc.), monitoring of land consumption, population concentration and access to services.
Cadastral Parcels	Single areas of Earth surface (land and/or water), under homogeneous real property rights and unique ownership, real property rights and ownership being defined by national law.	Protect state lands, reduce land disputes, facilitate land reform, agriculture, land management, taxation, disaster management, real Estate Market, Taxation, LPIS (Agriculture), Land consolidation, Infrastructure Management, Spatial Planning, Protection of Soil and Water, Statistics

Geospatial
Expected costs & benefits

Cost components	Cost components description	Magnitude of costs (range)
Infrastructural costs	Establishment of the API and bulk download, adaptation of the IT infrastructure to real time provision	Initial investment (one time only) depending on the solution, in between 250,000 and 3,000,000 EUR. For further developments, depending on the country and the size, an example is the costs for data storage device: 450,000 EUR (once off)
Data transformation costs	Costs related to data processing including data cleaning, preparation of metadata, aggregation, anonymisation, etc.	In between 100,000 and 200,000 EUR (yearly)
Operational costs	Costs related to data updates, replies to user requests, corrections of errors in the datasets, etc.	In between 150,000 and 350,000 EUR (yearly)
(Lost) income for data supplier	(Share of) revenue related to the provision of the HVD	Depending on country. (e.g. Sweden:90mln SEK)
Other costs	Any other costs such as legal advice on GDPR, training costs, etc.	Not Available
Negative impact on competition	The estimated impact of competition distortion vis-à-vis private organisations active in the domain.	Not available

Geospatial Lower intensity intervention – Recommended measures

	Description	Administrative units	Place Names	Addresses	Buildings	Cadastral parcels
Openness	License and terms of use	CC-BY 4.0				
	Format	GeoPackage; GeoJSON; INSPIRE requirements.	GeoPackage; CSV; GeoJSON; INSPIRE requirements.	GeoPackage; CSV; GeoJSON; INSPIRE requirements.	GeoPackage; GeoJSON; INSPIRE requirements.	GeoPackage; GeoJSON; INSPIRE requirements.
	Machine-readability	Mandatory				
	Availability of API, bulk download	Bulk download; INSPIRE distribution services; RestAPI (e.g. OGC API, ArcGIS RestAPI, Carto API).				Read-only mode (WMS service defined by INSPIRE).
Documentation	Metadata (dataset content description)	INSPIRE				
	Documentation (incl. structure and semantics)	INSPIRE; GeoDCAT-AP.				
	Data linking					
	Shared vocabularies/taxonomies	INSPIRE				
Completeness	Traceability	National Geodata Catalog and/or open data catalog				
	Update frequency and timeliness	Annual update	When necessary	When necessary	When necessary	When necessary
	Granularity	From municipalities to countries; sea-frontiers.	National coverage.	Partial National coverage (e.g. most populated cities).	Partial National Coverage (e.g. most populated cities); Level of scale 1:5000.	National coverage; Level of scale 1:5000.
	Key attributes	National identification code; identification code of the upper administrative level; official name; country code; name in multiple languages (only for countries with more than one official language).	Name; name in multiple languages (only for countries with more than one official language); category; latitude and longitude (INSPIRE)	Latitude and longitude; house number; suffix of the number; name of the street; name of the municipality; national identification code of the municipality; last update.	Footprint of the building; entrances; floors; type of use.	Geometry of cadastral parcels; type of particle; particle code; references to the administrative area to which the particle belongs.

Geospatial

Lower intensity intervention – Recommended measures

- Concerning **licences and terms of use**, the recommended licence for the lower intensity intervention is the Creative Commons Attribution 4.0 International **-** CC-BY 4.0. This recommended option will have no impact since this type of licences is widely used across Member States (MS). In the same time, this type of licence is also preferred by re-users.

- **APIs and bulk download**: Regarding the cadastral parcels - the recommendation is to ensure at least access through WMS service as defined by INSPIRE, in read-only mode, due to the difficulties in the release of this dataset. The WMS remains a data product which allows the creation of new georeferenced data and the view of the geo-data as a whole. The WMS protocol provides feature information (as XML) by identifying a point on a map. For the other datasets, the download through API, bulk download, OGC services listed by INSPIRE should be guaranteed.

- When it comes to **formats**, an option for the publication of datasets is to follow INSPIRE requirements. According to the stakeholders' suggestions, the recommendation is to use Geopackage and GeoJSON, which are two relatively recent open and low-cost formats. GeoJSON is commonly used, the re-users have shown a strong preference for new open standards like Geopackage. Using these formats increases the machine-readability of the data. The proposal includes the use of dedicated services for automatic conversion is added.

- **Granularity:** concerning Addresses and Buildings datasets the recommendation is to guarantee at least a partial coverage (e.g. most populated cities). Based on our interviews, the findings already show a diversified situation across the MS and main issues are related to the data ownership at local level, the costs of management and the frequency of update that strongly affect the lack of a full national coverage. However, the full national coverage is recommended for the rest of the datasets. The level of scale to be guaranteed for the Buildings and Cadastral parcels is 1:5000 or beyond. The increase of detail in the level of scale corresponds to a raise in the costs of the dataset production.

- The **metadata** should match the INSPIRE requirements as the transformation to geoDCAT-AP can be done automatically from this basis. It is important that the data are also displayed in geoDCAT-AP in order to integrate them automatically into national open data portals.

- When it comes to **key attributes**, the values identified correspond to the common characteristics available across the majority of the Member States, for each type of dataset in scope. This is due to the implementation of the INSPIRE directive. Therefore, the impact on the organisations will remain rather low and limited. The interviewed stakeholders highlighted the relevance of having names in multiple languages as a key attribute, because it significantly improves the reuse. Regarding the Buildings dataset, entrances and floors are recommended in the list of the key attributes, as based on these elements it's possible to calculate the height of the building. The type of use could be generic and very basic. The recommended attributes for the Cadastral Parcels are the basic ones needed to release cadastral data with respect to the GDPR and to guarantee a good level of reusability.

Geospatial Higher intensity intervention

	Description	Administrative units	Place Names	Addresses	Buildings	Cadastral parcels
Openness	License and terms of use	CC0				
	Format	GeoPackage; GeoJSON; INSPIRE requirements.	GeoPackage; CSV; GeoJSON INSPIRE requirements.	GeoPackage; CSV; GeoJSON; INSPIRE requirements.	GeoPackage; GeoJSON; INSPIRE requirements.	GeoPackage; GeoJSON; INSPIRE requirements.
	Machine-readability	Mandatory				
	Availability of API, bulk download	Bulk download; INSPIRE distribution services; RestAPI (e.g. OGC API, ArcGIS RestAPI, Carto API).				
Documentation	Metadata (dataset content description)	INSPIRE				
	Documentation (incl. structure and semantics)	INSPIRE / GeoDCAT-AP				
	Data linking					
	Shared vocabularies/taxonomies	INSPIRE				
Completeness	Traceability	National geodata Catalog and open data catalog.				
	Update frequency and timeliness	Annual update	When necessary	When necessary	When necessary	Continuous update (close to real-time).
	Granularity	From municipalities to countries; sea-frontiers.	National coverage	National coverage	National coverage; Level of scale 1:5000 (minimum requirement).	National coverage; Level of scale 1:5000 or beyond (1:2000).
	Key attributes	National identification code; identification code of the upper administrative level; official name; country code; name in multiple languages (only for countries with more than one official language).	Name; name in multiple languages (only for countries with more than one official language); category; latitude and longitude (INSPIRE)	Latitude and longitude; house number; suffix of the number; name of the street; zip code; name of the municipality; national identification code of the municipality; last update; type of position.	Footprint of the building; height; entrances; floors; type of use.	Geometry of cadastral parcels; type of particle; particle code; references to the administrative area to which the particle belongs.

Geospatial

Higher intensity intervention – Recommended measures

The main differences with the lower intensity intervention refer to the licence, the APIs and few changes on granularity and key attributes options for the datasets in scope:

- Concerning **licences and terms of use**, the recommendation is the use of CC0. According to our research, this type of licence is already adopted for the datasets in scope across several Member States. However, it continues to be not applied because it implies legal issues of compatibility and raises a lot of scepticism in terms of lack of attribution, and charging of responsibility.

- When it comes to the availability trough **APIs** and **bulk download**, the option for SPARQL Endpoint was initially considered for this higher-intensity policy option, but no longer available in this version. This would imply a more accurate implementation in the use of sharing vocabularies and metadata. This is already available in the Netherlands, for example. Also, the distribution and the download through APIs option is extended to the Cadastral Parcels.

- For **granularity**, the recommended level of scale is 1:5000 or beyond, for buildings and cadastral parcels datasets. In the case of Cadastral parcels, the range of the level of scale of the existing datasets varies from 1:10000 to 1:500. In most of the countries, the data is available at a scale of 1:2000. Concerning administrative units and place names, the scale depends on what needs be visualized.

- **Traceability:** The issues related to the interoperability between the geo-catalogues and open data catalogues can be solved by using the GEO-DCAT AP and DCAT-AP 1.0. However, the actual implementation of the open data catalogues encompasses the use of DCAT-AP 2.0, which results to be incompatible with the GEO-DCAT AP. This policy option should support a resolution of these challenges.

- **Key Attributes**: Concerning the Addresses in this intervention, the zip code is here added as a key attribute, although this information is privately owned in most of the cases and, could be significantly hard to obtain. Regarding the Buildings, the proposal of using dedicated services for automatic conversion is added. It is fundamental and very relevant in terms of value and reuse scenarios. This information is costly for data holders, if not collected from the beginning. Alternatively, the number of floors (levels) should be guaranteed. The type of use should be as detailed as possible. In this case, please see the INSPIRE recommendations, the information on the types of building are defined in their Annex III. The Cadastral Parcels' key attributes remain unchanged compared to the previous policy intervention.

Meteorological data

Ton Zijlstra

The Green Land

Meteorological

List of HVDs and their value

- Data holders and re-users see clear value chain for meteorological data
- Value created in large variety of sectors
- MS with open data policy show non-lineair increase in demand as well as benefits generally outweighing costs / new tax revenue outweighing costs by multiples.
- High friction comes from fees and mixed licensing (also in comparison with more uniform open US market)

Data sets
Observations data
Climate data (incl. validated observations)
Weather Alerts
Radar data
Numerical Weather Prediction model data

Meteorological
Expected costs and benefits

Expected costs	
Infrastructural costs	• Costs related to infrastructural investments, specifically **APIs**, are main concern, next to timely provisioning.
Data transformation	• Low costs expected related to data processing including data cleaning, preparation of metadata, aggregation, etc. unless uncommon file formats demanded
Operational costs	• Where third party cloud services are used, data volume and demand may mean increased marginal costs.
Other costs	• **Charging** for data is a common practice (13 MS charge for observations data) Marginal costs models regularly lack transparency. Experience suggests compensating revenue loss from general budget.

Expected benefits	
Increase in reuse benefits	• Increase overall number of users, and volume per user. Removing fees leads to **non-lineair jump in re-use**, and new commercial entrants.
Increase economic benefits	• Wide variety of sectors re-use data, Singular use cases can be bigger than total cost of transition. Over time **new tax revenue bigger** than costs of provisioning.
Increase in AI & innovation	• Re-use outside traditional meteorological value added services will grow. Novel uses emerge based on **real-time data** (e.g. IoT), or less used variables.
Increased social benefits	• More consumer centric services, informing **daily decisions** (e.g. mobility, sports)
Increased env. & climate change benefits	• Growing citizen science communities w.r.t. climate adaptation. **Strong role** in green energy (planning, and operations) Usage in agriculture to reduce water and pesticide usage and climate adaptation.

Meteorological

Lower intensity intervention

- High friction comes from fees, and dealing with loss of revenue is unavoidable even at low intensity interventions. 13 MS currently charge for observations data.
- Observations are the starting point of the meteo-data value chain
- Timeliness, completeness & highest available temporal/geographic resolution determine re-use value
- Weather alerts are at core of the public task, already published as information
- Validated observations feed into the climate record, adding to digital structured historical climate data

Data sets
Observations data
Structured historical climate data (incl. validated observations)
Weather Alerts
Radar data
Numerical Weather Prediction model data

Meteorological

Lower intensity intervention Recommended measures

Dimensions		Observations	Validated observations and digitised structured historical data	Weather Alerts
Openness-data specification	*License (terms of use)*	Creative Commons 0 or BY No terms of use	Creative Commons 0 or BY No terms of use	Creative Commons 0 or BY No terms of use
	Format	BUFR, NetCDF, ascii (RT), JSON (for hourly)*	NetCDF, JSON*	XML (CAP and/or RSS-Atom)*
	Machine-readability	Obligatory	Obligatory	Obligatory
	Availability of API, bulk download	Both API and bulk download	Both API and bulk download	Both API and bulk download
Documentation	*Metadata (dataset content description)*	Complete (*.csv document available)	Complete (*.csv document available)	Complete (*.csv document available)
	Documentation (incl. structure and semantics)	Complete and web-available	Complete and web-available	Complete and web-available
Completeness	*Update frequency and timeliness*	Every 5-10 minutes in real time for automated stations, hourly unvalidated, for the last 24hrs	Validated hourly (or better temporal resolution) published at least daily and daily average observations data ; structured historic data	As issued, or hourly
	Granularity	Per weather station, full temporal resolution	Per weather station, full temporal resolution	Alerts, 48hrs or more ahead
	Key attributes	All observation variables measured	All validated observation variables measured; daily averages per variable	

No specific recommendations for data linking, shared vocabularies, taxonomies, or traceability (except NWS or National open data portal w.r.t. alerts.)

