GRANULAR

KTA webinar on Data for Rural Areas

5 December 2024

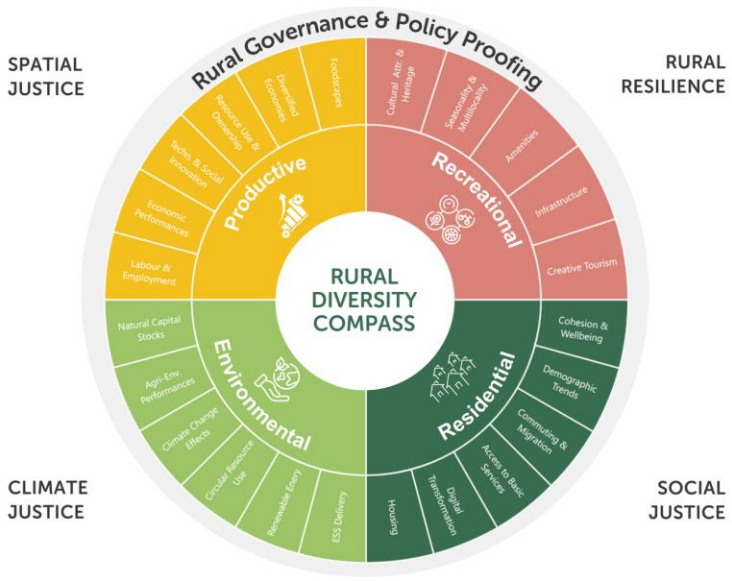
GRANULAR has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101061068. UK participants in the GRANULAR project are supported by UKRI- Grant numbers 10039965 (James Hutton Institute) and 10041831 (University of Southampton). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

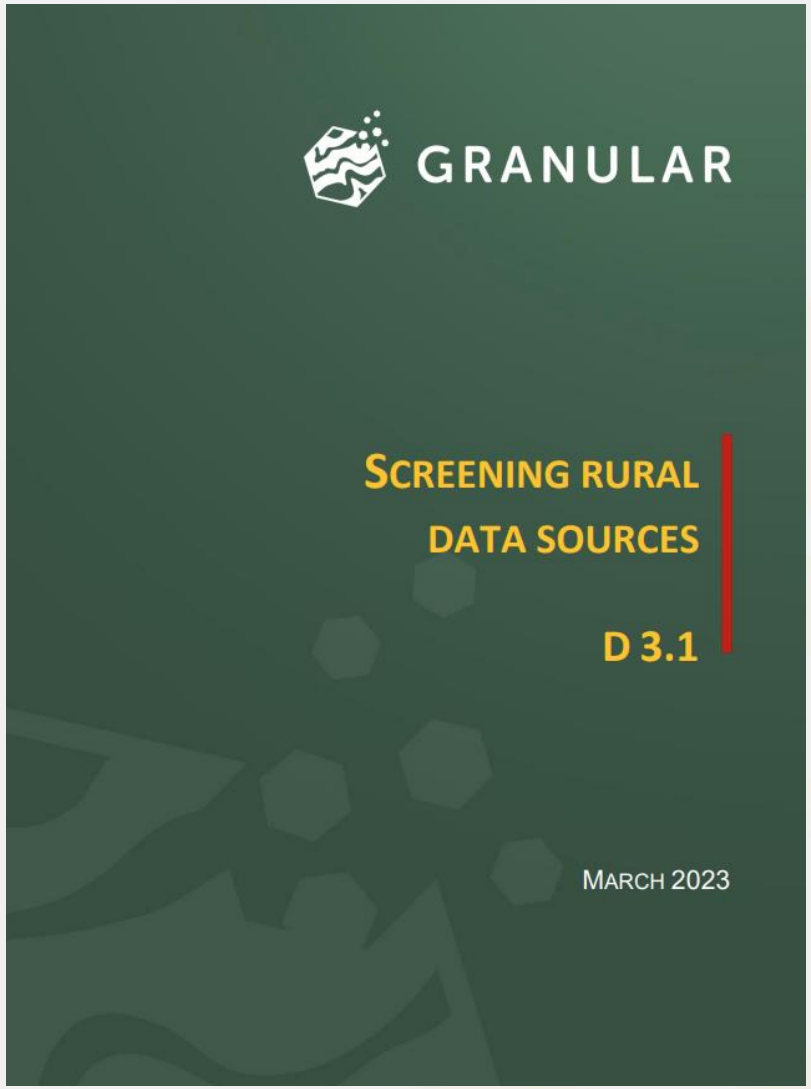
Co-funded by the European Union

BENCHMARK OF RURAL DATA, ISSUES, RESULTS

Data opportunities

- **Objective:** Generate and assess new data to better showcase the diversity of rural areas, their functional characteristics, challenges and opportunities.

