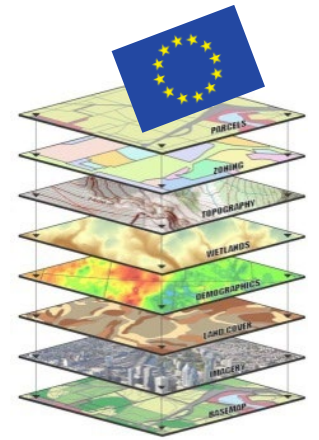# Creation of pan-European geospatial datasets -
- some updates

Joint EuroSDR & EuroGeographics Workshop

on Geodata Discoverability

*Eurostat – E.4 – GISCO – H.I. Reuter*

# GISCO – The Geographical Information System of the Commission

- Permanent service of the European Commission (Eurostat)

- Role: Answers needs of European Commission, Eurostat and the ESS for geographical information at EU/national/regional levels.

  - Provision of GIS (reference) data, services and software,

  - Support cartographic and spatial analysis activities,

  - Coordinates commission-wide GI activities,

  - Stimulate the use of GIS to support commission activities,

  - Support activities on the integration of statistical and geospatial information.

GISCO - the **G**eographic **I**nformation **S**ystem of the **CO**mmission
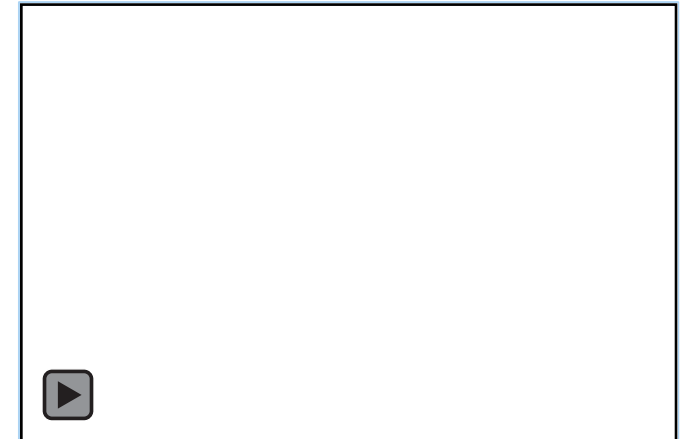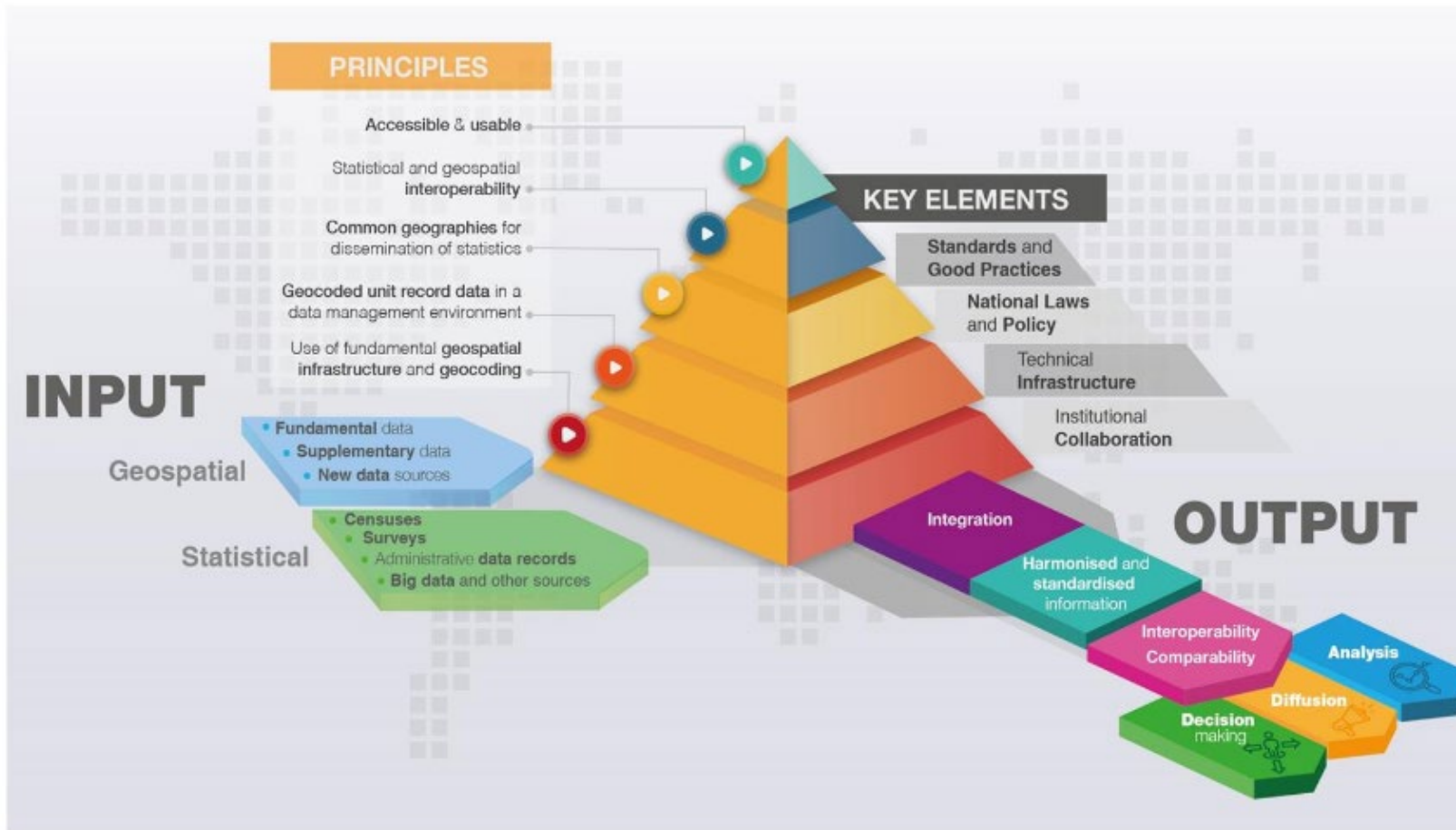
localise
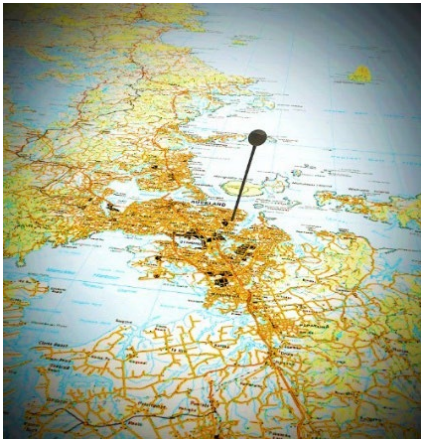
analyse
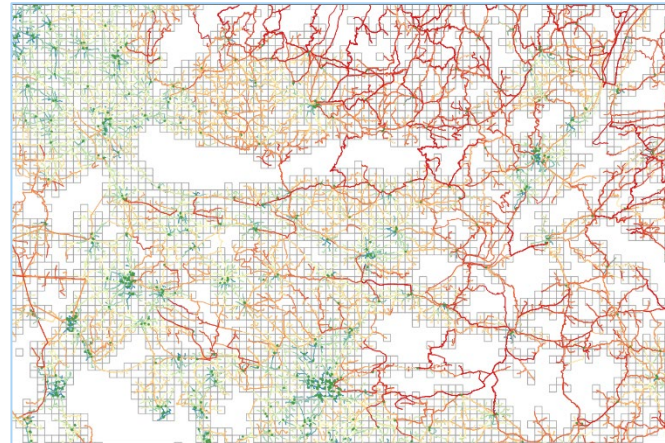
visualise