* Or other formats currently practiced in the field, using open standards

Meteorological
Higher intensity intervention

- Adds Radar data
- Adds Numerical Weather Prediction model data
- High volume, especially NWP ensemble data (depending on data format, and choice of parameters/levels)
- European Meteorological Infrastructure (EMI) might play a role for data provision
- Adds unstructured but digitised historical climate data (data transformation costs maybe involved)

Data sets
Observations data
Climate data (incl. validated observations)
Weather Alerts
Radar data
Numerical Weather Prediction model data

Meteorological

Higher intensity intervention Recommended measures

Dimensions		Observations	Climate	Weather Alerts	Radar	NWP model data
Openness-data specification	*License (terms of use)*	Creative Commons 0 or BY No terms of use	Creative Commons 0 or BY No terms of use	Creative Commons 0 or BY No terms of use	Creative Commons 0 or BY No terms of use	Creative Commons 0 or BY No terms of use
	Format	BUFR, NetCDF, ascii, json (for hourly)*	NetCDF,JSON*	XML (CAP and/or RSS-Atom)*	HDF5,JSON*	GRIB (or NetCDF)*
	Machine-readability	Obligatory	Obligatory	Obligatory	Obligatory	Obligatory
	*Availability of API, bulk downloa**d***	Both API and bulk download	Both API and bulk download	Both API and bulk download	Both API and bulk download	Both API and bulk download
Documentation	*Metadata (dataset content description)*	Complete (*.csv document available)	Complete (*.csv document available)	Complete (*.csv document available)	Complete (*.csv document available)	Complete (*.csv document available)
	Documentation (incl. structure and semantics)	Complete and web-available	Complete and web-available	Complete and web-available	Complete and web-available	Complete and web-available
Completeness	*Update frequency and timeliness*	Every 5-10 minutes in real time for automated stations, hourly unvalidated for the last 24hrs	Validated hourly (or better temporal resolution) published at least daily and daily average observations data; historic data	As issued, or hourly	Near real time in 5 minute intervals (or available shortest interval)	Every 6hrs, or better temporal resolutions, from the last 24hrs
	Granularity	Per weather station, full temporal resolution	Per weather station, full temporal resolution	Alerts, 48hrs or more ahead	Per radar station in the MS, and national composite	48hrs ahead or more in 1hr steps, national, at 2.5km/best available grid
	Key attributes	All observation variables measured	All validated observation variables measured		Reflectivity, backscatter, polarization. Precipitation, wind and echo-tops	Deterministic, and/or ensembles if available, for meteorologically relevant parameters and levels

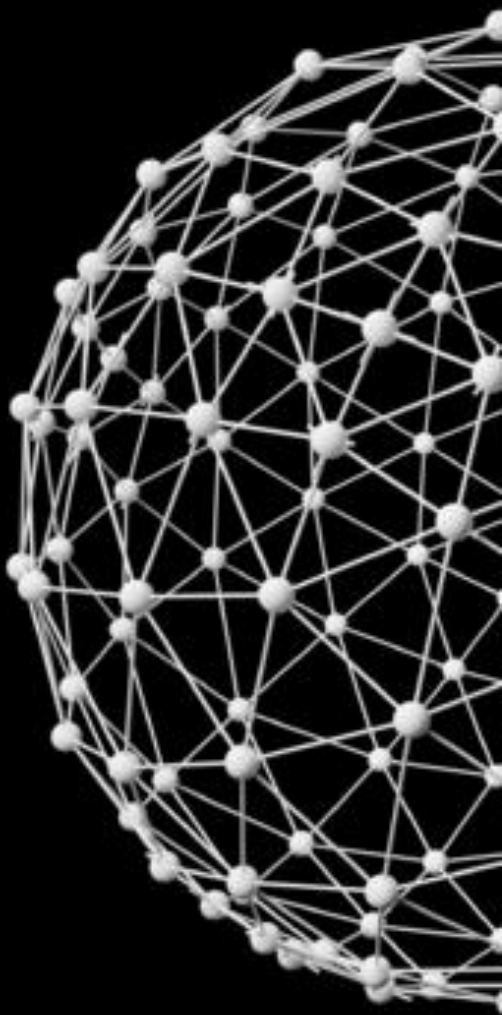
No specific recommendations for data linking, shared vocabularies, taxonomies, or traceability (except NWS or National open data portal w.r.t. alerts.)

* Or other formats currently practiced in the field, using open standards

Earth observation & Environment

Ton Zijlstra

The Green Land



Earth observation and environment

List of HVDs and their value

Earth Observation and Environmental extremely broad scope and fragmented (MS suggested hundreds of different datasets). Copernicus services depend on in-situ data from MS. Use cases broad across user groups and sectors, incl. public sector, but often hard to quantify (though many usage examples exist).

Main friction in fragmentation of data, diverse terms of use, fees.

All data within scope of INSPIRE and Environmental e-reporting priority data: adds open data requirement to INSPIRE / E-prio.

Earth observation	Environmental data	
Hydrography (I)	Air quality (prio)	Protected sites (I)
Land parcels (I), cover (II), use (III)	Biodiversity (prio)	Bio-geographical regions (III)
Elevation (II)	Emissions (prio)	Environmental monitoring facilities (III)
Geology (II)	Nature preservation (prio)	Habitats/Biotopes (III)
Ortho-imagery (II)	Noise (prio)	Natural Hazards (III)
Oceanography (III)	Waste (prio)	Soil (III)
Sea Regions (III)	Water (prio)	Species distribution (III)

Earth observation and environment

Expected costs and benefits

Extremely broad scope (MS suggested hundreds of different datasets), assessing costs and benefits hard to quantify. Re-use benefits often based on combinations of data from various subthemes, measurement data e.g. needs administrative data to be useful. No real 'hierarchy' of use value to be made.

Expected costs

Infrastructural costs	• Costs related to infrastructural investments, specifically **APIs**, are main concern.
Data transformation	• Low costs expected, compared to infrastructure, related to data processing including data cleaning, preparation of metadata, aggregation, etc. **unless significantly different from INSPIRE** requirements. Collating data from dispersed sources also a cost, yet its own benefit too.
Operational costs	• Rarely made explicit. Rising demand may increase bandwidth costs.
Other costs	• **Charging** for certain EO data is common, also between PSB's, with free data having re-use limitations applied. Experience suggests compensating revenue loss from general budget.

Expected benefits

Increase in reuse benefits	• **Increased demand** from citizens, research, businesses, and public sector. Removing friction in access to **in-situ data very important** to improve services e.g. Copernicus
Increase economic benefits	• **Wider variety of sectors** re-use data, (EARSC identified 100+ products and services in 22 market sectors).
Increase in AI & innovation	• **New services mixing EO/environmental data**. European Data Portal now lists over 40 such new services. Novel uses w.r.t. **execution of public tasks**.
Increased benefits for public services	• **Efficiency gains** for public sector bodies, including data fees. **Better public sector performance** of PSB's other than dataholder.
Increased social benefits	• Reporting obligations create demand w.r.t. transparency and **public engagement**.
Increased env. & climate change benefits	• Growing citizen science communities. **Strong role** in energy, agriculture, environment management and climate adaptation.

Earth observation and environment

Lower intensity intervention

Limits scope to Environmental e-reporting priority data including environmental measurements. Reporting obligations create their own re-use demand, but mostly re-use depends on combinations with other data now outside this scope.

Adds open data obligation to e-reporting priority data and Environmental monitoring facilities (III), removing non-commercial restrictions mostly.

Earth observation	Environmental data	
Hydrography (I)	Air quality (prio)	Protected sites (I)
Land parcels (I), cover (II), use (III)	Biodiversity (prio)	Bio-geographical regions (III)
Elevation (II)	Emissions (prio)	Environmental monitoring facilities (III)
Geology (II)	Nature preservation (prio)	Habitats/Biotopes (III)
Ortho-imagery (II)	Noise (prio)	Natural Hazards (III)
Oceanography (III)	Waste (prio)	Soil (III)
Sea Regions (III)	Water (prio)	Species distribution (III)

Earth observation and environment

Lower intensity intervention Recommended measures

Dimensions		E-reporting priority data and Environmental Monitoring Facilities	
Openness-data specification	*License (terms of use)*	Creative Commons 0 or BY (or equivalent open license) No terms of use	
	Format	Cf. existing INSPIRE specifications	
	Machine-readability	Obligatory	
	Availability of API, bulk download	Both API and bulk download, as prescribed in e-reporting priority list	
Documentation	*Metadata (dataset content description)*	Complete (INSPIRE)	
	Documentation (incl. structure and semantics)	Complete and web-available (INSPIRE)	
Completeness	*Update frequency and timeliness*	As collected, for EMF highest collected temporal resolution	
	Granularity	Highest collected temporal and geographic solution	
	Key attributes	All attributes mentioned in e-reporting priority data list and EMF locations, parameters measured and complete observations data	

No specific recommendations for data linking, shared vocabularies, taxonomies, or traceability, other than existing INSPIRE specifications

Earth observation and environment

Higher intensity intervention

Includes Environmental e-reporting priority data, and covers the listed INSPIRE themes in full. Adds open data requirements to INSPIRE (download) services. Allows the combinations found across varied use cases.

Removes restrictive terms of use and fees. INSPIRE's data harmonisation efforts extend to open data. Follows current/future INSPIRE standards.

Earth observation	Environmental data	
Hydrography (I)	Air quality (prio)	Protected sites (I)
Land parcels (I), cover (II), use (III)	Biodiversity (prio)	Bio-geographical regions (III)
Elevation (II)	Emissions (prio)	Environmental monitoring facilities (III)
Geology (II)	Nature preservation (prio)	Habitats/Biotopes (III)
Ortho-imagery (II)	Noise (prio)	Natural Hazards (III)
Oceanography (III)	Waste (prio)	Soil (III)
Sea Regions (III)	Water (prio)	Species distribution (III)

Earth observation and environment

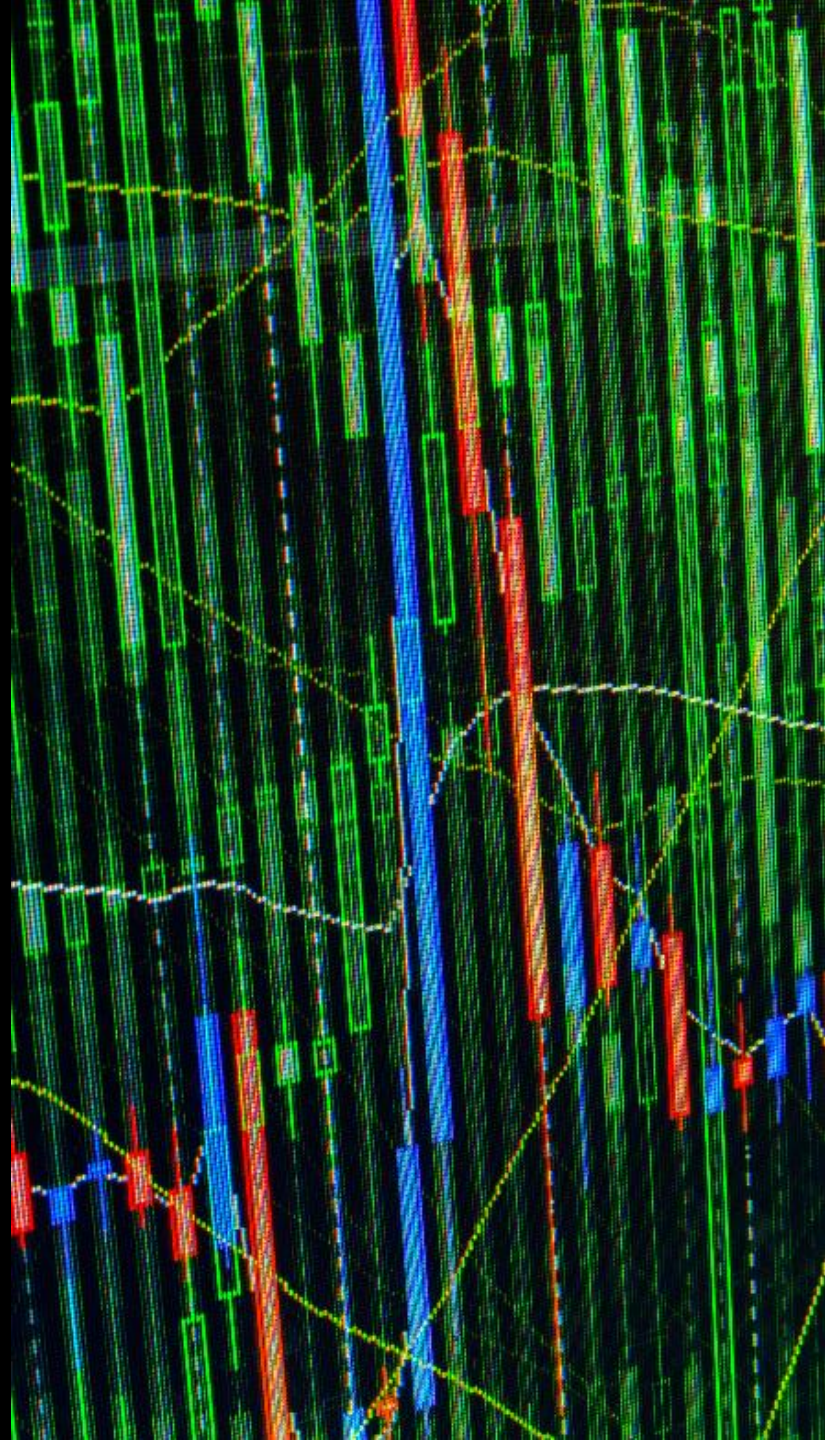
Higher intensity intervention Recommended measures

