



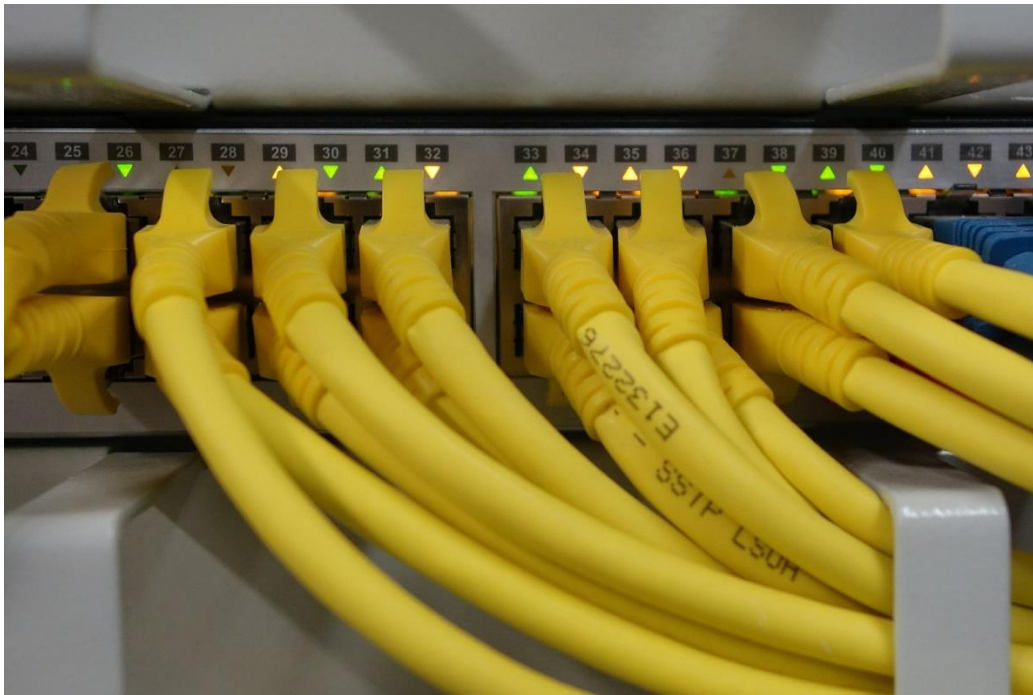
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727470



Report No ECHOES 2.2 - D2.2 ECHOES database description

Report

D2.2 ECHOES database description



Picture under creative common licence. © Michael Schwarzenberger at pixabay.com

Report

D2.2 ECHOES database description

KEYWORDS:

Database

Open access

User group

Data

VERSION

1

DATE

30.11.2018

AUTHOR(S)

Suvisanna Correia, Lassi Similä, Kalevi Piira, Lotta Kannari & Tiina Koljonen

QUALITY CHECK

Jaime Andres Rodrigues, Christian A. Klöckner & Stephan Schwarzsinger

PROJECT NO.

727470 (H2020)

NUMBER of PAGES/APPENDICES:

42/2

ABSTRACT

This report outlines the components of the open access database to be built in ECHOES. The points discussed include the user interface design, functionalities required, data content as well as different alternatives regarding the technical implementation of the database. An updated action plan, setting the milestones for developing the running solution, is presented as a conclusion. Key technical choices of the database are studied based on preliminary tests and coding of a relational (SQL) database structure, allowing flexible integration of research data that will be accumulated towards the end of the ECHOES project. The tests have been mainly built around the requirements set by the characteristics of ECHOES results and/or their draft versions, most notably, the multi-national survey implemented in 31 European countries. A user interface is designed, and it is considered to interact with the database through a map-based user interface with data presentation options based on main survey results and selected Key Performance Indicators (KPIs) as basic components. The KPIs will be selected based on selected SET Plan actions, such as Action 3.1 on energy consumers.

REPORT NO.

D2.2

ISBN

NA

CLASSIFICATION

public

CLASSIFICATION THIS PAGE

public

Document history

VERSION	DATE	VERSION DESCRIPTION
1	15.11.2018	First version for quality check
2	28.11.2018	Final version sent to European Commission

Table of contents

1	Introduction	5
2	ECHOES database development.....	7
2.1	<i>General overview.....</i>	7
2.2	<i>Summary of earlier project phases relevant for the database development.....</i>	7
2.3	<i>Objectives of database development in ECHOES</i>	8
	Satisfying external and internal needs	8
	Maximizing use and impacts while taking barriers into account	10
	Target audience and impact.....	11
2.4	<i>Scope of this document in the database development process</i>	12
3	Targeted database solution and its components	13
3.1	<i>Functionalities</i>	13
	Map application for exploring the European overview	13
3.2	<i>Users and access groups for database</i>	14
	User group narratives.....	14
	Access Groups for ECHOES database	15
3.3	<i>User interface</i>	15
	Data Clouds	17
	SET Plan Key Performance Indicators.....	19
3.4	<i>Database experiment with survey draft.....</i>	19
3.5	<i>Database for KPI values.....</i>	21
3.6	<i>Database system architecture and implementation technologies.....</i>	23
4	Data content.....	26
4.1	<i>Data collection and data types.....</i>	26
4.2	<i>Data services for audience (open access groups).....</i>	27
4.3	<i>Data for direct ECHOES research use</i>	27
5	Database implementation.....	30
5.1	<i>Schedule and major milestones.....</i>	30
5.2	<i>Partner responsibilities.....</i>	31
	Responsibilities explained in detail	31
	Data security and information protection.....	32
6	Summary and conclusions	33
7	References.....	35
8	Appendices.....	36

1 Introduction

As described in previous project documents and based on the project proposal (ECHOES 2016), “the overarching objective of ECHOES (Energy CHOICES supporting the Energy union and the Set-plan¹) is to unlock the policy potential of an integrated social science perspective bounded by central socio-cultural, socio-economic, socio-political, and gender issues that influence individual and collective energy choices and social acceptance of the energy transition in Europe. ECHOES will therefore foster the implementation of the European Strategic Energy Technology Plan (SET-Plan) and advance the Energy Transition, in addition to the decarbonisation of the EU’s future energy system” (Deliverable 2.1, Similä & Koljonen 2017).

There are three technological foci and three perspectives on “energy collectives” addressed in the ECHOES project. That is, “central to all research activities in ECHOES are the technological topics of a) smart energy technology, b) electric mobility, and c) buildings. ECHOES addresses the challenges by employing the innovative theoretical concept of “energy collectives” which covers determinants of energy choices from the perspective of (1) individual decision-making as part of collectives (micro), (2) collectives constituting energy cultures and life-styles (meso), and (3) formal social units (macro).” (Deliverable 2.1, Similä & Koljonen 2017)

Work Package 2 (WP2) of the ECHOES project (*Formulation of SSH database and SSH indicators*) aims to present the relevant data collected in ECHOES in an accessible, widely utilisable and visually attractive format. Hence, the “goal of the multidisciplinary ECHOES project’s Work Package 2 (WP2) includes collecting, synthesizing, and curating relevant data, in addition to formulating indicators related to individual and collective energy choices and behaviour” (Similä & Koljonen 2017). Summarized from Grant Agreement of the ECHOES project, the primary objective of WP2 is to create an open access database to contain all data produced in ECHOES. The ECHOES database, according to the Grant Agreement, will display how different national and EU governance frameworks affect decision-making processes and individual as well as collective choices. Regarding the data content, the Grant Agreement states WP2 will combine SSH relevant data from existing databases to an appropriate extent but will especially focus on creating and collecting new quantitative and qualitative data. Furthermore, the database collects all the quantitative and qualitative information from other ECHOES Work Packages (WP3-WP7) and creates an open source database, which will support national and EU level decision makers and stakeholders as well as modelling and the SSH research communities (ECHOES 2016).

In addition to the high-level description of the goals and foci of the ECHOES project and the database as described above, in practice, there are many specific questions to be solved in order to develop a highly usable and user-friendly database highlighting the results of the ECHOES project. The specific questions include:

- How to define or select the SSH data to be stored in the database appropriately and ensure its usability?
- Which software tools, data formats and technical architecture are needed for the implementation?
- Which different user groups are there? Do the requirements of these groups for the database differ and how?
- What kind of design of the user interface and what functionalities should be implemented?
- How can the data be stored with open access so that the General Data Protection Regulation is satisfied?
- Timeline: as the results of the ECHOES are to full extent only available at the end of the project, how to take this into account in database development?

This document is the “*ECHOES Database Description*”, Deliverable 2.2 of the ECHOES project, which sheds light on the specific database issues mentioned above and how they are tackled in the database design and implementation. Generally, the aim of the deliverable is bringing the description of the database to a more practical

¹ see <http://ec.europa.eu/energy/en/topics/technology-and-innovation/strategic-energy-technology-plan>

level compared to earlier efforts focusing on state-of-the-art of the databases and classification and analysis of their contents as e.g. in D2.1 (Similä & Koljonen 2017). The document is highly based on the technical development, specification and testing work done at VTT and in the ECHOES consortium over the year 2018. As another important event, a database workshop arranged in Rome in September 2018 collected views and suggestions from the consortium on database development on a broad basis. The suggestions are introduced in this document and, to selected extent, further used in forthcoming phases of the database development.

This document describes the key technology choices, data contents, as well as main building blocks of the database solution under development. The content of the document is described in the following:

1. **“Introduction”**: This Chapter puts the content of the deliverable into a general context, introduces the purpose of the ECHOES database, and describes the content of the document.
2. **“ECHOES database development”**: This Chapter introduces key design components of the database, their relation to project phases, summarizes relevant earlier task results, and discusses the internal and external needs as different drivers for the database development.
3. **“Targeted database solution and its components”**: This Chapter specifies the design choices for the ECHOES database (general target state of the database: including user interface, database architecture (technical parts), and functionalities). This part is based on ongoing specification work of the database completed in the ECHOES consortium so far.
4. **“Data content description”**: This Chapter describes national, regional, and EU level quantitative and qualitative data relevant for the database. Both data from other WPs as well as data from external sources for research purposes are discussed. Questions of data protection are also discussed here.
5. **“Database implementation”**: This Chapter introduces a concrete plan to timely implement the database. Schedule, major milestones, as well as partner roles and responsibilities are described.
6. **“Summary and conclusions”**: This Chapter summarizes the work conducted on the database development and concludes on the next steps.

According to current plans, the database will be opened up *to the public* only after the project has ended. For the database to achieve wide popularity and its benefit potential to be realized, it is important to pay attention to efficient communication and dissemination of the database. As one of the key goals of the ECHOES project is to foster the implementation of the SET-Plan through its results, one potential major event for such actions could be SET-Plan Conference in Helsinki in November 2019.

2 ECHOES database development

2.1 General overview

In WP2, VTT is responsible for creating the open access database, which will collect and formulate Social Sciences and Humanities (SSH) related quantitative and qualitative data from existing national, regional and EU-level databases and most importantly, the results from the other WPs of the project. An open access database to contain all data produced in ECHOES – with certain exceptions based on data protection issues – is created. The ECHOES database will display how different national and EU governance frameworks affect decision-making processes and individual as well as collective choices. The envisioned database solution and its components and implementation plan are described in detail in Chapters 3, 4 and 5.