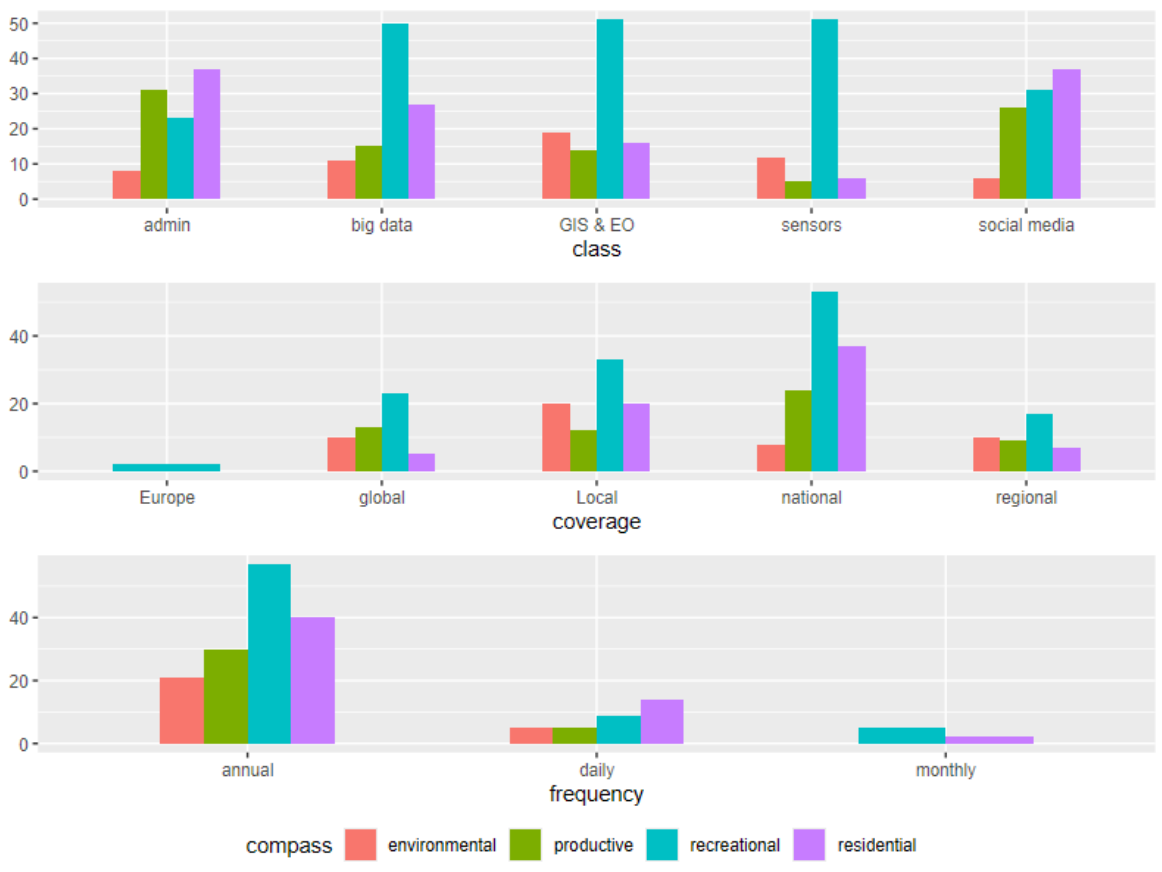


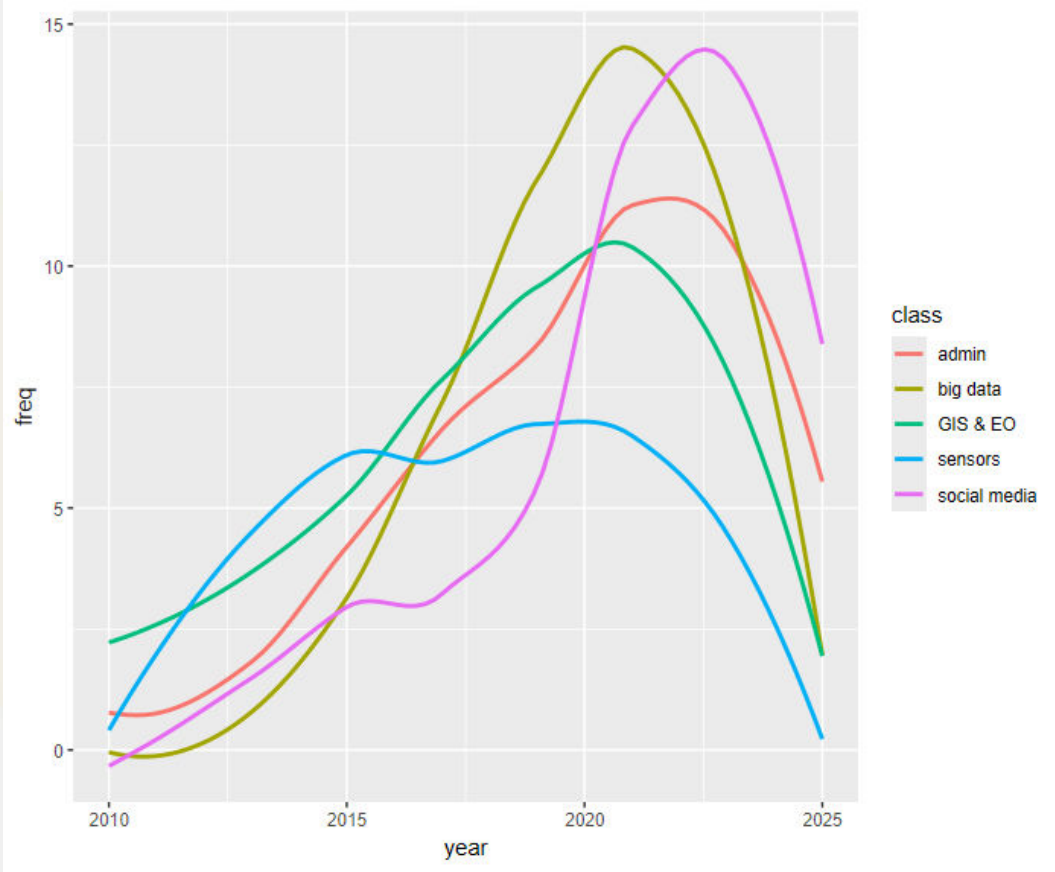


https://www.ruralgranular.eu/publications/

DATA AVAILABILITY

Data types and trends





WP3/4 et al., 2024. Systematic review of rural sustainability data. *In preparation*