Global Statistical Geospatial Framework

High Value Dataset Directive (HVD)

20.1.2023   EN   Official Journal of the European Union   L 19/43

COMMISSION IMPLEMENTING REGULATION (EU) 2023/138

of 21 December 2022

laying down a list of specific high-value datasets and the arrangements for their publication and re-use

(Text with EEA relevance)

THE EUROPEAN COMMISSION,

Having regard to the Treaty on the Functioning of the European Union,

Having regard to Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information ([1]), and in particular Article 14(1) thereof,

Whereas:

(1) It transpires from Directive (EU) 2019/1024 that an EU-wide list of datasets with a particular potential to generate socio-economic benefits and with harmonised re-use conditions is a significant enabler of cross-border data applications and services.

(2) The main objective of establishing the list of high-value datasets is to ensure that public data of highest socio-economic potential are made available for re-use with minimal legal and technical restriction and free of charge.

(3) Harmonising the implementation of the re-use conditions of high-value datasets entails the technical specification for making the datasets available in a machine-readable format and via application programming interfaces (APIs). Making high-value datasets available under optimal conditions strengthens the open data policies in the Member States, building on the principles of findability, accessibility, interoperability and reusability (FAIR principles).

(4) Annex I to Directive (EU) 2019/1024 lays down the themes of high-value datasets by listing six thematic data categories: 1) geospatial; 2) earth observation and environment; 3) meteorological; 4) statistics; 5) companies and company ownership; and 6) mobility.

Article 6

Entry into force and application

This Regulation shall enter into force on the twentieth day following that of its publi Union.