Particularly, key content of the ECHOES database is based on the multi-national survey implemented as a major ECHOES result in 31 European countries and Key Performance Indicators (KPIs) developed. According to its specification as exploitative result, the database “will allow external users to run queries in the main ECHOES survey data, explore key performance indicators (KPIs) based on the survey and statistical data and additional structural data. Country-based representation allows getting insights into the country-wise distribution of quantitative and qualitative ECHOES results and country comparisons. This will be provided in the form of a web-based frontend accessing the raw data stored in the database.” (Jimenez Iturriza 2018)

As its key content, “the ECHOES database will contribute a large dataset for drivers of consumer based energy choices from all 28 EU Member States and Switzerland, Turkey and Norway. This data is new and will focus particularly on the ECHOES technology foci and the ECHOES perspectives (identity factors, lifestyles, socio-economic drivers).” (Jimenez Iturriza 2018)

A large number of potential users and exploiters of the database (scientific community, policymakers, “general public”/consumers, as well as different business actors, primarily in energy sector) have been identified in different efforts of the project. To enable the relevant user groups to exploit the database, a tailored database solution is developed considering the needs of different user groups. Key components of the database solution include the user interface, structures for data stored, as well as technical and software solutions necessary for database construction.

2.2 Summary of earlier project phases relevant for the database development

The conceptualization of the ECHOES database started with analyzing the existing national and EU level databases with a relevance from the viewpoint of SSH data or energy transition disciplines. The primary goal of this phase was to identify existing gaps in quantitative data, which are needed in policy and other decision-making and by research communities. Results of this phase were documented in Deliverable 2.1 (D2.1) published and prepared in an early phase of the project in August 2017. The results of D2.1 were largely based on questionnaires of ECHOES consortium members and selected external experts, and desktop research on existing databases in the area. In order to add more external views to the first phase, an expert workshop was arranged in Brussels in October 2017.

D2.1 categorized Social Science and Humanities (SSH) and energy related data relevant to the project and studied characteristics of identified databases in the field. Overall, nearly 130 suggestions of relevant databases – or data sources – from different national, global and European contexts were classified and analyzed. According to results of D2.1, the variety of SSH data and databases are abundant in many substance areas per se, and providers of data portray a large variety. Yet, data sources were found to show very diverse formats and contents, and in many cases, search and data classification functions typical to ‘genuine’ databases were lacking from the data sources suggested. Particularly, national statistical agencies are frequent providers while ministries are a typical source of relevant data. Furthermore, the spectrum of other data providers ranges from research institutions to private companies to NGOs and industry organisations. Population data was assessed as the best-covered topic, whereas

areas of energy cultures, consumer behaviour, acceptance, and gender issues were assessed as areas with the weakest data coverage. Despite the results being characterized as indicative, they offer utilizable views preparing the work on the ECHOES database. Most importantly, as several general databases relevant for SSH and energy data were identified, the results of D2.1 strengthened the view that a key target in further design and implementation of the ECHOES database is in finding the most efficient solution for presenting the primary data achieved in ECHOES. This is due to the fact that the ECHOES database has a potential to contribute to fill the gaps identified especially in areas of energy cultures, consumer behaviour, acceptance, and gender issues. Hence, the importance of presenting the ECHOES results in a widely utilisable database was outlined. Furthermore, review work conducted in D2.1 provided an overview about different options for user interface solutions and functionalities included in the studied databases, which also feeds into the development of the ECHOES database solutions.

Decisive factors suggested in the expert workshop held in Brussels in October 2017 (Table 1) offer further guidelines for relevant data to be presented in the database. Many of the variable needs can be tackled through primary ECHOES data, whereas some should more naturally be based on external sources (especially economic factors including subsidies and costs). Thus, making choices about data content and structure supporting the factors presented in Table 1 is one of the key issues in the database development and in this report.

Table 1. Factors that drive energy choices.

General decisive factors that drive the energy choices supported by the workshop results including external participants. The list is non-exhaustive (Source: D2.1).
<ul style="list-style-type: none"> • Economic factors including subsidies and costs • Regional differences • Socio-demographic factors • Ease of a solution • Environmental and ecologic consciousness and lifestyles • Awareness and level of knowledge

2.3 Objectives of database development in ECHOES

Satisfying external and internal needs

Two key foci in database development have been identified during the course of the ECHOES project. That is, the main distinction can be described between

- (1) **"Public database"** – the primary goal is to present all the ECHOES results in an easily accessible open access database, presenting "the legacy of ECHOES" after the project
- (2) **"Internal database"** – to develop a data solution guaranteeing harmonized and timely input data directly for ECHOES analyses during the project.

One of the main goals of the public database is in storing all the data produced in ECHOES according to open access principles². For the purposes of developing the database, the public database includes more choices to be made and more open issues to be addressed. That is, there presumably are many different types of potential external users with their own needs. These needs are reflected in a larger variety of design options, and correspondingly, in a larger spectrum of technical specifications. Different options for external users may include fine-tuned user interfaces and user-friendly data structures, which may not be that emphasized in purely research oriented use. In comparison, research data for internal purposes typically satisfies tailored, highly specified needs.

² Regarding open data management, the European Commission has launched s.c. FAIR Data Management Principles (Findable, Accessible, Interoperable, and Re-usable).

Hence, the main efforts of the development work conducted have been concentrated on goal (1), i.e. “the public database”. In the database for external project users, important design objectives include the components described in Table 2. Thus, in the development work of the database, the following questions are to be answered.

Table 2. Components to be tackled in ECHOES database design

Component	Questions to be solved
Identification of different user groups and their needs	Which question would a user like the ECHOES database to answer? Which data does a user look for? Which ECHOES results interest him/her?
Functionalities to be implemented in the database	What kind of functionalities should the database include for certain users? What data, variables or indicators should be stored in the database to support this?
Design of the user interface	What kind of user interface [web site layout, menus, search functionalities] would satisfy the user requirements?
Visualisation of the content	What possibilities are there for the visualisation of research data and indicators?
Technical implementation of the database	Software, data repositories, interfaces, programming, data formats etc.

As described in Table 2, in addition to specifying the contents and structures of the database, appropriate technical solutions need to be developed in order to build a working database solution. These include software, programming, interface structure, data repositories, questions related to hosting of the database etc. The technical solutions evolve with the availability of project data/results and they will take their final form during the lifetime of the project. However, the different options have been mapped and preliminarily studied to a certain extent. These issues are discussed in sections 3.5 and 3.6. As another example, the Zenodo platform has been studied as a repository for open access data to store the qualitative and quantitative raw data (e.g. survey data, anonymized interviews, literature surveys, etc.) and make it available for secondary use (see D2.1). While the Zenodo platform is a potential solution for storing raw data, as a part in the ECHOES database solution, the variety of functionalities and user interfaces designed during the database development process indicate that the ECHOES database calls for more tailored solutions. Therefore, the development work has not solely been built on the Zenodo platform but primarily on technologies allowing more versatile implementation.

For internal ECHOES project needs (2), data collection and processing has a more tailored nature than solutions developed for the general public. There are differences in the time schedule, as well. As the internal data aims at serving the needs to produce the project results, a solution is needed before the public database launch that is planned to take place soon after the project or near its end, while the date of the launch has not been set at this point. A working version of the database will be built well in advance the end of the project, allowing a possibility to test and review the database even before it contains all ECHOES data. Regarding internal project needs, specification of data and questions on its collection and delivery: “who, what, when, how” need to be solved. The “Internal ECHOES database” is built to directly serve specified research needs within the project. However, there may be synergies between the internal and external targets. For example, a solution serving internal project needs could be designed to serve the external needs, too, such that the internal version could be used as a testbed for the public ECHOES database development. In addition, external data collected for internal use may serve the goals of developing indicators. If the internal database is open to project partners only, the external data might be more easily manageable from the viewpoint of permissions, etc., allowing a more flexible test use to a restricted group of users. However, if the internal project needs can be served through a lighter solution than a “genuine” database, a more tailored form of data delivery for specified cases – such as Excel spreadsheets – would not serve as a full-scale test environment and it calls for a separate effort.

In development of the ECHOES database for compilation of data primarily from external sources, there are wide possibilities for its contents. There are many data types potentially relevant for the internal ECHOES database already mentioned in the Grant Agreement and D2.1 (see Table 3).

Table 3. Quantitative and qualitative data (Source: D2.1).

Task 2.1: Quantitative data	Task 2.2: Qualitative data
Compilation of national energy and emission taxes and supports (incl. consumer and prosumer supports), consumer and prosumer energy prices, size of energy bills for different income groups, population structure, share of urban/rural population, mobility and building related data, etc.	Compilation of data on acceptance of new energy technologies and services, vested interests, level of awareness as well as cultural, and religious aspects, level of democracy/transparency of policy making processes (government autonomy), the design of administrative and legal approval processes (strength of the bureaucracy), geopolitical issues, national implementations of relevant EU Directives, national implementations of building codes, interest group strength, alliance structures (interest groups-politicians-bureaucracy), etc.

However, a need to specify the content of external data to be compiled in accordance with Table 3 has become evident as the database development has taken steps forward in the project. For example, the possibilities to present “mobility and building related data” can be enormous. The most important variables for data collection have been specified as those to be used directly in ECHOES analyses, and concrete compilation has concentrated around these variables (see Chapter 4). The data collected for ECHOES analyses has been specified from a wider list of options described in Table 3. Data collection conducted up to the point of this deliverable was prioritised so that the variables collected can be used in analysis of the multi-national survey, which is due to have results in November 2018. However, the external data can simultaneously be used to test and design the user interface for the database. Especially, a part of the ECHOES project combining variables such as national energy taxes or subsidy schemes with data collected in the multi-national survey and testing their explanatory power on decision-making, were used to select a prioritized external dataset to be considered in the database. The external data was compiled until November 2018 as described in Chapter 4. However, an option to enhance the external data collection towards the end of the project, to be presented in final version of the database is not ruled out.

Maximizing use and impacts while taking barriers into account

The ECHOES results help identifying best practice governance strategies for different types of formal social organizations, targeting specific technologies and desired changes in practices or behaviour. These results shall be available, visualized and retrievable in the ECHOES database. The material contained in the ECHOES database may include any type of data collected, synthesized, and formulated; quantitative and qualitative data related to individual and collective energy choices and behavior.

WP2 will combine SSH relevant data from existing databases to an appropriate extent but will especially focus on creating and collecting new quantitative and qualitative data. The database collects all the quantitative and qualitative information from WP3-WP7 and creates an open source database, which will support national and EU level decision makers and stakeholders as well as modelling and the SSH research communities.

To ensure high impact and usefulness, the ECHOES database is based on open access principles. Quantitative and qualitative data includes descriptive and comparative data units involving policy tools in effect, regarding energy choices for formal social units. Concerns, acceptability, ambivalence and involvement of the targeted stakeholders will be assessed according to the variables.

Whereas opening up all the data produced in the project and making it available works as a key criterion for designing the ECHOES database, certain barriers that need to be considered in development of the database have been identified, including

- (1) privacy of the respondents in surveys and/or interviews, including compliance with GDPR
- (2) copyright issues preventing storing of external data in the database.

Especially, these issues mean that tailored solutions have to be used in cases where the barriers become restrictive. Thus, other ways to represent the results in these cases are under special interest. There are several potential means such as aggregating the data and/or using different visual representations such as word clouds, semantic clouds or different types of charts. How to actually deal with the issues (1) and (2) in different ECHOES cases while allowing a sufficient and clear representation of the data, is one of the key questions in dealing with the database development.

Target audience and impact

The ECHOES database will be designed to be used by several levels of policymakers, such as EU, municipalities or governments of member states, in guiding their decisions, as well as for developers of future projects and further research.

