D5.2
Interrelations among different types of citizens in different geographic contexts

Prepared by: GESIS

Contributions from: UNIBO, TEC

Deliverable nature	Report
Dissemination level (Confidentiality)	Public (PU)
Delivery date	2022-07-29
Version	1.0
Total number of pages	52
Keywords	GIS, energy citizenship, spatial analysis
Cite as	Abel, D., Jünger, S., Lieth, J., Massari, M., Boulanger, S.O.M. (2022). Interrelations among different types of citizens in different geographic contexts. D5.2 of the Horizon 2020 project GRETA, EC grant agreement n°101022317, Cologne, Germany.
Project contact	Helinä Melkas, email: lut.Greta@lut.fi; Dennis Abel, email: dennis.abel@gesis.org; Stefan Jünger, email: stefan.juenger@gesis.org

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101022317.

Disclaimer and acknowledgement

The views expressed in this document are the sole responsibility of the authors and do not necessarily reflect the views or position of the European Commission or the European Climate, Infrastructure and Environment Executive Agency. Neither the authors nor the Agency nor the GRETA consortium are responsible for the use which might be made of the information contained in here.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101022317.

Executive summary

The overall objective of WP5 is to investigate the relationship between geographical levels and the main factors influencing the emergence and growth of energy citizenship and to examine under what conditions energy citizenship contributes to broader decarbonization policy goals.

Task 5.2 aims to create a GIS-based analytical tool to study the emergence of energy citizenship. This task answers to the need to create a bridge between data collected from case studies, their re-use in scenario predictions (WP4), and their framing for the purpose of Community Transition Pathways (CTPs) (T5.3). The task creates a GIS-based tool to georeference and spatialize the collected data to achieve this objective. This task accounts for variations within and between geographical levels by studying their spillover, crossover, and cross-level interaction effects on the local, regional, national, and supranational levels. For this purpose, instruments such as the georeferenced survey data, and other auxiliary data sources, will be used as developed within WP4.

The deliverable is structured in five main sections: following the introduction, section two summarizes and synthesizes previous work on the geographical dimension of energy citizenship and deduces relevant theoretical considerations for the GIS-based tool; based on this theoretical work, the third section outlines the general purpose of the tool, the schedule and the links to other work packages within GRETA; the fourth section discusses the data integration and pre-processing steps which are necessary to integrate primary and secondary data into the tool; considering the concept of the tool and the identified data opportunities and limitations, section five reports the scoping of the technical implementation, and discusses spatially explicit analytical approaches and the feasibility for integration into the tool.

Project information

Grant agreement No.	101022317
Acronym	GRETA
Full title	GReen Energy Transition Actions
H2020 Topic	H2020-LC-SC3-2020-NZE-RES-CC
Project URL	www.projectgreta.eu

Document information

	Number	Title
Deliverable	D5.2	Interrelations among different types of citizens in different geographic contexts
Work package	WP5	Community Transition Pathways for decarbonization across geographical levels
Task	T5.2	Cross-level analysis of energy citizenship through a GIS-based tool

Delivery date	Contractual: M15, Actual: M15
Nature	☒Report ☐Other ☐ORDP
Dissemination level	☒Public ☐Confidential
Authors (partners)	Dennis Abel, Stefan Jünger, Jonas Lieth (all GESIS), Martina Massari, Saveria O.M. Boulanger (both UNIBO)
Reviewers (partners)	Martina Massari, Saveria O.M. Boulanger (both UNIBO)
Summary (for dissemination)	This report describes the development of a GIS-based analytical tool to study the emergence of energy citizenship. This task bridges the data collected from case studies and the multinational survey, their re-use in scenario predictions, and their framing for Community Transition Pathways. The task creates a GIS-based tool to georeference and spatialize the data collected to achieve this objective. This task accounts for variations within and between geographical levels by studying their spillover, crossover, and cross-level interaction effects on the local, regional, national, and supranational levels. For this purpose, instruments such as the georeferenced survey data, and other auxiliary data sources, will be used.
Keywords	GIS, energy citizenship, spatial analysis

Version	Date	Description
0.1	2022-06-27	First draft
0.2	2022-07-10	Second draft
0.3	2022-07-15	Finalized version sent to internal review
0.4	2022-07-28	Revised version after internal review
1.0	2022-07-29	Final version

Table of contents

List of figures

List of tables

Abbreviations and acronyms

CTP	Community Transition Pathways
DoA	Description of Action
ESS	European Social Survey
EU	European Union
GIS	Geographic Information System
ISSP	International Social Survey Programme
NLP	Natural Language Processing
NUTS	Nomenclature of Territorial Units for Statistics
OPCE	Open Portfolio for Civic Energy Empowerment
PV	Photovoltaic
Q	Quartal of a Year
STM	Structural Topic Model
T	Task in the GRETA Project
WP	Work Package in the GRETA Project

Glossary with key definitions used in the deliverable

Georeferencing	Assignment of geocoordinates to data
Geocoding	Conversion of indirect spatial references into geocoordinates
GIS	Framework or software solutions to process, analyze, and visualize geospatial data

1 Introduction

1.1 Description of the deliverable

The call under which GRETA was funded — LC-SC3-CC-1-2018-2019-2020: Social Sciences and Humanities (SSH) aspects of the Clean-Energy Transition — mentioned the necessity "to understand in what kind of environments collaborative goal setting and commitment can take place, how relevant decisions are made and any trade-offs between competing goals are addressed" to answer to the question "Is energy citizenship more likely to emerge locally, or at regional, national or supranational levels? For what reasons?". The DoA describes Task 5.2 role in creating a GIS-based analytical tool to study the emergence of energy citizenship. This task answers to the need to create a bridge between data collected from case studies, their re-use in scenario predictions (WP4), and their framing for the purpose of Community Transition Pathways (CTPs) (T5.3). The task creates a GIS-based tool to georeference and spatialize the data collected to achieve this objective. This task accounts for variations within and between geographical levels by studying their spillover, crossover, and cross-level interaction effects on the local, regional, national, and supranational levels. For this purpose, instruments such as the georeferenced survey data, and other auxiliary data sources, will be used as developed within WP4.

The GIS-based tool facilitates the geographic assessment of energy citizenship emergence. Technically, this goal is achieved by linking several data sources based on geographic locations and identifiers for a joined analysis. For example, such an integrated approach allows the following examinations:

- Solar radiation potential in an area and the willingness of citizens to invest in solar panels;
- Natural disasters (e.g., heat waves, floods) as causes of climate change and people's change in attitudes toward renewable energies and climate change mitigation measures;
- The role of geographically unevenly distributed infrastructures like charging stations for the adoption and diffusion of electric cars.

1.2 Scope of the deliverable

Deliverable 5.2 directly results from the work performed in Task 5.2. This document provides a concept that defines the theoretical, analytical, and technical foundations for the GIS-based tool. The main goal of Task 5.2 is to offer a solution to report project output in an interactive and user-friendly way. The document summarizes the results from the tool's initial planning and scoping process and outlines the subsequent production steps until April 2023.

This deliverable follows the theoretical framework outlined in T5.1, where the definition of geographical levels of energy citizenship emergence is provided, and its determinants, barriers, and drivers are explicit. It is also linked with WP3, which deals with data gathering and background studies on the six case studies and the multinational survey. Significant input for the tool will be provided from the case studies, the quantitative modeling tasks in WP4, and in particular, T4.6, which provides the analytical basis for the tool.

1.3 Approach and output

Task 5.2 consists of several sub-tasks with different methodological approaches and outputs.

The following section on the spatial determinants of energy citizenship is based on a mixed-methods literature review in conjunction with T4.6 and previous deliverables of the project (specifically of T5.1).

Extensive desk research was performed to identify the data input for the project and the relevant analytical approaches. The identified datasets were categorized according to a pre-defined analytical framework and structured accordingly. The resulting database was saved in an excel-file and will be deposited in the project repository on Zenodo (see below).

The technical implementation of the tool was developed based on standard development guidelines for data tools (Fay et al., 2022; Sievert, 2020; Wickham, 2021). The task focuses on the tool's programming, which is performed in the programming language and software environment "R". The code for this tool will also be stored in the project repository on Zenodo.

We support open and reproducible research. All primary research steps (computational literature review, desk research on existing data sources, and tool programming) were documented. GRETA established the Open Portfolio for Civic Energy Empowerment (OPCE) as an Open Access repository on Zenodo (https://zenodo.org/communities/greta). The entire tool script and relevant anonymized and aggregated datasets will be stored in this repository.

2 Energy citizenship emergence as a spatial process

2.1 Geospatial perspectives in social science energy research

D5.2 addresses the role of spatial perspectives on energy citizenship emergence and explores the relevance of GIS-based approaches for understanding these dynamics. This project speaks to a current social science energy research trend, which has been described as a "spatial adventure" over the last three decades (Bridge, 2018; Castán Broto & Baker, 2018). Rather than just describing the geographical variation of technologies and policies, this spatial turn is characterized by a growing number of social science research that recognize energy systems' complex socio-spatial configurations. As Bridge (2018, p. 11) puts it, taking "space seriously in social science energy research leads researchers to ask different questions about energy systems, and admits alternative sites, actors and practices as legitimate objects of research. In this way, thinking about space can bring into view the analytical limits (and social consequences) of more conventional frameworks that treat space as an unproblematic substrate on which technical, economic and/or political action unfolds."

Abel et al. (2022) have systematically reviewed this "spatial turn" as part of task 4.6 of GRETA. The review shows that within this spatially-explicit research, current topics related to 1.) energy transition issues, including 2.) public attitudes and acceptance, 3.) innovation networks, the diffusion of technologies, and 4.) issues of energy access, poverty and justice are becoming increasingly salient. Figure 1 shows the shares of these topics among a large corpus of research articles[1] on energy issues and spatial perspectives.

[1] The study identified 7879 relevant studies out of a comprehensive corpus of 9232 research articles between 1980 and 2022 and employed natural language processing (NLP) in order to classify topics and geolocations. This subsection uses the methodology and data from this publication of Abel et al. (2022) and complements it with a specific perspective on energy citizenship and community energy. For the full research design and methodology, see Abel et al. (2022). The limitations of this approach should be noted: 1.) The sample of journal articles has been selected from one web catalogue; 2.) The sample is restricted to journal articles and excludes, for example, books; and 3.) The sample includes only articles written in English.

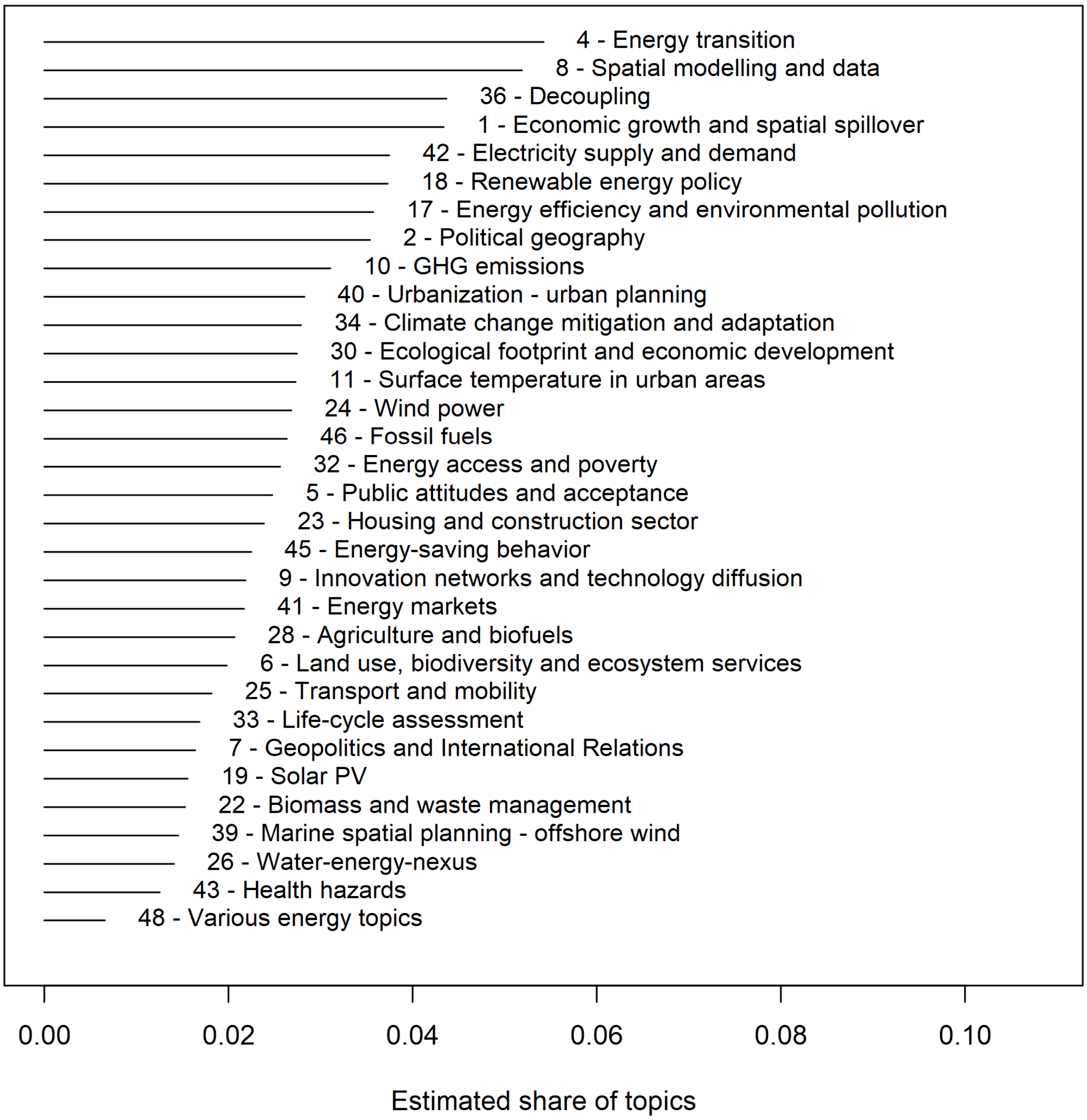


Figure 1: Relative shares of topics in a corpus of research articles between 1980 and 2022

Source: Abel, Lieth and Jünger (2022).