Dimensions		E-reporting priority data and Environmental Monitoring Facilities	INSPIRE Themes
Openness-data specification	*License (terms of use)*	Creative Commons 0 or BY (or equivalent open license) No terms of use	Creative Commons 0 or BY (or equivalent open license) No terms of use
	Format	Cf. existing INSPIRE specifications	Cf. existing INSPIRE specifications
	Machine-readability	Obligatory	Obligatory
	Availability of API, bulk download	Both API and bulk download, as prescribed in e-reporting priority list	INSPIRE download services (e.g. WFS or OGC:API) and bulk download
Documentation	*Metadata (dataset content description)*	Complete (INSPIRE)	INSPIRE, or DCAT-AP2
	Documentation (incl. structure and semantics)	Complete and web-available (INSPIRE)	INSPIRE
Completeness	*Update frequency and timeliness*	As collected, for EMF highest collected temporal resolution	INSPIRE
	Granularity	Highest collected temporal and geographic solution	INSPIRE
	Key attributes	All attributes mentioned in e-reporting priority data list and EMF locations, parameters measured and complete observations data	All features within INSPIRE theme's scope

No specific recommendations for data linking, shared vocabularies, taxonomies, or traceability, other than existing INSPIRE specifications

Statistics

Cristina Moise
The Lisbon Council



Statistics

List of HVDs and their value

Official statistics become high-value datasets based on their use - who needs them and why, how and where are they applied. In general, stakeholders and re-users agree that **all official statistics** should be considered high-value datasets.
Due to the broad selection of datasets within the official statistics, and considering the inputs received from stakeholders, we propose a short list of datasets as potential high-value datasets covering social, macro-economic and business statistics.

Social statistics reflects the countries' population social conditions through the perspective of demography, labour market, poverty and inequality and health aspects. **Demography** concentrate on the overall structure and trend of population development. It information information on *population*, and vital statistics - *births* (including *fertility rates*) and *deaths* (including *mortality rates*) statistics. **Labour market data** provide another set of information about of the socio-economic environment within a country. It helps identifying economic needs and niches of opportunities, and reflects the labour market dynamism via *employment* and *unemployment* rates, *job vacancies* and *labour force potential*, pointing out to both hotspots for development and structural issues within economies. **Poverty** and **inequality** data reflects the population social struggle and overall well-being. Material and social deprivation, at-risk-of-poverty and low work intensity households are important aspects of societal development of countries. An aging population needs better healthcare services. Thus, health data and, in particular, current healthcare expenditure provide relevant information on costs, general services and providers at countries' level.

Macro-economic statistics provides relevant information about the countries' development trend with the help of national accounts, price statistics and government finances datasets. The economy performance of a country is reflected in the **national accounts** data through GDP and main aggregates, key indicators on corporations and key indicators on households statistics. **Price statistics** brings in addition information about the economic stability. The transparency and accountability of the public spending is an important aspect of a country's economy. **Government finances data** provides information on efficiency and effectiveness use of public money, and help identify the government spending patterns.

Business statistics often covers the business sector's economic performance, reflected by data about short-term business statistics, trade and tourism. Business communities are an essential components of the countries' economic development, and **short-term business data** and **trade data** provide the relevant information to monitor their competitiveness and performance. In some of the Member States, the **tourism flows** constitute an important share of economy.

Statistics
Expected costs and benefits

Expected costs

- It is difficult to identify specific costs for particular datasets as in most cases it is part of the regular activity of the organization
- **Infrastructural costs** (*related to portals, APIs, Servers (could), etc.*) depend significantly on the current IT development structure and country size. Often such an overall estimate will include both on-time investments (e.g. development and implementation of a solution), and yearly maintenance costs of the infrastructure itself. Exact values were difficult to gather.
- **Data transformation costs** (*related to data processing including data cleaning, preparation of metadata, aggregation, anonymisation, etc.*) are also part of current activities of an NSI, and often there are no breakdowns by datasets or parts of them.
- **Operational costs** (*related to data updates, replies to user requests, corrections of errors in the datasets, etc.*) are another category that is part of the day-by-day of the NSIs, and a breakdown by specific costs was difficult to provide.
- **Other costs** (*related to legal advice on GDPR, training costs, etc.*) were also difficult to quantify

Expected benefits

- Statistics have different types of users: media and general public, international policies and organisations, decision makers, analysts, non-governmental organisations (NGOs), the civil society, producers of statistics and the scientific community and researchers.
- Better outcomes from the **decision makers** (public sector, business sector or individuals) by making use of high-quality information to perform their function.
- Increase in transparency and improvement of the accountability of public and other bodies from the **citizens** side
- Better understanding of the society by providing relevant information while respecting the rights of people, at the same time
- Enabling new **research** and **analysis** on the basis of a comprehensive evidence-base, leading to innovation and improved economic and social outcomes.

Statistics

General context and datasets overview

As part of the European Statistical System, the statistics are both standardised and harmonised across Member States.

Also, the datasets in scope are freely available in all Member States.

Category	Sub-category	Datasets, key variables and breakdowns
Social statistics	**Demography**	The datasets in scope are **population, fertility, mortality, births** and **deaths**. *Key variables* are population, births and deaths, life expectancy, crude rates for births and for deaths (including infant mortality), fertility rates (total, by mother's age), median age, old age dependency.
	Labour market	The datasets in scope are **employment, unemployment** and **potential labour force**. *Key variables* are employment and unemployment rate, types of employment (full-time, part-time, self-employed, temporary), duration of unemployment (including long-term unemployment), job vacancies, persons looking for jobs.
	Poverty and inequality	The datasets in scope are **poverty** and **inequality**. *Key variables* are people at-risk of poverty and social exclusion, severe material deprivation, material and social deprivation, households with low work intensity, income quintile, Gini coefficient of equivalised disposable income.
	Heath statistics	The dataset in scope is **current health expenditure**. *Key variable* is the current health expenditure.
Macro-economic	**National accounts**	The datasets in scope are **GDP and main aggregates, key indicators on corporations** and **key indicators on households**. Key variables include gross domestic product at market prices, gross value-added, final consumption expenditure of households, exports and imports, taxes and subsidies, gross capital formation (total and non-financial corporations, households), gross operating surplus and mixed income of non-financial corporations, financial sector assets and liabilities, compensation of employees, households sector assets and liabilities employment, disposable income and savings of households, gross national income, net lending/borrowing (total economy, financial and non-financial corporations, households).

Statistics

List of HVDs and their value

Category	Sub-category	Datasets, key variables and breakdowns
Macro-economic	**Price statistics**	The dataset in scope is **Harmonised Index of Consumer Prices (HICP)**. *Key variables* are HICP monthly index and rate of change (monthly and annually), HICP annual (rate of change).
	Government finances	The datasets in scope are **government expenditure and revenue** and **consolidated government gross dept**. *Key variables* are general government revenue, expenditure, net lending/borrowing of the general government, government gross debt.
Business statistics	**Short-term business statistics**	The datasets in scope are **industrial production, industrial producer price index** and **production in construction**. *Key variables* are industrial production index, industrial producers price index, domestic and non-domestic producers price indexes (euro and non-euro areas), production in construction index.
	Trade statistics	The datasets in scope are **retail trade volume** and **EU international trade in goods (imports and exports)**. *Key variables* are nights spent at tourist accommodation establishments, participation in tourism, tourism trips and expenditure made by EU residents, tourism nights spent by EU residents.
	Tourism statistics	The dataset in scope is **tourism flows in Europe**. *Key variables* are retail trade volume index, statistical values, net mass and supplementary quantity of international trade.

The policy options will focus more on the measures for publication rather than the number of data fields and/or datasets to be included as high-value datasets.

Statistics

Lower intensity intervention

	Description	Demography, labour market, GDP, government finances, business statistics, social statistics
Openness	Licence and terms of use	CC-BY 4.0
	Format	CSV, XML (SDMX), JSON
	Machine-readability	Mandatory
	Availability of API, bulk download	Simple structured APIs and partial bulk download available
Documentation	Metadata (dataset content description)	Simple structured file (e.g. using official statistics baseline standard to provide basic information only)
	Documentation (incl. structure and semantics)	Complete and web available
	Data linking	N/A
	Shared vocabularies/taxonomies	N/A

- The impact on the data holders' organisations is considered to remain **low**, as many of the recommendation are already available in the Member States
- The formats proposed are already commonly used by Member States (CSV in 22 out of 27 countries, XML/SDMX in 15, JSON in 10). Often, more than one format is provided
- The main cost driver is the development of APIs, which will mostly impact the countries that do not provide them already (12 Member States)

Statistics

High intensity intervention

	Description	Demography, labour market, GDP, government finances, business statistics, social statistics
Openness	Licence and terms of use	CC BY 4.0
	Format	CSV, XML (SDMX), JSON
	Machine-readability	Available
	Availability of API, bulk download	**Complex APIs (including customised clients) and bulk download available**
Documentation	Metadata (dataset content description)	**Well-developed structured file (i.e. description of the statistical data, as well as descriptions of the statistical concepts, methodologies and information on data quality)**
	Documentation (incl. structure and semantics)	Complete and web available
	Data linking	**Recommended, but not mandatory**
	Shared vocabularies/taxonomies	**Controlled vocabularies and taxonomies DCAT-compatible**

- The impact on the data holders' organisations is assumed on the **medium to high**, based on their current development
- Main cost drivers in terms of development and maintenance:
    - Shared vocabulary / taxonomies – important for integration with existing statistical data portals and expanding the use and re-use of the datasets (medium)
    - Data linking – facilitate to connect information from different sources, increasing the level of discoverability and re-use of the datasets; is cost intensive (high impact);
    - APIs remain an important cost driver as application with higher complexity might result in an increase impact on the organisations (medium to high)

Mobility

Laura Doumbouya
Deloitte

Mobility
Datasets in scope

The existing regulatory acquis covers a wide range of datasets pertaining to the "mobility" thematic area. Certain datasets covered by EU legislation were out of scope of this study due to ongoing EU-level policy developments and/or due to the existence of restricted access regimes.

Out of scope:

- Due to ongoing policy developments:
  - *Datasets covered by the Intelligent Transport Systems (ITS) Directive including*
    - the Rail Interoperability Directive
    - the Vessel Traffic Monitoring and Information Systems (VTMIS) Directive
- Due to restricted access regimes:
  - *Datasets covered by the Regulation on air traffic management (ATM) and air navigation services (ANS)*

In scope:

- Transport networks datasets, including those identified under the INSPIRE data theme:
  - *Common transport networks elements*
  - *Road, rail, water, air and cableways transport networks*
- Inland waterway and river infrastructure data identified under the River Information Services (RIS) Directive and its implementing legislation
  - *Static, urgent/dynamic infrastructure data and electronic navigational charts*

Mobility
Expected costs & benefits

Expected costs

INLAND WATERWAYS

- Current modalities for the provision of inland waterway infrastructure data across the EU are fairly aligned with the *to be* recommendations
- Full harmonisation will be undertaken in the context of the CEF funded VisuRIS COMEX

TRANSPORT NETWORKS

- Currently, these datasets are generally available free of charge, in a harmonised format and provided through large scale IT infrastructure.
- The establishment of APIs and the updates of the datasets will be the main cost drivers.

Expected benefits

Complementing existing efforts in promoting the reuse of such data through the obligation of providing these for free, in harmonised and machine readable formats and both through bulk download and APIs.

- Complementing the RIS Directive which does not define a harmonised format, availability of API/bulk-download, timeliness of the data, etc.

- Complementing the INSPIRE Directive by adding an 'Open Data' layer to its provisions, including the availability of APIs.