In addition to different levels of policymakers, several other potential user groups have been identified in the work, as initial ideas of the target audience were already sketched in the project plan. The following list of database users is based mainly on the results of the workshop held at the General Assembly in Rome in September 2018:

- **Scientific community:**
 - Energy-related behaviour researchers
 - Energy economics researchers
 - Energy technology researchers
 - Social science researchers
 - Researchers in energy technology innovations and diffusion
 - Researchers in energy, climate and innovation policies
- **Policymakers:**
 - European Commission
 - European Parliament
 - National Regulatory Authorities
 - National and regional (e.g. cities, municipalities) policy levels
- **Industrial/commercial community:**
 - Utilities
 - ESCOs
 - Distributed renewable energy producers
 - Car-sharing companies
 - Energy, housing, mobility associations for companies, private consumers, etc.
 - Smart energy technology providers / producers
 - Electric mobility producers / providers
 - Energy cooperatives (housing, small communities, etc.)
 - Grid operators
- **Consumers, citizens**

- Other
 - Media
 - NGOs
 - Regulators
 - Energy collectives

As can be concluded from the above, the ECHOES database potentially serves many purposes and groups. Serving the different needs is reflected in the design of the database and is discussed in Chapter 3.

2.4 Scope of this document in the database development process

This document describes the development plan for the ECHOES database, including goal, technology choices, schedules, key components and the choice options, as well as partner responsibilities. In particular, it increases the level of concreteness specifically related to the ECHOES project database development. Thus, it works as a natural addition to D2.1, whose target is more generic and focused on reference databases external to the project or 'state-of-the-art' of the databases in the fields of SSH and energy transition related data.

Despite giving a broad overview of relevant comparable databases and data, D2.1 did not go into detail about technical implementation of the database. It identified data produced in ECHOES on a high level, identified some barriers such as related to data privacy issues, legal issues, and copyright issues. Concretization of data to be produced by the other ECHOES WPs after D2.1 has allowed further development of technical solutions, data transfer etc. and parallel preliminary design of user interfaces etc. Thus, the work completed after the release of D2.1 allows this document to take a step further in the ECHOES database specification. In summary, key areas and questions discussed in this document are presented in Table 4.

Table 4. Components of the database worked on.

Component	Specifications/ Questions	Milestones/ schedule	Main approaches
1. User interface	user groups, access groups, visualisation, functionalities	identification of key user groups, interface designed	ECHOES database workshop at General Assembly, specification and design work done in the consortium
2. Data content and structures	variable specification, table definition, aggregation method, queries	main survey data and structure used in testing, external data prioritised and collected, tests on database conducted	ECHOES database workshop at General Assembly External data search, completing ECHOES tasks Discussions between key partners
3. Technical implementation of the database	Software, interfaces, data repositories etc.	working software solution	Discussions between key partners, VTT development work

Project phases completed have concentrated around components 1 and 2, whereas preliminary work has been conducted related to details on technical implementation of the database. Especially, multi-national main survey drafts conducted in the ECHOES project in 31 European countries have allowed testing, preliminary coding and consideration of data structures. However, coding and implementation of the database will be fully executed in the forthcoming phase of the project.

3 Targeted database solution and its components

This Chapter specifies the design choices for the ECHOES database: namely the general target state of the database, including functionalities, user interface and technical description of the database architecture. The development plan described in this Chapter is based on specification work completed within the ECHOES consortium.

Research results from other ECHOES WPs will have an important role in framing the final decisions about handling and presenting data in the forthcoming ECHOES database. However, due to their independent schedules, complete research results from other WPs have not been available by the time of writing this report and designing and collecting ideas for the functionalities needed for the database. Therefore, it is important to notice that the current database plan has areas that are subject to further dialogue between the ECHOES partners, and allow some flexibility in the plan describing the database. The database description is expected to dynamically evolve until the end of the project to complete parts that accommodate for forthcoming research results and other potentially emerging issues. However, generic principles and guidelines are proposed in this report, based on design work conducted so far. Especially, draft versions of parts of the “Main survey” from WP 4 and WP5 have been utilised in the preliminary design.

3.1 Functionalities

Functionalities of a database refer to the services the database performs and to the components related to these actions. The fundamental functionalities of a database are storing data and retrieving it by making queries into the stored data. There is an extensive variety of possibilities how these and additional tasks can be implemented. The database workshop at the ECHOES General Assembly in Rome, September 12th, 2018, discussed various specifications, user needs and ideas for the ECHOES database development in order to arrive at a common understanding of how the ECHOES database should function.

The questions to answer regarding database functionalities to be implemented include:

- What kind of functionalities should the database include for certain users?
- What data, variables or indicators should be stored in the database to support this?

This Chapter outlines the functionalities of the ECHOES database and discusses primarily the database service for various end-user groups. End-users interact with the user interface, which is described later in this Chapter. The user interface is also often referred to as the front-end of the entire database service, and it is usually based on a web browser application. After interaction, the user interface calls the server that performs database operations. The server side, also referred to as back-end, is where the actual database would be located. However, functionalities related to database management are not discussed here, as they are limited to interest of administrator users, and will be designed at a later stage of database development.

The main functionality available at the front-end of the ECHOES database would thus be a data search in the data contained in the database. Assimilated from earlier plans and the results of the database workshop in Rome, the prospective content and related functionalities of the ECHOES database are outlined next and in Appendix A.1, whereas the complete database workshop results are documented in Appendix A.2.

Map application for exploring the European overview

The open area of the ECHOES database website is envisioned to offer a quick overview to SSH aspects of the European energy transition. There is a growing common understanding in the consortium that the ECHOES front-end would be implemented as an interactive map application, that offers a number of possibilities for exploring data and brings Europe into highlight as the focus of the content. Some possibilities of a map application include

embedded passive content, available e.g. by mouse-over, a simple search using one term, and advanced search options to select countries, years or other variables. Any of these could be displayed on the map in the context of one or more countries being examined.

Results that are accessible to all public could be primarily presented in a map application, other visualisations and embedded tables that show the variables and their values. The map application is envisioned to include a heat map of ECHOES results or other variables, to compare one variable for all countries for quick overview. A heat map is a two-dimensional representation of data, often a map. On a heat map, values of data are represented by colours. Colours of a heat map help communicate relationships between data values that would be harder to understand numerically, e.g. in a table.

In addition, a spider web chart application would allow a user to compare from two to five countries against several selected variables. In addition, up to four charts could be displayed on a page simultaneously. Screen tests will help determine the acceptable parameters for web charts to maintain high-quality of representation. The variables could be related to individuals, such as awareness, acceptance in one's region or social group, age, socio-economic status or skill level regarding any of ECHOES technological foci. These could be also related to the structural variables, e.g. for buildings, type of house, energy choices, energy expenditure and consumption, or prices and demand of specific technologies in a country.

In addition, the open area could include details about survey methods, data generation, various documents for download and links. More data and information would be available to registered users on equally open access principles, while allowing the database administrators a better overview of that group of users. Different user groups and access possibilities are described and discussed next in part 3.2., and more details and possibilities for user interface development are covered in 3.3. Data content and possibilities are discussed in Chapter 4.

3.2 Users and access groups for database

User group narratives

The database workshop organised in Rome in September 2018 collected assumptions about various user groups that might benefit from the ECHOES database. Different user groups could interact with the ECHOES database in many ways. In the following, the main cases are summarized.

Policymakers at different levels of public decision-making face extensive amounts of information and need an access to preprocessed, clearly organised and relevant, reliable information. From the European level to regions, national and local levels, information that is directly relevant to that particular area and level, is presumably preferred.

In the energy sector, there are two large groups to consider: the **public administration** that oversees that the actors in this sector fulfill their legal obligations, and **private companies** that are looking at their opportunities, development trends and customer satisfaction. The latter is true also for other industries and commerce connected to technologies in buildings and construction, electric mobility and smart technology.

Universities and other research institutes may benefit from the information provided in the ECHOES database. Research professionals would tend to search detailed information and original data sets related to their particular field of interest, perhaps primarily in energy, social science, economics, spatial planning, sustainability issues and policy studies related to any of these.

Citizens act in the civil society in various roles. As actors in climate movement, other types of NGOs, or professionally, e.g. as a teacher, they may be interested in understanding the energy transition and how to influence it. In addition, individual citizens share and discuss information and influence each other through social media. As

consumers making energy decisions, they may buy energy efficient or smart devices. They may also invest in renewable energy production or energy efficiency either individually or collectively, and may benefit from accounts of peer experiences, while this database is not directly intended for consumer advice.

Access Groups for ECHOES database

Based on information gathered about potential database users, four major user types have been identified for the ECHOES database. These groups are described below in Table 5 as Access Groups. Each Access group relates to certain user interactions and access rights to the ECHOES database, for the purposes of the technical definition for database development and later for database management. The purpose of the ECHOES database is to provide open access to all results to be accumulated in ECHOES project and the Access groups will be defined to follow this principle.

Table 5. Access groups for ECHOES database

Access group	Access area	Type of data or role
All users	Open access	Public content Data search retrieves data in predefined queries and creates visualisations.
Registered users	Registered user area, free of charge	All above and public data available in multiple formats for download with code book.
ECHOES project partners	ECHOES data area	All above and shared background data for Echoes research. Results from the ECHOES project, if some parts not public. Raw data
Administrator users	Full access	All above and database management

As the first group, **all users** can interact with the ECHOES database's user interface, retrieve data with queries, explore visualisations and view other public content. In addition to this, **registered users** will be allowed to access and download the same content in multiple formats for e.g. research purposes free of charge. User registration is a common practice that allows database providers to gain better control of data downloads and use of data provided. It also allows the ECHOES consortium to evaluate the impact of the data better. The third group is foreseen to be internal **ECHOES partner organizations**, who would be able to access also data that has not been released yet, e.g. due to an embargo period. This group may also serve as pilot users during the testing stage of the database. Finally, **administrator users** need to have rights to perform full database management actions, while they have less interest in the content of the database.

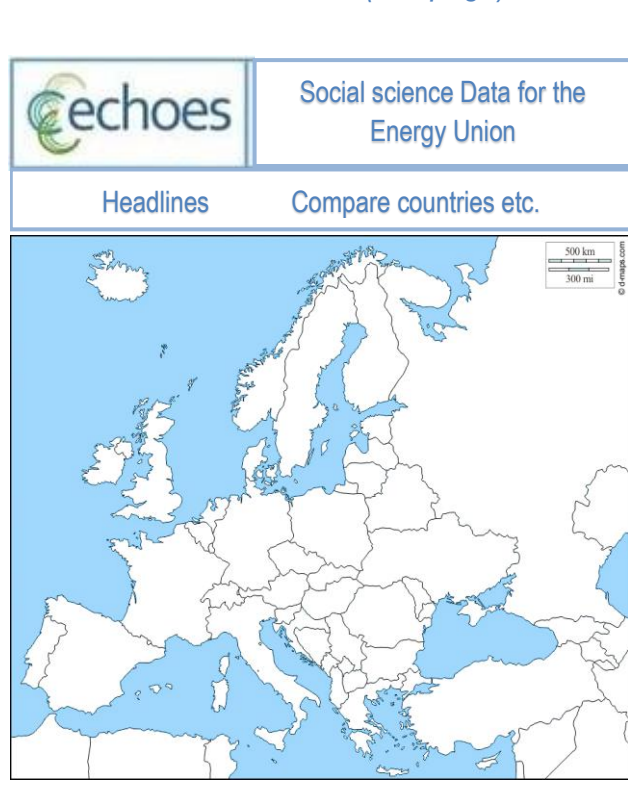
3.3 User interface

The user interface provides end-users with the capacity to interact with the content of a web service page, including a database service such as envisioned for ECHOES. To engage interested users, the user interface needs to be informational, easy to use and interesting. Previous work in ECHOES, in particular in D2.1., has gathered views about main topics to be covered in the database and its user interface. Figure 1 shows one possible solution for the ECHOES user interface, adapted from ideas originally presented by ECHOES project coordinator Christian A. Klöckner, NTNU.