By detecting the specific locations of places referred to in these articles, it can be shown that most case studies are situated in Western Europe and the UK (see Figure 2). Eastern and Southern European countries are less likely to be featured in the articles.[2] Scandinavian countries feature surprisingly little, although often depicted as sustainability pioneers in Europe. The global distribution of research topics among

[2] One confounding reason might be that only articles written in English were considered in the collection. At the same time, UK is the only native English-speaking country in the sample which makes this circumstance as the sole cause of the findings rather unlikely.

geographical areas is also highly diverse. European case studies are strongly associated with research themes on the energy transition, innovation networks, technology diffusion, and renewable energy policy. In contrast, aspects of energy access and poverty, energy efficiency, and public attitudes and acceptance are relatively less covered in Europe compared to other world regions.

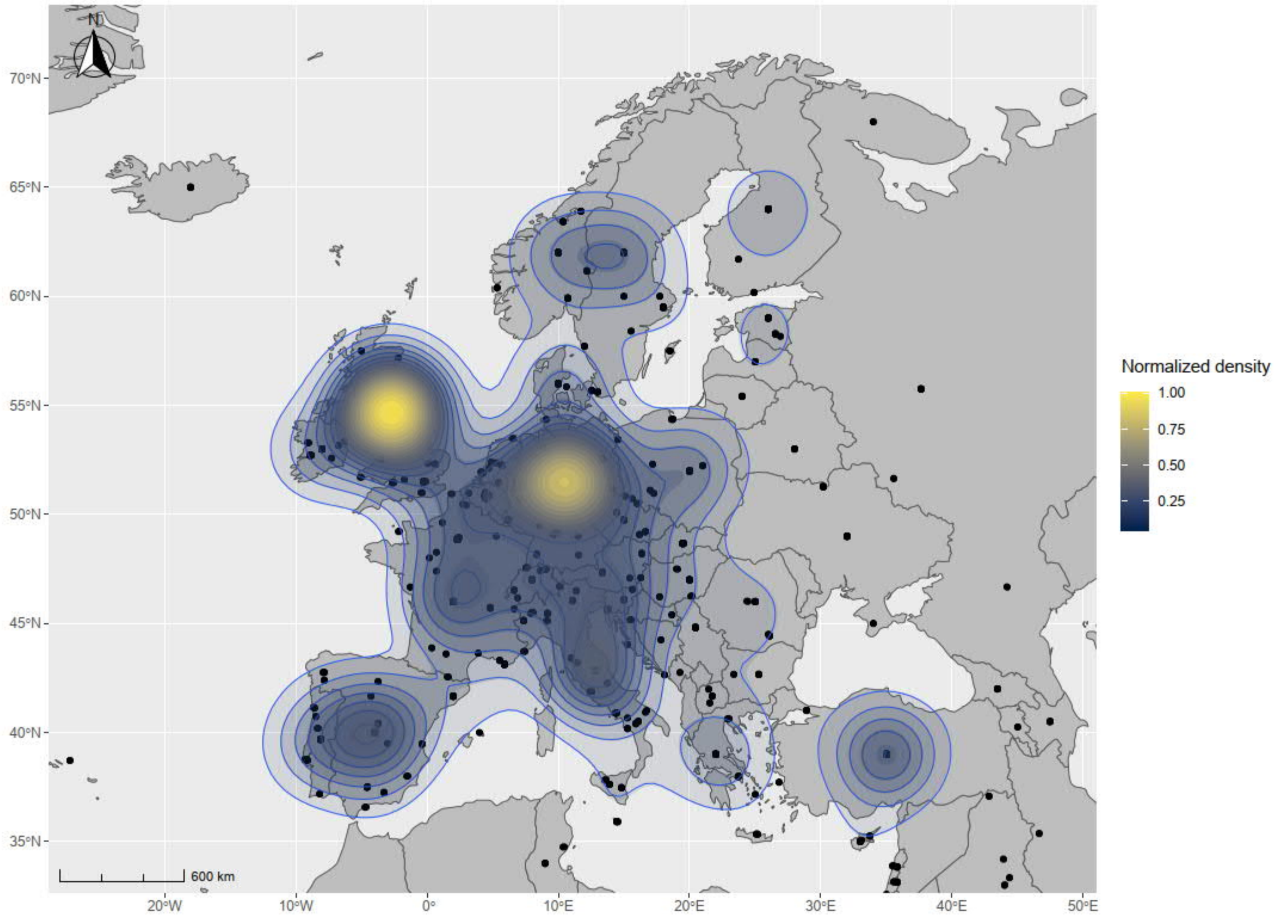


Figure 2: Geographical distribution of geolocated articles in Europe
Source: Abel, Lieth and Jünger (2022).

We complement the findings from Abel, Lieth, and Jünger (2022) by exploring the corpus regarding links to energy citizenship and community energy. Zooming into specific considerations of these themes in the corpus, the share of articles with references to energy citizenship, community energy, energy cooperatives, and general participation is relatively low. Based on a community-energy-dictionary[3], we identified 222 out of 7879 articles with such a thematic orientation (2.82%). Figure 3 shows the distribution over time. Although it is evident that the community theme underwent a steady increase over time, it becomes clear that the "spatial turn" in social science energy research has been largely unaware of energy citizenship and community energy aspects.

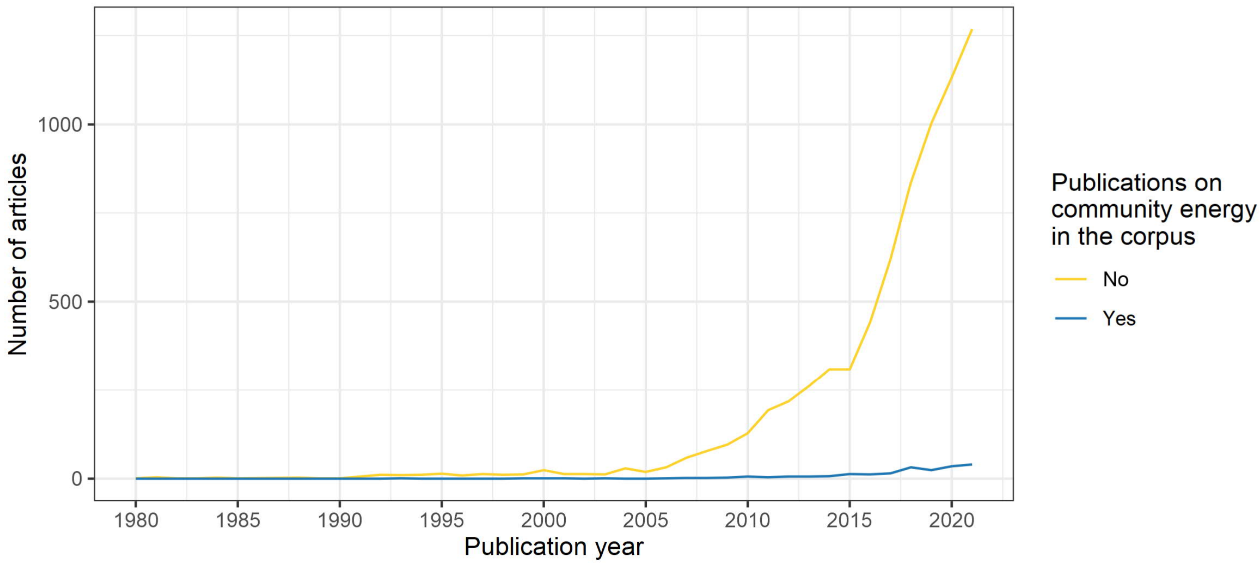
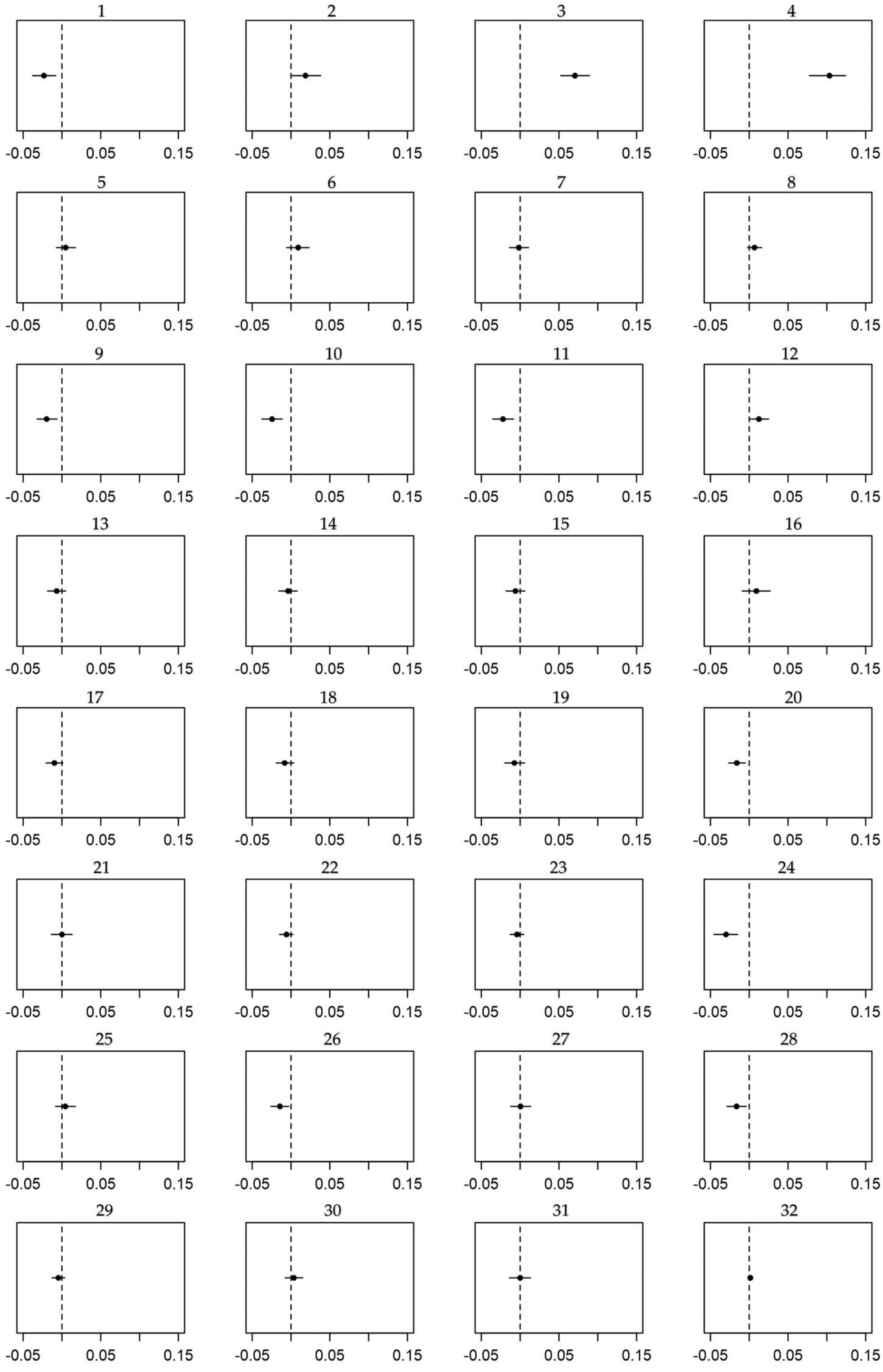


Figure 3: Community energy articles in the corpus

Focussing on these 222 studies on the cross-section between spatial analysis and energy citizenship and community energy, we then explored the relationship to the 32 general topics in the corpus. To calculate the effect of citizenship and community themes, we included a binary indicator for this in a structural topic model (STM). The output of this effect is shown in Figure 4. We find clear connections to topics regarding "political geography", "energy transition", "public attitudes and acceptance" and "renewable energy policy".

[3] We constructed a rather strict dictionary only including words which are highly likely connected to the topics of energy citizenship and community energy. This biases the analysis towards false negatives but reduces the risk of false positives. Articles which had at least one word in the title or abstract from the dictionary were coded as “yes”. The dictionary included the terms: "energy communit*", "communit* energy", "energy cooperative*", "energy citizen*", "citizen energy", "citizen engagement", "citizen ownership", "participation" and "participatory".



1 - Economic growth and spatial spillover, 2 - Political geography, 3 - Energy transition, 4 - Public attitudes and acceptance, 5 - Land use, biodiversity and ecosystem services, 6 - Geopolitics and International Relations, 7 - Spatial modelling and data, 8 - Innovation networks and technology diffusion, 9 - GHG emissions, 10 - Surface temperature in urban areas, 11 - Energy efficiency and environmental pollution, 12 - Renewable energy policy, 13 - Solar PV, 14 - Biomass and waste management, 15 - Housing and construction sector, 16 - Wind power, 17 - Transport and mobility, 18 - Water-energy-nexus, 19 - Agriculture and biofuels, 20 - Ecological footprint and economic development, 21 - Energy access and poverty, 22 - Life-cycle assessment, 23 - Climate change mitigation and adaptation, 24 - Decoupling, 25 - Marine spatial planning - offshore wind, 26 - Urbanization - urban planning, 27 - Energy markets, 28 - Electricity supply and demand, 29 - Health hazards, 30 - Energy-saving behavior, 31 - Fossil fuels, 32 - Various energy topics

Figure 4: Effect of community-covariate in STM models

Finally, we were interested in outlining which research disciplines within the social sciences addressed citizenship and community themes from a spatial perspective and which methodological approaches, including GIS-based approaches, were applied. Accessing corpus metadata and utilizing the methodology-dictionaries from Abel, Lieth, and Jünger (2022), we calculated a logistic regression on the community theme (Table 1). The output reveals that spatial perspectives on energy citizenship and community energy can be mainly attributed to the disciplines 1.) Geography, 2.) Environmental Sciences & Ecology, and 3.) Science & Technology studies. Generally, we can conclude that a highly inter- and transdisciplinary research community has addressed the energy community theme. Concerning methodological considerations, it also becomes evident that quantitative approaches are underrepresented in this theme, and no apparent connection to GIS-based applications is visible.