Mobility

Lower intensity intervention – Datasets in scope (1/2)

INSPIRE features under the Transport Networks data theme, including:

Common	Road	Rail	Water		Air		Cableways
• Access Restriction • Condition Of Facility • Maintenance Authority • Marker Post • Owner Authority • Restriction For Vehicles • Traffic Flow Direction • Transport Area • Transport Link • Transport Link Sequence • Transport Link Set • Transport Network • Transport Node • Transport Object • Transport Point • Transport Property • Vertical Position	• E-Road • Form Of Way • Functional Road Class • Number Of Lanes • Road • Road Area • Road Link • Road Link Sequence • Road Name • Road Node • Road Service Area • Road Service Type • Road Surface Category • Road Width • Roundabout • Speed Limit • Vehicle Traffic Area	• Design Speed • Nominal Track Gauge • Number Of Tracks • Railway Area • Railway Electrification • Railway Line • Railway Link • Railway Link Sequence • Railway Node • Railway Station Area • Railway Station Code • Railway Station Node • Railway Type • Railway Use • Railway Yard Area • Railway Yard Node	• Beacon • Buoy • CEMT Class • Condition Of Water Facility • Fairway Area • Ferry Crossing • Ferry Use • Inland Waterway • Marine Waterway • Port Area • Port Node • Restriction For Water Vehicles • Traffic Separation Scheme • Traffic Separation Scheme Area	• Traffic Separation Scheme Crossing • Traffic Separation Scheme Lane • Traffic Separation Scheme • Traffic Separation Scheme Separator • Water Link Sequence • Water Node • Water Traffic Flow Direction • Waterway • Waterway Link • Waterway Node	• Aerodrome Area • Aerodrome Category • Aerodrome Node • Aerodrome Type • Air Link • Air Link Sequence • Air Node • Air Route • Air Route Link • Airspace Area • Apron Area • Condition Of Air Facility • Designated Point • Element Length • Element Width	• Field Elevation • Instrument Approach Procedure • Lower Altitude Limit • Navaid • Procedure Link • Runway Area • Runway Centreline Point • Standard Instrument Arrival • Standard Instrument Departure • Surface Composition • Taxiway Area • Touch Down Lift Off Area • Upper Altitude Limit • Use Restriction	• Cableway Link • Cableway Link Sequence • Cableway Link Set • Cableway Node

Mobility

Lower intensity intervention – Datasets in scope (2/2)

Inland waterway and river infrastructure-related data under the RIS Directive and its implementing legislation, including:

Static	Dynamic / Urgent	Inland electronic navigational charts (ECDIS)
• Fairway characteristics *(02 ILW/01-21 & 33)* • Long-time obstructions in the fairway and reliability *(02 ILW/22-27)* • Rates of waterway infrastructure charges *(02 ILW/34)* • Other physical limitations on waterways *(03 LO/1; 04 LC/1-5 & 8-9; 05 BR/01-06)* • Regular lock and bridge operating times *(03 LO/2 &4-9; 04 LC/6-7; 05 BR/6)* • Location and characteristics of ports and transhipment sites *(PO/3-65; 07 PT/1-2)* • List of navigation aids and traffic signs • Navigation rules and recommendations	• Water depths contours in the navigation channel *(02 ILW/14)* • Temporary obstructions in the fairway *(02 ILW/22-25)* • Present and future water levels at gauges *(02 ILW/15-16)* • State of the rivers, canals, locks and bridges *(02 ILW; 03 LO; 04 LC; 05 BR)* • Restrictions caused by flood and ice *(02 ILW/24-25; 03 LO/6-7)* • Meteorological data (incl. wind direction)Short term changes of lock and bridge operating times *(03 LO/2; 05 BR/6)* • Short term changes of aids to navigation	• Waterway axis with kilometres indication *(02 ILW/01-07)* • Links to the external xml-files with operation times of restricting structures *(03 LO/2; 05 BR/6)* • Location of ports and transhipment sites *(06 PO/58-59; 07 PT/1-2)* • Reference data for water level gauges relevant to navigation *(02 ILW/15-16)* • Bank of waterway at mean water level • Shoreline construction *(02 ILW/9-12; 04 LC/1-4; 05 BR/3-6)* • Contours of locks and dams *(02 ILW/15-16)* • Boundaries of the fairway/navigation channel *(02 ILW)* • Isolated dangers in the fairway/navigation channel under and above water *(02 ILW/22-25)* • Official aids-to-navigation (e.g. buoys, beacons, lights, notice marks)

Mobility

Lower intensity intervention – Recommended modalities *(INSPIRE transport network data)*

Dimensions		Road transport	Rail transport	Water transport	Air transport	Cableways
Openness-data specification	License (terms of use)	CC-BY 4.0 No terms of use				
	Format	GML, GeoPackage, GeoJSON				
	Machine-readability	Mandatory				
	Availability of API, bulk download	Webservice, (OGC) API and bulk download				
Documentation	Metadata (dataset content description)	Complete (*.csv document available)				
	Data linking	Links to national INSPIRE Geoportals and datasets				
	Documentation (incl. structure and semantics)	Complete and available				
	Shared vocabularies	INSPIRE data specifications are recommended but not mandatory.				
	Taxonomies	N/A				
Completeness	Traceability	N/A				
	Update frequency and timeliness	When necessary				
	Granularity	From local to national, including links with cross-border networks, if and where available				
	Key attributes	Any national identification code; latitude and longitude				

Mobility

Lower intensity intervention – Recommended modalities *(inland waterway infrastructure data)*

Dimensions		Static	Dynamic/Urgent	Electronic Navigational Charts
Openness-data specification	License (terms of use)	CC-BY 4.0 No terms of use		
	Format	CSV, XML, (geo)JSON	CSV, XML, (geo)JSON	XML, .000, WMS
	Machine-readability	Mandatory		
	Availability of API, bulk download	Bulk download mandatory. Webservices and API recommended.	Web service, (OGC) API and bulk download mandatory.	
Documentation	Metadata (dataset content description)	Complete (*.csv document available)		
	Data linking	N/A		
	Documentation (incl. structure and semantics)	RIS Index Encoding Guide 3.0 developed by the Joint Task Force on the RIS Index	Commission Implementing Regulation (EU) 2018/2032 of 20 November 2018 for NtS	Commission Implementing Regulation (EU) 2018/1973 of 7 December 2018 for Inland ECDIS
	Shared vocabularies	RIS Index Encoding Guide 3.0	NtS Standard 4.0	Inland ECDIS Standard 2.5
	Taxonomies	N/A		
Completeness	Traceability	N/A		
	Update frequency and timeliness	When necessary	Daily to (near) real time	Monthly (for shallow sections)
	Granularity	Individual waterway km level		National waterway network and cross-border nodes
	Key attributes	ISRS Location Code		

Mobility

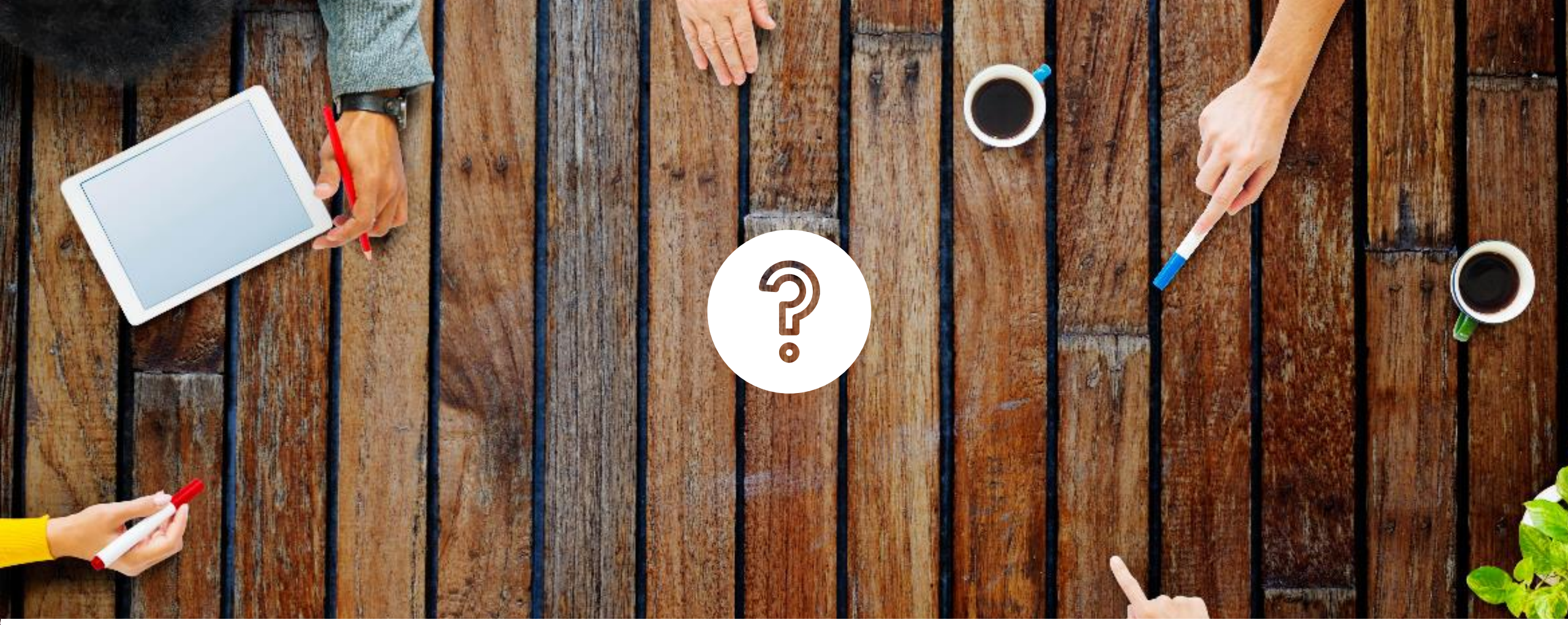
Higher intensity intervention – Datasets in scope

Inland waterways and river infrastructure data and INSPIRE Transport Networks (cf. lower intensity option)

+

***All* other national transport network datasets held by authorities beyond INSPIRE datasets for all modes of transport.**

		Road transport	Rail transport	Water transport	Air transport	Cableways
Openness-data specification	License (terms of use)	CC-BY 4.0 No terms of use				
	Format	Formats as recommended in PSI HVD measures for transport networks under INSPIRE (see lower intensity option) but not mandatory.				
	Machine-readability	Mandatory				
	Availability of API, bulk download	Webservice, (OGC) API and bulk download				
Documentation	Metadata (dataset content description)	Metadata as recommended in PSI HVD measures for transport networks under INSPIRE (see lower intensity option) but not mandatory.				
	Data linking	Links to national INSPIRE Geoportals and datasets as recommended in PSI HVD measures for transport networks under INSPIRE (see lower intensity option) but not mandatory.				
	Documentation (incl. structure and semantics)	Documentation as recommended in PSI HVD measures for transport networks under INSPIRE (see lower intensity option) but not mandatory.				
	Shared vocabularies	Shared vocabularies as recommended in PSI HVD measures for transport networks under INSPIRE (see lower intensity option) but not mandatory.				
	Taxonomies	N/A				
Completeness	Traceability	N/A				
	Update frequency and timeliness	When necessary				
	Granularity	Granularity as recommended in PSI HVD measures for transport networks under INSPIRE (see lower intensity option) but not mandatory.				
	Key attributes	Key attributes as recommended in PSI HVD measures for transport networks under INSPIRE (see lower intensity option) but not mandatory.				



3. Opinions, questions and answers – Part I

Derilinx

Pierre Baviera

CEO

Perspectives on themes, impact and innovation from a service delivery perspective

Derilinx

1

Health

Mobility

Climate and Sustainability

2

Significant move towards Impact and Return on Investment

- Linkages
- Access
- Exploration and Visualisation
- Harmonisation / Standardisation
- Quality

3

Open Innovation and collaboration

- Across Public Service Bodies
- With Private Sector

DATA.GOV.IE

SMART DUBLIN

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

iati International Aid Transparency Initiative

EAST•WEST MANAGEMENT INSTITUTE

Derilinx

Thank You!

pierre@derilinx.com
+353 (0)86 086 14562
www.derilinx.com
@derilinx

Estonia – Ministry of Economic Affairs and Communications

Ott Velsberg

Government Chief Data Officer

CEEP - European Centre of Employers and Enterprises

Katinka Kraus

Policy officer – Public Services

Your voice.
Your interests.
Your future.

Katinka Kraus

Policy Officer

European Centre of Employers and Entreprises
providing Public Services and SGIs

Data Intelligence Offensive - Austria

Peter A. Bruck

Secretary General

"PSI High Value Datasets webinar" 4. Sept 2020
DIO – Austria

DIO is based on the following insights:

- Data sharing and exchange are key in order to generate value and strengthen EU
- Functioning data markets are the essential prerequisite for data to be used as a strategic resource for innovation and value creation.
- The value added by data analytics, artificial intelligence, and machine learning is significantly increased by efficient data trading.
- High Value Data sets collected by one organisation/institution/enterprise need combined with other data and traded.
- The use, value creation and increase in competitiveness depend largely on the scope and quality of the available data.

Success factors for HVD sharing and trading are:

- **Trust and ethics**
- **Security and data sovereignty**
- **Data quality and interoperability**
- **Technologies and innovation**
- **A strong future through community**

DIO view on key issues which need to be addressed

- Lack of **trusted** and **secure platforms** for sharing personal and industrial data
- Lack of **privacy-aware analytics** methods for secure sharing of personal data and industrial data
- Lack of **ICT** and **Data skills** seriously limits the capacity of Europe to respond to the **digitization** challenge
- Insufficient involvement of **SMEs** and lack of access to **data** and **technology**
- **IT standardization** faces new challenges as technologies converge and **federated** systems arise, creating new gaps in **interoperability**
- Advance the **state of the art** w.r.t. **scalability**, **computational** efficiency of methods
- Analysis and **address privacy/confidentiality threat models**
- **incentive models for the sharing of data assets, incl. high value data sets**

PSI Alliance

Georg Hittmair

Legal and Public Affairs

PSI Alliance

Definition of High Value datasets and resulting GDPR concerns

- Current process: Implementing act to further specify Annex I of the directive (EU) 2019/1024 - **legislative procedure**
- **Opening clause** in GDPR Article 86, Recital 154 **is applicable**
- Personal data in documents…. should be able to be publicly disclosed…if the disclosure is provided for by Union or Member State law.
- **Such laws** should reconcile public access to official documents and the reuse of public sector information with the right to the protection of personal data and **may therefore provide for the necessary reconciliation** with the right to the protection of personal data pursuant to this Regulation.