The main element in Figure 1 is an interactive map of Europe that could offer a selection of possibilities to explore quantitative results from ECHOES studies along with other selected variables. These might contain a summary of

available data for KPIs defined in the WP2, including selected KPIs from the SET Plan Actions (especially Action 3.1 related to energy consumers) and some key results from the ECHOES survey.

Figure 1. User interface idea adapted from original image by Christian A. Klöckner, NTNU.

<i>User interface (webpage)</i>	<i>Database functionalities and components</i>
 <p>The screenshot shows a webpage header with the 'echoes' logo and the title 'Social science Data for the Energy Union'. Below the header are two buttons: 'Headlines' and 'Compare countries etc.'. The main content area features a map of Europe with a scale bar indicating 500 km and 300 mi. The map is overlaid with a grid, and a small inset shows a spider chart comparing two countries.</p>	<p><u>ECHOES main survey</u></p> <ul style="list-style-type: none"> - the database stores the quantitative data from the ECHOES database -provides some key results automatically on mouse click (on top of a country) -Some options to query results, more advanced search available for registered users - one variable heat maps - spider chart to compare two countries against several variables <p><u>Download data</u></p> <ul style="list-style-type: none"> -available for registered users - other collections according to availability <p><u>External data</u></p> <ul style="list-style-type: none"> -Available for ECHOES research group only <p><u>Understanding energy decisions</u></p> <ul style="list-style-type: none"> -This part of the database that stores the qualitative data as strings -tagged strings may be searched e.g. by keywords - further interactive analysis options may be sought, implementation subject to availability and feasibility of options - psychological experiments' results included <p><u>- KPI Data</u></p> <ul style="list-style-type: none"> -interfaces to existing, possibly EU wide databases that contain some of the KPIs <p><u>Scientific literature</u></p> <ul style="list-style-type: none"> -data collection based on D3.1 and others as defined -standard/recognized formats -searchable

The central elements offered in Figure 1 aim at presenting the research results of the ECHOES project in an accessible, user-friendly manner. In accordance with this view, different options for exploring and presenting data in the user interface are sought as part of database development and some are described in Figure 1. A map application offers a number of interactive options for exploring data, along with access to more detailed query options and complete datasets for registered users.

The ECHOES main survey consists of quantitative data that can be stored and handled with standard relational database technologies. The action of retrieving data from a relational database is called a query. A query procures data from data tables according to the conditions that a user selects in the user interface. One-variable queries could be used to show comparisons on several countries, either on the map application or in separate or embedded charts and tables. A query variable could be e.g., country, region, gender, age groups, socio-economic groups, just to name a few possibilities. With additional data processing, their relation to the available KPIs could also be explored. Possible key results may be also selected for static presentation on the map. In addition, psychological experiments produce quantitative data that may be handled best in this group.

Under the title of "Understanding energy decisions" resides the part of the database that contains qualitative data. Traditional relational databases have been designed for handling compact quantitative data, whereas qualitative data containing long text strings may call for another type of approach. Qualitative data, including primarily interview texts could be made available in two ways: Either by predefined analyses, e.g., semantic networks (see Figure 2), or by integrating appropriate algorithms to perform interactive analysis in the raw text corpus. This approach depends on the availability and feasibility of algorithms that could be implemented within the ECHOES database

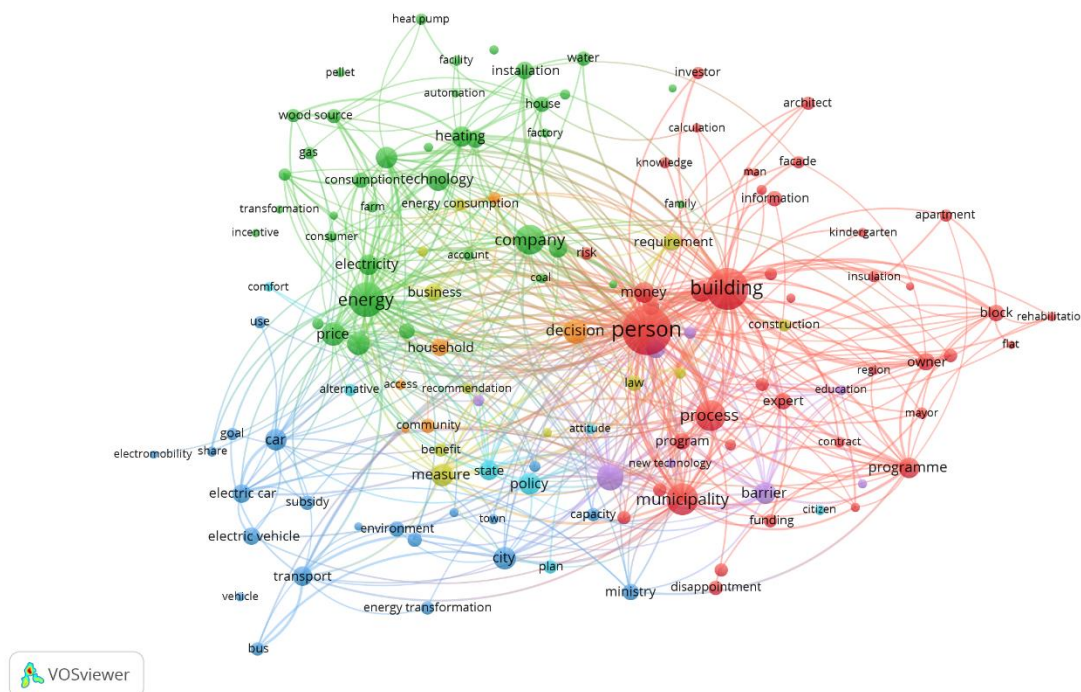
platform. Full texts with user-defined interactive analyses may be possible by e.g. using tagged strings that may be queried. However, at country level at least, only predefined analyses or other static documents could be made available due to data protection requirements. According to our evaluations, the raw text from interview transcripts cannot be made available for direct open access without violating GDPR, as people phrase answers in such way that it is possible to identify them by linking the answers and other documents on the net. For interested users, the database is envisioned to give registered access to further content free of charge, in line with open access principles. The available content could contain full ECHOES data and results, advanced search functions and possibly other relevant downloadable material, e.g. presentations, articles, or videos. However, these would be viewed as additional content and their provision would be subject to availability.

Furthermore, as a service provider and in order to fulfil the Open Data obligations of the ECHOES project, the user interface would provide access to the raw-data for quantitative datasets and related meta-data for download for registered users. In addition, the literature database may be made available, however, it would take an considerable effort to fit all references into a standard literature database format. Thus, the consortium will decide if the benefit of making the results of the extensive literature research available are big enough given the additional effort necessary, and given that in addition literature researches age quickly.

Data Clouds

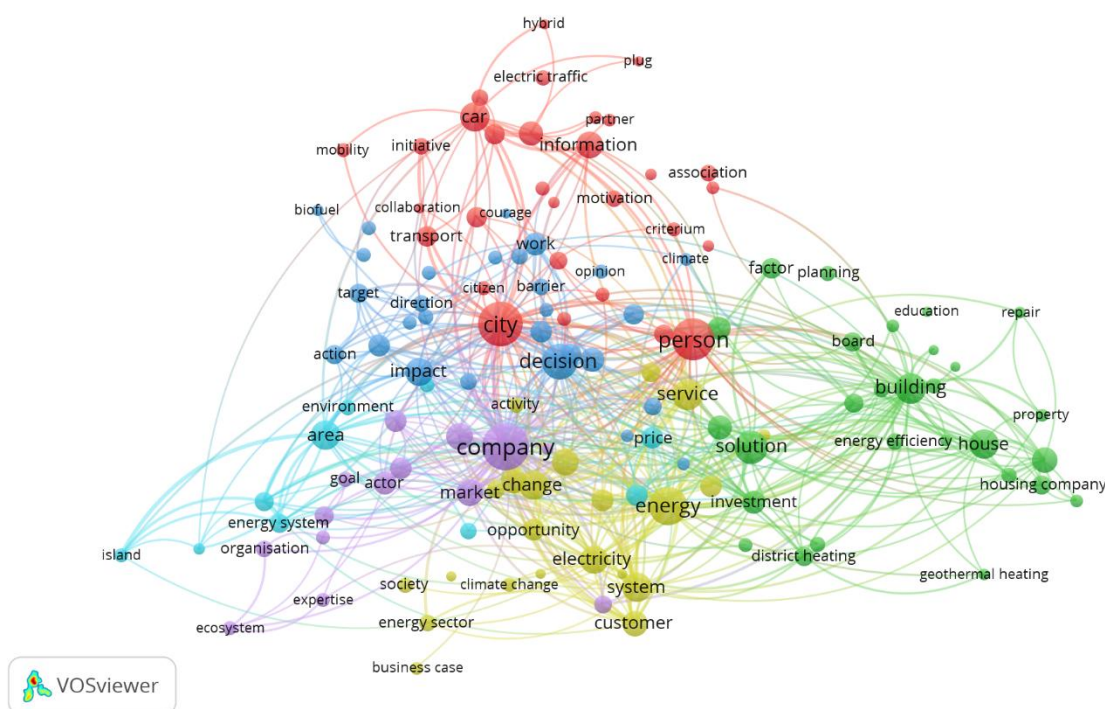
As one option for visualizing ECHOES results in text format, semantic maps (see Figure 2 and 3) have been considered. The option is illustrated in exemplary semantic maps, also referred to as clouds or networks. They have been produced using VOSViewer software. The visualisations are based on individual and focus group interviews concerning collective decision-making related to energy decisions in Finland and Bulgaria, produced in ECHOES WP6.2. That is, the information derived from the interviews and focus groups is inserted in the software and it produces the map-type representation, demonstrating the often represented concepts and their relations to other expressions. This kind of visualisation could offer fast introduction to text-based data while securing the privacy of interviewees.

Figure 2. Data Clouds can depict the most common expressions in decisions-making interviews. Results from Bulgaria. Source: Izaskun Jimenez Iturriza, Tecnalia



For example, comparing Figures 2 and 3 reveals thematic differences in decision-making interviews made in Bulgaria and Finland. In Bulgaria (Figure 2), the interviews have evidently discussed persons and buildings. Words like energy, company and municipality follow in weight, referring to their frequent occurrence. However, in Finland (Figure 3), the most common concepts appear to be company, person, city and decisions. Assuming the interviews in these two countries were carried out following similar study design, these results indicate that there may be a major structural difference in housing and management of buildings in the two countries studied. While these differences could lead to speculations of their reasons, the explanations can be found in the material from the interviews. The high frequency of “company” in Finland arises from interviews that were carried among residents of so-called housing-companies, where apartment owners are shareholders, who own independently sellable real estate.