Table 1: Relationship between community theme and article characteristics in corpus (heteroskedasticity-robust standard errors in parentheses)

	Dependent variable:
	Community (Logit model)
Research areas (reference: "Other")	
Environmental Sciences & Ecology	0.529**
	(0.203)
Business & Economics	-0.084
	(0.263)
Energy & Fuels	0.081
	(0.333)
Geography	0.683**
	(0.270)
Science & Technology - Other	0.461*
	(0.212)
Quantitative methodology	-0.613***
	(0.162)
GIS application	0.179
	(0.284)
Constant	-3.672***
	(0.171)
Observations	7,879
Log Likelihood	-994.584
Akaike Inf. Crit.	2,005.168
Note:	*p<0.1; **p<0.05; ***p<0.01

Summarizing these findings, we draw two main conclusions:

1. Social science energy research has been increasingly addressed through a spatial "lens", and this development is strongly connected to topics related to energy

transition issues. The geographical distribution and topical coverage, however, are unevenly distributed. In contrast to other regions, crucial topics such as energy justice and poverty, and public acceptance are underrepresented in Europe.

2. Generally, there is a research gap for energy citizenship issues within spatial perspectives. In particular, we have identified a large gap for mixed-methods GIScience, which integrates qualitative and quantitative data (Adelfio et al., 2019; Muenchow et al., 2019; Talen, 2000).

T5.1 has addressed the need to systematically conceptualize the geographical levels of energy citizenship emergence. In the next subsection, we synthesize the output from T5.1 and connect it to the development of the GIS-based tool.

2.2 Space, place, and energy citizenship: key findings from T5.1

The taxonomy of geographical levels and drivers for energy citizenship emergence (D5.1; Massari et al., 2022) provides a framework for defining the emerging characteristics of energy citizenship at different geographical levels: local, regional, national, supranational, and the transversal virtual level. These levels answer to "where" energy citizenship establishes and support the understanding of motivations behind the engagement of citizens in energy processes and the context of operation. The interaction with the context allows citizens to choose to engage in (different forms of) energy actions. In this regard, the geographical dimension of proximity with several domains and dimensions is not only meant as physical contiguity. Still, it is intended as "closeness", elements that enable links and stimulate collaboration networks thanks to daily (or sporadic) contact and exchange. The lens of proximity entails several domains, such as:

- The spatial domain describes the physical proximity between people (to each other) and people and the sources of energy (mainly related to availability). It also includes the interdependency between the municipality, its underlying (infra)structures, and citizens. Moreover, it provides a reflection on the natural environment and ecosystem of a specific context which can include different types of energy production resources and the climatic dimension.
- The policy and governance domain defines management and planning measures (along with more general measures such as taxation and technological innovation) taken at the institutional level.
- The social domain identifies the relationalities connecting the members of local communities considered energy citizens and the proximity between people and institutions. These relational ties frame the modalities within which energy citizenship forms and grows. They can be intentional, issue-oriented, purpose-, project- or practice-led.
- The technological domain entails the proximity with technologies and digital solutions to involve people in setting up an energy community, informing with data

on better energy behaviors, enabling small-scale smart solutions, and capitalizing on investments in renewables.
- The economic domain, with the market, sustainability, and economic factors, expresses the convenience conditions for the energy system, utilities, and end-users, supported by the proximity among producers, distributors, and consumers.

The proximity domains characterize the geographical levels and play essential roles in triggering the birth of different forms of energy citizenship but mostly in supporting citizens' engagement to grow and proceed in their pathway. Through the analysis of the GRETA case studies read through the lenses of space and engagement dimensions of energy citizens, the deliverable D5.1 was able to draft some categorization of the different geographical levels as follows.

This taxonomy highlighted five geographical levels where energy citizenship might emerge, be recognized, and act. The local level allows for spatial and social proximity based on daily contact and closeness of interests (including economic). It is the level that allows one to see closely and make visible the energy, therefore permitting better literacy on the topic. One risk, however, relates to the lack of diversity in an overly narrow local community, a variety that allows for openness to alternative directions, new ideas, and a window into more significant trends and urgencies. This advantage seems to be enabled by the regional level, which can intermediate between levels and is characterized as facilitating policy frameworks that are more strategic, broad, and prescriptive, allowing for incursions of bottom-up experiences. Consequently, the regional level seems to facilitate the cooperation of multiple actors around a theme and better control even in distributing resources at the territorial level. The regional level dialogues closely with the national level, a dimension in which the various proximity domains converge in a framework predominantly related to policies. Here, policymakers' commitment appears fundamental in creating an enabling context for energy citizenship. However, the political commitment of the national level necessarily incorporates the orientations and agendas of the supranational level. It also contributes to building an environment conducive to the emergence of energy citizenship through networking, mentoring, and advocacy with various actors. This environment is only effective if it includes all categories of actors and energy citizens, even the most vulnerable, geographically distant, or politically difficult to reach and recruit. Finally, the geographical layers are transversally supported by the virtual level, the layer that is most capable of voicing needs and trends. It provides a (public) space for education and training where citizenship can be upscaled, and citizens can access funding mechanisms, foster replication, and bridge individual gaps.

The taxonomy draws a multi-layered structure of relations and interdependencies that need to be considered while approaching the topic of energy citizenship. It suggests looking at energy citizenship from a local-global networking perspective. This is useful to connect leveraging mechanisms, provide solutions to barriers and bottlenecks, and contribute to better policy recommendations.

3 Purpose and schedule of the GIS-based tool

The DoA describes Task 5.2 role in creating a GIS-based analytical tool to study the emergence of energy citizenship. This task answers to the need to create a bridge between data collected from case studies, their re-use in scenario predictions (WP4), and their framing for the purpose of Community Transition Pathways (CTPs) (T5.3). The task creates a GIS-based tool to georeference and spatialize the data collected to achieve this objective. This task accounts for variations within and between geographical levels by studying their spillover, crossover, and cross-level interaction effects on the local, regional, national, and supranational levels. For this purpose, instruments such as the georeferenced survey data, and other auxiliary data sources, will be used as developed within WP4. In light of this definition of objectives, we deduce detailed targets and functionality principles:

Targets

1. Georeferencing and spatializing of major project data output;
2. Aggregating individual-level data on energy attitudes and behaviors onto higher-level spatial resolution (municipalities, regions, countries);
3. Linking of this data with context factors that explain drivers and barriers to energy transition actions: socio-economic factors, demographics, and geophysical indicators;
4. Enhancing the database by accessing and integrating secondary datasets;
5. Visualization of geographical patterns and mapping of spatial interdependencies, clusters, and networks;
6. Coverage of proximity domains identified in T5.1: geospatial, policy, social, technical, and economical.

Functionality

1. High share of visualization instead of numeric data representation;
2. Interactive usability of the tool, which includes the ability to adjust major input factors and scales;
3. Inclusive representation of project output based on case studies, the multinational survey, and modeling output from WP4;
4. Possibility to continuously update project data and allow visualization through different timespans.

Production schedule

T5.2 is scheduled from August 2021 to April 2023. Until the deliverable of D5.2 in July 2022, the project primarily covered the initial planning and scoping (see Figure 5). Two internal consortium meetings were held in November 2021 and February 2022 to elaborate on the purpose and functionality, the timeline, the definition of the end users, the definition of the content structure, the design of the user interface (UI), and the screening of the data requirements and links to other work packages. The scoping process has been finalized, although continuous feedback from all later stages is expected to reconsider some aspects of the scoping process.

The development of the tool is scheduled for Q3 and Q4 2022. This development will be complemented with a project workshop in Q4 to discuss a first prototype and the possibilities of integrating the data output from WP3 and WP4. Q1 2023 will be concerned with the testing and deployment of the tool and includes a second workshop which will include a user journeys mapping. A final project meeting in Q2 2023 will be held to discuss long-term storage issues and the continuous usage after the project end. Sections 4 and 5 discuss major outcomes from the scoping process concerning the data input and technical implementation.

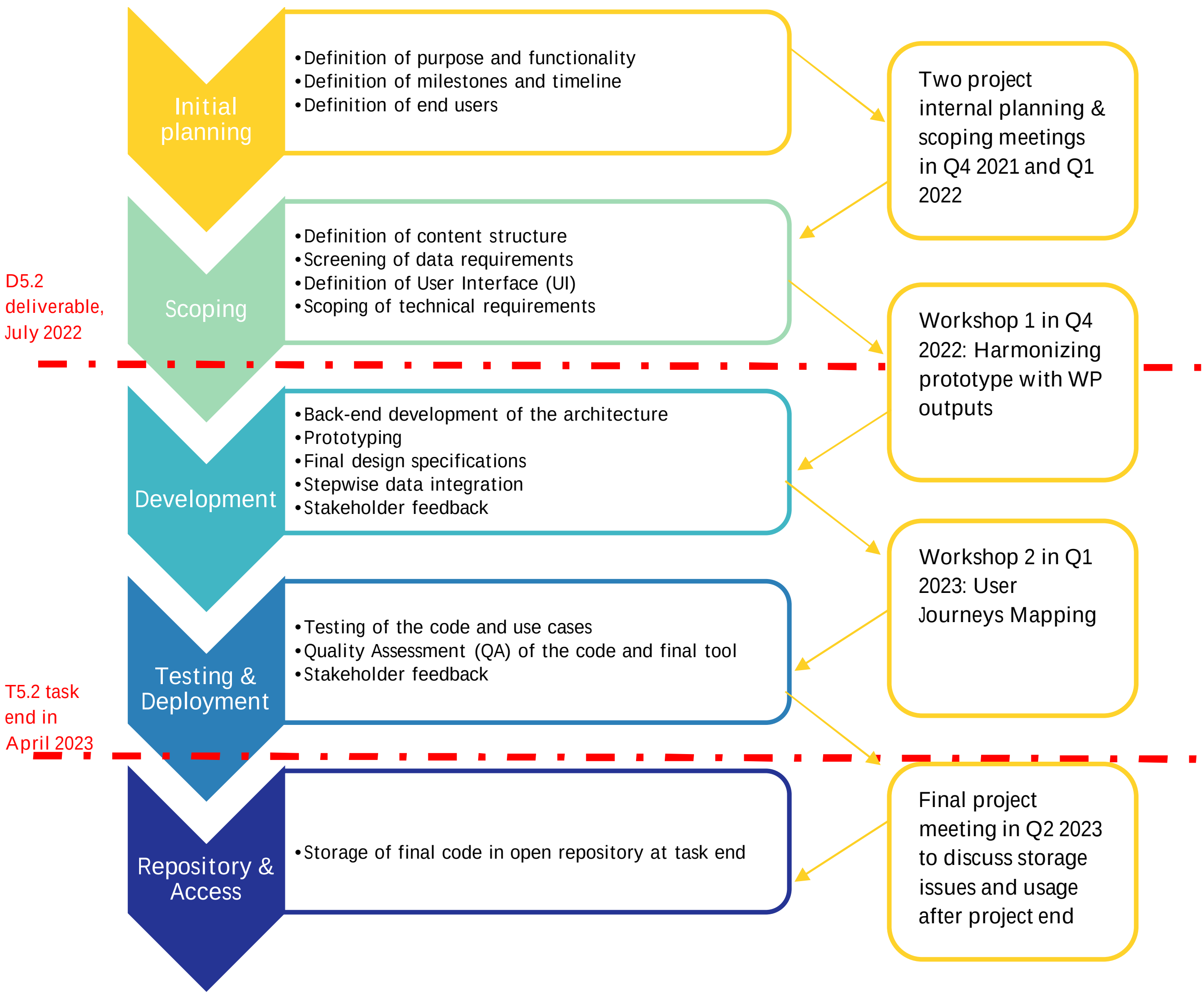


Figure 5: Production process

Links with other tasks within the GRETA project

The GIS-based tool relies on other tasks within GRETA for theoretical and data input and informs several tasks in WP5 and WP6 (see Figure 6). Regarding conceptual issues, this deliverable follows the theoretical framework outlined in WP1, where the definition of energy citizenship is provided, and its determinants, barriers, and drivers are made explicit (T1.1). In particular, T1.4 delivers input for the alignment of this framework with GIS modeling. Furthermore, the taxonomy of levels for energy citizenship emergence (T5.1) has laid the theoretical groundwork for T5.2. T5.1 developed a proximity domain and geographical levels matrix, which directly fed into the work of T5.2 and has been operationalized accordingly. The primary data input will be generated through the multinational survey and the case studies (T3.3 and T3.4). WP4 uses this data for several modeling purposes at local, regional, national, and supranational scales. The output from these tasks in WP4 will be fed to the GIS-tool. T4.6 is strongly linked to T5.2 as it represents an interface to selecting, harmonizing, and spatializing data from the case studies, the multinational survey, and the modeling exercises for usage in T5.2. Whereas T4.6 is responsible for the analytical parts, T5.2 is accountable for the technical implementation and transfer of the results into a user-friendly tool.

Considering the output of T5.2, it will directly inform the development of energy citizenship contracts (T5.3), providing a platform for monitoring their development and implementation in the case studies and assessing community pathways across geographical levels (T5.4). Through this connection, T5.2 also informs the synthesis of research results and policy recommendations in WP6. Considering the relevance of T5.2 to inform policymaking and stakeholders, this influences the selection of end users for the tool. Next to the research community, we expect a heterogeneous user group of various stakeholders. This requires user-specific data preparation and calls for a user-friendly design.