PSI Alliance

High Value datasets and „normal PSI datasets“

- Current process according to Article 14: Commission lays down a list of specific high-value datasets….. among the documents to which this Directive applies
- **= Directive applies to all the datasets mentioned in the workshops**
- Member states are obliged to release all those documents, not just the HVDs
- Even API, charging and machine readable format are requested in softer provisions
- Definition of HVDs dos not mean release yes/no, but which way to release

Contact:

Dr. Georg Hittmair
Compass-Verlag GmbH
Matznergasse 17, 1140 Wien
E-Mail: georg.hittmair@compass.at
Tel.: 0043 1 98116 103

FEBIS

Stephanie Verilhac Marzin

Institutional and Regulatory Affairs Coordinator

Comments on High Value Datasets

Scope

- Include **all businesses and business information** (e*g sole traders, company directors, managers, Ultimate Beneficial Owners..)* and not only registered companies
- Include **all information** that is made available through any type of business register
- Use as **a basis for common data layer facilitating the roll out of sectoral data spaces,** so it should be set **on maximum data availability** and not minimum
- **Open data** and no restrictive license nor terms of use or unjustified database rights

Comments on High Value Datasets

Benefits

- Use of business information in assessing and managing **TRADE CREDIT** benefits the whole economy, especially in economic turmoil times
- It is essential to have **updated information** in getting **creditworthiness assessments** which help prevent business failures
- Increase **CORPORATE TRANSPARENCY** and **AML** or **FRAUD detection** by leveraging data
- Ensuring the **legitimate right to information** in business context prevails

Lunch break

4. Presentation of the study results – Part II

Macro-economic impact assessment

Eike-Christian Koring
Deloitte

Macroeconomic impact assessment

Assumptions made by relevant literature

- Our impact modelling assumptions are based on relevant literature and the findings of our study research so far. However, several studies such as the latest study by the European Commission (2020) on the economic impact of open data point out that the exact reuse of open data and in particular PSI Data is barely documented. In addition, reuse can be endless and network effects can even multiply the economic value.
- The market size of PSI data is therefore defined as the market size of products, services, and content improved or enabled by PSI Data. However, which share of this value is attributed to open data can only be estimated.
- In order to determine the market value of PSI, the Vickery 2011 study provided a main base for the prediction of the direct impact.

Study	Context	Assumptions about data economy/ open data
Vickery 2011, Review of Recent Studies on PSI Re-Use and Relate Market Developments	Part of the impact assessment of the first EU PSI Directive. Looks at the impact of the adoption of the PSI Directive in 2006, extrapolates sectoral + national impact research for the entire EU, based on GDP and ICT	Predicts a direct economic benefit of open data of 40 bio. EUR in the EU Predicts an indirect economic benefit of 140 bio. EUR in the EU. Predicts socio-economic benefits of 40 bio. EUR in the EU. Predicts EU market for government data in 2010 at about 32 bio. EUR.
McKinsey 2013, Open Data	Large scale report about the economic potential of data and open data with macro-economic estimates.	Estimates the total economic potential in the EU at 900 bio. Dollar, applying a bottom-up approach.
EU 2015, Creating Value through open data	Macro-economic research into economic potential in the EU for 2016-20, building on Vickery 2011.	Predicts a direct economic value of open data of 75.7 bio. EUR in 2020. Predicts public sector savings in the EU of 1.7 bio. In 2020. Predicts an indirect value of open data in the EU from 265-286 bio. EUR in 2020.
EU 2017, Open Data Maturity in Europe	Yearly EU Data Portal Study	Predicted 325 bio. EUR of potential contribution of open data for 2016-2020 with 30.000 new jobs created in 2020.
Deloitte 2018, Reuse of Public Sector Information	Evaluation assessed the performance of the PSI Directive, whether it still responds to the stakeholders' needs and expectations and whether it fits the purpose of the next years.	Study based on the Vickery Study 2011 and assumptions of Eurostat and the Data Monitoring Tool. Assessment identified problem areas of re-use of PSI Data, among them costs of data re-use, availability, exclusive agreements

Macroeconomic impact assessment

Our method: Top-down approach

- The Economic Impact of HVDs (it's contribution to GDP) is estimated based on a **top-down approach**.
- The baseline has been calculated based on the extrapolation of the Vickery 2011 study and the forecasts of the European Data Market Monitoring Tool. The monitoring tool provides three forecast scenarios at the 2025 Horizon: a baseline scenario, a challenge scenario and a high growth scenario. The baseline scenario was chosen as the relevant scenario.
- In the year 2020, the outbreak of Covid-19 massively affected the European economy. Expected figures have been corrected to take into account the impact of this crisis.

First step

Estimate the value of the relevant economy (to be addressed) for the baseline. The **share in GDP/GVA** will be used as **main measurement indicator** in this regard.

Second step

Secondary sources and studies including e.g. the **EU Data Market Study** and **Vickery 2011** provide a baseline for the Economic Value of the Data Economy and relates it to the GDP.

We will use this data as a baseline.

Third step

The **economic impact** of the different **policy options/packages (on GDP)** will be measured for each indicator on an incremental basis (vs. the baseline).

Macroeconomic impact assessment

Our method: Top-down approach

Determination of baseline

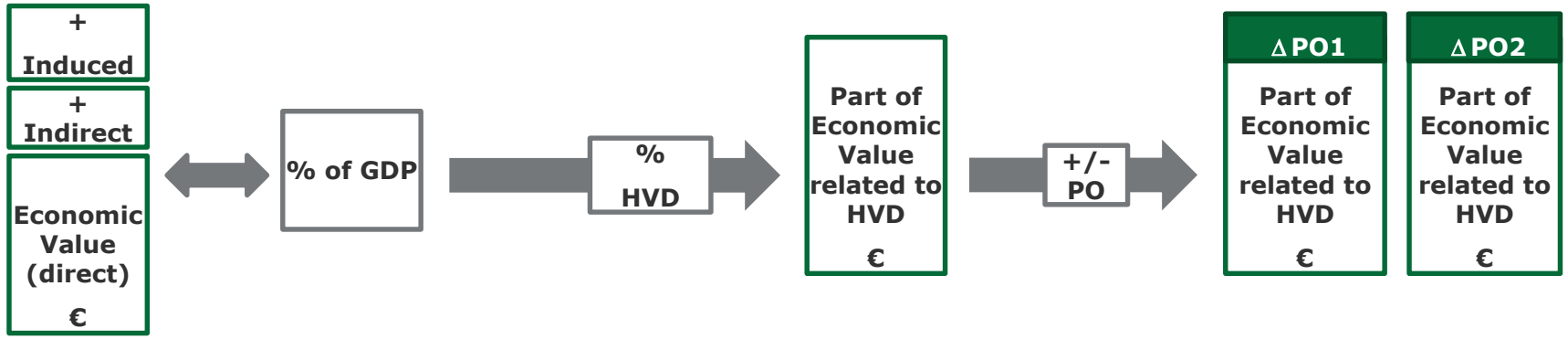
- Based on results of the studies e.g. EU Data Market Study / Vickery.
- Taking into account results of secondary studies and findings in D2/D3/D4.
- Extending the forecast to 2028 (2023 + 5 years).

Share attributable to HVD

- Estimating the share attributable to HVD.
- Based on secondary sources and preliminary results the value attributable to HVD is estimated.

Impact of Policy Options

- Based on the preliminary findings and results in the **six HVD domains**, the economic impact for the policy options is estimated.
- The result is presented as incremental impact on the Economic Value / share in GDP.



Macroeconomic impact assessment– Baseline Scenario

Market size of PSI related to the Data Market Value of the EU Data Economy

EU Data market value and growth rate [%] according to the EU Data Monitoring Tool 2020 (until 2025), OECD estimates afterwards – **baseline scenario**

EU Data market value and growth rate [%] according to the EU Data Monitoring Tool 2020 (until 2025), OECD estimates afterwards – **high growth scenario**

HVD | Baseline and Economic Impact

M€	2020	2021	2022	2023	2024	2025	2026	2027	2028
EU Data Monitoring Tool 2020 - baseline									
Data revenues	71 050	75 866	81 008	86 499	92 362	98 623	100 144	101 711	103 321
Data market value	62 244	65 795	69 584	73 628	77 948	82 564	83 837	85 149	86 497
Data market value - direct impact	54 081	58 481	63 239	68 385	73 948	79 965	81 198	82 469	83 775
grow th rate %	*(7.1%)*	*8.1%*	*8.1%*	*8.1%*	*8.1%*	*8.1%*	*1.5%*	*1.6%*	*1.6%*
EU Data Monitoring Tool 2020 - high growth									
Data revenues	71 050	80 943	92 215	105 055	119 684	136 350	138 453	140 620	142 846
Data market value	62 244	69 320	77 236	86 097	96 020	107 139	108 791	110 494	112 243
Data market value - direct impact	54 081	62 005	71 090	81 505	93 447	107 139	108 791	110 494	112 243
grow th rate %	*(7.1%)*	*14.7%*	*14.7%*	*14.7%*	*14.7%*	*14.7%*	*1.5%*	*1.6%*	*1.6%*
% above baseline	-	*6.0%*	*12.4%*	*19.2%*	*26.4%*	*34.0%*	*34.0%*	*34.0%*	*34.0%*
Market size PSI - Vickery	48 649	51 811	55 179	58 765	62 585	66 653	67 681	68 740	69 828
Assumption: baseline grow th rate: 7%	*(7.1%)*	*6.5%*	*6.5%*	*6.5%*	*6.5%*	*6.5%*	*1.5%*	*1.6%*	*1.6%*
% PSI on data market value	90%	89%	87%	86%	85%	83%	83%	83%	83%

forecast based on EU Data Monitoring Tool → *OECD GDP forecast* →

Baseline scenario PSI growth: Estimated market size of the PSI sector based on Vickery 2011. Growth estimations until 2020 according to predictions of the EU Data Monitoring Tool.

Assumptions from 2021 onwards: **baseline growth rate of 6,5% in the PSI sector** according to relevant literature and findings in previous deliverables. From 2026, the PSI market size is estimated to follow the growth path in line with the OECD estimates.

Macroeconomic impact assessment – Baseline Scenario

Market shares (% of PSI market) and direct impact of the six thematic areas

Based on existing literature and on the study research so far, we estimated the **market share of the six thematic areas in % of the PSI Market for the baseline scenario**.

Consequently, we could estimate the **direct impact in Mio. EUR of each thematic sector for the years 2023-2028**.

HVD | Baseline and Economic Impact

M€	2020	2021	2022	2023	2024	2025	2026	2027	2028
Market size PSI - Vickery	48 649	51 811	55 179	58 765	62 585	66 653	67 681	68 740	69 828
Assumption: baseline growth rate: 7%	*(7.1%)*	*6.5%*	*6.5%*	*6.5%*	*6.5%*	*6.5%*	*1.5%*	*1.6%*	*1.6%*
% PSI on data market value	90%	89%	87%	86%	85%	83%	83%	83%	83%
	forecast based on EU Data Monitoring Tool →						*OECD GDP forecast* →		
Market share [% of PSI Market]									
1. Company and company ownership	6%	6%	6%	6%	6%	6%	6%	6%	6%
2. Geospatial	34%	34%	34%	34%	34%	34%	34%	34%	34%
3. Meteorological data	20%	20%	20%	20%	20%	20%	20%	20%	20%
4. Earth observation and environment	15%	15%	15%	15%	15%	15%	15%	15%	15%
5. Statistics	16%	16%	16%	16%	16%	16%	16%	16%	16%
6. Mobility	9%	9%	9%	9%	9%	9%	9%	9%	9%
Total	***100.0%***	***100.0%***	***100.0%***	***100.0%***	***100.0%***	***100.0%***	***100.0%***	***100.0%***	***100.0%***
Market share									
1. Company and company ownership	3 016	3 212	3 421	3 643	3 880	4 132	4 196	4 262	4 329
2. Geospatial	16 541	17 616	18 761	19 980	21 279	22 662	23 011	23 372	23 742
3. Meteorological data	9 730	10 362	11 036	11 753	12 517	13 331	13 536	13 748	13 966
4. Earth observation and environment	7 200	7 668	8 166	8 697	9 263	9 865	10 017	10 174	10 335
5. Statistics	7 784	8 290	8 829	9 402	10 014	10 664	10 829	10 998	11 173
6. Mobility	4 378	4 663	4 966	5 289	5 633	5 999	6 091	6 187	6 285
Total	**48 649**	**51 811**	**55 179**	**58 765**	**62 585**	**66 653**	**67 681**	**68 740**	**69 828**

Economic impact assessment – Policy interventions

Growth rates for lower and higher intervention option (1/2)

Assumptions on growth rates for lower and higher intervention option for the thematic areas company for 2023-2028.