Figure 3. Data could showing the most common expressions and their relations interviews on collective decisions-making in Finland. Source: Izaskun Jimenez Iturriza, Tecnalia



SET Plan Key Performance Indicators

The user interface proposal in Figure 1 suggests presenting SET Plan Key Performance Indicators on top of the interactive European map. The KPIs that might come into question in connection with key research results arising from ECHOES research are yet to be specified. However, the most relevant KPIs regarding ECHOES may be ones being developed under EU SET Plan Action 3.1 that focuses on energy consumers, but the KPIs from the other two Actions may also include opportunities. The indicators deal with the following areas of action:

Areas of proposed energy consumer Key Performance Indicators

- Involvement of citizens in energy production: Self-production of energy
- Consumer behaviour preferences – smart appliances
- Participation in demand response
- Market participation
- Impact on household energy bills
- Societal readiness level

Indicators for energy consumer actions may inform e.g. on issues such as household PV and bioenergy consumption, participation in energy cooperatives, smart metering and use of energy savings mobile applications, or availability of demand response choices in energy markets. In the following steps of WP2, indicators will be formulated based on the collected data, which can be used for EU level monitoring of the progress of the Energy Union, especially related to solidarity, security, and confidence.

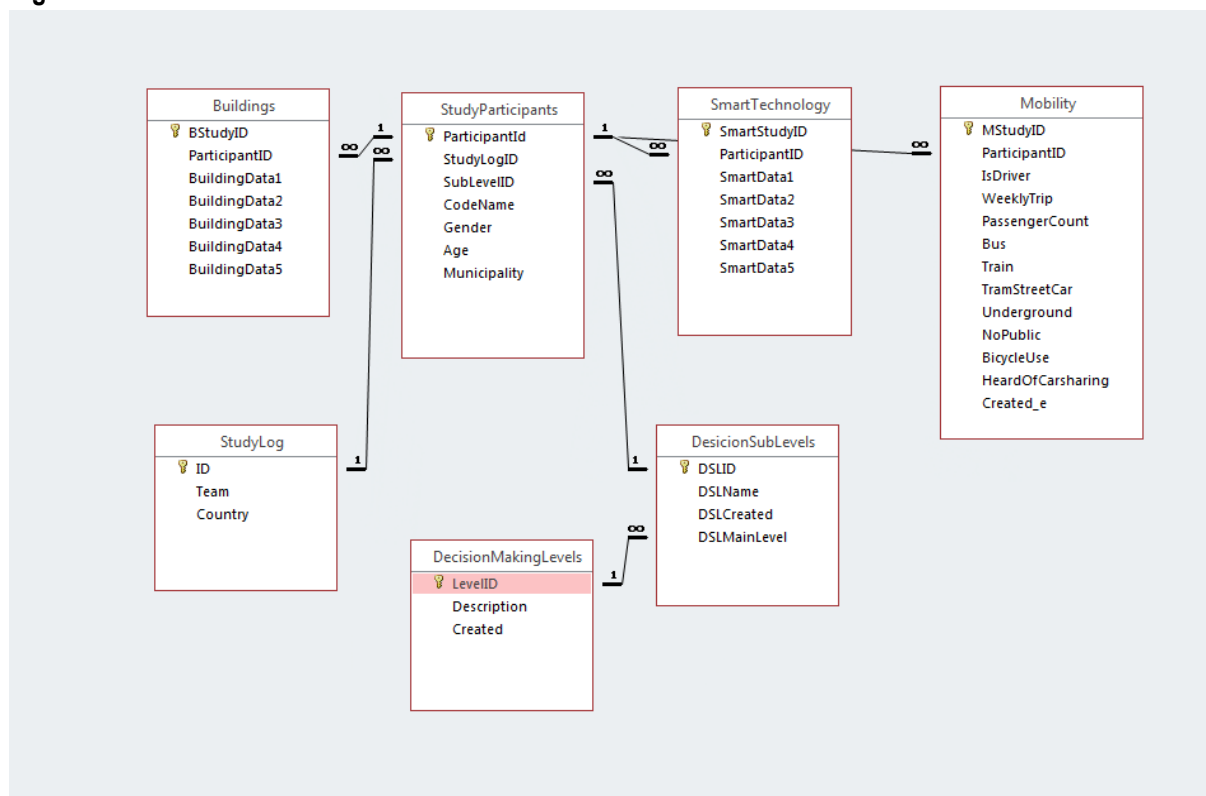
3.4 Database experiment with survey draft

To examine the fit of a relational (SQL) database structure for research data that will be accumulated in ECHOES project, a small experimental database has been developed. MS Access was used as a visual planning tool in this

experiment, which was carried out through June and July in 2018. However, it is worth noting that since MS Access is a desktop application, it is suitable for testing, but not sufficiently robust for a public database service. The server based database development plan is described in detail in Chapter 3.6.

Figure 4 shows an example of a relational database, which consists of several data tables. The relation map shows how data is organised in a relational database. The data tables contain all data included in the database and are neither seen nor can they be directly accessed by a user. The underlying structure defines the relations between the data tables and directs what kind of queries can retrieve data from the tables.

Figure 4. Database schema shows data tables and their relations.



The database in this example is set to organise a content that might consist of a number of ECHOES studies. This example has been developed solely for demonstration purposes. Each of the three technological foci, buildings, smart technology and mobility, are represented. In addition, data of study participants is organised in a table, along e.g. with the country and decision-making level related to the respective study and respondent. The relations between tables determine e.g. that for each country, there could be several studies, but each study has to be identified to one country only. These kinds of rules are set in the database design process, and these decisions define what kind of queries can be made in the database. Figure 5, below, shows an example of a data table, which contains survey responses to questions about different types of residential heating sources. The key columns on the left connect each response to a unique study and respondent.

Figure 5. An example illustrating survey results regarding heating forms in residential buildings; a data table from data schema in Fig 4.

All Access Obj...	Search...	Tables	Relationships	Buildings	DecisionMakingLevels	Mobility	StudyLog	StudyParticipants
		Attitude1						
		Attitude2						
		Attitude3						
		Buildings						
		DecisionMakingLevels						
		DecisionSubLevels						
		Mobility						
		SmartTechnology						
		StudyLog						
		StudyParticipants						

BStudyID	ParticipantID	BData1	BData2	BData3	BData4	BData5	Click to Add
1	1	1 district	electric	geothermal	air source heat		
2	2	3 district	electric	geothermal			
3	3	4	electric				
4	4	2 district			air source heat		
5	5	4 district			air source heat		
6	6	5	electric		air source heat		
7	7	4	electric				
8	8	5		geothermal			
9	9	0 district					
10	10	0	electric	geothermal	air source heat		
11	11	0 district					
12	12		electric		air source heat		
(New)							

A query example with results is shown in Figure 6. Each of the three technological foci was included in this example. As shown on the table on right, a query tool assists in selecting the desired variables for a query. The variables could include e.g. the country, municipality and the type of residential heating from a survey, if a survey included such questions. The table on the left side shows the results, which in this case present some types of residential heating in Finland.

Figure 6. An example of a query to data table and results on residential heating

All Access Obj...	Search...	Tables	Relationships	Buildings	DecisionMakingLevels	Mobility	StudyLog	StudyParticipants
		Buildings						
		DecisionMakingLevels						
		DecisionSubLevels						
		Mobility						
		SmartTechnology						
		StudyLog						
		StudyParticipants						

Country	Municipality	BuildingData1	BuildingData2	BuildingData3
Finland	district	electric	electric	
Finland	district	electric	electric	
Finland	district	electric	electric	
Finland	district			
Finland	1	electric	electric	
Finland	4	electric	electric	
Norway	5			

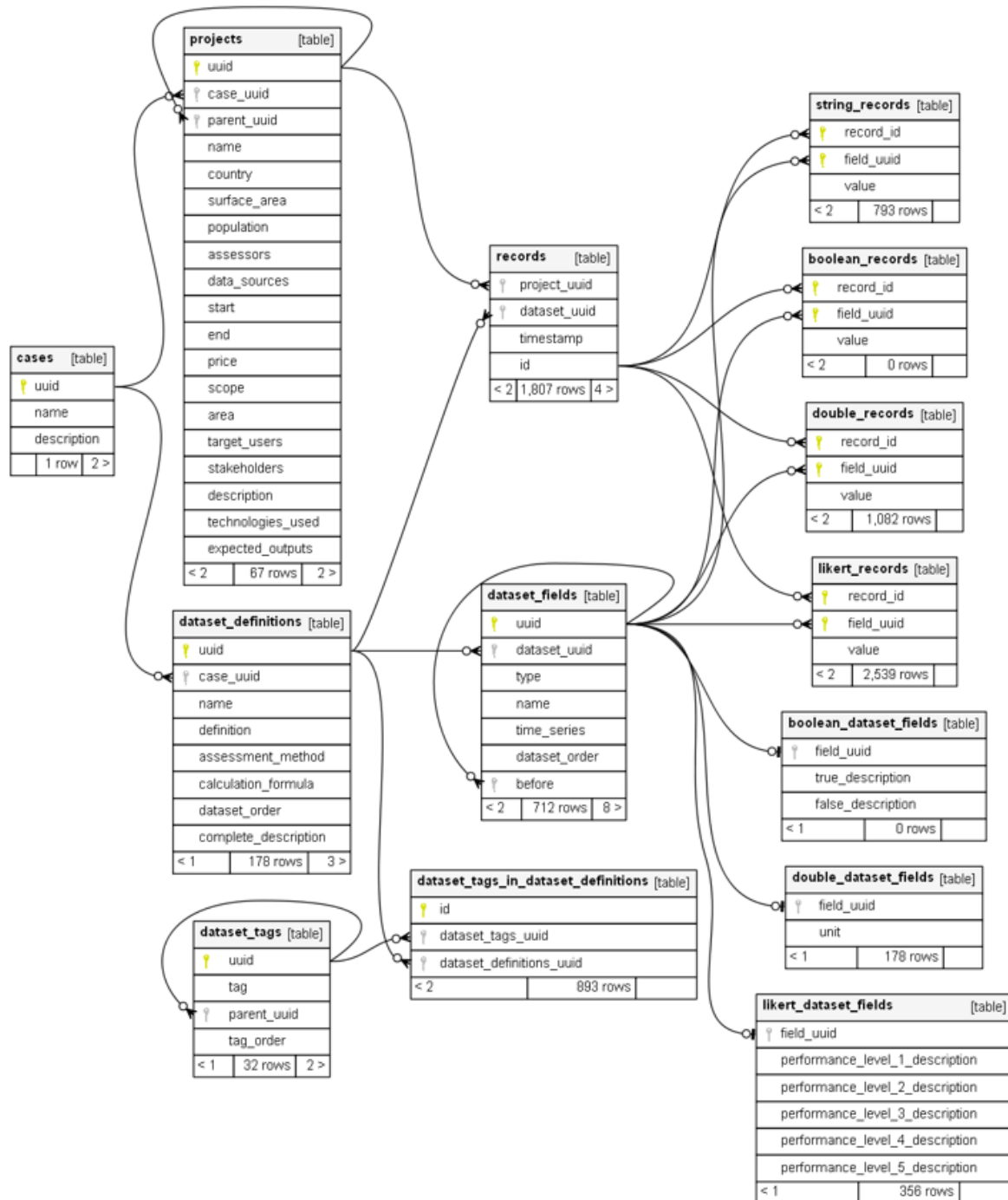
Field:	Country	Municipality	BuildingData1	BuildingData2	BuildingData3
Table:	StudyLog	StudyParticipants	Buildings	Buildings	Buildings
Sort:					
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:					
or:					

The data content and results shown here were created for testing and demonstration purposes and were not included in any actual research within ECHOES project. In addition to the experiment, draft versions of parts of the “Main survey” from WP 4 and WP5 were fitted into a relational structure with confirming results. Relational databases are appropriate for handling and presenting quantitative data, including data that uses some type of quantitative scale, such as a Likert rating scale. Therefore, it is seen as the technology of choice for ECHOES database. The forthcoming ECHOES database service may actually consist of several such databases, which may contain data from different studies and different topics among research results accumulated within ECHOES project.

