# High-value datasets
# A European Strategy for Data

EuroGeographics webinar

19 May 2020

*Jiri PILAR*
*Unit G1 - Data Policy and Innovation*
*DG CNECT, European Commission*

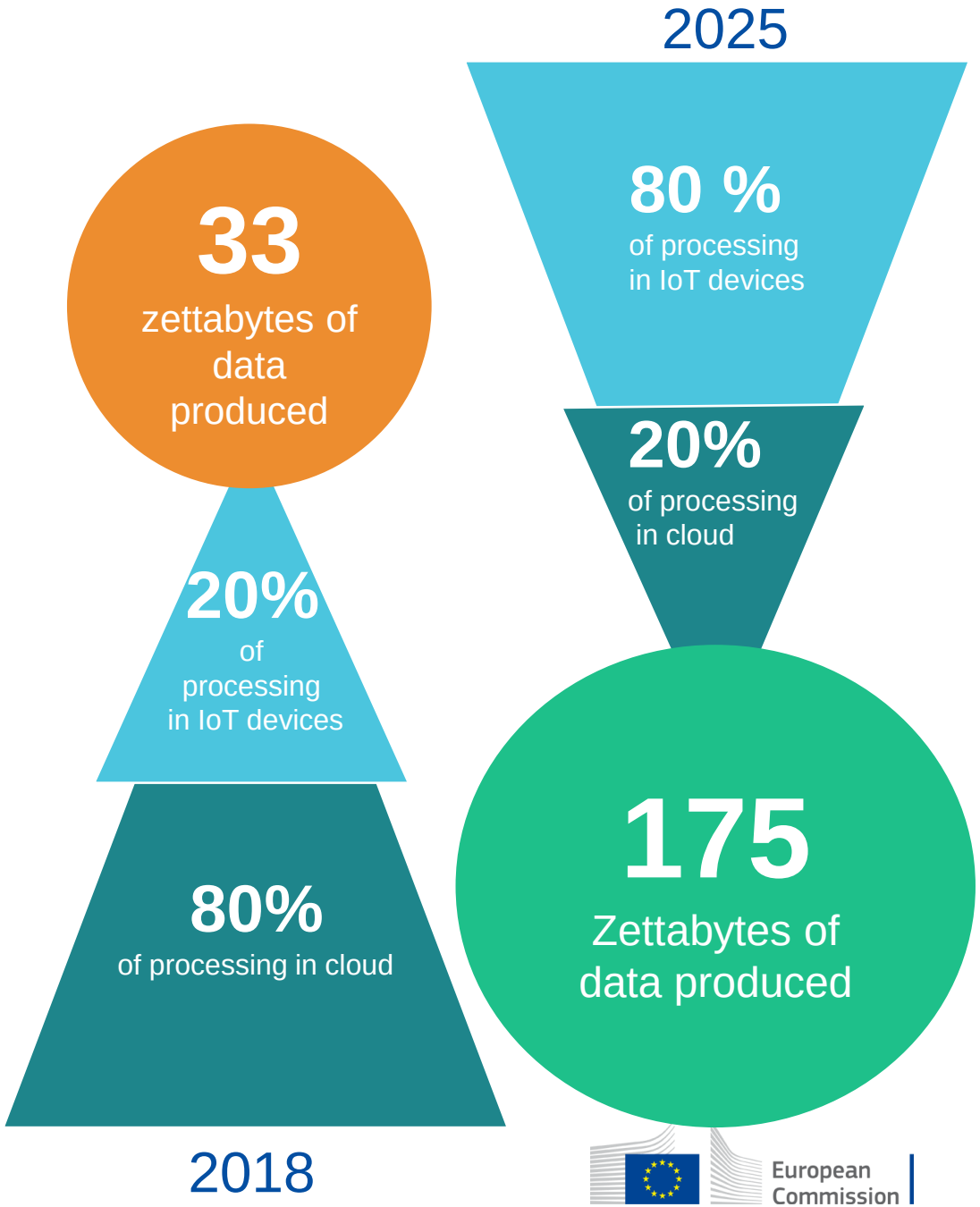
*“I want European businesses and our many SMEs to access high quality data and create value for Europeans – including by developing Artificial Intelligence applications.”*