It shall apply from 16 months after entry into force of this Regulation.
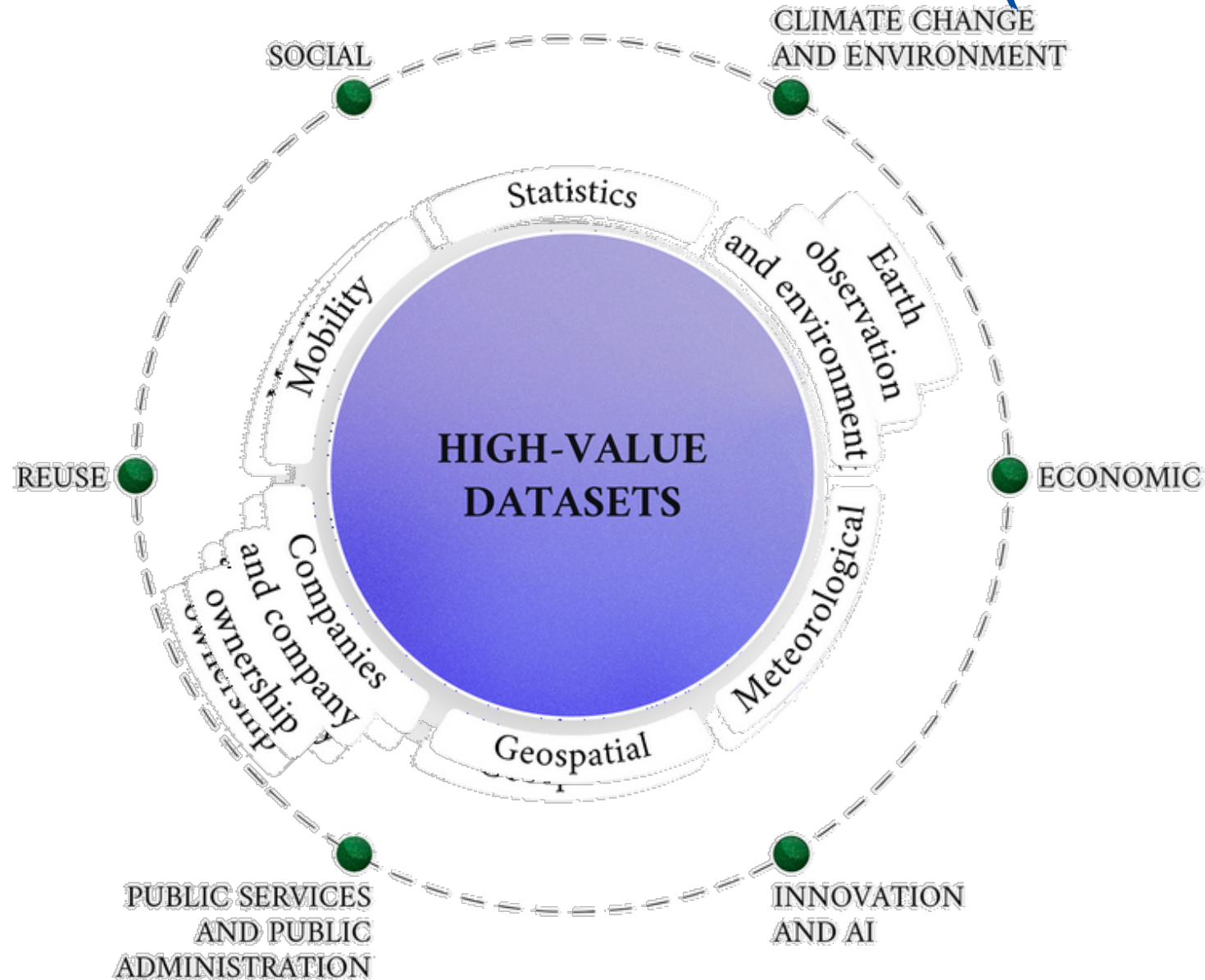
Publication: 20.01.2023
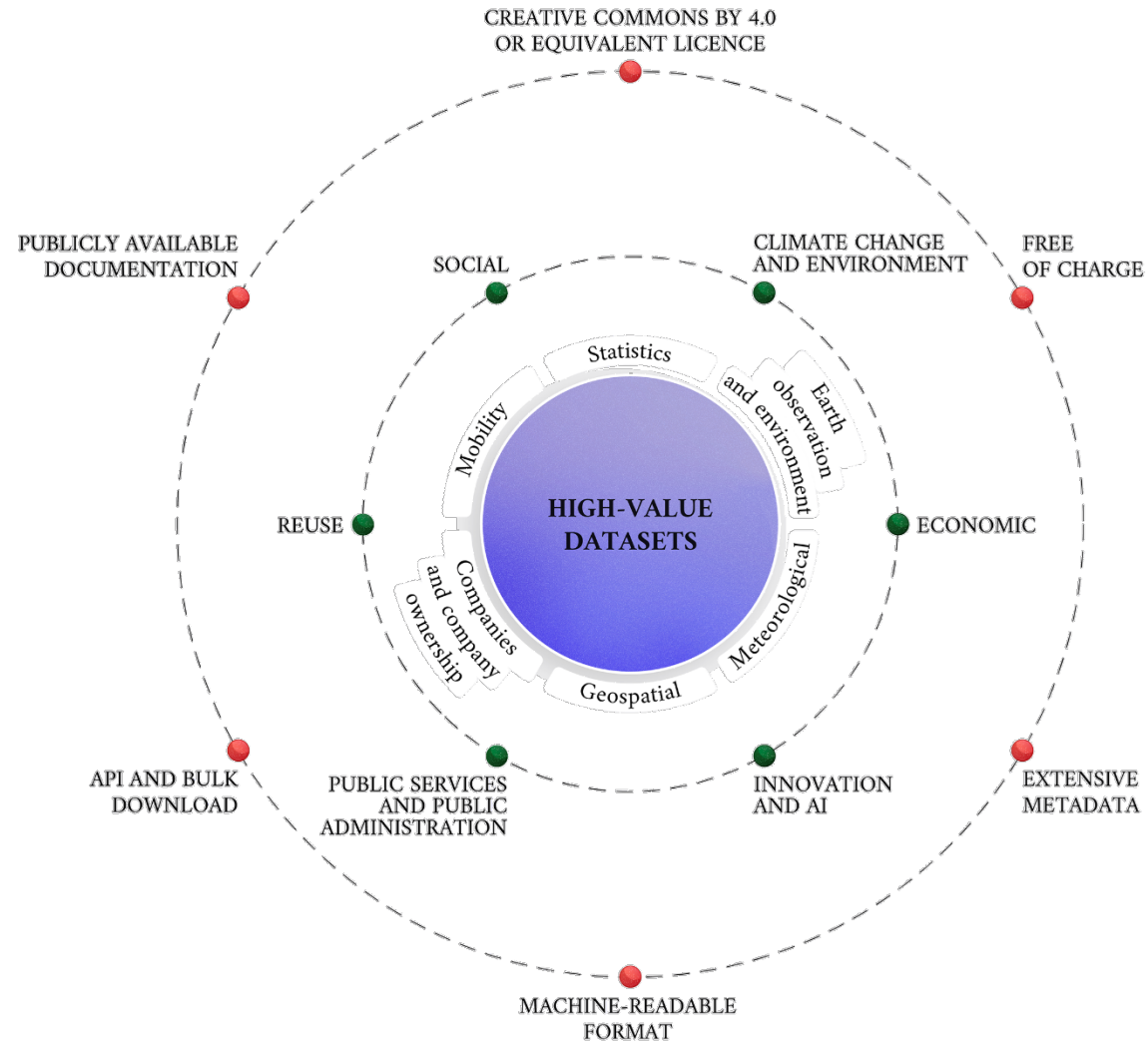Deadline: 20.June 2024