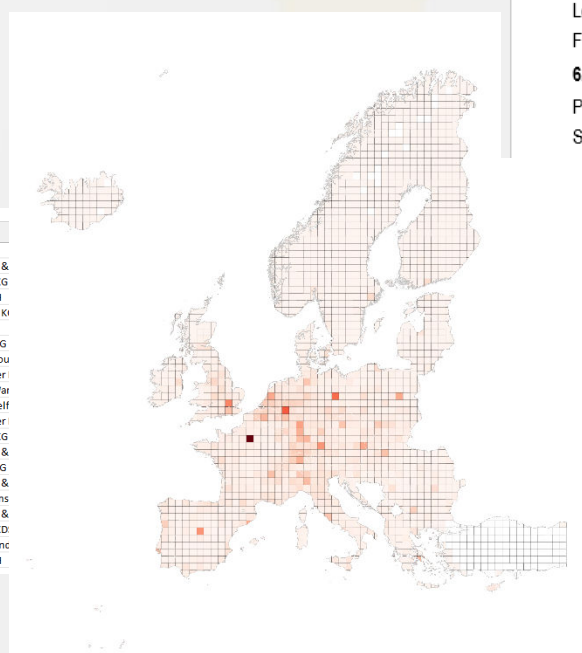
PRODUCTIVE

“labour” OR “employment” OR resource* OR “ownership” OR foodscapes OR “diversified economies” OR “economic performances” OR “technological innovation”

- Data examples
  - Regional economy indicators (Open Street Maps)
  - National Statistics (Eurostat – EU SILC, …)
  - National Chambers of Commerce

- Methods
  - GIS
  - Surveys
  - Statistical analysis
  - Web-scraping

	A	B	C	D	E	F	G	H	I
	Zentrale	Region	Markttyp	Partnerfirma	Regie	Organisation	Geschaeftsart	VKF	MA	Inhaber
2	Aldi Süd	Aldi-Dormagen	Aldi Süd		R	Aldi	Discounter			ALDI Einkauf GmbH &
3	Lidl	Lidl-Leverkusen	Lidl		R	Lidl	Discounter			Lidl Stiftung & Co. KG
4	Penny	Penny-West	Penny		R	Rewe	Discounter			PENNY Markt GmbH
5	Rewe Regionen	Rewe-West	REWE City	Rahmati	S	Rewe	Supermarkt (400 bis 799 VKF)			S-Mart GmbH & Co. K
6	Rewe Regionen	Rewe-West	REWE City	REWE Regiebetriebe [R	R	Rewe	Supermarkt (400 bis 799 VKF)			REWE Markt GmbH
7	Rewe Regionen	Rewe-West	REWE	REWE SEH	S	Rewe	Supermarkt (400 bis 799 VKF)			REWE Wintgens OHG
8	Netto-Marken-Discoun	Netto Schels	Netto		R	Edeka	Discounter			Netto Marken-Discou
9	Edeka Regionen	Edeka Rhein-Ruh	Edeka	Hetzenegger	S	Edeka	Supermarkt (400 bis 799 VKF)	600		Markus Hetzenegger
10	Edeka Regionen	Edeka Rhein-Ruh	Edeka	Feldmann-Höner	S	Edeka	Supermarkt (400 bis 799 VKF)			Feldmann-Höner Wa
11	Norma	Norma	Norma		R	Norma	Discounter			Norma Lebensmittelf
12	Edeka Regionen	Edeka Rhein-Ruh	nah & gut	Hetzenegger	S	Edeka	Supermarkt (400 bis 799 VKF)	300		Markus Hetzenegger
13	Lidl	Lidl-Leverkusen	Lidl		R	Lidl	Discounter			Lidl Stiftung & Co. KG
14	Aldi Süd	Aldi-Dormagen	Aldi Süd		R	Aldi	Discounter			ALDI Einkauf GmbH &
15	Rewe Regionen	Rewe-West	REWE	REWE SEH	S	Rewe	Supermarkt (400 bis 799 VKF)			REWE Wintgens OHG
16	Aldi Süd	Aldi-Dormagen	Aldi Süd		R	Aldi	Discounter			ALDI Einkauf GmbH &
17	NN	NN	Alnatura		R	Sonstige	Biomärkte			Alnatura Produktions
18	Aldi Süd	Aldi-Dormagen	Aldi Süd		R	Aldi	Discounter			ALDI Einkauf GmbH &
19	Kaufland	Kaufland	Kaufland		R	Markant	Verbrauchermarkt (2500 bis 4999 VKF)			Kaufland Vertrieb KD
20	Edeka Regionen	Edeka Rhein-Ruh	Marktkauf	EDEKA Regiebetriebe [F	R	Edeka	SB-Warenhaus (ab 5000 VKF)			Marktkauf Einzelhanc
21	Penny	Penny-West	Penny		R	Rewe	Discounter			PENNY Markt GmbH



LOCAL FOOD PRODUCERS – QUESTIONNAIRE

I. GENERAL SECTION:

1. Where is your enterprise/farm located?_____ **zip code** _____

2. When did your business first become established?_____

3. How long have you been running your business? _____

4. What is the legal form of your business?

- ❑ Individual farm
- ❑ Registered business activity
- ❑ Civil Company
- ❑ Limited liability company
- ❑ Public limited company
- ❑ General partnership
- ❑ Limited partnership
- ❑ Association/ Foundation
- ❑ Other, what kind_____

5. What is your educational background?

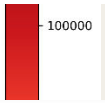
Level: ❑ Elementary ❑ Essential vocational ❑ Secondary education/high school ❑ Higher

Field: ❑ Agricultural ❑ Economical ❑ Artistic ❑ Other, what kind_____

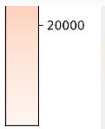
6. How many people, including yourself, work in the enterprise/farm?