*Thierry Breton,*
*Commissioner for the Internal Market*

European Commission

# Europe has everything to play for

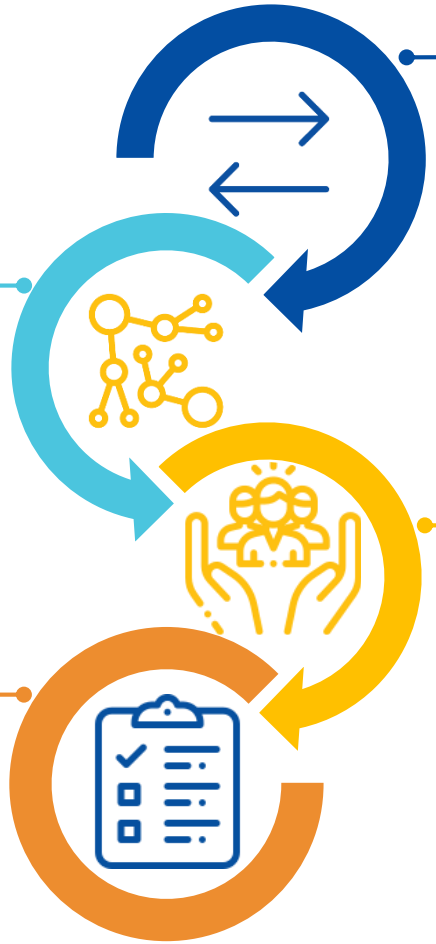
- Data can transform all sectors of the economy and is crucial for AI
- Personal and non-personal data can be a source of innovation for new products and services
- Data can contribute to tackle societal challenges such as climate change, health, mobility, etc.
- Data can make our lives and work easier and better




European Commission

# European Strategy for Data

A common European data space, a single market for data



Data can flow within the EU and across sectors

Availability of high quality data to create and innovate

European rules and values are fully respected

Rules for access and use of data are fair, practical and clear & clear data governance mechanisms are in place

European Commission

# What are the problems?

**Not enough data available for reuse**

- More public sector data can be made available
- Low uptake of voluntary data sharing among companies
- No clarity on the use of private sector data for the common good

**No real user empowerment**

- imperfect data portability mechanisms

**Lack of European data processing & storage solutions**

**Absence of comprehensive data governance approaches**

- To address legal and technical barriers within and across sectors (e.g. standardisation & interoperability)

**Skills shortage and low data literacy**

**Fragmentation of the single market**

European Commission

# Deploying the strategy through 4 Pillars

### A cross-sectoral governance framework for data access and use

including a legislative framework for the governance of European data spaces and other cross- sectoral measures for data access and use

### Enablers

Total investments of € 4-6 billion in a High Impact Project on European data spaces and federated cloud infrastructures

### Competences

Empowering individuals, investing in digital skills & data literacy and in dedicated capacity building for SMEs.

### Rollout of common European data spaces

in crucial economic sectors and domains of public interest, looking at data governance and practical arrangements.

International Aspects

European Commission

# Common European data spaces

Rich pool of data (varying degree of accessibility)

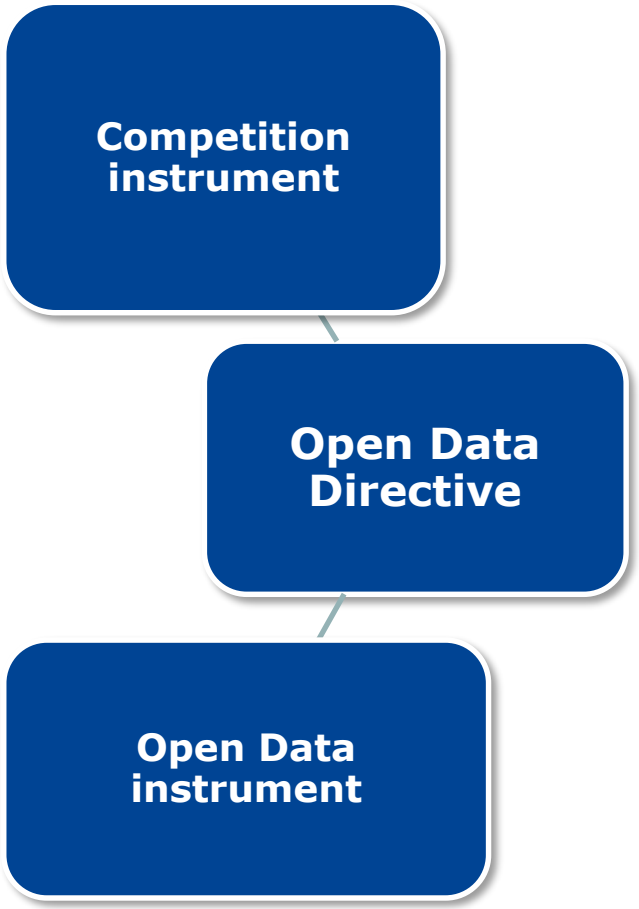
Free flow of data across sectors and countries

Full respect of GDPR

Horizontal framework for data governance and data access


- Technical tools for data pooling and sharing
- Standards & interoperability (technical, semantic)
- Sectoral Data Governance (contracts, licenses, access rights, usage rights)
- IT capacity, including cloud storage, processing and services

# Open Data Directive – basic facts




Introduces a minimal set of rules to make more data from the public sector easier and cheaper to use for innovation

The Directive does not apply to:

- data not publicly accessible under access to information regimes
- information protected by third party intellectual property rights

Application of the Directive must fully respect the Union and the national data protection rules

**Latest revision in 2019:** now entitled '**Directive on Open Data and the re-use of public sector information**'

# High value datasets – key points

Datasets listed in the implementing act to be made available for free, in machine-readable formats, via APIs and (where relevant) as bulk downloads