Legal text: EN

European Commission

High Value Dataset Directive (HVD)

FEATURES OF HIGH-VALUE DATASETS

● = *Macro characteristics*
● = *Mandatory technical requirements*

CREATIVE COMMONS BY 4.0
OR EQUIVALENT LICENCE

PUBLICLY AVAILABLE
DOCUMENTATION

SOCIAL

CLIMATE CHANGE
AND ENVIRONMENT

FREE
OF CHARGE

Statistics

Earth observation and environment

Mobility

HIGH-VALUE
DATASETS

REUSE

ECONOMIC

Companies and company ownership

Meteorological

Geospatial

API AND BULK
DOWNLOAD

PUBLIC SERVICES
AND PUBLIC
ADMINISTRATION

INNOVATION
AND AI

EXTENSIVE
METADATA

MACHINE-READABLE
FORMAT

https://data.europa.eu/en/publications/datastories/high-value-datasets-overview-through-visualisation

European
Commission

	Granularity*	Key attributes	Geographical coverage
Administrative units	• All generalisation levels available up to the level of scale 1 : 5 000	• Unique identifier • Unit type (administrative or maritime	Single or multiple
Buildings	• All generalisation levels available up to the level of scale 1 : 5 000	• All generalisation levels available up to the level of scale 1 : 5 000	
Cadastral parcels	• All generalisation levels available up to the level of scale 1 : 5 000	• National identification code • All generalisation levels available up to the level of scale 1 : 5 000 • Unique identifier • Geometry	when
Geographical names	N/A		
Reference parcels	• Equivalent at least to 1 : 10 000 and, as from 2016, to 1 : 5 000 as referred to in Article 70(1) of Regulation (EU) 1306/2013	• Equivalent at least to 1 : 10 000 and, as from 2016, to 1 : 5 000 as referred to in Article 70(1) of Regulation (EU) 1306/2013	
Addresses	N/A	• Unique identifier	
Agricultural parcels	• Equivalent at least to 1 : 10 000 and, as from 2016, to 1 : 5 000 as referred to in Article 70(1) of the Regulation (EU) 1306/2013	• Equivalent at least to 1 : 10 000 and, as from 2016, to 1 : 5 000 as referred to in Article 70(1) of the Regulation (EU) 1306/2013 • Thoroughfare (street) • Name	

Granularity*

Key attributes

Geographical coverage

* If datasets are not available at the scale indicated in the table, but are available at higher spatial resolution(s), they shall be provided at the available spatial resolution.

• Administrative units (e.g. municipality, province, country)

• Postal descriptor (e.g. post code)