3.5 Database for KPI values

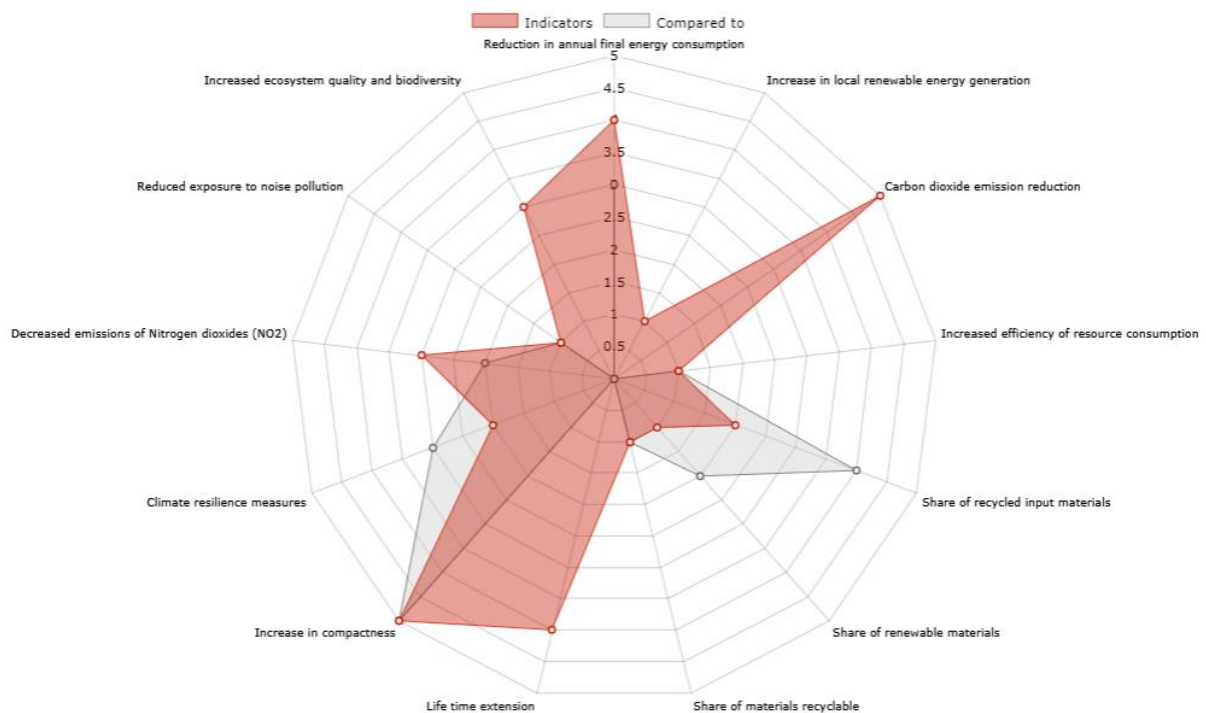
A very early draft of a database schema for ECHOES related KPI (Key Performance Indicator) data is shown in Figure 7. The Figure does not describe the KPIs to be developed in ECHOES content-wise but the technical database structure allowing flexible calculation of KPIs once defined.

Figure 7. Draft database schema for ECHOES KPI data



This database structure makes it possible to store KPI values, KPI description assessment methods and description of KPI calculation formula to the SQL database. The schema also supports time series and makes it possible to build visualisation APIs for e.g. spider web based visualisation (Figure 8).

Figure 8. Example of KPI visualisation on a spider web chart.



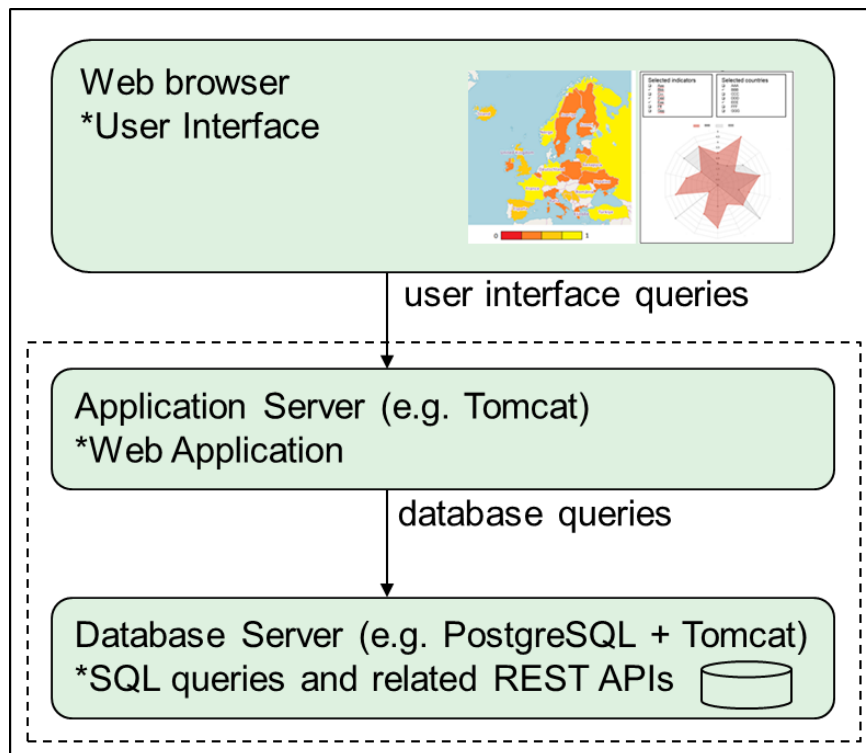
3.6 Database system architecture and implementation technologies

An overview of the database and related Web application system architecture is shown in Figure 9. As discussed earlier in Chapter 3.1, the main parts of the ECHOES database system architecture are:

- A database server hosting the SQL database, a web server and needed database query and API codes
- An application server including a web server and hosting the ECHOES end user application which main functions are the browser based user interface and related use of ECHOES database REST API queries
- An end users web browser for using the ECHOES database based Web application

Figure 9 specifies some technology choices for the ECHOES database.

Figure 9. An overview of the database and related Web application system architecture



The database related implementation will be based on a PostgreSQL database (relational database) and the coding (for SQL queries and related APIs) will be carried out using Java programming language, running on an open source based Apache Tomcat Web server.

The basic idea is to implement the ECHOES database as a service, which in principle enables the use of database queries in different kinds of end user devices and their respective applications. Next, the REST queries send back the query result as JSON response. Initial draft examples of these survey data related REST APIs are presented in Table 6.

Table 6. Draft examples of survey data related REST APIs

HTTP method	Endpoint	Endpoint parameters	Description
POST	/carsharing	will be defined later	Query survey related car sharing data
POST	/energysavingattitude	will be defined later	Query survey related energy saving attitude data
POST	/mobility	will be defined later	Query survey related mobility data
POST	/motivation	will be defined later	Query survey related motivation data
POST	/tempriseimpact	will be defined later	Query survey related temperature rise impact data

For each of the needed database queries, the parameters to be handled through the REST APIs need to be defined. This means that in the development process, the terms data end-users can handle in database queries need to be



defined in advance in order to make the queries possible. The database schema and more specific database queries and related REST APIs will be specified and implemented later.

The Web application for the ECHOES database will be a Web browser based application running on a Web server, which may be either the same or different Tomcat server as the one containing database related codes, and it will be implemented using the Angular framework. This Web application uses the ECHOES database queries via related REST APIs.

4 Data content

4.1 Data collection and data types

Data collection within ECHOES project has been described in detail in Deliverable D3.2 (Klößner et al. 2017), and data management principles are outlined in Deliverable D8.1 (Vesely et al. 2017). Database development will follow the provisions set out in these reports.

Types of data created in the ECHOES project include

- Anonymized interview transcripts
- Database of relevant literature
- Policy list analyzed
- Data-input for meta-analyses
- Experimental data
- Interview transcripts
- Main survey data

In addition to data created within the project (see the list above from Similä & Koljonen (2017)), there will be data for research purposes from external sources, which may contain national, regional, and EU level quantitative and qualitative data.

Principally, data that is foreseen to be stored in the ECHOES database consists of either numbers, text strings or different types of documents that may contain qualitative or quantitative data, as documented in project Deliverable D3.2. The first two types can be contained in relational data tables, while documents require a different type of storage, which needs to be considered. In addition, data protection and privacy need to be considered to define, what kind of access can be admitted to certain document sets, such as interview transcripts and translations. Further, it is important to review database users' data needs against these conditions as well.

According to D3.2, the multinational survey is the most substantial data collection activity in ECHOES. Including more than 17.000 participants in all 28 EU member states, plus Norway, Turkey, and Switzerland, the survey is implemented by November 2018 with a comprehensive quantitative questionnaire instrument including a choice experiment. The survey includes sections relevant for all three ECHOES perspectives. Due to its acknowledged significance for the ECHOES, the survey was taken as a primary test infrastructure for designing the database.

In D3.2, the data collection activities of ECHOES are described in 12 different data sub-types, providing valuable information for designing the database.

- Document studies
- Literature search
- Local surveys to study questions
- Quantitative experiments
- Interviews
- Case studies
- Netnography
- Workshop
- Focus groups
- Site visits
- Discussion events

According to D3.2, many of the data types (see list above) are qualitative in nature and are structurally comparable to interviews or documents in their requirements for databases as discussed in Chapter 3.3. However, especially quantitative experiments (WP4), and local surveys (WP2, 4, 6), may include numerical components which could be suitable for quantitative representations similar to those considered for the main survey. Thus, intensive dialogue between the partners conducting these parts of the research is needed (WP2, 4, 6) in order to define the most efficient database solution for the respective data type.

4.2 Data services for audience (open access groups)

Open access will be provided in two different user groups: public material open for everyone and more detailed data open for registered users. The data provided for these groups may be defined in further detail after decisions will be made regarding the use of research results, and the database is moving from the development stage toward implementation.

In the database workshop arranged in Rome, some responses considered how data could be organised within the ECHOES database. The ideas are presented in the following:

- Exploitable database (.xls, spss, map)
- Year
- Country
- Decision maker category
- Geographical scope: international, regional, national, local/municipal, urban/rural
- Technological focus (buildings, smart technology, mobility)
- Stakeholder types
- Energy or emission sector (i.e. transport, buildings, industry, agriculture, etc.)

In addition, the Grant Agreement proposes certain ways for grouping data. These groupings could be applied to documents by attaching relevant keywords that may be used to search them.

4.3 Data for direct ECHOES research use

Additional data from external sources has been sought for analysis purposes within ECHOES consortium and will be made available for all its' members. These data are publicly available but property of institutions other than ECHOES consortium members and often subject to restrictions of use. In order to avoid copyright issues, this material shall be available for internal use in the analyses among the project partners only.

Since the list of data types to be included in the database in Grant Agreement (Table 3) is very generic, during the database development, an effort was made to define the most important data used directly in ECHOES analyses. In response to this, data collected for this purpose, especially in order to support analyses of the main survey, is described in Table.

Table 7. External data for ECHOES research analyses.