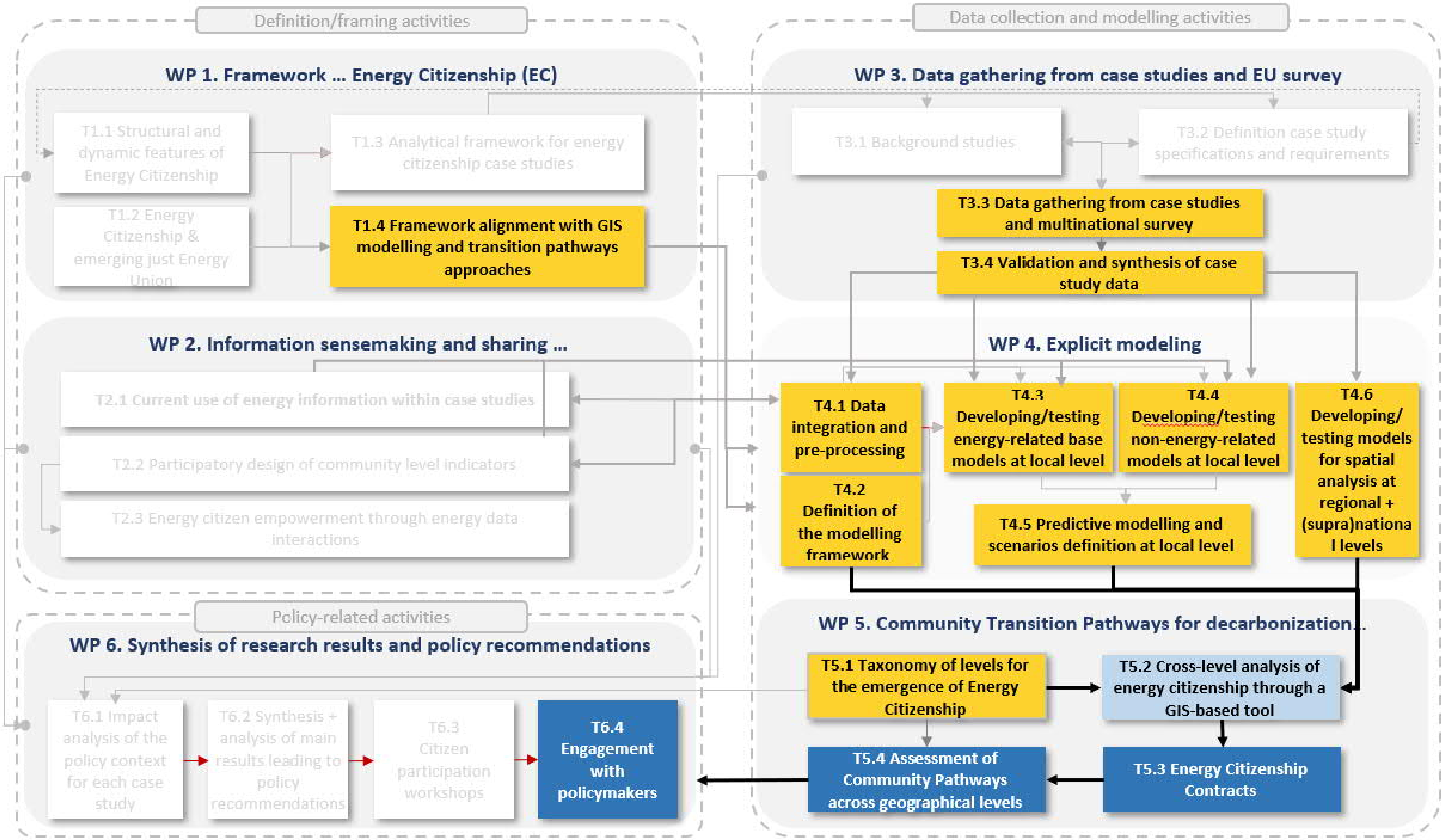


Figure 6: Role of T5.2 within GRETA WP structure (yellow – input, blue – output)

4 Data integration and pre-processing

4.1 Data integration and augmentation process

The GRETA project structure is highly diverse and includes qualitative output from six case studies, survey data from up to 10,000 respondents, and simulation results from WP4. This heterogeneous data structure represents a challenge for integration and augmentation processes. To prepare and comprehensively process this data, we follow the "GISualization" framework proposed by Adelfio et al. (2019). The multi-layer framework was developed to adaptively and iteratively manage complex data integration tasks in mixed-method approaches for "knowledge production in urban transformation processes" (p. 163). It is also described as a "collaborative communication platform" that supports the navigation between various data formats and methodological approaches and is therefore highly suitable for transdisciplinary projects like GRETA. It is furthermore suitable for stakeholder involvement and community participation. GIS-based visualization represents the core of the framework but is integrated with other additional methodological approaches.

	LAYERS	DATA COLLECTION/SOURCES	DATA ANALYSIS
QUALITATIVE	ENDOGENOUS QUALITATIVE DESCRIPTIONS	Mapping, observations, photos, videos	Geo-ethnography, Narratives, Site analysis, Socio-spatial and functional interpretation
QUALITATIVE	EXOGENOUS QUALITATIVE DESCRIPTIONS	Interviews, workshops...	Content analysis, Cross-compared interviews, Thematic analysis...
	EXOGENOUS QUALI-QUANTITATIVE DESCRIPTIONS	Literature	Thematic/content analysis, frequencies, descriptive stats and radar graphs
QUANTITATIVE	ADVANCED STATISTICS	Local, municipal, regional statistics...	Entropy index, Moran I, Hot spots, Cluster analysis, ...
QUANTITATIVE	BASIC STATISTICS	Local, municipal, regional statistics...	Basic demographics, density measures, means and ratios...

Figure 7: Data collection, data analysis methods, and tools for data management for the five analytical layers of the GISualization (source: Adelfio et al., 2019)

Figure 7 shows the five primary layers of the framework, the potential data sources, and analysis techniques. Quantitative data is provided based on basic descriptive statistics and advanced input transformations. Secondary data sources on socio-economic and demographic characteristics, survey data on attitudes and behavior, and geophysical variables represent crucial data input for the tool. We extend this understanding with advanced statistics from project output generated through the multinational survey and modeling exercises in WP4. The qualitative layers of the GISualization framework are composed of endogenous and exogenous descriptions supplied through interviews, workshops, observations, participatory mapping, photos, and videos. These data sources will be primarily provided by the six case studies. Figure 8 represents our proposed data integration tree for T5.2 derived from the GISualization framework.

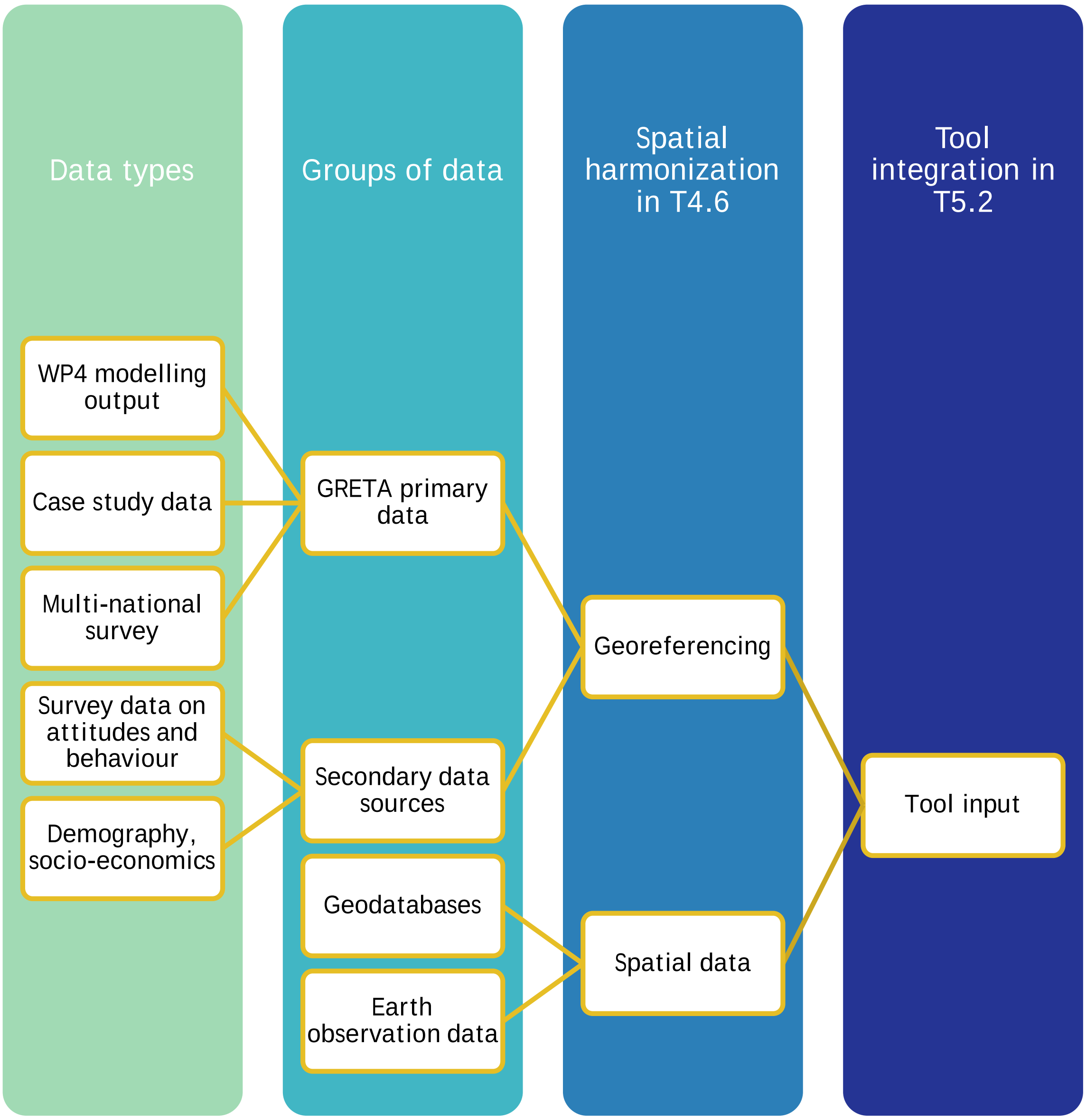


Figure 8: Data integration tree for T5.2

The primary data output is subject to the development of the GRETA WPs 1-4. We refer to the task-specific deliverables for a detailed discussion of these data sources. The following subsection will discuss major secondary data sources and the potential for enhancing the spatial analysis and integration in T5.2.

4.2 Secondary data sources

In order to identify the secondary data input for the project, extensive desk research was performed. The identified datasets were categorized according to a pre-defined analytical framework and structured accordingly. The resulting database was saved in an excel-file and will be deposited in the project repository on Zenodo. In this subsection, we will discuss the tool's existing survey data sources as a crucial input source.

The multinational survey conducted within the GRETA project will yield crucial insights on energy citizenship attitudes and behavior across the EU. The GIS-based tool will include additional data from existing European surveys to complement this survey, draw conclusions over time, and map national and EU-wide patterns of citizens' energy behavior and attitudes. We have identified relevant survey projects and suitable items within these surveys based on comprehensive desk research. Furthermore, we have classified these surveys according to their general theme and year of implementation, methodology, the number and type of respondents, and geographical scope and level of georeferencing. The linked excel-file in the GRETA repository on Zenodo reports the entire list. Table 2 summarizes a subsample.

Table 2: Survey programmes with relevant items for GRETA research (selection)

Survey name	General theme	Year	Scope
Comparative National Elections Project - Germany 2017	Political	2017	National
Comparative National Elections Project - Italy 2018	Political	2018	National
Comparative National Elections Project - Portugal 2015	Political	2015	National
Comparative National Elections Project - Spain 2015	Political	2015	National
Comparative Study of Electoral Systems - CSES Module 5 Advance Release 3	Political	2015-2021	Global
DNB Household Survey	Socio-economic	2021	National
EIB Climate Survey Edition III 2020-2021	Environ	2021	EU
Eurobarometer 94.2 (2020)	Socio-economic	2020	EU
European Election Studies - Voter Study 2019	Political	2019	EU
European Social Survey 2016: Attitudes to climate change - ESS Round 8	Environ	2016	Europe (23 countries)
European Values Study - Wave 5	Values	2017	Europe
Flash Eurobarometer 388 (Attitudes of Europeans Towards Waste Management and Resource Efficiency)	Environ	2013	EU
Flash Eurobarometer 420 (Attitudes of Citizens towards Shale Gas in Selected European Regions)	Environ	2015	EU
Flash Eurobarometer 441 (European SMEs and the Circular Economy)	Environ	2016	EU
Flash Eurobarometer 456 (Small and Medium Enterprises, Resource Efficiency and Green Markets, wave 4)	Environ	2017	Global
Flash Eurobarometer 463 (Europeans' Satisfaction with Passenger Rail Services)	Infra-structure	2018	EU

Flash Eurobarometer 467 (The Use of the Collaborative Economy)	Socio-economic	2018	EU
International Social Survey Programme: Citizenship II - ISSP 2014	Political	2014	Global
International Social Survey Programme: Environment III - ISSP 2010	Environ	2010	Global
International Social Survey Programme: Role of Government V - ISSP 2016	Political	2016	Global
LISS Panel - State of the Environment and Environmental Policy	Environ	2021	National
Special Eurobarometer 513 (Climate Change)	Environ	2021	EU

Regarding the geographical scope of the survey data, we identified EU-wide and national data sources, including items on energy attitudes and behavior. The resolution of the georeferencing, however, is often limited. Most surveys include variables that measure the observation's regional level, either as Eurostat NUTS-1/-2 classifications or, in cases of national surveys, as country-specific regional coding. Only a few surveys gather and publish highly detailed geolocations (e.g., GESIS Panel). Due to data protection issues, these surveys can only be accessed through on-site secure data centers. For the publicly available surveys, these limitations restrict options for linking these surveys with spatial points of interest and geographical variables (see Section 5.4.3).

