Analysis of the European energy system
under the aspects of flexibility and technological progress

Call: H2020-LCE-21-2015

Grant Agreement Number: 691685

Deliverable

D2.1 Data Management Plan (DMP)

Deliverable type:	ORDP: Open Research Data Pilot
WP number and title:	WP2: Data Warehouse and Model Coupling
Dissemination level:	Public
Due date:	Month 6 – 31th October 2016
Lead beneficiary:	ESA²
Lead author(s):	**Robert Kunze**
Reviewers:	Claudia Hawke (TUD), Witold-Roger Poganietz (KIT)

This project REFLEX has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 691685.

Document history

Version	Date	Author/Editor	Description
0.1	20.10.2016	Robert Kunze	Initial version, sent to reviewers
0.2	27.10.2016	Robert Kunze	Revised initial version, sent to reviewers
1.0	01.11.2016	Robert Kunze	Published version with reviewers comments incorporated

TABLE OF CONTENTS

ABBREVIATIONS

DMP	Data management plan
DSM	Demand side management
DWH	Data warehouse
EIM	Exploitation and Innovation Manager
EMS	Energy models system
LCA	Life Cycle Assessment
RES	Renewable Energy Sources

LIST OF TABLES

SCOPE OF THE DOCUMENT

This document provides the draft version of the Data Management Plan (DMP) for the REFLEX project according to the Open Research Data Pilot (ORD pilot) under Horizon 2020. The purpose of the DMP is to support the data management life cycle of all data that will be collected, processed or generated by the project.

The document structure and contents are based on the Guidelines on FAIR Data Management in Horizon 2020 (Version 3.0, 26 July 2016) and on the Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020 (Version 2.1, 15 February 2016). The document was generated by using the Digital Curation Centre's DMP online tool. The following sections outline the types of collected and generated data, how these data will be exploited and made accessible for verification and re-use and how data will be curated and preserved upon closure of the project. Changes are reserved – the DMP will be continuously updated during the project.

ADMIN DETAILS

Project Name: REFLEX (Horizon 2020 DMP)
Grant Agreement No.: 691685
Principal Investigator / Researcher: Prof. Dominik Möst, TU Dresden

Project Description:

The future energy system is challenged by the intermittent nature of renewables and requires therefore several flexibility options. Still, the interaction between different options, the optimal portfolio and the impact on environment and society are unknown. It is thus the core objective of REFLEX to analyse and evaluate the development towards a low-carbon energy system with focus on flexibility options in the EU to support the implementation of the SET-Plan. The analysis are based on a modelling environment that considers the full extent to which current and future energy technologies and policies interfere and how they affect the environment and society while considering technological learning of low-carbon and flexibility technologies.

For this purpose, REFLEX brings together the comprehensive expertise and competences of known European experts from six different countries. Each partner focusses on one of the research fields techno-economic learning, fundamental energy system modelling or environmental and social life cycle assessment. To link and apply these three research fields in a compatible way, an innovative and comprehensive energy models system (EMS) is developed, which couples the models and tools from all REFLEX partners. It is based on a common database and scenario framework. The results from the EMS will help to understand the complex links, interactions and interdependencies between different actors, available technologies and impact of the different interventions on all levels from the individual to the whole energy system. In this way, the knowledge base for decision-making concerning feasibility, effectiveness, costs and impacts of different policy measures will be strengthened, which will assist policy makers and support the implementation of the SET-Plan. Stakeholders will be actively involved during the entire project from definition of scenarios to dissemination and exploitation of results via workshops, publications and a project website.

Nature: Research and innovation actions based on energy systems modelling

Research Questions:

1) How do current and future energy technologies and policies interfere?
2) What will be an optimal combination of flexibility options to cope with the future flexibility needs?
3) How do these technologies and policy measures affect the environment, economy and society?