• Date of last update

European Commission

Stacking GIS layers

Country

Statistical Units

Administrative Units
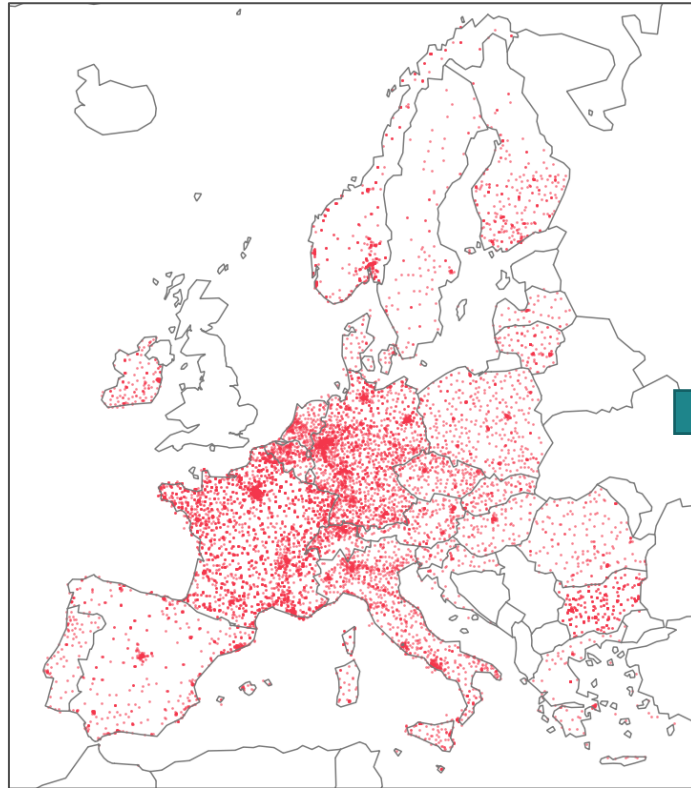
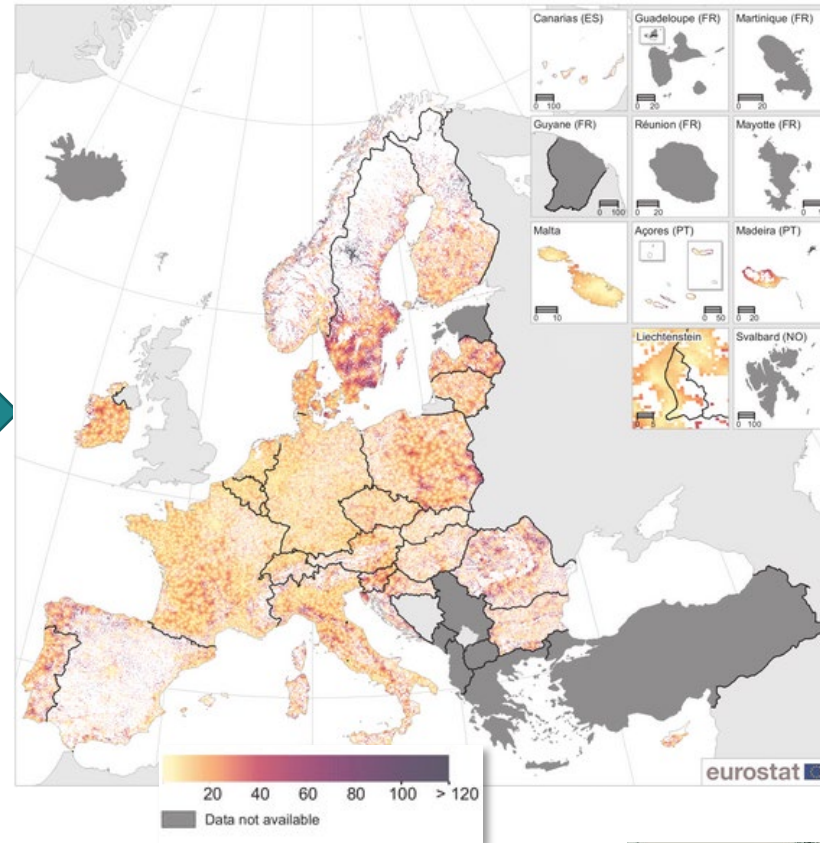Cadastral Parcel

Building

Address

European Commission

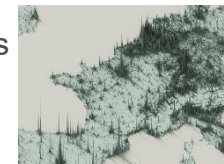# Application Example for pan European GI Datasets

AD,AU,US

TN, PD,SU

SU

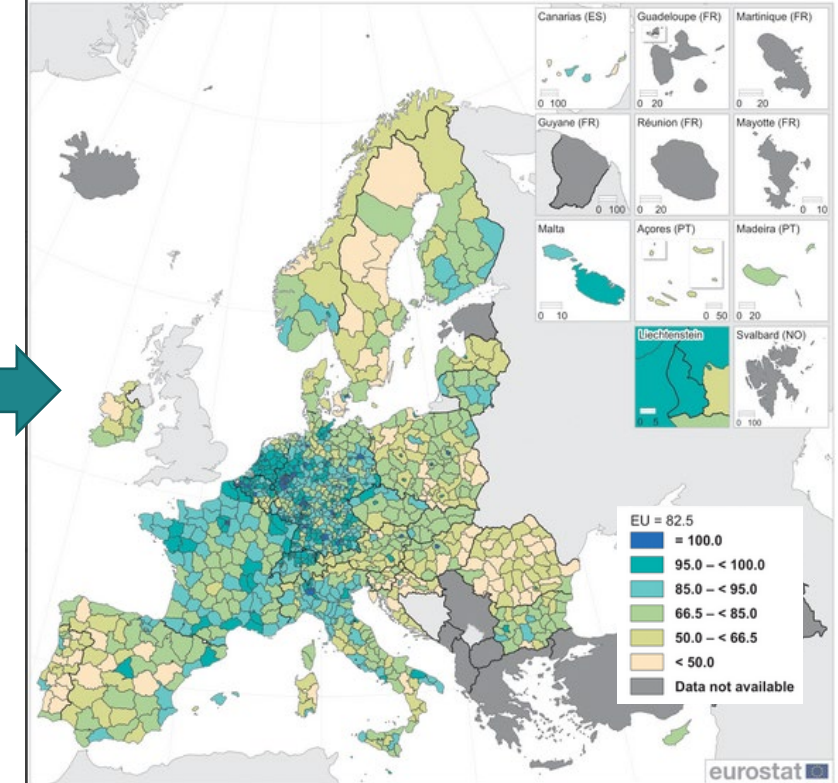Data: Utility Services (US) - Healthcare from MS responsible authoritative organisation

Analytics: Routing, Population, Utility Services

Statistics: Reporting at NUTS level, DEGURBA

EU = 82.5
= 100.0
95.0 - < 100.0
85.0 - < 95.0
66.5 - < 85.0
50.0 - < 66.5
< 50.0
Data not available

20 40 60 80 100 > 120
Data not available

eurostat

European Commission

Prod   https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/

Establishing pan European Datasets Workflow

Analysis / Dissemination

Computer Systems

INSPIRE Geoportal

data.europa.eu

Kohesio: discover EU projects

Analysis in Incoming WIKI
Import into Database
 QAQC Cycles,
 Integration, Generalisation,
 Topological Checks,
 Feedback   with MS

………..