- In **company and company ownership**, the lower intensity intervention option includes only two datasets in scope which would be, as described in the specific section above, the low-hanging fruits as the implementation would require little effort from member states. Hence, only a slight change of the growth rate in comparison to the baseline is expected and therefore a growth rate of 7.0% for the lower intensity intervention option was assigned. The higher intensity intervention option would include a wide scope of datasets and consequently a medium to major change of the growth rate in comparison to the baseline. Considering that the company and company ownership high value data set market is relatively small, a growth rate of 8.5% was assumed.

- For the lower intervention option for **geospatial** high value data, only limited requirements of data will be necessary which requires only minimum changes in the current options. 4-5 high value datasets will be in scope with this option. It can be concluded that this intervention option would result in a slight to medium change in the growth rate of the geospatial sector, also because of the importance of geospatial data and its expected impact to affected stakeholders. Taking this into account, a growth rate of 7.5 % for the lower intervention option of geospatial high value data was assumed. The higher intervention option is the far-reaching option, whereas the main differences to the lower intensity option refer to the licence, the APIs and few changes on the granularity and key attributes. Considering this medium change in comparison to the baseline, a growth rate of 8.5% was attributed to the higher intensity intervention option.

- The **meteorological** lower intensity intervention option encompasses three datasets, namely observations, climate and digitised structured historical climate data. This intervention stays close to what is already common, but leaves room for Member States to adopt more advanced technology. Hence, this intervention option would go along with a medium change of the growth rate in comparison to the baseline. A growth rate of 8.0% was therefore attributed to the lower intensity intervention option of meteorological high value data. In the meteorological higher intensity option, three additions are made, unstructured historical data, radar data and numerical weather prediction model data. As all data sets are highly voluminous, this intervention option can be considered as a major change in the growth rate in comparison to the baseline. Therefore, a growth rate of 9.5% was attributed to the higher intervention option of meteorological datasets.

Economic impact assessment – Policy interventions

Growth rates for lower and higher intervention option (2/2)

Assumptions on growth rates for lower and higher intervention option for the thematic areas company for 2023-2028.

- The lower intensity intervention option of **earth environment and observation** includes an open data obligation to the environmental reporting and observation data which means mostly removing reuse restrictions. This constitutes an expected medium change of the growth rate in comparison to the baseline, for which a growth rate of 8.0% was assigned. The higher intensity intervention option involves that the scope of data involved is broadened by adding additional INSPIRE themes relevant to earth and observation. This intervention option implies a medium to major change in comparison to the baseline. A growth rate for the higher intensity intervention option of 9.0% was therefore assigned.

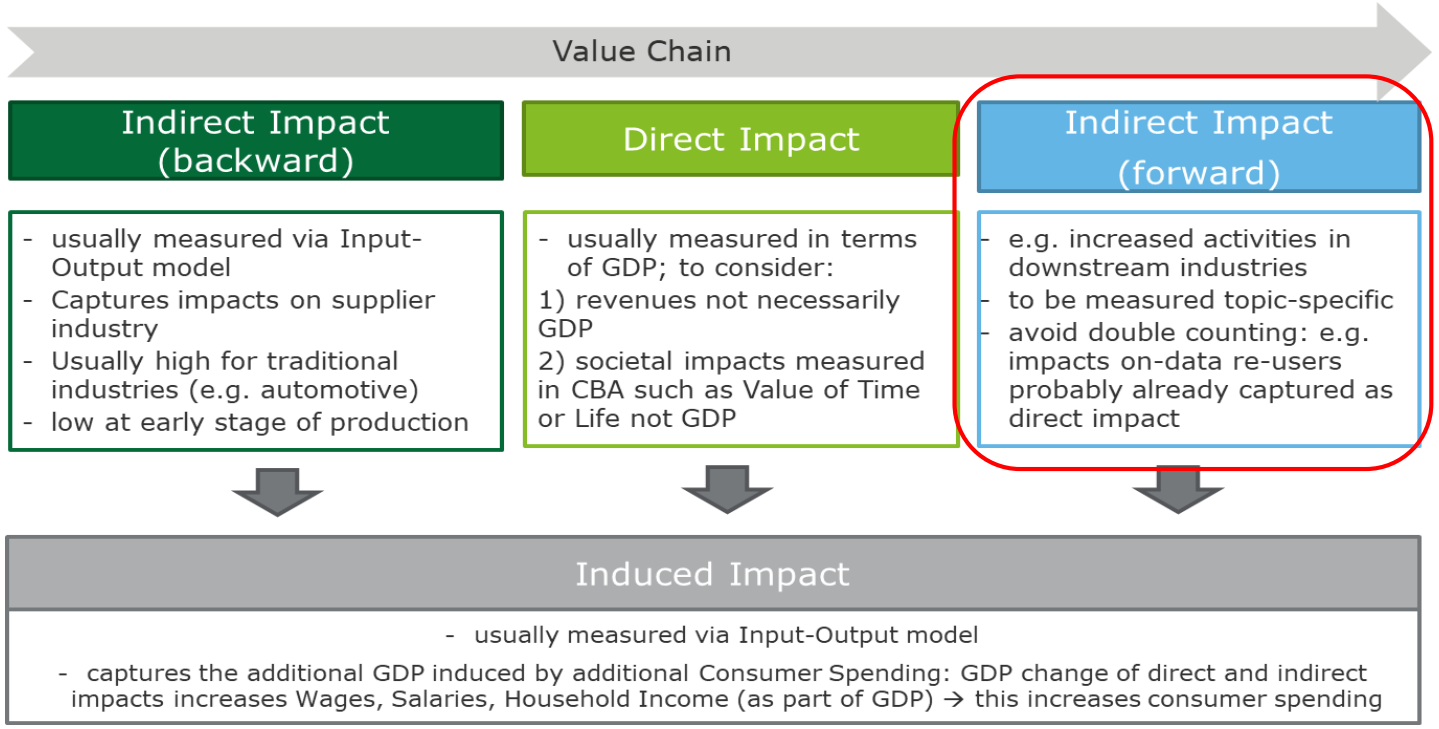
- For **statistics**, the lower intensity intervention option requires only a set of minimum changes to the current publication options available which is expected to have low to no impact on the institutions and stakeholders affected. Hence, no to slight changes of the growth rate of statistics in comparison to the baseline is expected for the lower intensity intervention option. Therefore, a growth rate of 7.0% for the lower intensity intervention option for statistics was assumed. The higher intensity intervention option implies new changes in terms of measures for publication. Improved APIs and metadata files, development of controlled vocabularies and taxonomies are expected which therefore constitutes a medium change in comparison to the baseline, also because the market of high value data in statistics is already quite advanced relative to the other thematic areas. A growth rate of 8.0% for the higher intensity intervention option in statistics is consequently assumed.

- For **mobility**, the lower intervention option covers, as stated in chapter 3.1, only the bare minimum of the datasets and 2 out of 4 categories. This signifies a slight to medium change of the growth rate in comparison to the baseline. A growth rate for this slight to medium change of 7.5% was therefore assigned for the lower intervention option of mobility. The higher intervention option, in contrast, is more ambitious than the previous one as all four categories of datasets are included in the high value data list. Hence, this can be interpreted as a medium change in the growth rate in comparison to the baseline and therefore a growth rate of 8.0% for the higher intervention option in mobility was assigned.

Macroeconomic impact assessment

General methodology and indicators: Multipliers to measure indirect and induced impacts

- In order to fully reflect on the overall impact on the economy, indirect/induced impacts can been calculated based on the estimates of the direct impacts.
- The analysis focuses on the indirect (forward) impact on downstream industries, which is considered to be the major indirect impact.
- To measure these impacts with regard to the PSI/HVD economic activity, results of the EU Data Monitoring have been analyzed. As a result, for the indirect (forward) impact a magnitude between ca. 2.5 and 3.0 has been identified.

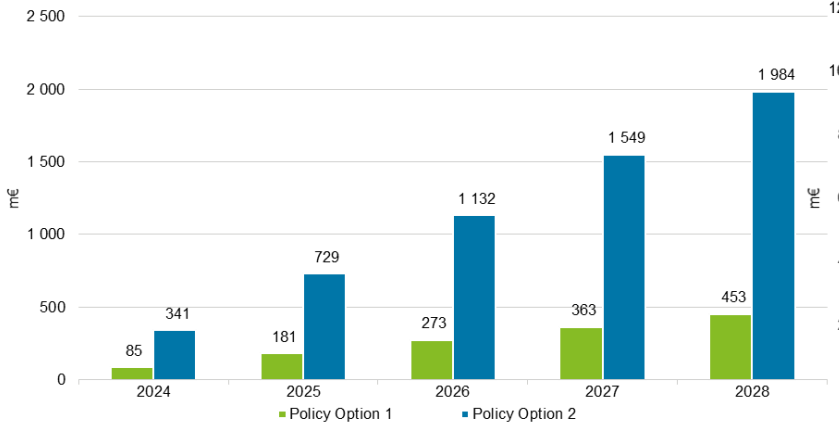


Macroeconomic impact assessment

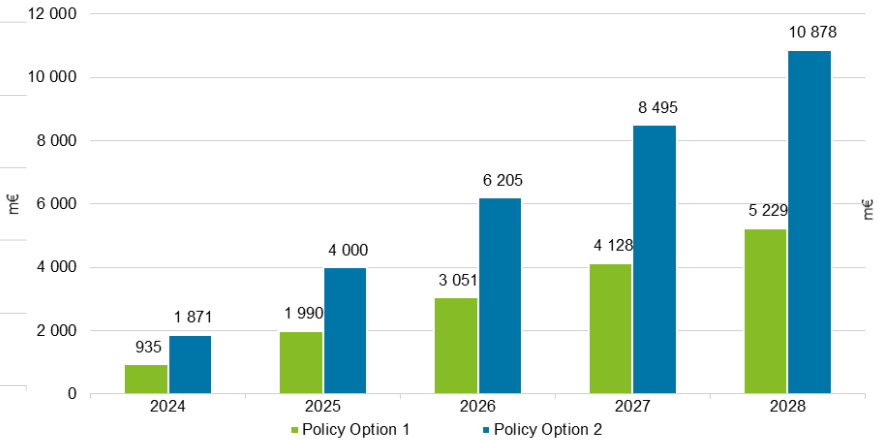
Preliminary results – total impact by policy option

Based on our top down-approach, the calculations yield the following **results for the total impact by thematic area and policy option for the years 2024-2028**:

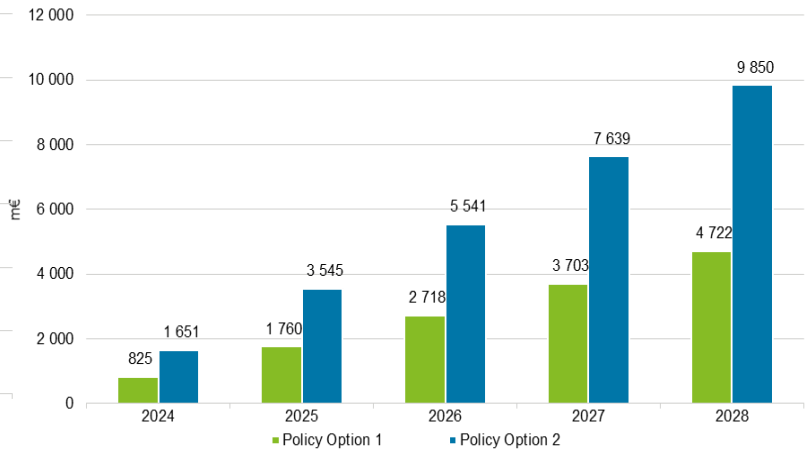
Economic Impact total - 1. Company and company ownership



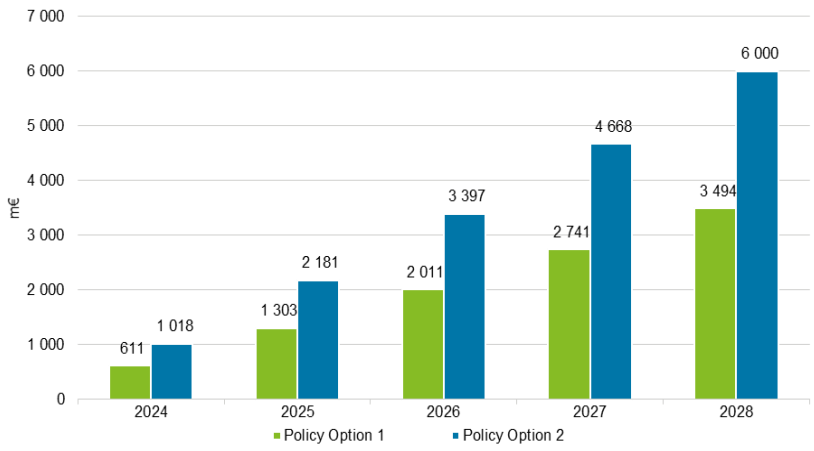
Economic Impact total - 2. Geospatial



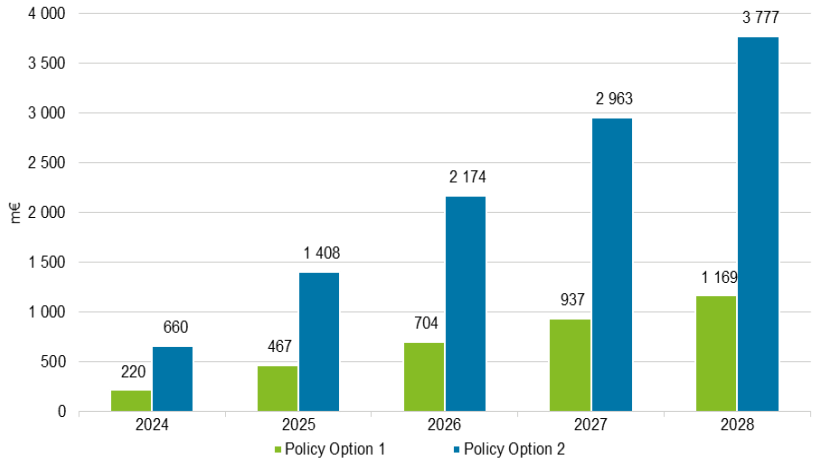
Economic Impact total - 3. Meteorological data