Purpose:

Support for the implementation of the SET-Plan:

- Analysing the impacts of technological development and innovation on the energy system and its dynamics
- Comparative assessment of the impacts and the sustainability performance of all relevant energy technologies
- Assessing the related impacts on the environment, society and economy - Analysing of technology policy measures in the framework of the SET-Plan
- Understanding the complex links/interactions/interdependencies between the different actors, the available technologies and the impact of the different interventions on all levels from the individual to the whole energy system
- Providing model based decision support tools for the different actors in the energy system in order to facilitate handling the complex system

Funder: European Commission (Horizon 2020)

1 DATA SUMMARY

The core objective of the REFLEX project is to analyse and evaluate the development towards a low-carbon energy system in the EU up to the year 2050. The focus is laying on flexibility options to support a better system integration of Renewable Energy Sources (RES). The analysis and the assessment of REFLEX are based on a modelling environment that considers the full extent to which current and future energy technologies and policies interfere and how they affect the environment and society.

1.1 PURPOSE OF DATA COLLECTION AND GENERATION

The purpose for data collection and their preparation within REFLEX is to provide the needed input data for the applied mathematical energy system models. The model pool of the REFLEX partners contains bottom-up simulation tools and fundamental system optimisation models on national and therewith also on European level as well as approaches for Life Cycle Assessment (LCA). Typically, one model cannot cover all aspects of an energy system or the implications of specific policies. Each of these different models focuses on a specific sector or aspect (heat, electricity, mobility, environmental/ social impacts etc.) of the energy system.

For analysing and answering the given research questions (see Admin Details above), the different models and approaches will be coupled to a so called integrated Energy Model System (EMS). The analysis by applying the EMS allows to perform an in-depth and at the same time holistic assessment of the system transformation and shall contribute to the scientific underpinning of the SET-Plan. The final result data of the EMS helps to understand and investigate the complex links, interactions and interdependencies between the different actors and technologies within the energy system as well as their impact on society and environment.

The results of a model based analysis depend not only on the chosen methodology, but also on the quality of the data used. For a consistent analysis within the EMS in REFLEX, a common database with harmonised datasets is needed. It will be implemented in a form of a Data Warehouse[1] (DWH). The DWH of the REFLEX project will contain four groups of data (see Table 1 for an overview), which will be explained in more detail in the following sections.

[1] The DWH consists of central databases with tools for data processing optimized for analysis purposes. It contains needed data for the project from several heterogeneous sources in consolidated/harmonized form. The DWH will be managed by and hosted at a server of the project partner ESA² in Karlsruhe (Germany).

Table 1: Groups of data within the DWH of the REFLEX project

Data group	Description/Contains
Existing model input data	- Data collected, generated or purchased from commercial providers by a project partner before the start of the REFLEX project - Input data collected, generated or purchased from commercial providers by a project partner in the context of projects on behalf of other clients run in parallel to the REFLEX project
Collected and generated new model input data	- Data collected from public available sources or purchased from commercial providers by a project partner in the context of REFLEX - Data collected through surveys conducted by a project partner in the context of REFLEX - Data generated based on existing, new collected or new purchased data by a project partner in the context of REFLEX
Generated intermediate model output data	- Intermediate results of the model applications for data exchange between the models or for further assessments within the project; due to this the 2nd and the 3rd data group could overlap
Generated final result data of the EMS	- Final results of REFLEX generated by model applications, e.g. CO2-emission, energy demand, technology impact evaluation etc.

1.2 EXISTING MODEL INPUT DATA

Energy system models need many different input data for modelling the real world. Each of the applied models in REFLEX has already been used as stand-alone applications. Therefore each model has its own database with already existing data. By this input data are rather model specific and an un-conditional application over several models is limited. They originate from own previous work and own assumptions of the project partners as well as from literature and have been developed over many years. Some of this existing data will be re-used in the REFLEX project, if they are up to date or if no better data are available.

Not all of the existing data are relevant for the DWH. Only data that

- are needed for more than one model within the EMS and therefore have to be harmonized

and/or

- are needed to validate the results presented in scientific publications (so called "underlying data")

will be included in the DWH.