Majority of work

Eurostat

European Institutions

European Citizen/Member states/regions

Human Intelligence

Local Contacts
Reports

EUROPEAN STATISTICAL SYSTEM

eurogeographics

NSI/NMCAs

Social media

Research projects
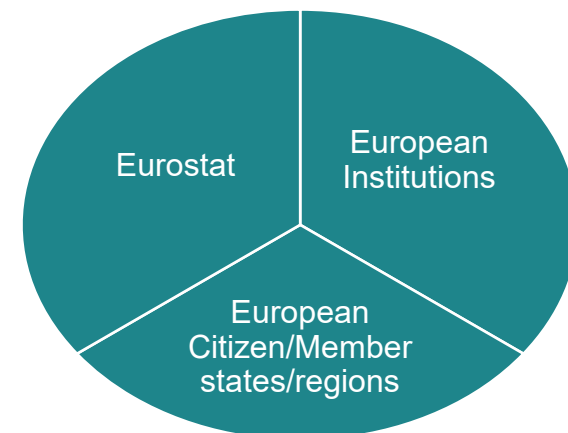
GISCO ReferenceDB
Website   *Tercet*
Services: *Download*,
AddressAPI, *ID*, *VectorT*,
OGC-API

European Commission

Data & systems:
GIS reference data – AU-SU-CP- BU-AD
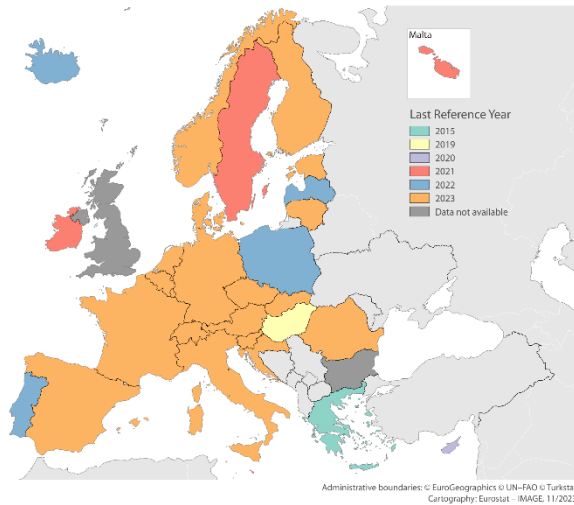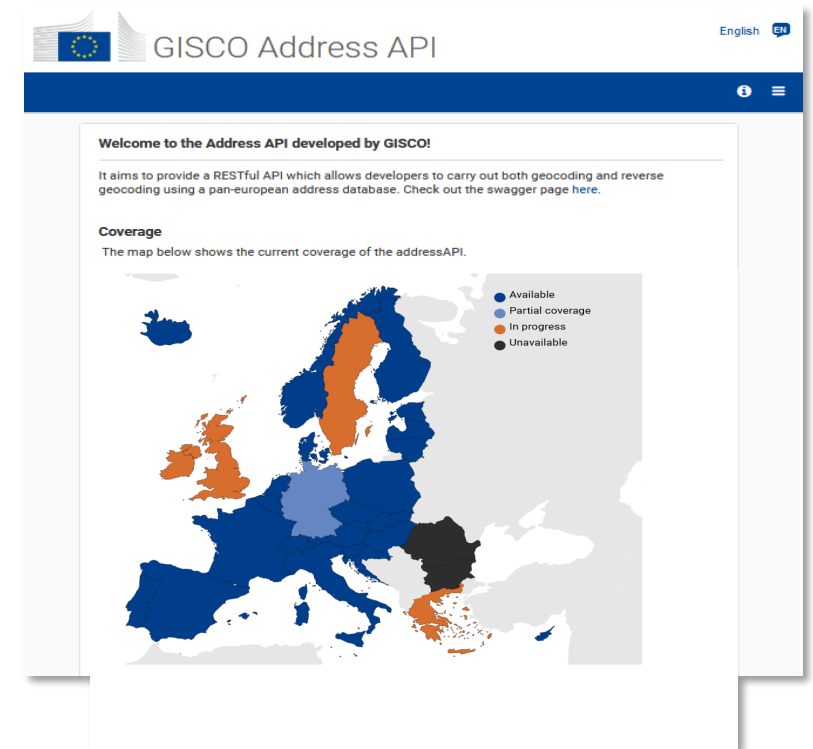

Administrative Units

Currently processing
for 2023/2024
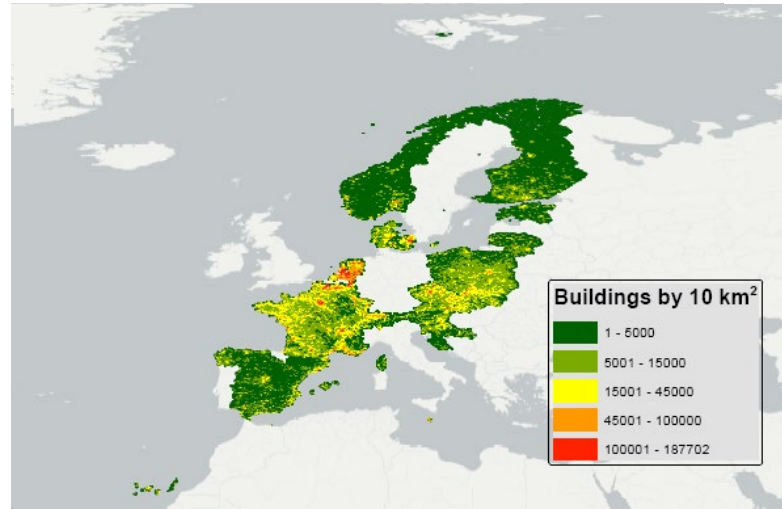
Building Units