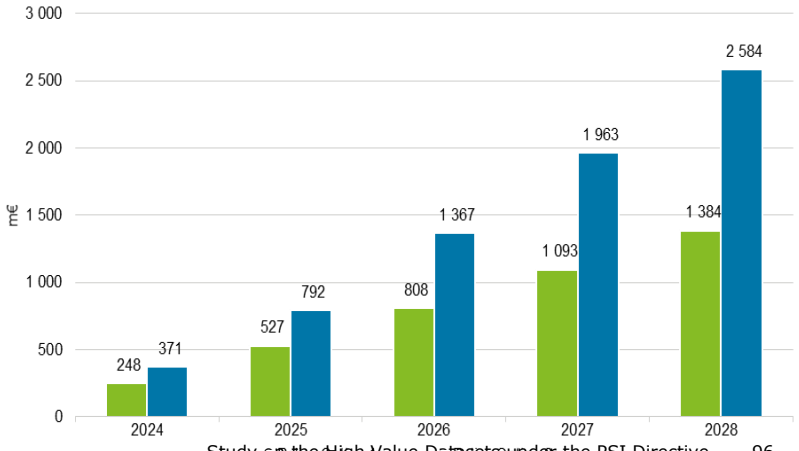
Economic Impact total - 4. Earth observation and environment



Economic Impact total - 5. Statistics



Economic Impact total - 6. Mobility



Comparison of policy options

Eike-Christian Koring
Deloitte

Multi-Criteria Analysis (MCA)

Ranking the three intervention options with four assessment criteria

We used the MCA for two reasons:

First, it is an alternative to the Cost-Benefit Analyses (CBA) performed in the previous phases.

Second, it is particularly **relevant** at the following stages of our **Impact Modelling**:

- At the stage of **assessing the economic, social and environmental dimensions** of each of the three intervention options, including possible trade-offs between these dimensions
- At the stage of **comparing the policy options**: against the main criteria of Effectiveness, Efficiency and Coherence as well as Proportionality and Feasibility.

Multi-Criteria Analysis (MCA)

Explanation of the criteria

Effectiveness	Effectiveness looks at the extent to which different options would achieve the objectives of the intervention, looking for evidence of why, whether or how these changes are linked to the intervention.
Efficiency	Efficiency is defined at the costs and benefits of a policy option as they accrue to different stakeholders.
Coherence	The coherence of each option with the overarching objectives of EU policies, both internally and with other EU interventions.
Proportionality	EU action should be relevant and necessary and should offer value beyond what Member State action alone can deliver and not go further than is necessary to resolve the problem or meet the policy objective.
Feasibility	The options should respect the principle of conferral (legal feasibility). They should also respect any obligation arising from the EU Treaties (and relevant international agreements) and ensure respect of fundamental rights. Options that would clearly fail to garner the necessary political support for legislative adoption and/or implementation could also be discarded (political feasibility).

Source: Better Regulation Guidelines (European Commission, 2017)

Multi-Criteria Analysis (MCA)

Our non-compensatory approach

In our **input matrix** we used five criteria with different weighing which aims to reflect the reality of the importance of the different criteria.

Assessment criterion	Weight	Direction	Performance value
Effectiveness	0.30	1	Qualitative: -3 / +3
Efficiency	0.20	1	Qualitative: -3 / +3
Coherence	0.10	1	Qualitative: -3 / +3
Proportionality	0.10	1	Qualitative: -3 / +3
Feasibility	0.30	1	Qualitative: -3 / +3

Multi-Criteria Analysis (MCA)

Our non-compensatory approach – Step 1: Weighted Performance

The weightings, direction and performance of the three intervention options was then determined separately by PSI-sector for each of the four criteria. The importance of the criteria is reflected in the respective weights.

Input matrix				Policy Options			
				1. Company and company ownership			
Criteria	Weight	Direction	Performance value	P1 (lower intensity intervention)		P2 (higher intensity intervention)	
				Performance	Weighted performance	Performance	Weighted performance
Effectiveness	0,30	1	-3 / 3	1,50	0,45	2,50	0,75
Efficiency (B/C-ratio)	0,20	1	Benefit-/Cost-ratio or Score -3 / + 3	2,50	0,50	1,50	0,30
Coherence of the policy options	0,10	1	-3 / 3	2,00	0,20	2,50	0,25
Compliance with the proportionality principle	0,10	1	-3 / 3	2,00	0,20	1,50	0,15
Legal and political feasability	0,30	1	-3 / 3	2,50	0,75	-1,00	-0,30
				10,50	2,10	7,00	1,15
				2. Geospatial			
Criteria	Weight	Direction	Performance value	P1 (lower intensity intervention)		P2 (higher intensity intervention)	
				Performance	Weighted performance	Performance	Weighted performance
Effectiveness	0,30	1	-3 / 3	1,75	0,53	2,75	0,83
Efficiency (B/C-ratio)	0,20	1	Benefit-/Cost-ratio or Score -3 / + 3	3,00	0,60	1,75	0,35
Coherence of the policy options	0,10	1	-3 / 3	3,00	0,30	1,75	0,18
Compliance with the proportionality principle	0,10	1	-3 / 3	2,50	0,25	3,00	0,30
Legal and political feasability	0,30	1	-3 / 3	3,00	0,90	2,00	0,60
				13,25	2,58	11,25	2,25
				3. Meteorological data			
Criteria	Weight	Direction	Performance value	P1 (lower intensity intervention)		P2 (higher intensity intervention)	
				Performance	Weighted performance	Performance	Weighted performance
Effectiveness	0,30	1	-3 / 3	2,00	0,60	3,00	0,90
Efficiency (B/C-ratio)	0,20	1	Benefit-/Cost-ratio or Score -3 / + 3	3,00	0,60	2,00	0,40
Coherence of the policy options	0,10	1	-3 / 3	2,00	0,20	3,00	0,30
Compliance with the proportionality principle	0,10	1	-3 / 3	3,00	0,30	3,00	0,30
Legal and political feasability	0,30	1	-3 / 3	1,00	0,30	1,00	0,30
				11,00	2,00	12,00	2,20

Multi-Criteria Analysis (MCA)

Our non-compensatory approach – Step 1: Weighted Performance

The weightings, direction and performance of the three intervention options was then determined separately by PSI-sector for each of the four criteria. The importance of the criteria is reflected in the respective weights.

				4. Earth observation and environment			
Criteria	Weight	Direction	Performance value	P1 (lower intensity intervention)		P2 (higher intensity intervention)	
				Performance	Weighted performance	Performance	Weighted performance
Effectiveness	0,30	1	-3 / 3	0,50	0,15	3,00	0,90
Efficiency (B/C-ratio)	0,20	1	Benefit-/Cost-ratio or Score -3 / + 3	3,00	0,60	2,00	0,40
Coherence of the policy options	0,10	1	-3 / 3	2,00	0,20	3,00	0,30
Compliance with the proportionality principle	0,10	1	-3 / 3	2,50	0,25	2,00	0,20
Legal and political feasability	0,30	1	-3 / 3	2,50	0,75	2,50	0,75
				10,50	1,95	12,50	2,55
				5. Statistics			
Criteria	Weight	Direction	Performance value	P1 (lower intensity intervention)		P2 (higher intensity intervention)	
				Performance	Weighted performance	Performance	Weighted performance
Effectiveness	0,30	1	-3 / 3	1,50	0,45	2,50	0,75
Efficiency (B/C-ratio)	0,20	1	Benefit-/Cost-ratio or Score -3 / + 3	1,75	0,35	2,90	0,58
Coherence of the policy options	0,10	1	-3 / 3	2,50	0,25	2,50	0,25
Compliance with the proportionality principle	0,10	1	-3 / 3	1,00	0,10	2,00	0,20
Legal and political feasability	0,30	1	-3 / 3	2,00	0,60	1,75	0,53
				8,75	1,75	11,65	2,31
				6. Mobility			
Criteria	Weight	Direction	Performance value	P1 (lower intensity intervention)		P2 (higher intensity intervention)	
				Performance	Weighted performance	Performance	Weighted performance
Effectiveness	0,30	1	-3 / 3	1,00	0,30	2,50	0,75
Efficiency (B/C-ratio)	0,20	1	Benefit-/Cost-ratio or Score -3 / + 3	1,50	0,30	1,50	0,30
Coherence of the policy options	0,10	1	-3 / 3	3,00	0,30	2,50	0,25
Compliance with the proportionality principle	0,10	1	-3 / 3	3,00	0,30	2,50	0,25
Legal and political feasability	0,30	1	-3 / 3	3,00	0,90	1,50	0,45
				11,50	2,10	10,50	2,00

Multi-Criteria Analysis (MCA)

Our non-compensatory approach – Step 2 and 3: Outranking and Permutation Matrix

We prepared an "**outranking matrix**": Options are compared pairwise. For each comparison, all weights are summed for the criteria where Option A is favoured over Option B as indicated by the weighted performance of each criterion.

The policy options were subsequently ranked in 2! (=2) different ways but the aim is to find the permutation with the maximum likelihood score. Hence, the optimal ranking is one with the maximum likelihood score.

Policy ranking permutation					
Rank #	Policy ranking permutation		Policy pairings	Coefficients of policy pairings	Final score
1. Company and company ownership					
1	P1	P2	P1-P2	0,60	0,60
2	P2	P1	P2-P1	0,40	0,40
2. Geospatial					
1	P1	P2	P1-P2	0,60	0,60
2	P2	P1	P2-P1	0,40	0,40
3. Meteorological data					
2	P1	P2	P1-P2	0,20	0,20
1	P2	P1	P2-P1	0,40	0,40
4. Earth observation and environment					
2	P1	P2	P1-P2	0,30	0,30
1	P2	P1	P2-P1	0,40	0,40
5. Statistics					
2	P1	P2	P1-P2	0,30	0,30
1	P2	P1	P2-P1	0,60	0,60
6. Mobility					
1	P1	P2	P1-P2	0,50	0,50
2	P2	P1	P2-P1	0,30	0,30

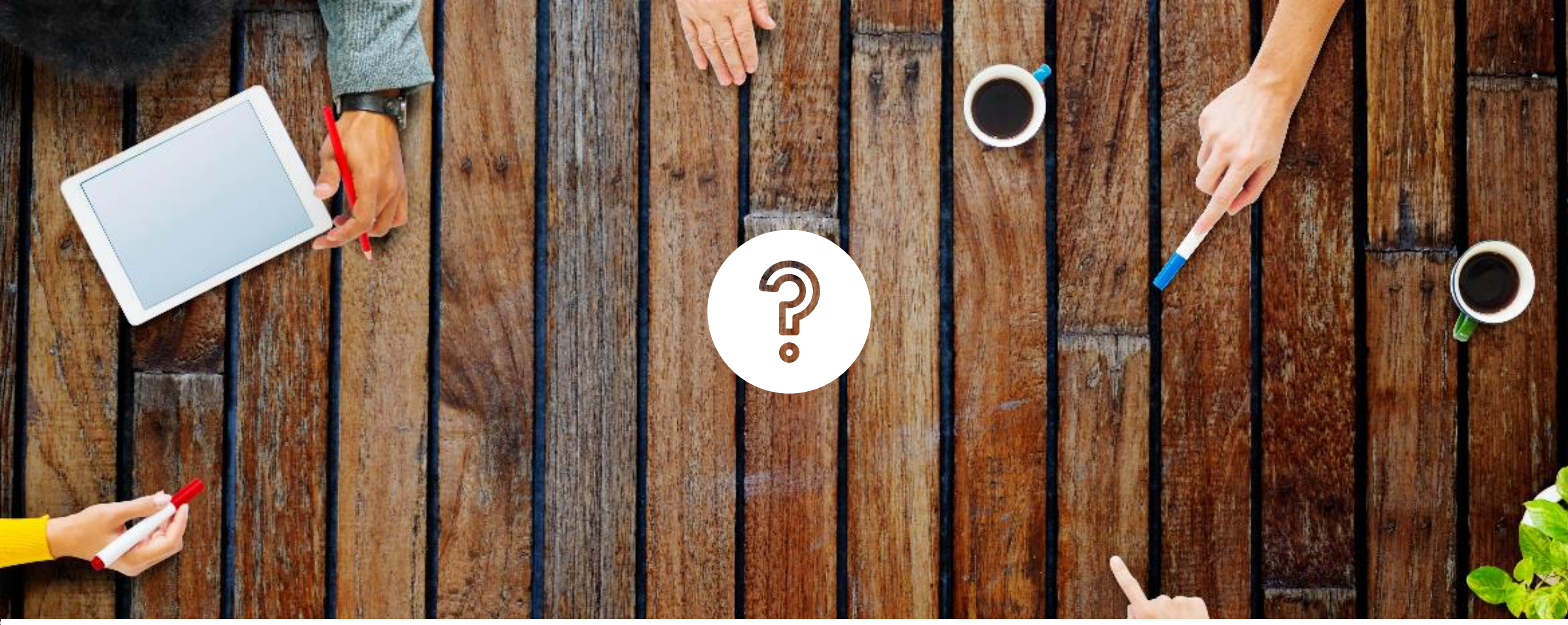
Policy ranking permutation with the final scores (maximum likelihood score) per PSI sector

Multi-Criteria Analysis (MCA)
Results and Conclusions

As a result, **Policy Option 1 (lower intensity intervention)** is identified as preferred option in the thematic areas of **Company & Company Ownership**, **Geospatial data and Mobility**.