The harmonization of input data is necessary to ensure a consistent analysis within the EMS. For the same information the same dataset (values) have to be used in all models. The consortium decides which of the existing datasets will be used. They will be included in the DWH and will be provided to all models before initializing the EMS run. Table 2 gives an overview of the existing and re-used input datasets/parameters.

Table 2: Existing essential (re-used) datasets in REFLEX

Dataset (category)	Sub-categories (Quantity)	Period (from until)	Spatial Reference	Source	Values per scenario (Quantity)
Buildings average life time of heating systems[1]	0	2015 2050	EU28	EU regulation, other projects, own assumptions	
Buildings compliance rate index[1]	0	2015 2050	EU28	EU regulation, other projects, own assumptions	
Buildings energy efficiency increase due to mandatory commissioning[1]	0	2015 2050	EU28	EU regulation, other projects, own assumptions	
Buildings energy efficiency investement cost (with national financial incentives)[1]	0	2015 2050	EU28	EU regulation, other projects, own assumptions	
Buildings minimum energetic standards[1]	0	2015 2050	EU28	EU regulation, other projects, own assumptions	
Buildings thermal renovation[1]	0	2015 2050	EU28	EU regulation, other projects, own assumptions	
Power plants availability	6	2010 2050	EU28	DIW 2013	
Power plants efficiency	11	2010 2050	EU28	DIW 2013	
Power plants emisson factor	3	2010 2050	EU28	UBA 2014	
Power plants interest rate	0	2010 2050	EU28	IEA et al. 2010	
Power lifetime of investment	6	2010 2050	EU28	IEA et al. 2010	
Power plants load change cost (depreciation)	0	2010 2050	EU28	DIW 2013, Traber & Kemfert 2011, own assumptions	
Power plants load change cost (fuel factor)	0	2010 2050	EU28	DIW 2013, Traber & Kemfert 2011	
Power plants operations management cost fixed	18	2010 2050	EU28	DIW 2013, VGB PowerTech 2011a	
Power plants operations management cost variable	11	2010 2050	EU28	DIW 2013, Traber & Kemfert 2011	
Power plants specific investment[2]	11	2010 2050	EU28	DIW 2013	
Power plants start-up cost (depreciation)	18	2010 2050	EU28	Traber & Kemfert 2011	
Power plants start-up cost (fuel factor)	18	2010 2050	EU28	Traber & Kemfert 2011	
Tertiary & residential sector tax energy carrier[1]	3	2015 2050	EU28	EU regulation, other projects, own assumptions	
Vehicles CO_2 standard[1]	2	2015 2050	EU28	EU regulation, own assumptions	
Vehicles fuel consumption factors	4	2015 2050	EU28	GHG-TransPoRD project, ASSIST project	

[1] Scenario-dependent [2] Only for technologies for which no experience curve data are available or will be generated

1.3 COLLECTED AND GENERATED NEW MODEL INPUT DATA

Some of the needed input data for the models will be updated or newly defined according to the research questions and the focus of the analysis within the REFLEX project. Therefor publicly and commercially available data will be used. Unavailable data will be generated by empirical surveys and/or appropriate assumptions. These data will be included in the DWH and in that way provided as harmonized datasets for all models within the EMS. This group includes data for:

- Scenario framework
- Demand side management
- Experience curves

1.3.1 DATA FOR SCENARIO FRAMEWORK

These data describe the overall framework for the model-based analysis and include the main macro-economic and societal drivers as well as techno-economic parameters and regulations/conditions of the political environment. Therefor the defined scenario storylines for REFLEX will be translated into quantitative model input parameters until the year 2050, which is the defined horizon for the analysis. They may be defined on a European level, or be further distinguished on a national, sectoral or technological level. The macro-economic trends and the societal drivers are likely to be based upon official projections provided by the European Commission (e.g. the upcoming Reference Scenario). All political assumptions will be elaborated considering current and past policy implementations and will be discussed with the European Commission and stakeholders. Table 3 gives an overview of the data for the scenario framework.