New datasets for AT, CH, HR

Call for contributions:
estat-gisco-panEU@ec.europa.eu

Annual update in process
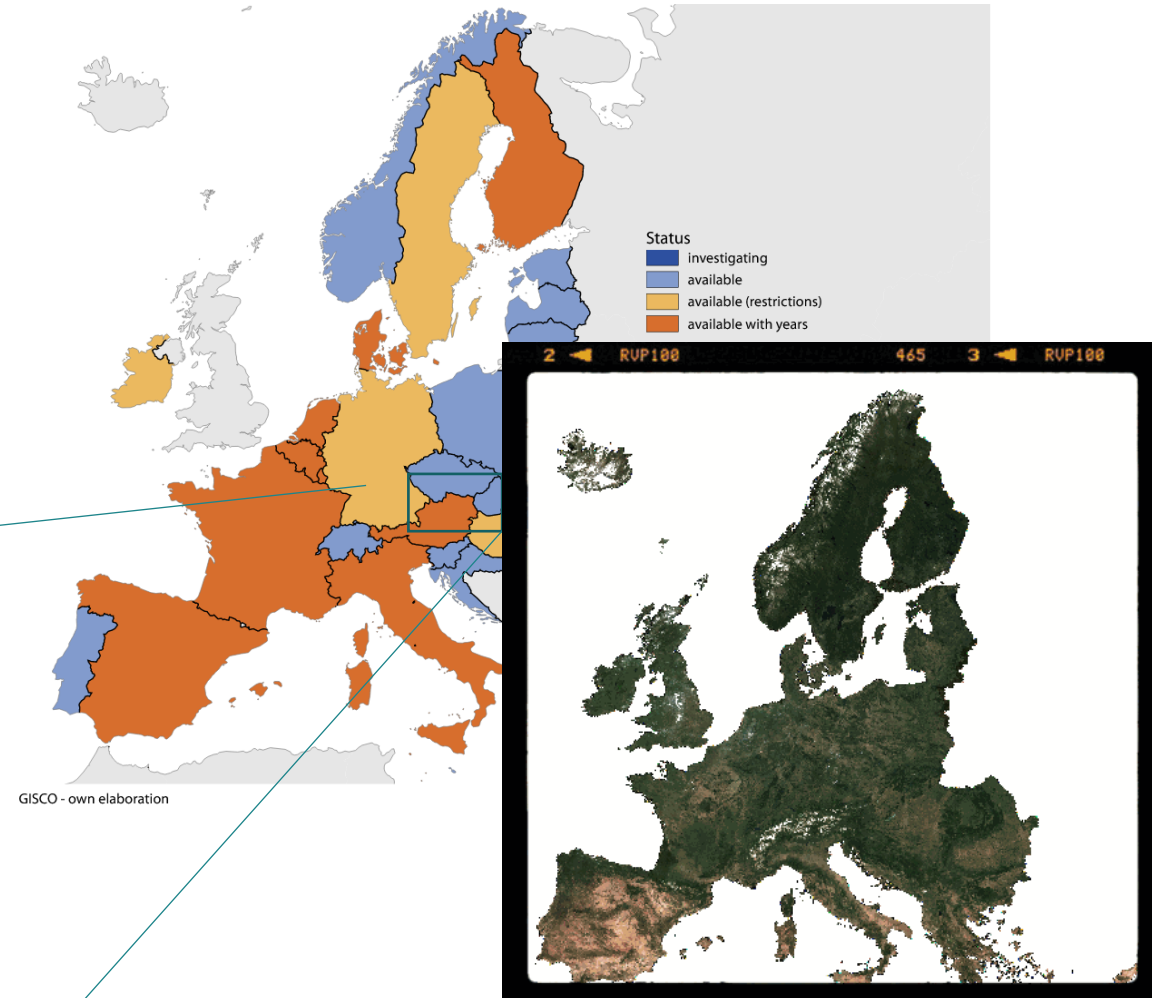Dissemination pending some
legal aspects

OrthoImagery (I)

- latest snapshot of OI from MS organisation

- If time series available – will be supported

- Depends on MS service performance + data requirements

- Up to 0.2m

OrthoImagery Status March 2022
Available in Development Instance

Status
- investigating
- available
- available (restrictions)
- available with years

GISCO - own elaboration

European Commission

GIS Observations

- *Generalized data* (250k scale vs 10k on ground) – not usable Requirement: Data provided according to EU core reference data specification

- *Data are not aligned* across datasets (AU,AD,PC, BD,TN) (road in the middle of a building, address point in a different AU unit), Requirement: QAQC in MS

- Quite *varying quality control* across themes and countries Requirement: Quality Control

- Richness of datasets differs extremely (from line geometries only without attributes to full stack dataset according to INSPIRE specifications) Requirement: Rich data content needs to be in distributed datasets

- Feedback loop with MS extremely important Requirement: Resources at MS level to fix issues