Permanently _____ including family members _____

Seasonally _____ including family members _____



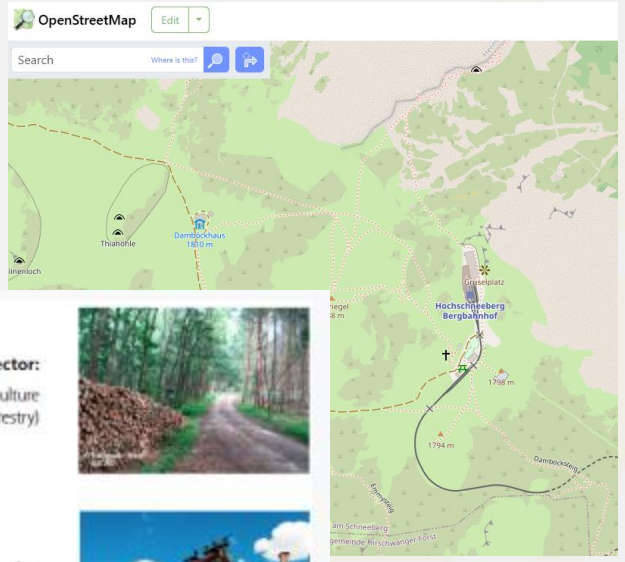

```
Number of points in each category in the filtered dataset:
key
amenity    2800117
shop       1939594
office      325995
tourism     315522
sport       143746
craft        97870
Name: count, dtype: int64
```

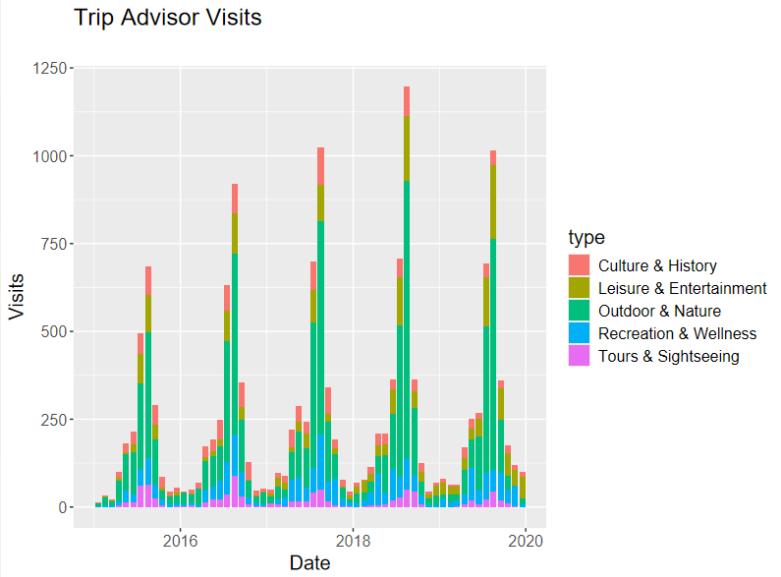


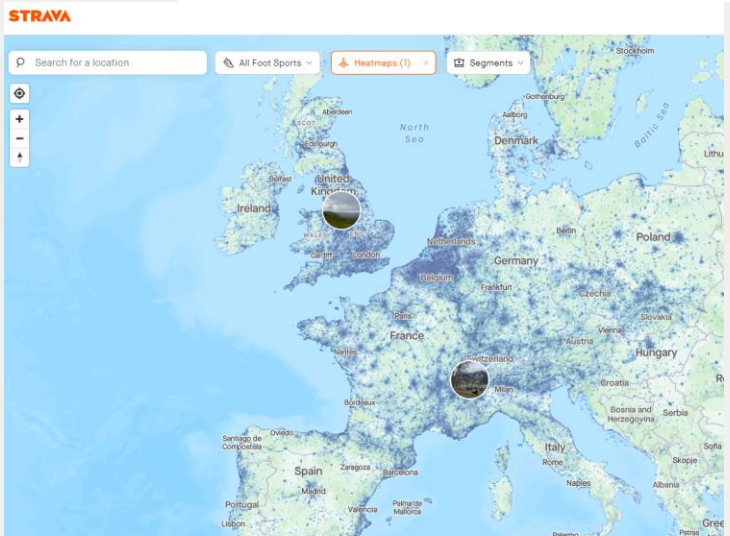
RECREATIONAL

amenities OR infrastructure OR “cultural attractiveness” OR heritage OR tourism* OR seasonality OR “multi-locality”

- Data examples
  - Open Street Maps (OSM)
  - Social media (TripAdvisor, Booking, Airbnb, etc.)
  - Facebook, Instagram, etc.
  - LUCAS - Eurostat
  - Strava
- Methods:
  - Web-scraping
  - Time series analysis
  - GIS/spatial statistics
  - Computer vision/VLMs