Type of data	Variable	Sources
Economic statistics	Gross domestic product (GDP)	Eurostat
Economic statistics	Gross domestic product per capita	UN Statistics

Energy and Climate Future	Prognosed climate change impact - Economic	https://www.nature.com/articles/nature15725 https://web.stanford.edu/~mburke/climate/map.php
Energy and Climate Future	Prognosed climate change impact - Climate	Eurostat: global mean temperature deviation. https://www.ncdc.noaa.gov/temp-and-precip/ghcn-gridded-products/
Electricity Taxes and Fees	Sum of all taxes and fees applicable to household consumers	Eurostat
Electricity Taxes and Fees	Sum of all taxes and fees applicable to household consumers	https://ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_price_statistics
Gasoline Taxes and Fees	Sum of all taxes and fees applicable to household consumers	IEA - purchase data: https://www.iea.org/statistics/prices/
Energy	Share of low voltage network which lies underground = LV < 1 kV ground cable	https://www3.eurelectric.org/media/113155/dso_report-web_final-2013-030-0764-01-e.pdf https://www3.eurelectric.org/powerdistributionineurope/
Climate	Average heating Degree Days	https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/joc.3959 https://ec.europa.eu/eurostat/web/products-datasets/product?code=nrg_chdd_a
Climate	Average cooling Degree Days	https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/joc.3959 https://ec.europa.eu/eurostat/web/products-datasets/product?code=nrg_chdd_a
Climate	Average annual precipitation	https://www.eea.europa.eu/data-and-maps/indicators/european-precipitation-2 https://www.eea.europa.eu/data-and-maps/data/external/e-obs-gridded-dataset ; https://www.eea.europa.eu/data-and-maps/indicators/european-precipitation-2/assessment
Taxes	Average sales tax	https://ec.europa.eu/taxation_customs/sites/taxation/files/resources/documents/taxation/vat/how_vat_works/rates/vat_rates_en.pdf European Commission - EU VAT rates
Energy	Energy consumption total	Eurostat
Energy	Energy consumption per capita	Eurostat
Energy	Electricity consumption total	Eurostat
Energy	Electricity consumption per capita	Eurostat

Energy	Renewable share	Eurostat
Energy	Residential electricity price	Eurostat

All the secondary data sources presented in Table contain quantitative, country level datasets, which will not refer to individuals directly or indirectly.

5 Database implementation

Whereas the preceding Chapters of this document have described the key design components needed in building a database, in this Chapter, implementation is discussed in more detail. Hence, a plan including a schedule, major milestones, partner roles, and responsibilities is presented.

5.1 Schedule and major milestones

Based on the specification of work and discussions related to the database, and in order to summarize, the preliminary “Action plan” drafted in D2.1 is updated in this deliverable. An action plan concretizes the next steps in ECHOES database development towards its finalization (Table 8). However, the Action plan is subject to constant development, discussions and VTT is open to further development ideas.

The basic idea of implementation is building the database solution based on the main survey, and specifying queries and schemas, a work already well-advanced based on draft versions. After the basic structures regarding the key content are implemented, more and more data will be stored and functionalities will be step-wise implemented in close collaboration with database developers and research result providers from WP3-WP7.

Table 8. Action plan to develop the ECHOES database.

Step	Responsible	Timeline	Notes
1 Introducing the results of D2.1, raising open questions, introducing the action plan	VTT	ECHOES General Assembly San Sebastian 9/2017	D2.1 on state-of-the art was finished 8/2017
2 Developing an internal questionnaire for specifying the ECHOES database architecture.	VTT (incl. database expert)	9/2017	Input from the San Sebastian meeting taken into account. The relevance of questions were verified with the VTT database expert
3 Answering the questionnaire and getting the data collections clarified	ECHOES partners.	10/2017	Data collections documented in D3.2
4 Developing an architecture of the database. Key components: <ul style="list-style-type: none"> User interface Functionalities Technical solutions (interfaces etc) 	VTT (incl. database experts), partners EI, NTNU	Started early 2018, continuing until the end of the project	Clarifying the role of data repositories (e.g. Zenodo) in the overall solution ECHOES General Assembly Database Workshop 9/2018 results used in design
5 Specifying data content (internal data, external data, prioritising from options)	VTT, partners producing result data	Started 8/2018 with survey draft, continuing until the end of the project	Several meetings between data providers and database designers
6 Coding a functioning test/pilot version of the database solution based on data available	VTT	Tests started early 2018, specification continues as more and more ECHOES result data available (esp. main survey, 11/2018)	VTT is responsible for the construction and programming of the database and its interface. VTT and NTNU provide the design and infrastructure for the database.

7 Publishing a test version of the database internally. Components to start with <ul style="list-style-type: none"> • Main survey • Literature reference data 	VTT (incl. database expert)	early 2019	VTT's server will host the database initially during the construction and piloting phase
8 Developing a test version of the database and iteration based on user feedback. Further components <ul style="list-style-type: none"> • KPI data • "Understanding energy decisions" 	VTT based on ECHOES partner feedback	early 2019-11/2019	A block "Understanding energy decisions" in the draft interface will be based on quantitative and qualitative ECHOES result data from WP3-WP7 to be specified in dialogue between the responsible researchers and database team.
9 Publishing the database externally	VTT, NTNU	11/2019	NTNU's server will host the final open-access ECHOES database.
10 Maintaining the database after the project	NTNU, to be confirmed	11/2019 onwards	

5.2 Partner responsibilities

Responsibilities explained in detail

The exploitation plan drafted for the ECHOES results describes the roles and responsibilities for partners needed to contribute in database development. They are defined as follows.

- EI provides the data from the survey; WP3, WP4, WP5, and WP7 partners have contributed to survey design
- VTT and NTNU provide the design and infrastructure for the database. VTT is responsible for the construction and programming of the database and its interface.
- VTT provides access to the secondary data.
- VTT selects and formulates the KPIs to be shown in the database based on the SET-Plan KPIs, like Action 3.1 on energy consumers.
- The partners conducting interviews in WP5 and WP6 contributed to the countrywise conclusions presented in the database frontend (IUE, JR, NTNU, VTT, TECNALIA, UACEG)
- The partners conducting the psychological experiments contributed to the experimental data (ULEI, ROMA3, NTNU, VTT, TECNALIA, IUE, UACEG)

Mirrored against the list of partner responsibilities above, it can be seen that efforts of building the database until today have mainly been conducted in collaboration between VTT, EI, and NTNU, as they have mainly related to the main survey and fundamental planning of the infrastructure. As the interface and functionality designs sketched have stressed the importance of demonstrating "Understanding energy decisions" in the ECHOES database and the role of WP4-WP6 in producing the results, the partners responsible for producing the results as listed above are needed to participate in a similar dialogue. That is, the nature of result data and possibilities for its representation in these work packages has to be studied on a similar level of detail as the main survey has this far been studied as the key component of the database. In preliminary testing, based on draft multi-national survey data structures, it has been clearly demonstrated that the formats and structures of the data have to be precisely defined in order to deliberate the functionalities and visualisations. Thus, in forthcoming development work considering the results



of additional Work Packages, a similar approach starting from the results of each Work Package and Task, allowing modular development, seems to be appropriate for building up the database.

Data security and information protection

Basic rule of the open access principle regarding the ECHOES result data is that all anonymized data that can be shared publicly will be made available through the project database for future use.

ECHOES data will be transferred between the named non-EU countries (Norway and Turkey) and countries in the European Union to allow for joined analyses and storage of all data in the common database. All data transferred between project partners, either within or outside the EU, will be restricted to anonymized data and decisions should be made, if there is justifications to carry out data transfer in encrypted form via secured channels. In such case, it is of high importance for all project partners involved in data transfer to the database.

VTT's server will host the database initially during the construction and piloting phase and then NTNU's server will host the final open-access ECHOES database. This allocation of tasks impacts the database implementation and hence calls for an intensive dialogue between VTT and NTNU and the partners providing the data.

6 Summary and conclusions

Building a database including ECHOES results and relevant external data calls for a careful plan, and there are several building blocks to be decided and several partners involved. Up to the point of this deliverable (D2.2) of the ECHOES project (November 2018), the following steps have been taken:

- Identifying and analysing relevant SSH and energy reference databases, identifying gaps and points to be highlighted in ECHOES database (D2.1)
- Specifying the data content in sufficient detail, defining the multi-national main survey as a key content and main effort of the project under deeper analysis (dialogue between the ECHOES partners, especially VTT and EI)
- Mapping data needs of ECHOES partners for ECHOES analyses, prioritising urgent needs for the ECHOES analyses, collecting the data accordingly (dialogue between the ECHOES partners)
- Mapping different technological options and structures needed to implement the database (VTT development work in collaboration with partners)

Clarifying roles between partners (ECHOES partners, General Assembly in September 2018 in Rome), part of the characteristics are open and depend on details of the results of ECHOES, as key outcomes. Main building blocks of the database are:

- Points to be highlighted in the ECHOES database solution based on ECHOES research and external workshop:
 - Energy cultures, consumer behaviour, acceptance, and gender issues.
 - Regional differences, socio-demographic factors, ease of a solution, environmental and ecologic consciousness and lifestyles, awareness and level of knowledge
- User interface
 - Map-based user interface with versatile search options
 - Allowing queries to be filtered, e.g. regionally, socio-demographically and by gender would support to implement desired characteristics for ECHOES database
- Key data content of the database:
 - Main survey data as a key component
 - KPIs. WP2 will formulate indicators based on the collected data, which can be used for EU level monitoring of the progress of the Energy Union, especially related to solidarity, security, and confidence.
 - Scientific literature data based on D3.1 (Bireselioglu et al. 2017), possibly project publications and other project documents
 - Interview text data in processed format
 - Psychological experiments and local surveys as supplementary options for quantitative functionalities
 - External data based on variables used in ECHOES analysis such as understanding the survey and interview data
- Key functionalities
 - Comparisons of key results
 - Raw data download
- Key technical choices of the database
 - Relational (SQL) database structure for research data that will be accumulated in ECHOES project
 - Data repositories as complementing platforms for the overall solution (e.g Zenodo data repository).
- The main parts of the ECHOES database system architecture are:

- Database server including SQL database, Web server and needed database query and API codes
- Application server including Web server and related ECHOES end user application, whose main functions are the browser based user interface and the use of ECHOES database REST API queries
- End users web browser for using the ECHOES database based Web application

As the main purpose of the ECHOES database is in presenting the project results and variables explaining the results, not all the details of the database can be described as the work is still on-going and the data yet unavailable. Most notably, a key component in design is the main survey, which has just recently provided the raw dataset in November 2018. Planning work has, however, been enabled by draft versions of the materials available and specifications of the data presented (D3.2), complemented with close co-operation between the ECHOES partners responsible for producing the data.