- Various slight variations of INSPIRE implementation – we still prefer INSPIRE implementations. Requirement: Harmonising INSPIRE implementations
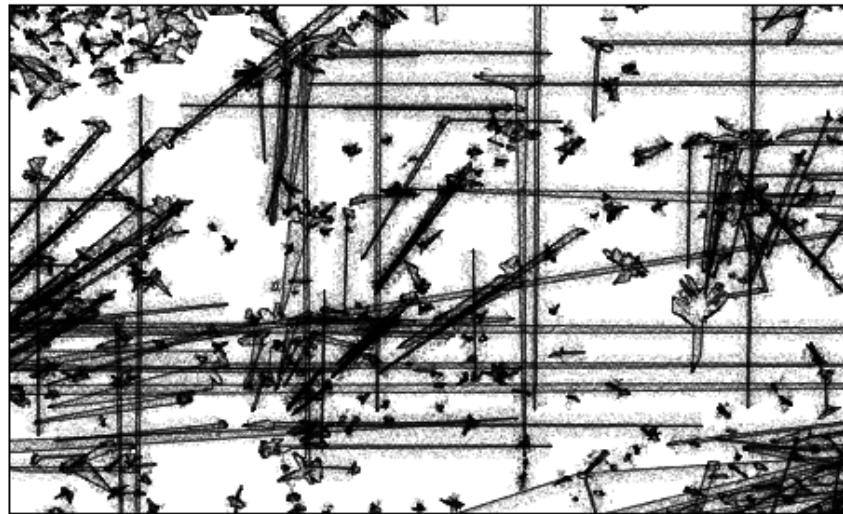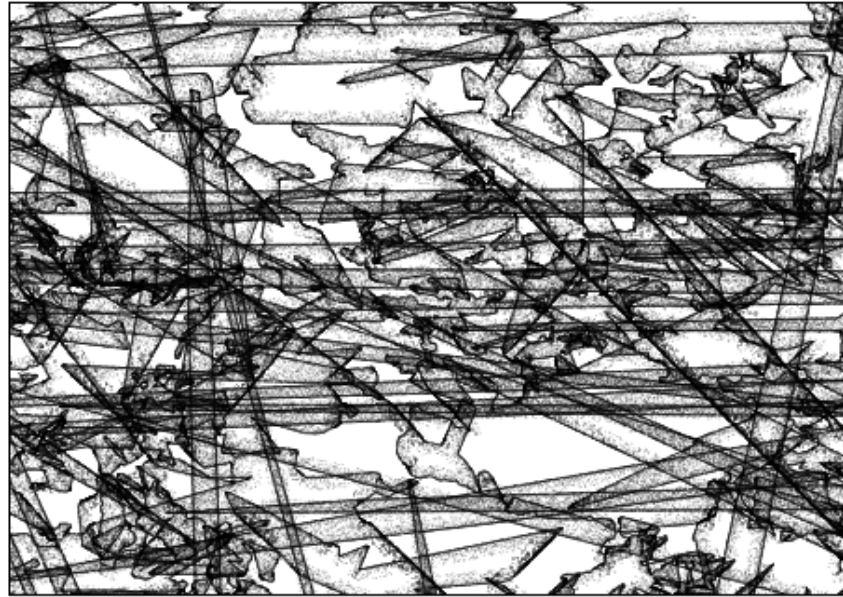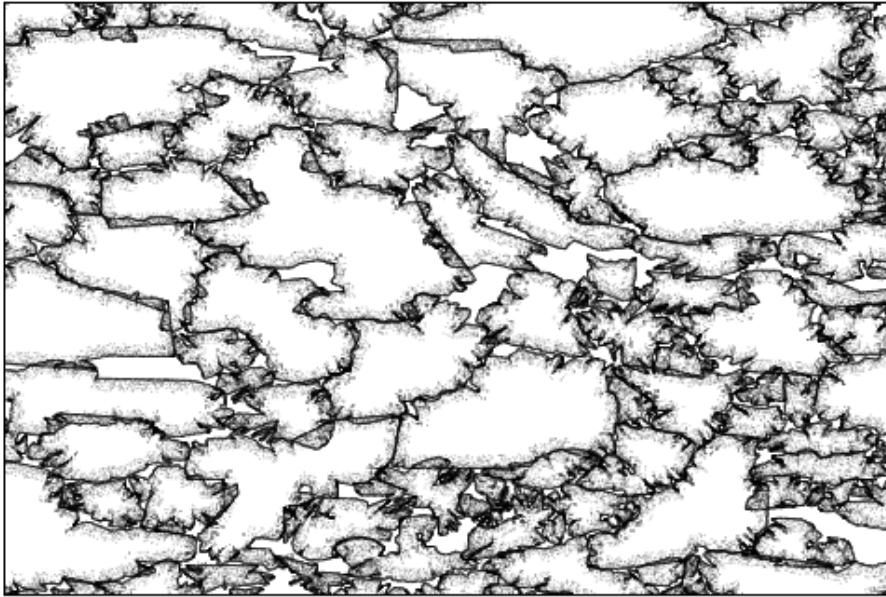
Various Examples of concave Hulls of Postal Codes across the European Union - top left: good positional dataset, top right: positional issues, bottom right: limited outliers on top of a good dataset with respect to postional accuracy, bottom right: positional issues together with

Conclusions

- Eurostat works on a range of topics for pan-European data sets

- Depending on the datasets a) European Databases are created or b) datahub pointing back to MS resources to provide services internally/externally

- Annual updates – In general good feedback iterations with data providers as they see their data used and can improve…… takes time + resources (e.g. funding available via REGIO/CONNECT/ESTAT)

- Progress is hampered by e.g. a) documented and available datasets, b) stability of services, c) performance of services and d) willingness of data providers.

- Code of conduct for Geospatial Data is required

European Commission

Thank you

ESTAT-GISCO@EC.EUROPA.EU

ESTAT-GISCO-PANEU@EC.EUROPA.EU