Policy Option 2 (higher intensity intervention) is the preferred option in the thematic areas **Meteorological Data, Earth Observation & Environment** and **Statistics**.

	Preferred PO (Equal Weight 0.20)
Company & Company Ownership	PO1
Geospatial Data	PO1
Meteorological Data	PO2
Earth Observation & Environment	PO2
Statistics	PO2
Mobility	PO1



5. Opinions, questions and answers – Part II

Finland - Ministry of Finance

Riitta Autere

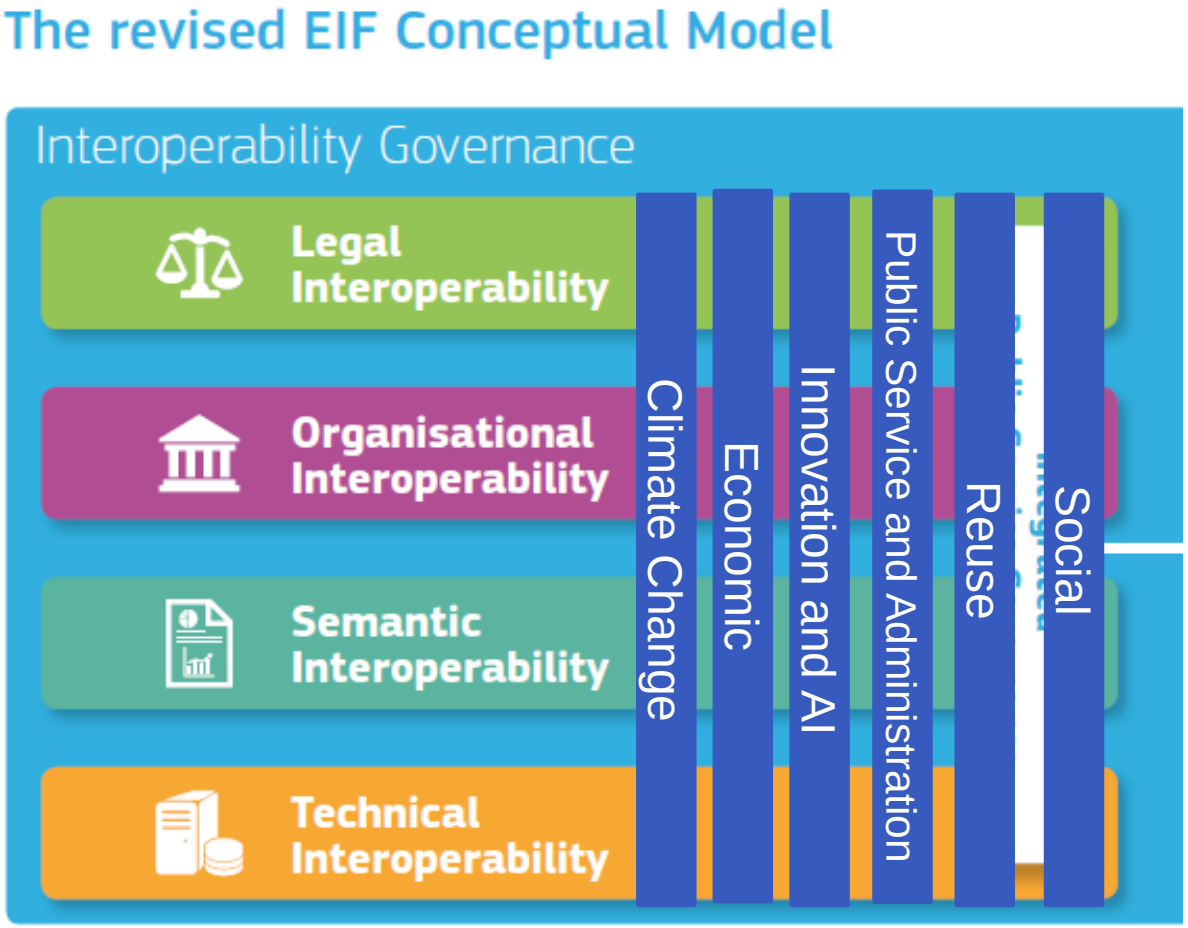
Ministerial Advisor – Public Sector ICT

MINISTRY OF FINANCE

The European Interoperability Framework (EIF) as a common adoptation tool for PSI HVDS

Riitta Autere, Ministerial Adviser & Olli-Pekka Rissanen, Chief Specialist
4.9.2020
PSI High Value Datasets webinar

The EIF Conceptual Model with PSI HVDS value framework's characteristics



- The **European Interoperability Framework (EIF)**, part of the Communication (COM(2017)134) from the European Commission adopted on **23 March 2017**.
- The framework gives specific guidance on how to set up interoperable digital public services.
- https://ec.europa.eu/isa2/eif_en

MINISTRY OF FINANCE

Why the EIF?

Underlying principle 1: subsidiarity and proportionality
Underlying principle 2: openness
Underlying principle 3: transparency
Underlying principle 4: reusability
Underlying principle 5: technological neutrality and data portability
Underlying principle 6: user-centricity
Underlying principle 7: inclusion and accessibility
Underlying principle 8: security and privacy
Underlying principle 9: multilingualism
Underlying principle 10: administrative simplification
Underlying principle 11: preservation of information
Underlying principle 12: assessment of effectiveness and efficiency

Some of the recommenadtions relevant for the PSI HVDS adoptation:

- Rec. 2: *Publish the data you own as open data unless certain restrictions apply.*
- Rec. 4: *Give preference to open specifiations, taking due account of the coverage of functional needs, maturity, and market support and innovation.*
- Rec. 6 & 7: *Reuse and share solutions …* and *Reuse and share information and data…*
- Rec. 12: *Put in place mechanisms to involve users in analysis, design, assessment and further development of European public services.*
- Rec. 16: *Use information systems and technical architecture that care for multilingalism…*
- Rec. 23: *Use a structured, transparent, objective and common approach to assessing and selecting standards and specifications. …*

MINISTRY OF FINANCE

Not see the wood for the trees? Could the EIF help us?

Legal Interoperability

Organisational Interoperability

INSPIRE LEGISLATION

Clean Vehicles Directive

First Company Law Directive 68/151/EEC

?

Semantic Interoperability

Technical Interoperability

Geospatial Low intensity intervention

Dimensions		Administrative units	Place Names	Addresses	Buildings
Openness	License and terms of use	CC-BY 4.0			
	Format	GML; GeoPackage; GeoJSON; INSPIRE recommendations.	GML; GeoPackage; CSV; GeoJSON; INSPIRE recommendations.	GML; GeoPackage; CSV; GeoJSON; INSPIRE recommendations	GML; GeoPackage; GeoJSON; INSPIRE recommendations.
	Machine-readability	Recommended			
	Availability of API, bulk download	Bulk download; INSPIRE distribution services; RestAPI (e.g. OGC API, ArcGIS RestAPI, Carto API).			
Documentation	Metadata (dataset content description)	INSPIRE			
	Documentation (incl. structure and semantics)	INSPIRE; GeoDCAT-AP.			
	Data linking				
	Shared vocabularies/taxonomies	INSPIRE			
Completeness	Traceability	National geodata Catalog or open data catalog.			
	Update frequency and timeliness	Annual update	When necessary	When necessary	When necessary
	Granularity	Boroughs; municipalities; districts; provinces; regions; national borders; sea-frontiers.	National coverage	Partial National coverage (e.g. most populated cities)	Partial National coverage (e.g. most populated cities)
	Key attributes	National identification code; identification code of the upper administrative level; official name; short name abbreviation; coordinate reference system used by the national government.	Name; category; latitude and longitude (WGS84) (INSPIRE).	Latitude and longitude (wgs84); house number; suffix of the number; name of the street; name of the municipality; national identification code of the municipality; last update; type of position.	Footprint of the building.

MINISTRY OF FINANCE

MINISTRY
OF FINANCE

FBNed | FamilieBedrijven Nederland

Albert Jan Thomassen

Executive Director

Open Ownership

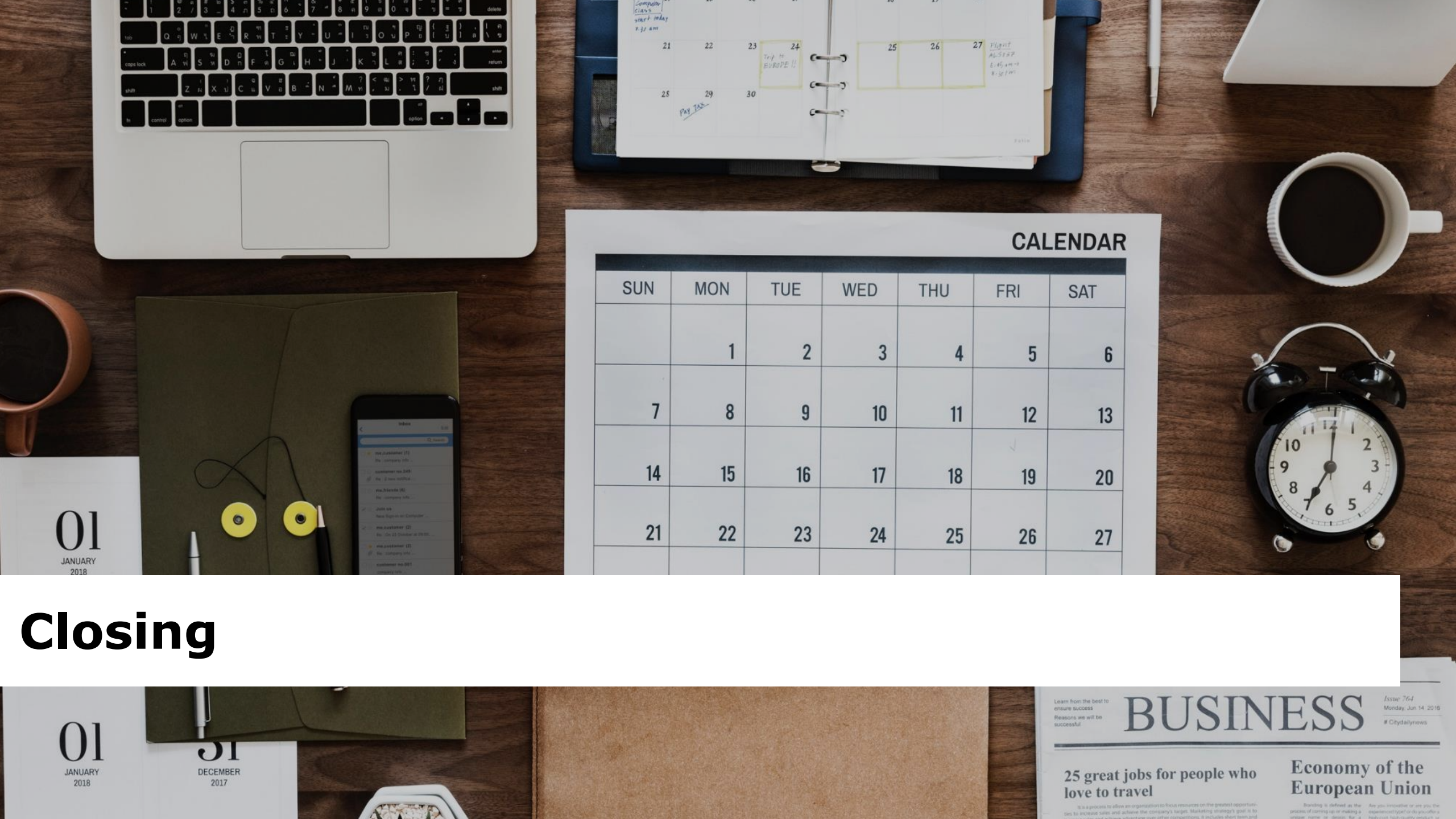
Thom Townsend

Executive Director

Open Q&A Session

Audience

Closing


CALENDAR
SUN MON TUE WED THU FRI SAT
01
JANUARY
2018
31
DECEMBER
2017
BUSINESS
25 great jobs for people who love to travel
Economy of the European Union

Deloitte.

open data institute

THE GREEN LAND

the Lisbon council
think tank for the 21st century

Thank you!

Deloitte.

IS 668746