Table 3: Data for scenario framework

Dataset (category)	Sub-categories (Quantity)	Period (from until)	Spatial Reference	Source	Values per scenario (Quantity)
GDP-Gross domestic product	0	2000 2050	NUTS 0 (EU28+NO+CH)	Capros et al. 2016 (Assumptions for NO+CH based on other Horizon 2020 projects)	330
POP-population	0	2000 2050	NUTS 0 (EU28+NO+CH)	Capros et al. 2016(Assumptions for NO+CH based on other Horizon 2020 projects)	330
Price electricity (initial average cost of gross electricity generation)	0	2000 2050	NUTS 0 (EU28+NO+CH)	(Assumptions for NO+CH based on other Horizon 2020 projects)	330
Price fossil energy carrier	3	2015 2050	EU28	Capros et al. 2016	24
ETS-CO_2-Price[1]	0	2015 2050	EU28	Capros et al. 2016, own assumptions	8
NON-ETS-CO_2-Price[1]	0	2015 2050	EU28	own assumptions	8
Vehicles CO_2 standard[1]	2	2015 2050	EU28	Capros et al. 2016	16

[1] Scenario-dependent

1.3.2 DATA FOR DEMAND SIDE MANAGEMENT

Relevant data for investigating system flexibility by Demand Side Management (DSM) are rarely available from public and commercial sources. In particular the available database for the tertiary sector with regard to DSM is incomplete. Therefore, an empirical survey on DSM in the tertiary sector will be conducted with the aim to improve the model input data and to fill data gaps. Based on the analysis of the collected specific empirical data, existing datasets will be extended as well as new datasets generated. The design of the survey will be established by the REFLEX partners and it will be conducted for 10 European countries by an international market research institute. The DSM data for further countries will be deduced from the survey results. Relevant model input parameters for modelling DSM options which should be deduced from the empirically ascertained data are given in Table 4.

Table 4: Parameters for modelling DSM options

Dataset (category)	Sub-categories (Quantity)	Period (from until)	Spatial Reference	Source	Values per scenario (Quantity)
DSM potential (share of flexible load per energy usage process)	50	2015 2050	NUTS 0 (EU28+NO +CH)	empirical survey, public and commercial sources	28000
DSM cost (activation cost per energy usage process)	50	2015 2050	NUTS 0 (EU28+NO +CH)	empirical survey, public and commercial sources	28000
DSM time of interfere (maximum load reduction time)	50	2015 2050	NUTS 0 (EU28+NO +CH)	empirical survey, public and commercial sources	28000
DSM number of interventions (frequency of DSM measures)	50	2015 2050	NUTS 0 (EU28+NO +CH)	empirical survey, public and commercial sources	28000
DSM shifting time (allowed points of time or time periods/frames for DSM)	50	2015 2050	NUTS 0 (EU28+NO +CH)	empirical survey, public and commercial sources	28000

1.3.3 DATA FOR EXPERIENCE CURVES

To enable endogenous modelling of technological developments and due to this to production cost reductions, experience curves for the most relevant technologies for each sector will be developed and implemented in the sectoral models. Special attention will be given to determination of uncertainty ranges of progress ratios (i.e. the slopes of the experience curves), as these can have a major impact on modelling results, especially for long-term modelling until 2050. In addition, especially for technologies that depend strongly on either the available geographical potential (e. g. wind onshore, offshore, biomass) or on raw material prices, decomposition of the experience curve using a multi-level experience will be performed. This allows determination of the most important factors behind cost development, such as variations in steel or oil prices, as well as scale effects. The needed empirical data for defining the experience will be surveyed by means of interviewing experts, specific survey methods and by analysing detailed statistics (e. g. construction, production and consumer price indices as well as installed capacities and cost developments in the electricity, heat and mobility sector). Table 5 gives an overview of the technologies for which experience curves will be developed.