Furthermore, the sampling strategies of most surveys are aimed at representativeness at the national level, which means that the results of single surveys do not produce representative results on the regional level even if the regional variable is included in the dataset. An example of a survey with a regional focus and sampling strategy on the regional level is Borz et al. (2022), which is stratified across 17 regions from 12 EU member states. Simple pooling of the existing surveys to increase the sample size across regions faces the challenge of different wordings of relevant questions. Thus, integrating and harmonizing heterogeneous survey items across a large geographical scope represents a considerable challenge for this project. Therefore, a crucial task in T4.6 is an investigation of the feasibility of systematic harmonization of these various surveys to create a spatially-representative sample. Harmonization to enhance the sample size and regional data coverage is an emergent discipline. Hoffmann et al. (2022) have shown the potential of this approach on the basis of pooled Eurobarometer surveys from 1,019,723 respondents over the period 2002-2019 to build a composite measure for public environmental concern. A crucial starting point for this task is identifying sets of items that refer to the same latent variables and constructs of energy citizenship.

Based on our survey collection, we identified more than 200 items that inform our knowledge of energy behavior and attitudes and more general environmental, social and economic attitudes and perceptions of citizenship. We classified these items according to several characteristics, including attitudes, behavior, and the role of the government, and also identified items employed in business surveys. The excel-file reports the entire item set, including the respective item code. Tables 3-5 cover a selection of questions classified according to the measurement of 1.) individual attitudes, 2.) individual behavior, and 3.) the role of government and business. These items will be checked for harmonization purposes in T4.6.

Table 3: Relevant survey items – individual attitudes (selection)

Survey	Question
BAR_ES	With what interest do you generally follow these news about climate change: with a lot of interest, a lot of interest, little or no interest?
BAR_ES	Do you think it has been very useful, quite, little or not at all useful to reach agreements and adopt solutions to the problems derived from climate change?
BAR_ES	On this issue of climate change there are usually different opinions. With which of the following phrases do you agree the most?
BAR_ES	Do you think that the impact of climate change can be reduced, or climate change stopped, or nothing can be done in the current situation?
BAR_ES	In your opinion, can it be said that we are in a climate emergency situation, in a serious but not an emergency situation, or do you think the situation is not serious and are you exaggerating when talking about the danger it poses?
BAR_ES	To what extent: a lot, a lot, a little or not at all, do you trust this series of actions to reduce or stop climate change?
EES_2019	What do you think of Environment?
EIBCS_2021	Do you feel that climate change is having an impact on your everyday life?
EIBCS_2021	Which of the following statements do you most agree with? The most significant way to stop or drastically limit climate change is through…
ESS_2016	How much have you thought about climate change before today?
ESS_2016	Do you think that climate change is caused by natural processes, human activity, or both?
ESS_2016	To what extent do you feel a personal responsibility to try to reduce climate change?
ESS_2016	How worried are you about climate change?
ESS_2016	Do you think that climate change is caused by natural processes, human activity, or both?
EVS_2017	How much do you agree or disagree with this statement? Q56.E Many of the claims about environmental threats are exaggerated
FEURO_388	How important is it for you that Europe uses its resources more efficiently?
FEURO_388	Which of the following actions do you think would make the biggest difference in how efficiently we use resources?
ISSP_2010	Generally speaking, how concerned are you about environmental issues?
ISSP_2010	And how much do you agree or disagree with each of these statements?: We worry too much about the future of the environment and not enough about prices and jobs today
LISS_2021	Do you agree or disagree with the following statements about the energy transition?
SEURO_513	QB2. And how serious a problem do you think climate change is at this moment? Please use a scale from 1 to 10, with "1" meaning it is "not at all a serious problem" and "10" meaning it is "an extremely serious problem".
SEURO_513	QB4.7. To what extent do you agree or disagree with each of the following statements? :-Adapting to the adverse impacts of climate change can have positive outcomes for citizens in the EU
SHE_2008	Are you worried about the environmental situation?
SHE_2008	Are you aware of any environmental campaigns within the last year (water, energy, recycling, etc.)?
SHE_2008	Did you encounter any environmental problems in your area during the year 2007?
TTS_2009	To what extent are you concerned about: climate change

Table 4: Relevant survey items – individual behaviour (selection)

Survey	Question
BAR_ES	I am going to read you a series of personal actions that can be done for the environment and to reduce pollution: taking into account your consumption habits, your lifestyle and your daily routines, could you tell me for each of them if you Do you do it regularly, do not do it regularly but would you be willing to do it, or do you not do it regularly and would not be willing to do it?
EIBCS_2021	To what extent, if at all, do you think your own behaviour can make a difference in tackling climate change?
EIBCS_2021	Irrespective of the COVID-19 crisis, regarding your own behaviour and climate change, would you say…:
EIBCS_2021	Regarding food, do you do any of the following to fight climate change?
EIBCS_2021	Assuming COVID-19 related restrictions on travel are lifted, are you planning to do any of the following to fight climate change?
EIBCS_2021	Among the following actions to fight climate change, what would be…
ESS_2016	In your daily life, how often do you do things to reduce your energy use?
ESS_2016	How likely do you think it is that limiting your own energy use would help reduce climate change?
EVS_2017	For each of the following voluntary organisations, please indicate which, if any, do you belong to. Please indicate whether you belong to… Q4.E Conservation, the environment, ecology, animal rights
EVS_2017	How much do you agree or disagree with this statement? Q56.A I would give part of my income if I were certain that the money would be used to prevent environmental pollution
EVS_2017	How much do you agree or disagree with this statement? Q56.B It is just too difficult for someone like me to do much about the environment
EVS_2017	How much do you agree or disagree with this statement? Q56.C There are more important things to do in life than protect the environment
EVS_2017	How much do you agree or disagree with this statement? Q56.D There is no point in doing what I can for the environment unless others do the same
FEURO_382	How often do you travel by … ?
FEURO_388	Which of the following actions are you undertaking to reduce the amount of household waste that you generate?
FEURO_388	What are the main reasons why you are not trying to reduce the amount of waste you generate? Would you say that…
FEURO_388	Do you sort the following types of waste, at least occasionally?
FEURO_388	What would convince you to separate more of your waste?
FEURO_388	What would convince you to separate at least some of your waste?
FEURO_388	Which of the following aspects do you consider most important when buying a durable product, like a washing machine or a fridge
FEURO_388	Would you buy the following products second hand?
FEURO_388	What prevents you from buying second hand products?
FEURO_388	There are emerging alternatives to buying new products. Have you ever done any of the following?
FEURO_388	What prevents you from buying a remanufactured product?
FEURO_388	What prevents you from leasing or renting a product instead of buying it?
FEURO_388	What prevents you from using sharing schemes?
FEURO_388	In your opinion, which of the following actions would be the most efficient in reducing littering?
ISSP_2010	How willing would you be to pay much higher prices in order to protect the environment?
ISSP_2010	And how willing would you be to accept cuts in your standard of living in order to protect the environment?
ISSP_2010	How much do you agree or disagree with each of these statements?: It is just too difficult for someone like me to do much about the environment
SEURO_513	QB5. Have you personally taken any action to fight climate change over the past six months?
SHE_2008	Did you participate in any of the following activities during the year 2007?

Table 5: Relevant survey items – role of government and business (selection)

Survey	Question
EIBCS_2021	In order to combat climate change, which areas of action should be prioritised?
ESS_2016	To what extent are you in favour or against the following policies in [country] to reduce climate change?
EUANDI_2019	The promotion of public transport should be fostered through green taxes (e.g. road taxing)
EUANDI_2019	Renewable sources of energy (e.g. solar or wind energy) should be supported even if this means higher energy costs
FEURO_388	How important is it for you that Europe uses its resources more efficiently?
FEURO_441	Has your company undertaken any of the following activities in the last 3 years?
FEURO_441	Over the last 3 years, what percentage of your company's turnover have you invested on average per year to undertake these activities?
FEURO_441	Should you decide to undertake activities related to the circular economy, what percentage of your company's turnover would you intend to invest on average per year?
FEURO_456	What actions is your company undertaking to be more resource efficient?
FEURO_456	Over the next two years, what are the additional resource efficiency actions that your company is planning to implement ?
FEURO_456	What impact have the undertaken resource efficiency actions had on the production costs over the past two years? The production costs have…
FEURO_456	Over the past two years, how much have you invested on average per year to be more resource efficient?
FEURO_456	Which type of support does your company rely on in its efforts to be more resource efficient?
FEURO_456	Does your company offer green products or services?
FEURO_456	How much did these green products or services represent in your annual turnover of the latest available fiscal year?
FEURO_456	For how long has your company been selling green products or services?
ISSP_2016	Listed below are various areas of government spending. Please show whether you would like to see more or less government spending in each area. The environment.
ISSP_2016	On the whole, do you think it should or should not be the government's responsibility to… ...provide industry with the help it needs to grow.
ISSP_2016	On the whole, do you think it should or should not be the government's responsibility to… ...impose strict laws to make industry do less damage to the environment.
SEURO_513	QB4.1. To what extent do you agree or disagree with each of the following statements? :-Promoting EU expertise in clean technologies to countries outside the EU can help create new jobs in the EU
SEURO_513	QB4.3. To what extent do you agree or disagree with each of the following statements? :-The costs of the damages due to climate change are much higher than the costs of the investments needed for a green transition
SEURO_513	QB4.4. To what extent do you agree or disagree with each of the following statements? :-Reducing fossil fuel imports from outside the EU can increase energy security and benefit the EU economically
SEURO_513	QB4.5. To what extent do you agree or disagree with each of the following statements? :-Taking action on climate change will lead to innovation that will make EU companies more competitive
SEURO_513	QB4.6. To what extent do you agree or disagree with each of the following statements? :-More public financial support should be given to the transition to clean energies even if it means subsidies to fossil fuels should be reduced

4.3 Indicator availability for proximity domains and geographical levels matrix

T5.1 of GRETA developed a comprehensive matrix for analyzing proximity domains and the respective geographical levels. A significant task for T5.2 is to translate the taxonomy of T5.1 into measurable geographic entities in light of the accessible data sources. Synthesizing our data review from Section 4.2, we identified possibilities and limitations for developing indicators for these domains. Table 6 describes the T5.1 matrix and includes color-coding for the indicators (green – good data availability, yellow – limited data availability, red – poor data availability). Limitations in data availability can largely be attributed to the fact that data only partially exist, often because these data are measured on the national level, and no harmonization between all EU countries has been implemented so far.

There are indeed ways to navigate data availability issues that could be considered. Categories in data could be collapsed or geographic attributes disaggregated to create comparable datasets, at least on the country level. Moreover, national or regional databases could be researched and integrated to compile harmonized datasets. All these efforts are resource-intensive and time-consuming, though. Due to the interdisciplinary nature of the data sources, they are also challenging to evaluate in the GRETA project. While in some cases, e.g., behavioral data, some steps in integrating data will be taken, data sources will mainly be exploited as-is.

Spatial domain

The spatial domain is relatively well covered. Urban structures, climatic features, and resource availability are documented in national and EU geodatabases. These databases provide detailed shapefiles and gridded data for land use, land cover indicators, and atmospheric variables. Measuring EU-wide energy infrastructure, however, is much more problematic. Datasets mainly exist on the national level; the data quality varies depending on the specific technology or infrastructure type and would require harmonization of data across member states. Research projects have created substantial EU-wide datasets for individual technologies, like onshore wind or residential solar PV. Some examples of EU-wide geospatial data sources that can be exploited in GRETA are:

- Climate Risks: Carter, J.G, Hincks, S, Vlastaras, V, Connelly, A, and Handley, J. 2018. European Climate Risk Typology. Available at: http://european-crt.org/index.html.
- Land cover and land use: Copernicus Land Monitoring Service (CLMS) 2022. Available at: https://land.copernicus.eu/.
- Heating and cooling degree days: Eurostat 2022. Available at: https://ec.europa.eu/eurostat/web/main/data/database.

- Transport networks (Airports and Ports): Eurostat 2022. Available at: https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data.

An emerging data source in recent years, partly included in the sources mentioned above, is earth observation data. Most notably, in our context, the EU has launched its very own earth observation program called Copernicus. This program comprises many sensors on earth and satellites that collect, amongst others, environment-related indicators on a fine-grained spatial and temporal dimension. With Copernicus, it is possible, e.g., to measure solar radiation and develop indicators for its potential to increase the installation of solar panels (see Section 6.3 for an example of how such information could be used within GRETA).

Policy domain

The policy domain covers political agendas, administrative structures, and regulatory frameworks. Administrative structures are well covered through the Eurostat Nomenclature of territorial units for statistics (NUTS). However, national statistical accounts do not systematically cover dimensions like political agendas and regulatory frameworks. In contrast, the measurement of these dimensions relies on research projects which provide evidence for these features. It should be noted, though, that the systematic coverage of the policy domain mainly refers to the national level. Data quality decreases with higher geographical resolution. Local level and large-scale datasets on the policy domain are rare. Examples of relevant datasets include:

- Political agenda and party positions: WZB Berlin Social Science Center 2022. The Manifesto Project. Available at: https://manifesto-project.wzb.eu/.
- Regional governance: Charron N, Dahlberg S, Holmberg S, Sundstrom A, Alvarado N and Dalli, C M (2020). The Quality of Government EU Regional Dataset, version Nov20. University of Gothenburg: The Quality of Government Institute, https://www.gu.se/en/quality-government
- Local Governance: Ladner, A. et al. (2019). The Local Autonomy Index (LAI). In: Patterns of Local Autonomy in Europe. Governance and Public Management. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-319-95642-8_9.
- Renewable energy regulatory frameworks: IEA/IRENA Renewable Energy Policies and Measures Database (2022). Available at: https://www.iea.org/policies/about.
- Environmental Performance: Wolf, M. J., Emerson, J. W., Esty, D. C., de Sherbinin, A., Wendling, Z. A., et al. (2022). 2022 Environmental Performance Index. New Haven, CT: Yale Center for Environmental Law & Policy. epi.yale.edu.