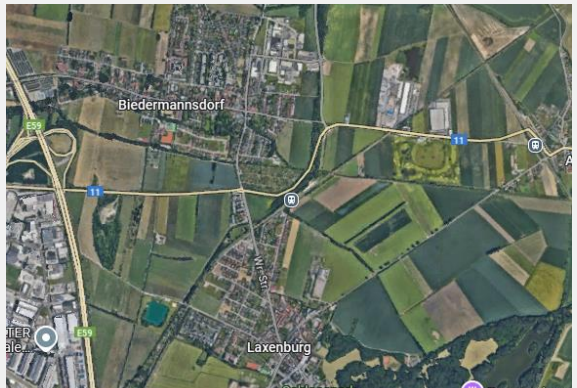


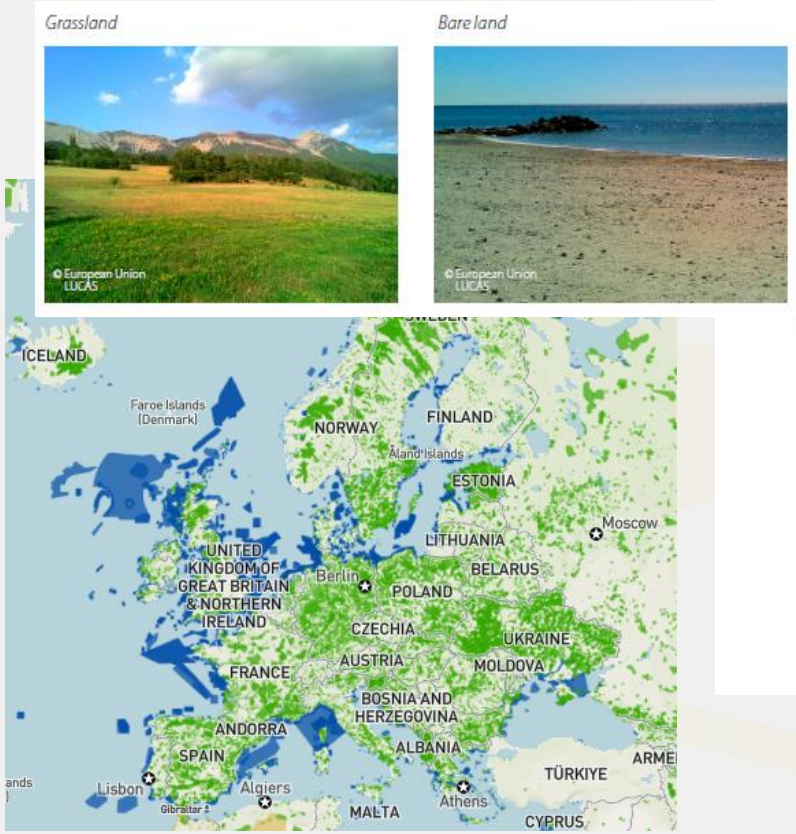


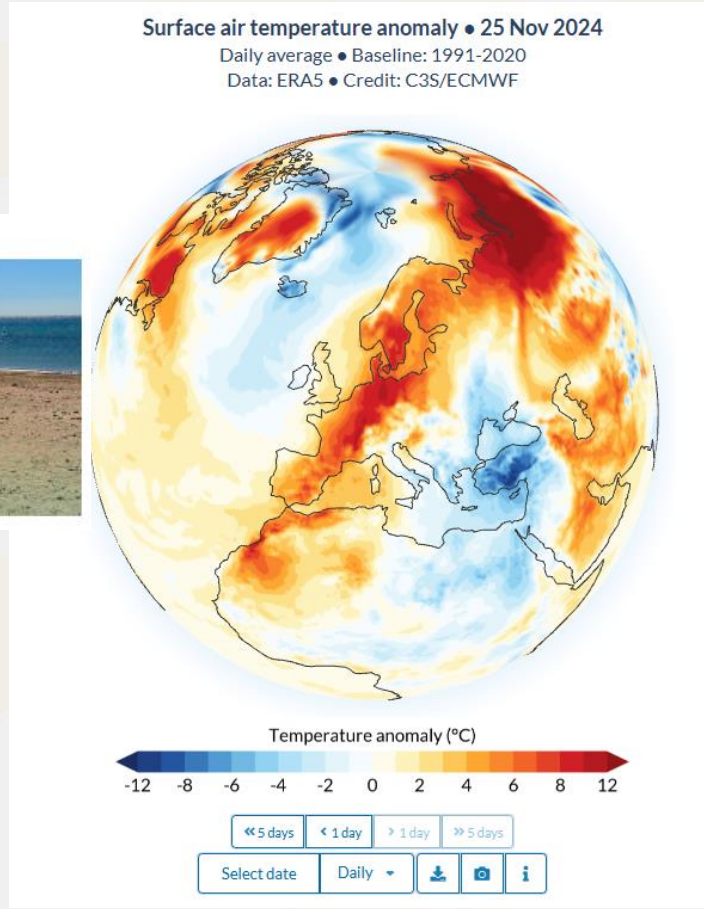
ENVIRONMENTAL

“natural capital” OR “agri-environmental” OR “agriculture” OR “circular resource use” OR “climate change effects” OR “renewable energy” OR “ecosystem services”

- Data Examples
  - Satellite data (Copernicus)
  - Land use / Land cover
  - Ecosystems (https://globalecosystemsatlas.org/)
  - LUCAS - Eurostat
  - Biodiversity
  - WCMC – global protected areas
- Methods
  - AI, ML, Deep Learning
  - LLMs, VLMs (SatCLIP, CLIP)