Table 5: Defined technologies per sector for which experience curves will be developed

Category	Technology
Electricity Generation	CCS (Membrane, Oxyfuel, Pre, Post)
Electricity Generation	CCGT (Gas)
Electricity Generation	EFCC (Gas)
Electricity Generation	Biomass: digestion
Electricity Generation	Biomass: gassification
Electricity Generation	Biomass: combustion
Electricity Generation	Concentrated Solar Power (CSP)
Electricity Generation	Geothermal: (dry, flash, binary)
Electricity Generation	Photovoltaics: modules (mono/poly, CdTe)
Electricity Generation	Photovoltaics: system level aspects
Electricity Generation	Photovoltaics: CdTe
Electricity Generation	Wind: onshore
Electricity Generation	Wind: offshore
Electricity Generation	CHP
Electricity Generation	micro-CHP

Table 5 (continuation)

Category	Technology
Electricity Generation	Pulverized Coal-fired
Electricity Generation	(P)FBC (Coal)
Electricity Generation	IGCC (Coal)
Electricity Generation	Gas Turbine (inlc. Mini- and Micro-Gas-Turbine)
Electricity Generation	Nuclear (generation 2, 3 and 4)
Electricity Generation	Steam turbine (coal/gas)
Electricity Generation	Photovoltaics: CIGS
Electricity Storage	Battery: Lithium-(Ion, Polymer, Air)
Electricity Storage	Battery: Redox-Flow
Electricity Storage	CAES
Electricity Storage	Flywheel
Electricity Storage	Pumped Storage Plants
Electricity Storage	Battery: Molten Salt
Electricity Storage	Battery: NiMH
Electricity Storage	Battery: NiCd
Heating/cooling	Electric boiler (P2H)
Heating/cooling	Heat Pump (air/air)
Heating/cooling	Heat Pump (air/water/ground)
Heating/cooling	Solar Thermal (process heat or large scale)
Heating/cooling	Air conditioning
Heating/cooling	Thermal Energy Storage: (sensible, underground, Phase Change, TCS, Ice)
Heating/cooling	Night-Storage Heaters
Heating/cooling	Electric boilers for district heating
Heating/cooling	Gas- or oil boiler
Heating/cooling	Fridges, freezers: cooling
Industry	Air separation: membrane
Industry	Air separation: conventional
Industry	Aluminium electrolysis
Industry	Chemical & mechanical pulp production
Industry	Electric arc and induction furnaces: copper, cink, etc.
Industry	Electrolysis (wet chemical): copper, cink, etc.
Industry	Electrolysis, chemical industry
Industry	ODC, chlorine electrolysis
Industry	Fischer-Tropsch-synthesis
Industry	Industrial CCS
Industry	Industrial heat
Industry	Cement and raw mill: cement production
Industry	Electric booster: container glas production
Industry	HIsarna process: steel production
Industry	ULCOWIN process: steel production
Mobility	Battery electric vehicles
Mobility	Fuel Cell vehicles
Mobility	Plug-in hybrid electric car
Mobility	Flexifuel car
Mobility	Overhead wiring trucks
Mobility	Biofuels (jet, vehicle, marine)
Mobility	CNG vehicles
Mobility	LPG car
Mobility	Diesel vehicles
Mobility	Gasoline car
Power to X	Power to Hydrogen
Power to X	**Power to Methane**
Power to X	Power to Methanol
Power to X	**Hydrogen synthesis, ammonia**
Electricity end-use	Lighting: LED
Electricity end-use	**Elevators/escalators**
Electricity end-use	Data centers
Electricity end-use	**Building Energy Management/Automation systems**
Electricity end-use	Ventilation
Electricity end-use	**Dishwashers, dryers, washing machines**
Other	Power Grids (HVDC point-to-point, also meshed systems)

In order to estimate the potential of alternative fuel technologies, both regarding Europe (e. g. MID in Germany or UKTS in UK) and regarding Asia and North America, mobility patterns and related market potentials will be derived from available mobility surveys. The analysis will be focused on major common driving patterns. The reason for the analysis of main global passenger car markets is to identify the global market penetration of electric vehicles, which will influence the demand for Li-ion batteries substantially in order to assess the future prices of electric vehicle batteries and fuel cells based on the learning curve theory. In this sense, the global automotive market (especially including Northern America and Asia) will be taken into account for investigating the uptake of alternative car technology in Europe.