Social domain

The social domain is comprised of dimensions that address issues of community, collaboration, knowledge and learning, and awareness. Harmonized, large-scale data for these issues are rare, and we expect this domain to rely heavily on qualitative case study data. Section 4.2.1 describes existing survey programs that could partially be utilized to generate large-n proxies for these dimensions, particularly individual-level awareness. Public databases on energy cooperatives, transnational city networks, NGO umbrella organizations and business associations can represent additional sources to map community and collaboration patterns. Examples include:

- Renewable energy cooperatives: REScoop.eu network 2022. Available at: https://www.rescoop.eu/network/map/.
- Transnational city networks: Covenant of Mayor for Climate & Energy Europe 2022. Available at: https://www.covenantofmayors.eu.
- Business networks: Carbon Disclosure Project (CDP) 2022. Available at: https://www.cdp.net/en.
- NGO networks: Climate Action Network Europe (CAN) 2022. Available at: https://caneurope.org/members/.

Technical domain

The technical domain is moderately well covered. Generally, there are only a few official bodies on technical energy infrastructure datasets. Most relevant data sets are published within volunteer and research projects or survey studies. Regarding energy systems, multiple interest groups are involved in creating open data repositories, especially concerning power plants and their capacity. OpenStreetMap provides valuable and openly available micro data on energy systems and technological readiness, e g., electric charging stations, and transmission lines. Detailed and large-scale data on home appliances and household energy consumption is still sparse and not openly available. Related insights can be found in supranational surveys that measure household behavior, e. g., ESS, ISSP, or Eurobarometer. Data on enabling structures is largely proprietary or published solely by local authorities. In general, various sources have done much work to provide openly accessible data on energy infrastructure. However, a large part of information remains uncollected or sealed in private databases. Data on energy systems is continuously compiled by some sophisticated projects. This data comes in aggregated form (capacity) and individual geometries (plant locations):

- OPSD: https://open-power-system-data.org/
- Global Power Plant Database: https://github.com/wri/global-power-plant-database
- GeoNuclearData: https://github.com/cristianst85/GeoNuclearData
- GEO: http://globalenergyobservatory.org/
- Our World In Data: https://github.com/owid/energy-data

Behavioral data might be able to act as an alternative to direct household-level consumption measurements. ESS, ISSP, and Flash Eurobarometer 388 collected behavior data for European households:

- ESS: https://ess-search.nsd.no/en/study/f8e11f55-0c14-4ab3-abde-96d3f14d3c76

- ISSP: https://www.gesis.org/en/issp/modules/issp-modules-by-topic/environment/2010
- Flash Barometer 388: https://data.europa.eu/data/datasets/s1102_388?locale=en

Smart meter data is very sparse and usually only available for micro case studies (i.e., a few households). The Smart Meter Data Portal is a database of many openly available smart meter case study data sets. OpenMeter is a German project that seems to work on a larger scale integration:

- https://www.openmeter.de/
- https://smda.github.io/smart-meter-data-portal/

Openstreetmap provides a lot of data on public energy infrastructure, most importantly electric charging stations, transmission lines, and facilities offering open wifi spots:

- https://openchargemap.org/
- https://openinframap.org/
- https://wiki.openstreetmap.org/wiki/Key:internet_access
- https://github.com/giacfalk/EV_charging_network_accessibility_analysis

Data on enabling structures is mainly made available locally, mostly in combination with open data portals. Wigle and Freifunk-Karte are two approaches that offer VGI on a larger scale:

- https://www.wigle.net/
- http://www.freifunk-karte.de/

Technological readiness can be measured by examining the EU's current fuel and fleet state. The EAFO recently reworked their online appearance and now offers data on alternative fuels and vehicle fleets in common formats (csv, xlsx):

- https://alternative-fuels-observatory.ec.europa.eu/

Economic domain

The economic domain is defined by the economic structure, incentive schemes for renewable energy, value redistribution mechanisms, and cluster economy features. Comprehensive databases like the Annual Regional Database of the European Commission's Directorate General for Regional and Urban Policy (ARDECO) provide various basic statistics on these domains, which are suitable to inform the GIS-based tool. In addition, we identified several indices which provide information on regional economic clusters and performance. Incentive schemes are furthermore covered by regulatory framework databases in the policy domain, and cluster economy features can be partially addressed with databases on energy cooperatives and NGO networks as discussed in the social domain (both see above). Examples of such data include:

- Gross Value Added (GVA) by NACE Sector: ARDECO Database 2022. Available at: https://knowledge4policy.ec.europa.eu/territorial/ardeco-online_en.
- Employment by NACE Sector: ARDECO Database 2022. Available at: https://knowledge4policy.ec.europa.eu/territorial/ardeco-online_en.
- Regional Competitiveness: DG Regional and Urban Policy 2022. EU Regional Competitiveness Index, 2019. Available at: https://ec.europa.eu/regional_policy/en/information/maps/regional_competitiveness/#map.

Table 6: Complete proximity domains and geographical levels matrix (green – good data availability, yellow – limited data availability, red – poor data availability)

Proximity domains	Dimensions	Geographical levels					Descriptors	Indicators To be completed with T5.2	Examples
SPATIAL DOMAIN		L	R	N	S	V			
SPATIAL	Urban structure	X					Types of structure (neighbourhood, block, building, others)		An example is the Marstal Fjernvarme (Denmark) case, where the urban structure allowed the community to group around a solar and heating district.
	Climatic area / region	X	X	X	X	X	Homogeneous climatic area (similar climatic conditions / necessities)		An example is the Tirano (Italy) case (regional level), where the location inside the Po valley, put participants in similar climatic conditions (especially related with air quality).
	Resource availability	X	X	X	X	X	Availability of natural resources to be exploited (eolic, solar, others)		An example of this, at the local level, is the Oborniki slaskie (Poland) case, where the locally available biomass triggered the creation of a community around its use.
	Energy infrastructure closeness	X	X	X			Presence of energy plants in an area; presence of existing energy infrastructures		*(under completion)*
	Cluster and activities closeness	X	X	X			Physical presence of associations, groups of firms, institutions, clusters, etc.	See economic domain	An example is the ENERGEIAKH KOINOTHTA KARDITSAS SYNPE (Greece) case, where the presence of an already existing energy cooperative with infrastructures, triggered the enlargement of the cooperative toward the creation of a proper energy community.
POLICY DOMAIN		L	R	N	S	V			
POLICY	Political agendas	X	X	X	X		Presence of strategies, goals, actions and an agenda on transition goals; presence of green procurements mechanisms		An example is the Region Emilia-Romagna (Italy) that produced the Labour and Climate Pact together with the local authorities, universities, enterprises, trade unions and non-profit sector, to agree to full employment and green transition.
	Administrative structure (across levels)	X	X	X	X	X	Hampering or enabling aggregating mechanisms and agency		France, for example, has introduced a territorial development strategy for renewable energies, jointly designed with all the stakeholders to facilitate the development of agreed, cooperative projects. Under the proposed method, the scheme would operate through energy transition committees, bringing in elected representatives, socio-economic actors and citizens.
	Regulatory framework	X	X	X	X	X	Structure and presence of norms habilitating or not people/tech aggregations		The Netherlands with the National Regional Energy Strategy (RES) Program helps regions to achieve the transition to natural gas-free. The RES establishes how the sustainable generation of energy can fit into the spatial planning and the electricity network, and how support for the measures can be created in society. The national RES Program supports the regions in creating these RESs by supporting and sharing knowledge. It further connects parts, highlights risks and threads and identifies linkage opportunities.
SOCIAL DOMAIN		L	R	N	S	V			
SOCIAL	Community dimension	X				X	Presence of community bonds or ties; presence of leaders or spokespersons		An example of it (at the local level) is the Solbyn Association (Sweden), where one of the first trigger was the presence of an housing cooperative with the aim to promote sustainable living.
	Collaboration mechanisms	X	X	X	X	X	Presence of active associations, third sector, social enterprises, community projects, practices, civic actions		*(under completion)*
	Knowledge and learning	X	X	X	X	X	Presence of specific training programs on energy; presence of network with other		An example of it (at the national-virtual level) is the SOM Mobilitat case, that operates on a network at European level to share good practices and resources with other sustainable mobility cooperatives through

							cities/regions in order to learn from peers; training on technical aspects, other		the creation of the first network of mobility cooperatives in Europe called REScoop Mobility, under the umbrella of the REScoop.eu cooperative federation.
	Energy Awareness	X	X			X	Presence of specific awareness goals under the energy citizenship configuration.		An example of it (at the national level) is the Enercoop (Frances) case, where a specific attention is given to the increase of awareness of citizens aroung the topic of energy and renewables.
TECHNICAL DOMAIN		L	R	N	S	V			
TECHNICAL	Energy system	X	X	X			Energy system structure, management and characteristics (e.g. hierarchical or distribute structure)		*(under completion)*
	Technological readiness	X	X	X		X	Readiness of the context in relation with technology, e.g. electric vehicle distribution, presence of the recharging stations, energy grid, etc.		*(under completion)*
	Technological appliances	X				X	Presence and distribution of devices (e.g. small-scale home devices; urban devices; etc)		An example of it (at the local level) is the Svalin co-housing complex (Denmark), where forms of innovation and energy awareness are triggered by the presence of innovative small appliances (e.g. colour lighting in relation with energy consumption)
	Enabling structures	X	X	X	X	X	E.g. supporting online platforms; digital twins, control rooms, sharing platforms, ICT coverage (e.g. wifi free connection hotspots; 3/4/5g networks)		An example of it (at the virtual level) is the SOM Mobilitat case, that has a platform for multiple actions, from the functioning of the project itself, to crowdfunding, to management.
ECONOMIC DOMAIN		L	R	N	S	V			
ECONOMIC	Energy system economic structure	X	X	X	X	X	E.g. energy price; subsidies for renewable energy transition; etc.		*(under completion)*
	Energy system enabling economics		X	X			Presence of incentives (national, regional or others); innovative business models; etc		An example of this is the Banister House Solar case (UK), where an innovative business model including leasing and the involvement of several financial actors have been put in place.
	Value redistribution / inclusion mechanisms	X	X	X		X	Presence of mechanisms related with energy poverty		A light example of this (at the local level) is the Magliano d'Alpi (Italy) case, where the EV charging colums are free of charge for residents. This has been done for increasing citizens trust and awareness on the energy topic.
	Cluster economy	X	X	X			Concentration of entities representing non-governmental organisations, traditional and socially-oriented enterprises and other institutions		An example of it (at the local level) is the SAS Ségala Agriculture et Energie Solaire Cooperative (France) where one of the first motivations around the creation of the energy community was the presence of the same economic cluster or participants, in this case agriculture and breeding.

5 Scoping of technical implementation

5.1 Building the tool

The tool development is based on the initial planning and scoping steps, which have been carried out to date. Following Fay et al. (2022), the building workflow of the tool is based on five consecutive and interlocked steps, which ensure the development of a robust and production-ready application: 1.) Design, 2.) Prototype, 3.) Build, 4.) Strengthen, 5.) Deploy. Sections 1-4 have laid the groundwork for the design phase. Follow-up workshops will map the user journey and provide valuable insights into the User Experience (UX), a crucial concept for building the application. Steps 2-5 will be executed in Q3 2022 until Q2 2023. The technical implementation requires selecting appropriate software and programming language and data processing and analysis techniques. The remaining sections, 5.2 to 5.4, will address these steps.

5.2 Open source software requirements

5.2.1 Software environment and programming language R

R is a statistical software and programming language that has risen significantly in popularity (R Core Team, 2022). Especially in the social sciences, R has proven to be one of the most prevalent software packages for data analysis (Trisovic et al., 2022). While R is a useful and complete tool by itself, the wide variety of so-called packages (software extensions) makes R suitable for most quantitative and qualitative social sciences purposes. The official package distribution archive CRAN offers packages ranging from data mining and manipulation to complex statistical inference and spatio-temporal analysis. In the GIS software spectrum, R has become an established choice for geospatial computations (Bivand et al., 2013) with packages like 'sf' and 'raster'. All published packages are open-source, collaborative, free of charge, and commonly pass through a peer-review process, either by their publishing journal (e.g., The R Journal) or by a community of volunteers (Anderson et al., 2021).

5.2.2 Shiny App UI

One of the main shortcomings of R is its lack of a graphical interface. The base functionalities only allow for very rudimentary graphical visualizations. While it is possible to create appealing illustrations using packages like 'ggplot2', R remains a code-based software with a steep learning curve. In order to reduce the technical requirements needed to use the GIS-based tool, the 'shiny' package will be used to produce applications with a straightforward user interface. Shiny allows the creation of web apps, dashboards, and other tools that are continuously connected to the computational power of R (see Figure 9). This means that (spatial) data analysis tasks can be performed using a graphical user interface but are still conducted using R on the server side. This can be especially useful for manually exploring and manipulating small-scale geographic datasets and moving R closer to graphical GIS software like QGIS.