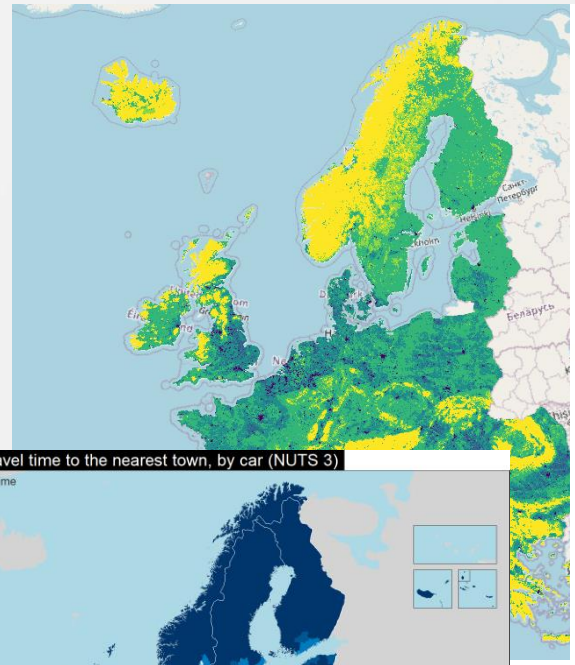


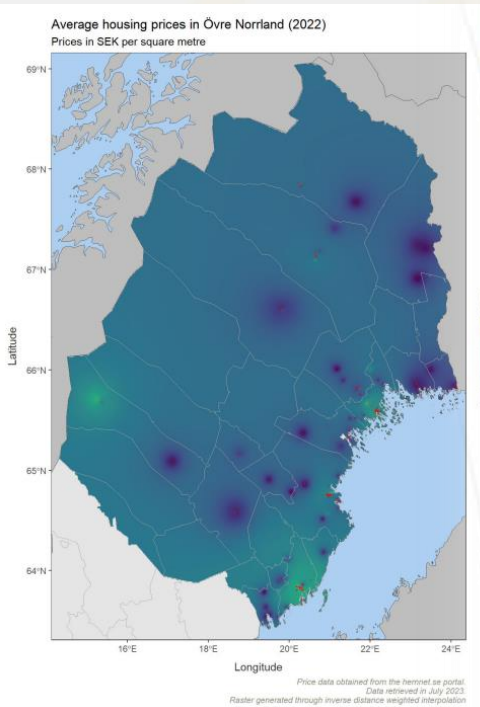
RESIDENTIAL

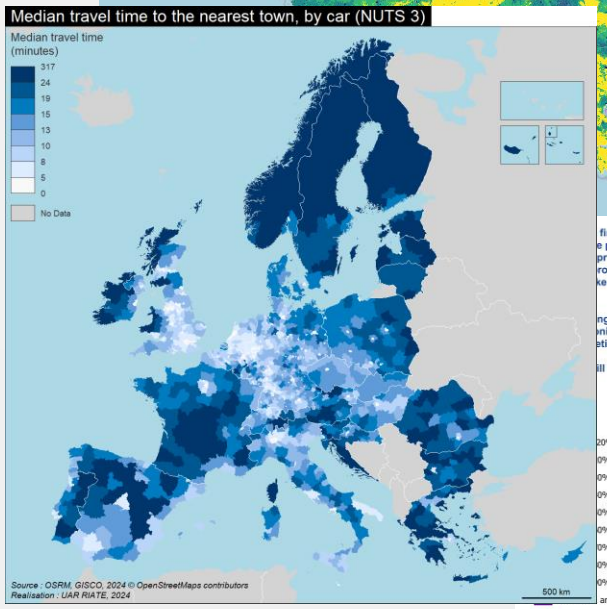
housing* OR “demographic trends” OR commuting OR migration OR “basic services” OR “digital transformation” OR cohesion OR wellbeing OR “well-being”

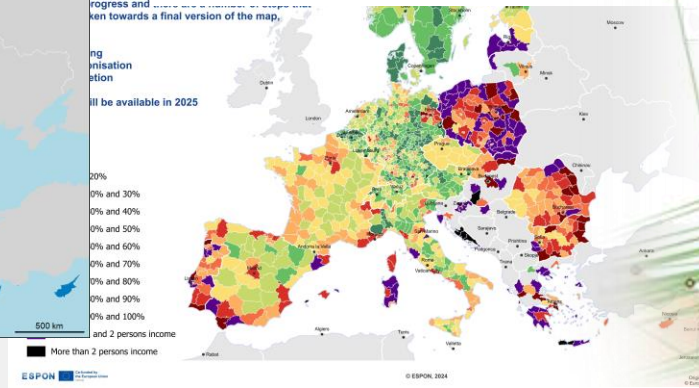
- Data examples
  - Housing prices
  - Rural attractiveness
  - Mobility (Facebook)
  - Accessibility (OSRM)

- Methods:
  - Web-scraping
  - OSM routing
  - Econometric models
  - Geo-statistics