In further development work, a restricted test version of the database with limited access and functionalities will be implemented step-wise, which will function as a core in the remaining database development of the project. The development work will be based on the main survey data, and as more and more result data will come in, more far-reaching tests of the database structure become possible. According to D3.2, the main data collection activities will happen in 2018 in WP4, WP5, WP6, and the multinational survey. There are also some data collection activities set to be finalised by early 2019 (D3.2), thus, the opportunities to broaden the tests related to development of the database grow significantly at that timeframe, motivating the planned launch of a test database in early 2019. Especially, a modular approach that allows step-wise implementation of the database, is suggested. A modular database development plan takes the schedules of data delivery flexibly into account starting from the main survey, and thus helps ensure a smooth database implementation

7 References

- Biresselioglu, M. E., Burov, A., Carrus, G., Demir, M. H., Dimitrova, E., Fritsche, I., Kaplan, M. D., Klöckner, C. A., Koksvik, G., Masson, T., Mutafchiiska, I., Panno, A., Røyrvik, J., Schwarzing, S., Skjølsvold, T., Tasheva, M., Tiberio, L., Velte, D. & Vesely, S. 2017. Social Science Perspectives on Electric Mobility, Smart Energy Technologies, and Energy Use in Buildings – A comprehensive Literature Review. Report No ECHOES 3.1 – D3.1 Literature Review. https://echoes-project.eu/sites/echoes.drupal.pulstartecnalia.com/files/ECHOES_D3.1_literature_report_1.pdf
- Jimenez Iturriza, I. 2018. ECHOES Plan for the Exploitation and Dissemination of Results (Update October 2018). Report No ECHOES-D8.6PEDR.
- Klöckner, C.A., Koljonen, T., Røyrvik, J., Carrus, G., Lettmayer, G., Biresselioglu, M.E., Kollmann, A. & Jimenez Iturriza, I. 2017. Data Collection Strategy Plan. Report No ECHOES 3.2. <https://echoes-project.eu/sites/echoes.drupal.pulstartecnalia.com/files/D3.2.pdf>
- ECHOES 2016. Grant Agreement number: 727470 — ECHOES — H2020-LCE-2016-2017/H2020-LCE-2016-RES-CCS-RIA.
- European Commission 2018. Key Performance Indicators (KPIs) <https://setis.ec.europa.eu/activities/initiatives/key-performance-indicators-kpis>
- Similä, L. & Koljonen, T. 2017. Report of existing data and databases as additional to new data collection and formulation. Report No ECHOES 2.1. <https://echoes-project.eu/sites/echoes.drupal.pulstartecnalia.com/files/D2.1.pdf>
- Vesely, S., Olgard Dalseth Røyrvik, J., Klöckner, C.A., Koljonen, T., Jimenez Iturriza, I. 2017. ECHOES Data management plan. Report No ECHOES 8.1 – D8.1 Data Management Plan

8 Appendices

A.1 Appendix 1. ECHOES functionalities and variables proposed in the WP2 Workshop, Rome, September 12th, 2018

Purpose	Functionality	Possibilities	Parameters <small>*subject to availability</small>
Exploring the European view	Map application of ECHOES results for quick overview	Heat map: Compare results on one variable for all countries (regions?) Spider web chart: Compare two countries against several selected variables	Awareness Acceptance Attitude vs region Attitude vs social group Prices, demand (of specific technologies?) Energy choices, expenditure, consumption Country Age Gender Type of house Income, socio-economic status Skills, knowledge level
	Examine results in detail	Retrievable data in sign-in user area, most versatile format / multiple formats Retrievable documents in sign-in area	Full or selected data
Quick country information		Country information sheets	
"Understand methods"			
Information and Support for collective and individual energy actions	"Find an expert"		Availability, location of supporting services
	Document / information collection		National incentives
			Commercial financing
			Current impact of planned RE choice (in one of ECHOES foci)
Policy outlook	Examine existing policies and regulations	Comparative policy documents analysis	Political frameworks, strategies Legislation Licensing Support schemes, grants

			Recommendations Subsidies Taxes Codes, e.g. buildings, spatial planning
Purpose	Functionality	Possibilities	Parameters *subject to availability
Echoes consortium area		Literature collection: peer-reviewed publications * ECHOES and others	Research on energy aspects related to: Social groups, Lifestyles & value systems, Mobility patterns, Types of built environment, Technological competence, Attitude & value systems, Transport systems, Energy behaviour connected to socio-economic factors
			Collective action in the energy and mobility field Change in individual attitudes
Research	Supporting data		Variables that are specific for research question
Supporting businesses			
Terms of use			
Technical support			

Table App. 1. ECHOES functionalities and variables proposed in WP2 Workshop, Rome, September 12th, 2018


A.2 Appendix 2. ECHOES WP2 Workshop notes, presentation

External ECHOES database (results, group A)

User groups	Functionalities, variables, indicators	Design, visualization
Local community people	<p>What kind of incentives are available in my country?</p> <p>What is the current impact of my eventual RE choice in one of ECHOES foci</p> <p>Are there services supporting people in bureaucracy?</p> <p>chang..?/company</p>	<p>Clean Simple Easy to use Tec. backup</p> <p>Social media Updates variations</p>
<p>Research/academia Full set of outcomes</p> <p>Academics- Exploitable database for generating reports/statistics</p> <p>Policy makers -recommend -survey -analyses -policy briefs</p> <p>Energy consultants -survey (orient. in market)</p> <p>Energy collectives (cooperatives, NGOs)</p> <p>Municipality ?</p>	<p>Academics Interactive maps for comparing parameters (prices, demand, awareness, consumption etc across countries)</p> <p>Academics Datasets regarding country-specific parameters (xls files)</p> <p>Indicators -Energy demand/intensity/user groups -Energy prices -Share of renewables -Acceptance levels -Awareness levels -Barriers -Motivators -Willingness to use renewables</p> <p>Policy makers Heat map for awareness, acceptance, attitude with respect to regions, social groups etc</p> <p>Types of projects -technology -process -economic sector -energy parameter -financials -project owner -geographic spread</p> <p>Information regarding different energy lifestyles -> too wake more efficient customer segmentation (from survey + WP5)</p>	<p>Design Exploitable databa (.xls, spss, map) -By year -By country -By decision maker category -By tech focus</p> <p>-Maps -Interactive maps -Comparisons</p> <p>Design Highlights/qualified summary, visual presentation of ... tipping points</p>
<p>Local town administrations</p> <p>Ad agency / communication dept in a company / utility Message to segments / strategy / message .. ?</p> <p>Facilitators Market promoters ->survey ->guidelines ->research</p> <p>Companies White papers as summaries of deliverables and important results</p> <p>Energy companies -ESCO -DSOs</p> <p>Industry Survey Identifying market demand?</p>	<p>Academics Coding obtained from raw data of interviews, focus groups</p> <p>EU Policymakers</p> <p>Policy makers -European -National -Regional -Local</p> <p>Covenant of mayors (JRC)</p> <p>Policy makers Heat map for awareness, acceptance, attitude with respect to regions, social groups etc</p> <p>NGOs</p> <p>? Profile -country -age -type of house -income -gender</p> <p>Energy company managers</p> <p>Energy service providers</p> <p>Tech company Sell product</p>	<p>-Energy consumption by type -demographics -relevance of indicators with decisions (by decision making group, by country, by gender etc. -comparative policy documents analysis</p> <p>Data Local / regional plans and strategies (oad practice supported by numerical evid.)</p> <p>New technol. dev. Cap. Building oport. Licensing Legislat.</p> <p>Financial sources -public incentives -commercial financing</p> <p>Attitudes (survey) / qualitative -attitudes -segment</p> <p>Country data Country-specific -Elec cars fund(?) -Grants</p> <p>Information regarding citizens attitudes towards smart technologies</p> <p>Information skills & knowledge of users</p> <p>Companies Database of legislations by country, subject, year</p>

External ECHOES database (results, group B)		
User groups	Functionalities, variables, indicators	Design, visualization
Government agencies National administration Angry white men National agencies Ministry of energy		Option to get access to sci-use only data Quantitative and qualitative data Details about survey methods / data generation Readable by different stats packages (R, SPSS,...) Raw data + Code Book
Journalists (media) Journalists Regulators (energy) Government organisations related with energy Social media "influences"	Understand methods "Find an expert" Country infosheets Cross-country comparisons Find definitions / meaning of "scientific" vocabulary	Maps displaying key findings Country-wise polls for download Tables/sheets Quantitative data on country map Maps by country Example ESPON
RESEARCHERS (policy) RESEARCHERS (spatial planning of DEVELOPMENT) RESEARCH INSTITUTIONS RESEARCHERS (social)	Variables that are very specific for research quest. Moderator variables (e.g. cult. orient.). Awareness & openness to innovations in... National incentive policies Political climate / dominant polit. Orientat.	Visualisation of links between variables (example indicator chain) Correlation charts + filtering Clips (video) ppts
Consumer associations Sustainability researchers Basic researchers	Original / raw data for further use Interaction effects btw. personal & country-level Socio-demographic data	Infographics by cluster (tematic) Search engine (compare countries # variables) Example URBAN AUDIT
NGOs Energy providers Car manufacturers NGOs & citizens Providers (E sector) Decision makers (policy design) Car manufacturers Energy companies Eco tech companies Environmental NGOs Construction sector companies Business / technology Electrodomeestic companies	Energy consumption (hour profile) Energy consumption -Plug standards (recharge technology) -population density Electrodomeestic used National/regional, local legislations & regulations on constr. (en. efficiency + spatial planning) Lifestyles & environmental awareness	Time series (kW, hour) Table Washing m. 1 TV's 3 Air cond.1 Generally speaking maps are cool

External ECHOES database (results, group C)

User groups	Functionalities, variables, indicators	Design, visualization
<div>Universities – research institutes</div> <div>Researchers</div> <div>Scientists</div>	<div>Research on energy aspects related to</div> <ul style="list-style-type: none"> - Social groups - Lifestyles & value systems - Mobility patterns - Types of built environment - Technological competence - Attitude & value systems - Transport systems 	
<div>Policymakers</div> <div>Public institutions</div> <div>Public administration</div> <div>European Commission</div> <div>Policymakers</div> <div>EC officials, SET Plan officials</div>	<div>On correct geographic level</div> <ul style="list-style-type: none"> -international -national -regional <div>Possibilities to compare e.g. between regions</div> <div>Energy behaviour connected to socio-economic factors</div> <div>Gender, age, socio-economic status</div> <div>Perceived drivers and obstacles (-> which political interventions have .. to succeed)</div>	<div>Overviewed</div> <div>Short and clean information</div> <div>Keywords links to the documentation</div> <div>Statistics</div> <div>Simple search / advanced search options</div>
<div>Private marketing agencies</div>	<div>Data about energy behaviour</div> <div>Web based application</div> <div>Technical support if necessary</div> <div>Clear terms of data use</div> <div>Regional differences in ECHOES variables for operational planning</div> <div>Regional differences in ECHOES variables for operational planning</div> <div>Data in helping marketing efforts</div>	<div>Thematic maps</div> <div>MAPS</div> <div>Zoom in and out</div>
<div>Energy companies</div> <div>Industry</div> <div>-Energy service companies (ESCOs)</div> <div>-Grid operators</div> <div>-Energy providers</div> <div>Company users, e.g. smart tech/EV companies for understanding market</div>	<div>-Market potential (e.g. maps)</div> <div>-User needs /demand -> specific products/functionalities (e.g. ranking)</div> <div>-Acceptance/adoption readiness / criteria</div> <div>-Political frameworks impacting market potential</div> <div>-target group characteristics (e.g. energy related spendings)</div> <div>-status of national implementation of EU acquis</div>	<div>(geographic/social group-related) RANKINGS</div> <div>(demand, advancement in policy implementation)</div>
<div>Consumers</div>		<div>Language?</div>
<div>USER GROUPS</div> <div>Public companies</div> <div>Private companies</div> <div>NGOs</div> <div>Universities</div> <div>Students</div> <div>Urban research</div> <div>Urban planning</div> <div>Mobility planning</div> <div>Urban & regional policy</div> <div>NGO sector</div> <ul style="list-style-type: none"> - Environmental - Social 	<div>Students</div> <div>-Policy basis fore energy related issues for their research</div> <div>-EU energy legislation</div> <div>-Statistics</div> <div>-Political, economic, social technical and environmental variables</div>	