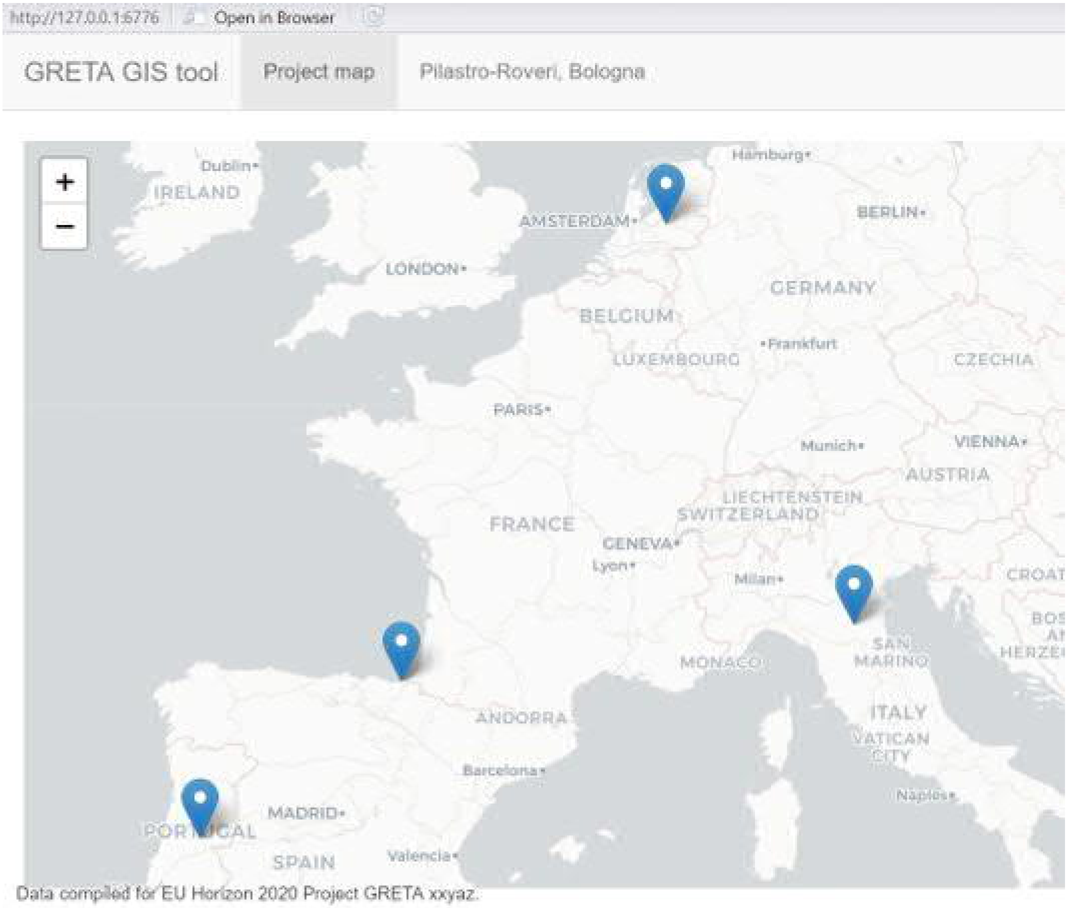


Figure 9: Snapshot from GRETA ShinyApp prototype

5.2.3 RMarkdown

A central selling point of R as a statistical software is that it is an open-source programming language. Unlike proprietary graphical interface software like SPSS (for statistical purposes) or ArcGIS (for spatial analyses), each step is (1) explicitly documented inside code files and thus entirely replicable and (2) verifiable due to the open-source nature of the software. The package 'RMarkdown' (and 'knitr') expands on these principles and facilitates the translation of R code to documents like reports or presentations. Consequently, R becomes a software that can be easily integrated into more comprehensible data visualization and presentation forms to produce a more reproducible type of data-based research.

5.3 Georeferencing and linking datasets

Georeferencing is the process of adding direct spatial identifiers to data. These identifiers are usually geocoordinates, which can be simple points on the earth's surface or depict more complex geometries, such as lines or polygons (Jünger, 2019). Generally, geometries refer to a feature in the real world – points could be house locations, lines could be streets, and polygons the outline of a city district. A feature at hand can also entail additional information, e.g., the condition of a building, the number of residents in the building, its heating type, and the like. Georeferencing is a technique that produces such data; georeferenced data are the output.

What is needed to accomplish the goal of receiving georeferenced data are tools for georeferencing, which are implemented as part of Geographic Information Systems (GIS). GIS relates to a framework and software solutions to process, analyze, and visualize geospatial data. One of the essential steps to get georeferenced data is the exploitation of geocoding, i.e., the conversion of indirect spatial identifiers (e.g., a location name) to geocoordinates. GIS often comprises automatic procedures for geocoding, but there are also services on the internet that can be used separately. For this purpose, users can rely on services such as Google's Geocoding API[4] or OpenStreetMap's Nominatim engine[5]. Moreover, if only a database of harmonized place names for the whole world is of interest, the project GeoNames[6] may be interesting. No matter what tool is used, geocoded and thus georeferenced data are the prerequisite for further processing the data.

In many applications, not only the geographic localization of citizens, cooperatives, and other agents is of interest, but also contextual information about these localities. Spatially identifiable information from georeferenced data thus can be used to link several datasets to one another. In a GIS, these techniques are known as spatial linking, spatial overlay, or spatial join. The resulting data resemble the original ones, yet are augmented with additional contextual information that can enrich the further spatial and substantial analysis.

Figure 10 shows how spatial linking is accomplished based on the example of a fictional survey respondent. The respondent's location is georeferenced and comprises a simple point in space. By projecting this point in one shared space with the information from other data, e.g., all the different stacked layers, we can relate it to all this information. For example, we can retrieve information about which city district the respondent lives in or gather environmental information structured in uniformly

[4] https://developers.google.com/maps/documentation/geocoding/start
[5] https://nominatim.org/
[6] http://www.geonames.org/

shaped grid cells. Spatial linking is a flexible method that can be combined with various geometric operations.

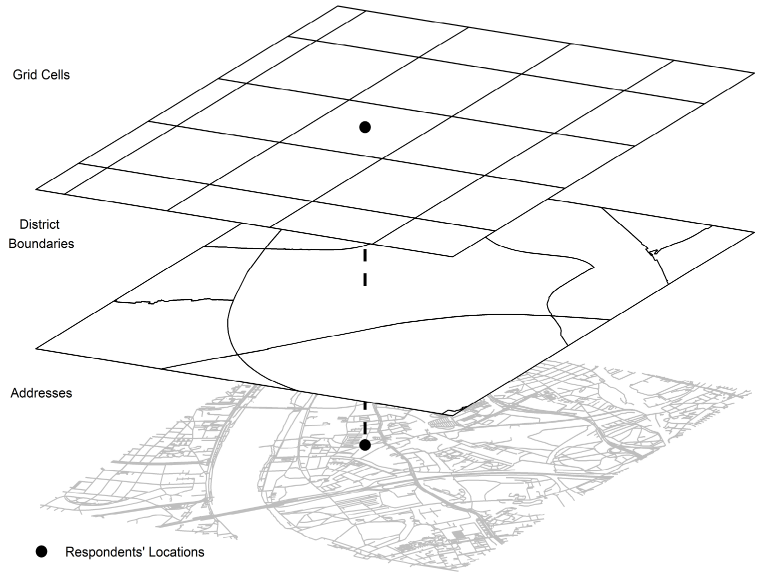


Figure 10: Visualization of spatial linking with different geospatial layers

In some cases, we do not only want to add information from an auxiliary dataset just by the specific location of a point (Figure 10, Panel 1). We can also consider adding information from surrounding areas, e.g., by drawing circular buffers around a point and extracting summary statistics for all attributes within this area. Alternatively, we may calculate distances to points of interest and use this as relevant information for the application. All these three options of spatial linking are shown in Figure 11 with an example of adding environmental information from road traffic noise and soil sealing to a point. There are many more operations to be considered. Yet, a comprehensive review would go beyond the scope of this deliverable (for an overview of applications to social science survey research, e.g., please refer to Jünger, 2019). Nevertheless, in GRETA, we expect to use an extensive range of techniques, depending on the input and data we aim to spatially link.

For example, we target using point data from the multinational survey. These points should be georeferenced location information of the survey respondents based on their private addresses in the best-case scenario. We can gather information about the solar radiation at the respondents' specific location from satellite data and can, therefore, estimate the solar energy potential. Having this information about solar energy potential, we can analyze if survey respondents with high values are more likely (or more willing) to install solar panels on the roof of their house. Such an analysis is only possible due to the georeferencing of the data and the spatial linking with other data.

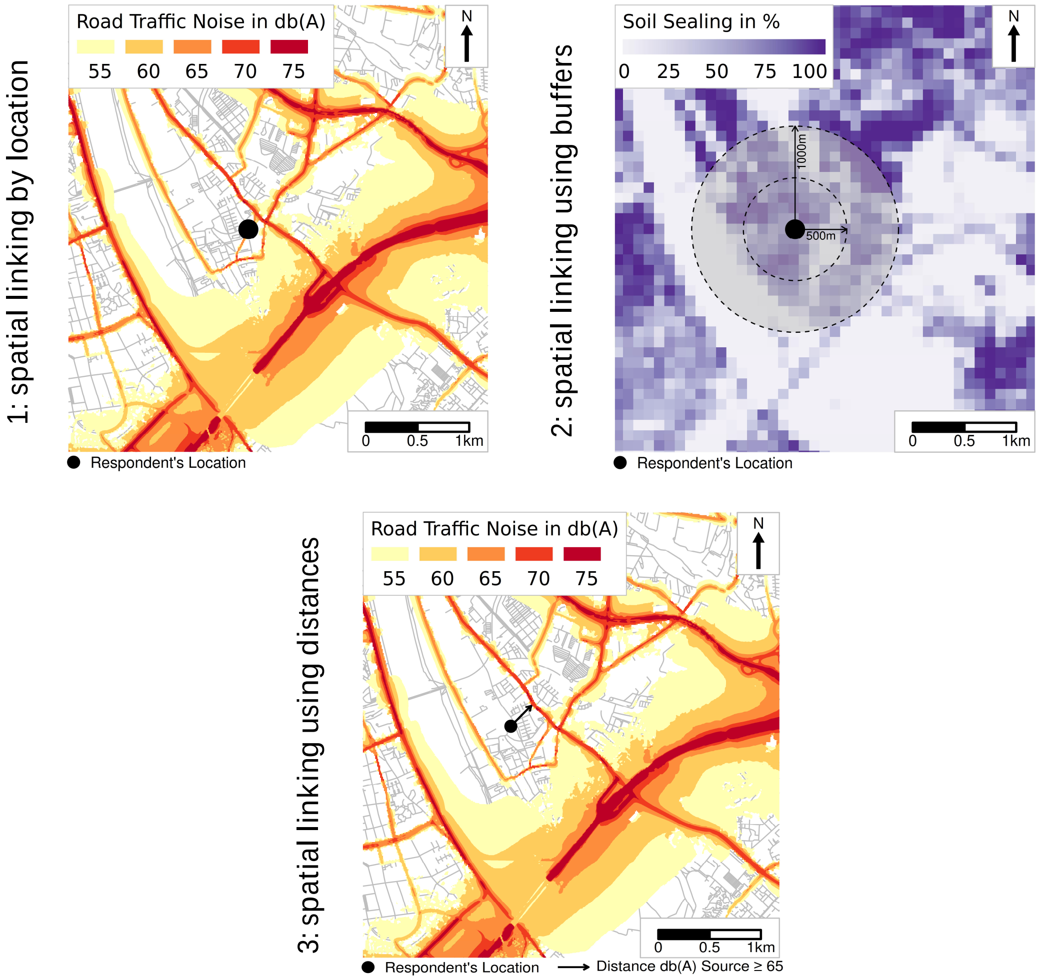


Figure 11: Variations of spatial linking (map sources: Jünger, 2019)

In the following subsection, we specifically describe our choice of the R programming language for the task of building a GIS-based analysis tool. We also introduce more GIS analysis techniques that can be exploited within the tool and for presenting the results in the GRETA project.

5.4 GIS-based analytical approaches

The R programming language does not support spatial analysis natively. However, due to the modular nature of R, its capabilities have increased gradually over the last two decades. Bivand (2006) reports a threefold increase of search results tagged with "spatial" on the official R website between 2001 and 2006. As of now, the exact search returns 15.000 results, a fifteenfold increase compared to 2006. Lovelace (2019) argues that R stands out in the GIS landscape due to the vastness of general statistical packages, which facilitates custom tasks in, e. g., spatial modeling. Substantial progress has been made due to the change from the old *sp* and *raster* packages to the newer *sf* and *terra* packages (Bivand, 2021). *sf* introduces a "tidy" (Wickham et al., 2019) workflow for handling spatial vector data that is based on C++ and is, therefore, able to interface the widely used GIS libraries GEOS, GDAL, PROJ, and S2 (Pebesma, 2018). The package supports the most basic functionalities of full-scale GIS software. It can read most of the standard GIS data formats, handle coordinate reference systems, create basic plots, and conduct geometric computations like binary predicates and unary operations. The *terra* package also takes vector data but is most prominent for its ability to efficiently process raster data (Hijmans, 2022; Bivand, 2021). Among various basic functions, e.g., merge, trim, mosaic, shift, or resample raster data, *terra* allows the cell-based functional manipulation of raster grids using a range of custom "apply-like" functions. Since *terra* supports vector and raster data, the package facilitates transformative operations between vector and raster data, such as cropping or value extraction by polygon and rasterization. The *stars* package applies a similar paradigm optimized to handle data cubes containing spatio-temporal raster or vector data (Pebesma, 2021). Finally, as claimed by Lovelace (2019), R is not restricted to just these GIS frameworks but can act as a bridge between different spatial analysis tools. R can directly interface ArcGIS (using the R-ArcGIS bridge), QGIS (*RQGIS*, Muenchow et al., 2017), GRASS GIS (rgrass7, Bivand, 2022), SAGA GIS (RSAGA, Brenning, 2018), geodatabases (DBI, rpostgis, R-SIG-DB, R Special Interest Group on Databases et al., 2021; Bucklin & Basille, 2018), Python (reticulate, Ushey et al., 2022), and JavaScript (leaflet, plotly, Cheng et al., 2022, Sievert, 2020). The following three subsections expand on the modular ecosystem of R and introduce additional packages to perform in-depth cartography, cluster analysis, modeling, and network analysis.