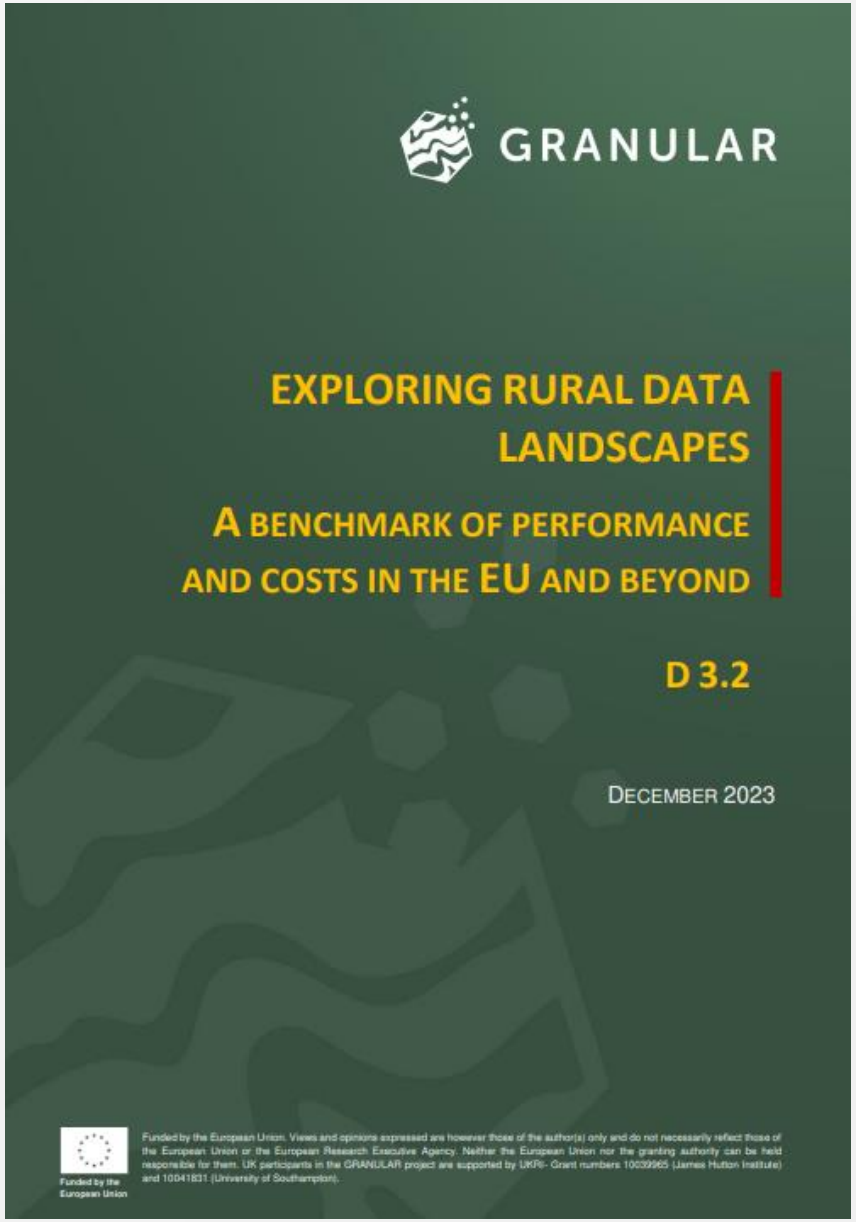








TASK 3.2 EXPLORE OPTIONS, PERFORMANCE AND COST (M06-15)



3 key inspirational elements / sections

1) The costs of and for rural data – a journey across national statistical offices and authorities in the EU and beyond
2) 9 Inspirational examples from Granular partners – Data, tools and approaches
3) 27 Data fiches

Deliverable - Submitted in DEC 2023

OVERVIEW OF NATIONAL DATABASES OPENLY AVAILABLE, CONTAINING DATA ABOUT RURAL DEVELOPMENT ISSUES

Country	Main database with rural-development relevant data	Data domains in open databases	Finest resolution of data free of charge
Bulgaria	The IS Infostat platform (https://infostat.nsi.bg) publishes data relevant for rural development. It includes business, demographic social, macroeconomic, environment energy and multi-domain statistics. Data on population and housing census is available free of charge. Paid databases provide users access to more detailed data at lower levels after disaggregation. Grid data for population (total, age groups and sex) for 2011, 2021 is free of charge.	Demography, energy and health.	Grid data for population: Total population, Age groups 0-14; 15-64 ;65+ and sex, for 2011 and 2021.
Croatia	Several databases are available but not strictly connected with rural development issues. The so-called PC-Axis databases are available at https://web.dzs.hr/PX-Web_e.asp?url=%22Eng/Archive/stat_databases.htm%22. Some data available is at municipality level. Grid 1000 is accessible, free of charge, at the GeoSTAT - Web GIS portal of the Croatian Bureau of Statistics (https://geostat.dzs.hr). The available grid-level data is on population – number of populations, number of population by large age groups, population by educational attainment, population by activity, business register (active business entities). Tourism data on accommodation capacities and tourist arrivals and nights. The PC-Axis databases https://web.dzs.hr) has some data by municipalities. Some census data by settlements is available at https://podaci.dzs.hr/en/statistics-in-line/. All data published online is free of charge. If special data processing is needed for grid-level data, it is charged according to the subject's hourly rate.	Agriculture, demography, economy, energy, environment, tourism / recreation, transport.	1km grid. Available data free of charge at the GeoSTAT - Web GIS portal for the Croatian Bureau of Statistics https://geostat.dzs.hr/?lang=en.
Cyprus	The main database by Statistics Cyprus is CYSTAT-DB, available at https://cystatdb.cystat.gov.cy/pxweb/en/8.CYSTAT-DB/. It contains data on agriculture, livestock, fishing, business register, construction, education, energy, environment, external trade, health, industry, information society, innovation, labor market, living conditions, social protection, national accounts, population, price indices, public finance, research and development, services, tourism and trade.	Demography, Health, Education, ICT Usage.	1000m for grid-level data.
Finland	The main database published by Statistics Finland is called StatFin (https://www.stat.fi/tup/statfin/index_en.html). StatFin is freely accessible and includes data on population, economy, housing, transport, tourism, consumption, prices, wages and salaries, energy, enterprises etc. The Paavo database (https://www.stat.fi/tup/paavo/index_en.html) contains data by postal code area on the population structure, education, income, housing, workplaces, households; life stage etc. There is also a grid-level database with grid sizes of 250 m x 250 m, 1 km x 1 km and 5 km x 5 km (https://www.stat.fi/tup/ruututietokanta). The grids cover the whole of Finland, but this is not for free and charged based on number of licenses and grid size. Data is available on the areas; population structure, level of education, income of inhabitants and households, size and stage in life of households, buildings and dwellings, workplaces, and main activities of inhabitants. Population structure grid data 1km x 1km is free of charge.	Agriculture, demography, economy, energy, environment, housing, infrastructure, mobility, tourism / recreation, transport.	Population structure grid data 1km x 1km.