## Exceptions:

- Free availability requirement shall not apply to public undertakings if there is a risk of competition distortion
- In case of a substantial impact on the budget of the public bodies involved, free availability can be delayed by up to 2 years

# Thematic categories of high value datasets

- ***Geospatial***
- ***Earth observation and environment***
- ***Meteorological***
- ***Statistics***
- ***Companies and company ownership***
- ***Mobility***

An Implementing Regulation planned for 2021 will define the list of specific high-value data sets within the 6 thematic categories set out in Annex I among the documents to which the Directive applies

*Examples in recital 66*:

"the thematic categories listed in the Annex could inter alia cover postcodes, national and local maps (Geospatial), energy consumption and satellite images (Earth observation and environment ), in situ data from instruments and weather forecasts (Meteorological), demographic and economic indicators (Statistics), business registers and registration identifiers (Companies and company ownership), road signs and inland waterways (Mobility)."

Categories can be extended (delegated act).

# Progress so far and next steps

- The Directive was published in the Official Journal of the EU on 26 June 2019
- 2 years of transposition in Member States
- Work towards the definition of the List of HVDs (2019-21)
  - Public Sector Information Group (expert group twice a year)
  - Open Data Committee (comitology: for the Implementing Regulation)
  - Public consultation – closes on 31 May 2020
  - Impact assessment – supported by the ongoing study till September 2020
  - Dedicated actions (workshops, focus groups)
  - Proposal of the list of HVDs to be submitted to the Open Data Committee in Q1 2021

# High value datasets in scope. Where we started

## Geospatial

| 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.<br>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<br>Economic, social and environmental analysis<br>Cultural identity and heritage<br>Mapping and navigation<br>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.<br>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.<br>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. |
| Hydrography | Hydrographic elements, including marine areas and all other water bodies and items related to them, including river basins and sub-basins [from INSPIRE Data Specifications on Hydrography] | Mapping physical objects, Reporting, Modelling & spatial analyses |
| Land Use | A collection of areas for which information on existing (present or past) land uses is provided. Territory characterised according to its current and future planned functional dimension or socio-economic purpose (e.g. residential, industrial, commercial, agricultural, forestry, recreational). | land planning,<br>analysis of land consumption,<br>ecological network mapping,<br>greenhouse inventory reporting |
| Digital Terrain Model | Digital Terrain Models (DTMs) | High-resolution light detection and ranging (lidar) data are used in energy infrastructure siting, design, permitting, construction, and monitoring to promote public safety through the reduction of risks. For example, lidar data are used to identify safe locations for energy infrastructure by analyzing terrain parameters. |
| Digital Surface Model | Digital surface model (DSMs) | |
| 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,<br>disaster management, real Estate Market, Taxation, LPIS (Agriculture), Land consolidation, Infrastructure Management, Spatial Planning, Protection of Soil and Water, Statistics |

While identifying the HDVs in scope, our research (desk+interviews to data holders + expert interviews) suggested that these are the main ones to be included in the Geospatial category.

For example, the following datasets are covered by other thematic areas:

- **Land use (Earth Observation and Environment)**
- **Digital terrain models (Earth Observation and Environment)**
- **Digital surface models (Earth Observation and Environment)**
- **Hydrography (Mobility)**
- **Transport networks (Mobility)**

European Commission

# High Value datasets in scope of our analysis - We need your feedback

## Geospatial

| 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.<br>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<br>Economic, social and environmental analysis<br>Cultural identity and heritage<br>Mapping and navigation<br>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 |

**Key questions:**

- *Do you have additional feedback on the proposed list?*
- *How do you use these datasets?*
- *What are the main encountered barriers for reuse in your experience?*
- *Where do you get the data (provide an example)?*
- *Would it be possible to provide some use cases for different datasets that show their potential?*

European Commission

# Public Open Data: 2020 call

Call **CEF-TC-2020-2** Public Open Data indicative dates:
Opens 30/6/2020 – deadline for submission 5/11/2020

- Budget: 3 M€
- Co-funding rate: up to 75% of eligible costs
- Indicative duration of the actions: up to 36 months

Proposals must be submitted by:

- consortia composed of a minimum of two entities
- at least one must be a public sector body
- based in one or more Member States and/or EEA countries participating in the CEF Telecom programme

Innovation and Networks Executive Agency

# Other supporting activities

- **Open Data digital infrastructure: European Data Portal and EU Open Data Portal**

- **Digital Europe Programme (DEP)**: Specific Objective 2 'Data for Artificial Intelligence (AI)' will strengthen core AI capacities in Europe, including data resources. Calls will focus on, inter alia, making specific datasets interoperable and fit for AI applications. Activities could cover, for example:
  - **curation;**
  - **semantic annotation;**
  - **harmonisation of metadata;**
  - **facilitating publication in machine-readable formats and accessibility through APIs.**

# Thank you very much for your attention

## For further questions:

**email:** CNECT-G1@ec.europa.eu
Unit G1 of DG CONNECT

## Websites with more information:

https://ec.europa.eu/digital-single-market/en/open-data

https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=celex:32019L1024

https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en