5.4.1 Data visualization and mapping

Data visualization of geographic data poses a different challenge than ordinary data visualization. Geographic space is highly complex, subjective by nature, and inherently three-dimensional. In order to create informative and tangible maps, geodata needs a sensible projection, orientation, generalization, and map design. R packages can assist in an illustrative mapping process. First, and most importantly, the *ggplot2* package provides various geographic mapping helpers and complete integration of spatial objects created by *sf* and *stars*. This package provides the baseline for all mapping (and general plotting) endeavors in R. It can be extended by several auxiliary packages, such as *ggspatial* and *ggsn* for adding spatial orientation (such as north arrows or scales), *ggmap* for spatial context (i. e., base maps), *ggrepel* for readable place annotations, or packages like *classInt*, *RColorBrewer* and *viridis* for designing insightful symbologies. The packages *tmap*, *chloroplethr*, *cartography,* and its successor *mapsf* are mapping frameworks based on *ggplot2*, but incorporate cartographic design principles in a more intuitive workflow. While these approaches work well for presenting static maps, e.g., for reports, this deliverable's goal is to develop a GIS-based tool that requires more interactive types of visualization. For this purpose, the packages *leaflet* and *plotly* interface the identically-named JavaScript libraries to create interactive maps for web visualizations. Usually, to create interactive illustrations, an entirely different skill set is required. The *leaflet* and *plotly* packages are entirely integrated into R and need no knowledge of JavaScript. Interactive maps are (web) applications that depict dynamic geometries or raster grids on a base map. Users can pan and zoom on the map and select and manipulate geometries. Roth (2013) argues that interactivity promotes the visual thinking of exploratory and analytical science. While static maps are simplified based on the author's subjective perception of what is important, interactive maps allow users to freely explore a thematic and effectively remove cartographic generalizations and geographic boundaries. Perdana et al. (2018) and de Mendonça et al. (2012) show that interaction in spatial data visualization improves the explorative and analytical performance of professionals and policymakers.

5.4.2 Cluster analysis

Cluster analysis describes the challenge of grouping observations based on their similarity. The definition of group and similarity differs greatly depending on the applied clustering algorithm. Consequently, the amount of R packages that can perform or support cluster analysis is innumerable. The *cluster* package comes pre-installed with R and implements the basic clustering algorithms introduced by Kaufman & Rousseeuw (1990) and Struyf et al. (1997). Despite its age *cluster* is still predominantly used for cluster analysis and incorporates partitioning and hierarchical clustering methods. It is supplemented by the *stats* package, which also comes pre-installed and holds functions to perform k-means and hierarchical clustering. The *cluster* and *stats* packages form a base for other clustering approaches, most notably the *factoextra* package (Kassambra & Mundt, 2020). *factoextra* mainly supports cluster analyses by producing "publication ready plots" (Kassambra, 2020) that are enhanced by auxiliary approaches (e.g., principal component analysis or correspondence analysis) and validation measures (e.g., silhouette coefficients or gap statistics). R also has capacities to enable more specific types of cluster analysis, such as density-based (Hahsler et al., 2019) or model-based clustering (Scrucca et al., 2016). A more exhaustive catalog of clustering packages is given by Leisch & Gruen (2022).

5.4.3 Network analysis

Following the "GISualization" framework and expanding the methodological toolset beyond pure GIS-based approaches, we also consider network approaches to study and visualize project output in the tool. Network analysis has become increasingly popular in the social sciences to visualize and study the complex relationships between entities. As a visualization tool, network approaches allow the visualization of complex clusters of nodes and edges. The application possibilities are extensive and include the identification of key actors and links, the measurement of strengths of relationships, structural properties and communities, diffusion patterns, and network evolution over time (Ognyanova, 2021). Beyond visualization, inferential network approaches have generated various indicators to measure network characteristics like centrality and homophily (Cranmer et al., 2017). Network approaches are very well established in R packages and compatible with our development approach. *igraph* (Nepusz, 2022) is the most comprehensive collection of network analysis tools for visualization. *ggraph* (Pedersen & RStudio, 2021) and *tidygraph* (Pedersen, 2022) transpose network analysis into the tidyverse and the grammar of graphics logic (Wickham, 2010) of *ggplot2*

(Wickham, 2016). More advanced statistical analysis can be conducted with the *sna*-package, which offers graph-level indices, structural distance and covariance methods, and network regression (Butts, 2022). The *ergm*-package provides an integrated toolset for analyzing networks based on exponential-family random graph models (ERGMs) (Handcock et al., 2022). The *statnet*-package (Handcock et al., 2019) includes *ergm* and offers a wide range of management, exploration, analysis, and visualization features of network data.

5.5 Repository and long-term storage

GRETA established the Open Portfolio for Civic Energy Empowerment (OPCE) as an Open Access repository on Zenodo (https://zenodo.org/communities/greta). The entire tool script and relevant anonymized and aggregated datasets will be stored in this repository.

6 Conclusion & outlook

Task 5.2 aims to create a GIS-based analytical tool to study the emergence of energy citizenship. In this deliverable, we have described previous work in the project concerning the general purpose of the tool and its data integration needs. As the development will be finalized later in the project (April 2023), in the last step, we have shown the scoping and the technical implementation, including analytical approaches that will enter the tool. While this development follows a rather strict technical path, the overall goal is to give analytical insights into variations within and between geographical levels of energy citizenship emergence.

Apart from the scientific relevance and its impact on the research community, these findings will also inform other missions of GRETA. First and foremost, this refers to citizenship contracts and their implementation in the case studies (T5.3), community pathways across geographic levels (T5.4), and GRETA's policy recommendations (WP6). The GIS-based analytical tool will also be helpful for various users and stakeholders not directly connected to the project. All the development will take place in the open following Open Source publishing strategies. Thus, the tool's conceptual groundwork and technical stack will live on to provide further projects with an accessible solution to present complex and multidimensional data.

References

Abel, D., Lieth, J., & Jünger, S. (2022, June). Mapping the Spatial Turn in Energy Transition Research: A Computational Literature Review. Conference Paper presented at the Council for European Studies (CES) - 28th International Conference of Europeanists.

Adelfio, M., Kain, J.-H., Stenberg, J., & Thuvander, L. (2019). GISualization: Visualized integration of multiple types of data for knowledge co-production. Geografisk Tidsskrift-Danish Journal of Geography, 119(2), 163–184. https://doi.org/10.1080/00167223.2019.1605301

Alboukadel Kassambara, & Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. https://CRAN.R-project.org/package=factoextra

Brenning, A., Bangs, D., & Becker, M. (2018). RSAGA: SAGA Geoprocessing and Terrain Analysis. https://CRAN.R-project.org/package=RSAGA

Bivand, R. (2006). Implementing Spatial Data Analysis Software Tools in R. Geographical Analysis, 38(1), 23–40. https://doi.org/10.1111/j.0016-7363.2005.00672.x

Bivand, R. (2021). Progress in the R ecosystem for representing and handling spatial data. J. Geogr. Syst., 23(4), 515–546. https://doi.org/10.1007/s10109-020-00336-0

Bivand, R.. (2022). rgrass7: Interface Between GRASS Geographical Information System and R. https://CRAN.R-project.org/package=rgrass7

Borz, G., Brandenburg, H., & Mendez, C. (2022). The impact of EU Cohesion Policy on European identity: A comparative analysis of EU regions. European Union Politics, 23(2), 259–281. https://doi.org/10.1177/14651165221076444

Bridge, G. (2018). The map is not the territory: A sympathetic critique of energy research’s spatial turn. Energy Research & Social Science, 36, 11–20. https://doi.org/10.1016/j.erss.2017.09.033

Bucklin, D., & Basille, M. (2018). rpostgis: Linking R with a PostGIS Spatial Database. The R Journal, 10(1), 251–268. https://doi.org/10.32614/RJ-2018-025

Butts, C. T. (2022). sna: Tools for Social Network Analysis (2.7). https://CRAN.R-project.org/package=sna

Castán Broto, V., & Baker, L. (2018). Spatial adventures in energy studies: An introduction to the special issue. Energy Research & Social Science, 36, 1–10. https://doi.org/10.1016/j.erss.2017.11.002

Cheng, J., Karambelkar, B., & Xie, Y. (2022). leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. https://CRAN.R-project.org/package=leaflet

Cranmer, S. J., Leifeld, P., McClurg, S. D., & Rolfe, M. (2017). Navigating the Range of Statistical Tools for Inferential Network Analysis. American Journal of Political Science, 61(1), 237–251. https://doi.org/10.1111/ajps.12263

Fay, C., Guyader, V., Rochette, S., & Girard, C. (2022). Engineering production-grade shiny apps (First edition). CRC Press, Taylor & Francis Group.

Friedrich Leisch, B. G. (2022). CRAN Task View: Cluster Analysis & Finite Mixture Models. Comprehensive R Archive Network (CRAN). https://cran.r-project.org/web/views/Cluster.html

Giraud, T., & Lambert, N. (2016). cartography: Create and Integrate Maps in your R Workflow. Journal of Open Source Software, 1(4), 54. https://doi.org/10.21105/joss.00054

Giraud, T., & Lambert, N. (2017). Reproducible Cartography. In Advances in Cartography and GIScience (pp. 173–183). Springer. https://doi.org/10.1007/978-3-319-57336-6_13

Hahsler, M., Piekenbrock, M., & Doran, D. (2019). dbscan: Fast Density-Based Clustering with R. J. Stat. Soft., 91, 1–30. https://doi.org/10.18637/jss.v091.i01

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., Krivitsky, P. N., Bender-deMoll, S., & Morris, M. (2019). statnet: Software Tools for the Statistical Analysis of Network Data (2019.6). https://CRAN.R-project.org/package=statnet

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., Krivitsky, P. N., Morris, M., Wang, L., Li, K., Bender-deMoll, S., Klumb, C., Bojanowski, M., Bolker, B., Schmid, C., & Cheng, J. (2022). ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks (4.2.2). https://CRAN.R-project.org/package=ergm

Hijmans, R. J. (2022). terra: Spatial Data Analysis. https://rspatial.org/terra/

Hoffmann, R., Muttarak, R., Peisker, J., & Stanig, P. (2022). Climate change experiences raise environmental concerns and promote Green voting. Nature Climate Change, 12(2), 148–155. https://doi.org/10.1038/s41558-021-01263-8

Jünger, S. (2019). Using Georeferenced Data in Social Science Survey Research. The Method of Spatial Linking and Its Application with the German General Social Survey and the GESIS Panel. GESIS Series, 24. Cologne: GESIS – Leibniz Institute for the Social Sciences: GESIS für Sozialwissenschaften. https://doi.org/10.21241/ssoar.63688

Kassambara, A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. https://CRAN.R-project.org/package=ggpubr

Kaufman, L., & Rousseeuw, P. J. (1990). Finding Groups in Data. Wiley Online Library. Advance online publication. https://doi.org/10.1002/9780470316801

de Mendonça, A. L. A., & Delazari, L. S. (2012). Remote Evaluation of the Execution of Spatial Analysis Tasks with Interactive Web Maps: A Functional and Quantitative Approach. Cartographic Journal, 49(1), 7–20. https://doi.org/10.1179/1743277411Y.0000000020

Muenchow, J., Schäfer, S., & Krüger, E. (2019). Reviewing qualitative GIS research—Toward a wider usage of open-source GIS and reproducible research practices. Geography Compass, 13(6), e12441. https://doi.org/10.1111/gec3.12441

Muenchow, J., Schratz, P., & Brenning, A. (2017). RQGIS: Integrating R with QGIS for Statistical Geocomputing. The R Journal, 9(2), 409–428. https://doi.org/10.32614/RJ-2017-067

Nepusz, T. (2022). igraph: Network Analysis and Visualization (1.3.3). https://CRAN.R-project.org/package=igraph

Ognyanova, K. (2021). Network Visualization with R. www.kateto.net/network-visualization

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal, 10(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Pebesma, E. (2021). stars: Spatiotemporal Arrays, Raster and Vector Data Cubes. https://CRAN.R-project.org/package=stars

Pedersen, T. L. (2022). tidygraph: A Tidy API for Graph Manipulation (1.2.1). https://CRAN.R-project.org/package=tidygraph

Pedersen, T. L., & RStudio. (2021). ggraph: An Implementation of Grammar of Graphics for Graphs and Networks (2.0.5). https://CRAN.R-project.org/package=ggraph

Perdana, A., Robb, A., & Rohde, F. (2018). Does Visualization Matter? The Role of Interactive Data Visualization to Make Sense of Information. Australasian Journal of Information Systems, 22. https://doi.org/10.3127/ajis.v22i0.1681

Robin Lovelace, J. N. (2019). Geocomputation with R. Taylor & Francis. https://doi.org/10.1201/9780203730058

Roth, R. E. (2013). Interactive maps: What we know and what we need to know. Journal of Spatial Information Science, 2013(6), 59–115. https://doi.org/10.5311/JOSIS.2013.6.105

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

R Special Interest Group on Databases, Hadley Wickham, & Kirill Müller. (2021). DBI: R Database Interface. https://CRAN.R-project.org/package=DBI

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. The R Journal, 8(1), 289–317. https://doi.org/10.32614/RJ-2016-021

Sievert, C. (2020). Interactive Web-Based Data Visualization with R, plotly, and shiny. Taylor & Francis. Advance online publication. https://doi.org/10.1201/9780429447273

Struyf, A., Hubert, M., & Rousseeuw, P. (1997). Clustering in an Object-Oriented Environment. Journal of Statistical Software, 1, 1–30. https://doi.org/10.18637/jss.v001.i04

Talen, E. (2000). Bottom-Up GIS. Journal of the American Planning Association, 66(3), 279–294. https://doi.org/10.1080/01944360008976107

Tennekes, M. (2018). tmap: Thematic Maps in R. J. Stat. Soft., 84, 1–39. https://doi.org/10.18637/jss.v084.i06

Ushey, K., Allaire, JJ., & Tang, Y. (2022). reticulate: Interface to ’Python’.

Wickham, H. (2010). A Layered Grammar of Graphics. Journal of Computational and Graphical Statistics, 19(1), 3–28. https://doi.org/10.1198/jcgs.2009.07098

Wickham, H. (2016). Ggplot2. Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H. (2021). Mastering Shiny: Build interactive apps, reports, and dashboards powered by R. O’Reilly.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Lin Pedersen, T., Miller, E., Milton Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the Tidyverse. Journal of Open Source Software, 4(43), 1686. https://doi.org/10.21105/joss.01686