9 INSPIRATIONAL EXAMPLES / GOOD PRACTICES

GRANULAR

France: M
scale – Ca
Orientales

Louise Chasse
Claude Balagu

COSTS

BUDGET OF THE WHOLE INITIA

Galicia – Spain: Tele
rural areas of Ouren

María Isabel Doval Ruiz & Bre

COSTS

The extracted data

Scotlan
results

Jonathan H

COSTS

Nordi

Mats Stj

COSTS

Pall

COSTS

Data Collection

Global: Geo-Wiki Earth Observation & Citizen Science

Ivelina Georgieva (IIASA)

The Geo-Wiki platform provides anyone with the means to engage in monitoring of the Earth's surface by classifying satellite, drone or ground-level imagery. Since its inception, Geo-Wiki has grown rapidly, with currently over 22,000 registered users having contributed more than 18 million image classifications from around the world.

type of geography for which there was limited data available.

27 DATA FICHES

GRANULAR

DATA FICHE

Human powered mobility by Strava Metro

https://metro.strava.com/

Ian McCallum & Martin Hofer (IIASA)

INDICATOR CLASS: Residential

Data

Resolution: Class: vectors
Type: Shapefiles
License: https://metro.strava.com

Spatial

Resolution: NA
Extent: Global
Granularity: street-level
Coordinate System: WGS 1984 (EPSG: 4326)

Temporal

Resolution: Annual
Extent: 2015 - present
Frequency: Updated monthly by users who agree to opt in to making their data public

DESCRIPTION

The **Strava** dataset is the **largest collection of human-powered transport information** in the world. Metro aggregates, de-identifies and contextualizes this dataset to help make communities better for anyone on foot or a bike.

Strava works with urban planners, trail networks, city governments and safe-infrastructure advocates to understand mobility patterns, identify investment opportunities and evaluate the impact of infrastructure changes free of charge.

1

USER COSTS

Free & open Access

Since 2020, Metro is free to any organization that shares the mission to make cities better for cyclists and pedestrians. COVID-19 has accelerated the world's need for Metro – for smarter and more sustainable design of our cities, and for giving back to the communities that support millions of Strava athletes around the world. If you believe in human-powered transport and think you can make an impact, apply now. If your application is successful, you will be granted free access.

DATA INFRASTRUCTURE

Information not available.

DATA GOVERNANCE & MANAGEMENT

Staff costs: Using the online dashboard and UI would mean costs would be low, and anyone with basic internet skills could query and visualize the data.

Data analysis needed: If a user chooses to download the data, e.g. Shapefiles, this will then require GIS skills to work with and further analyze the data. This should however be fairly basic for any GIS analyst. Onin total would likely be enough to develop basic tools to work with and process the data.

Quality assurance: Strava is performing QA on their end, hence little effort is required for this from the user.

DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE

https://metro.strava.com/ https://www.strava.com/

How to cite

Dataset: https://metro.strava.com

Co-funded by the European Union

UK Research and Innovation

2

Innovation & use of open data advocated in D3.2.

Inspirational examples => ongoing investments / work with new & open datasets => foster innovation / maximize societal & economic benefits of open data.

Demand for complex & granular data, in line with the Directive's focus on providing valuable information tailored to user needs.

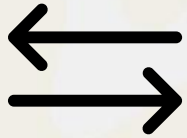
Collaboration, knowledge sharing & peer-to-peer learning among public sector bodies to enhance data availability and reuse. *(Directive's aim to minimize costs for reusing public sector information).*

Capacity building & training in local territories & public sector bodies

Quality assessment & validation to ensure reliable & usable data for decision-making (*Directive's encouragement of high-quality public sector information*).

Sustainability, accessibility => need for capacity-building programs and streamlined data access processes *(Directive's aim of ensuring sustainable and accessible public sector information).*

Cost dimensions: good open data available! YET for utilization, personnel costs, quality assessments, and data management needed



SUGGESTIONS FOR GRANULAR

- **Continue working with presented data & inspirational examples**
- **Living & Replication Labs** may a**nalyze presented data** types and survey results for **potential replication or further exploration**.
- **Explore datasets like the Strava dataset**, possibly through applying for free access, to enhance understanding.
- Reflect on l**ocal needs and circumstances, and engage with authors and contacts** for further discussions.
- **Align actions with the Directive's goals** of improving data availability, promoting collaborative efforts, and tailoring practices to local conditions.