Emission degression data will also be considered in REFLEX. However, the content, nature and scope of these data are still under discussion and will be determined during the project.

1.4 GENERATED INTERMEDIATE MODEL OUTPUT DATA

By coupling the different approaches of the RFELEX partners, the systems boundaries of each stand-alone model will be partly disbanded and most exogenous parameters of each model will become endogenous variables of the EMS. This will be done by using the relevant output data of one model as input data of another model.

To achieve a stable final state of the EMS within each REFLEX scenario storyline, several iterations with all models are performed. The generated results during these iterations are needed for the data exchange between the different models. Therefore these data will be included in the DWH and will be provided to each model by using a data interface. Table 6 shows the relevant datasets for the data exchange within the EMS.

Table 6: Datasets as intermediate results for data exchange between models within the EMS

Dataset (category)	Sub-categories (Quantity)	Period (from until)	Spatial Reference	Source	Values per scenario/iteration (Quantity)
Price electricity (hourly, wholesale and retail prices incl. taxes)	0	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	12,264,000
Demand electricity (hourly)	50	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	12,264,000 - 245,280,000
Demand electricity for mobility (yearly)	7	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	1,050
Demand district heating (yearly)	3	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	450
Power plants installed capacity and operating (yearly and hourly)	22	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	5,280 - 67,200
Power plants emissions (yearly)	132	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	15,840
Power plants demand energy (yearly)	22	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	2,640
Mobility demand energy (yearly)	56	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	6,720
Mobility emissions (yearly)	56	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	6,720

1.5 GENERATED FINAL RESULT DATA OF THE EMS

After achieving a stable state based on several iterations with all models within the EMS within each REFLEX scenario storyline, the result data of the different models will be collected and combined within the DWH to the final result data of the EMS. These data will be analysed to derive the key findings and are the basis for answering the research questions of the REFLEX project (see ADMIN DETAILS). Table 7 gives an overview on the major result data of the EMS.

Table 7: Major result data of the EMS

Dataset (category)	Sub-categories (Quantity)	Period (from until)	Spatial Reference	Source	Values per scenario (Quantity)
Price electricity average yearly (wholesale and retail prices incl. taxes)	0	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	120-240
Demand electricity (yearly)	50	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	6000
Demand district heating (yearly)	3	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	450
Power plants installed capacity (yearly)	22	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	2640
Power plants operation (yearly)	22	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	2640
Power plants emissions (yearly)	132	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	15,840
Net transfer capacities between countries installed (yearly)	143	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	572
Net transfer capacities between countries operation (yearly)	143	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	572
Mobility demand energy (yearly)	56	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	6,720
Mobility emissions (yearly)	56	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	6,720
Life cycle environmental and resource impacts	19	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	2280
Life cycle human health (damage / toxicity)	2	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	240
Life cycle societal impacts (risk level)	5	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	600
Costs external	1	2015 2050	NUTS 0 (EU28+NO+CH)	Applied energy system models	120

1.6 DATA UTILITY

REFLEX collects and generates a certain amount of research data. On one hand, these data are necessary to meet the objectives of the project. On the other hand most of the collected and generated data will be useful for further research and even for the energy industry.

1.6.1 DATA FOR SCENARIO FRAMEWORK

The collected and prepared data for the scenario framework are tailored to the aim and scope as well as the specific research questions of the REFLEX project and the applied models to answering them. They are thus primarily useful as underlying data to ensure the transparency of the generated results and the comparability of the project outcomes to other studies with similar analysing scope. However, the data will be used for updating the existing databases of the different models and will be used for further model based research of the project partners outside of REFLEX.

1.6.2 DATA FOR DEMAND SIDE MANAGEMENT

The empirical survey on DSM aims to provide the needed model input data for the REFLEX project. With the collected specific empirical data the database for investigating system flexibility by DSM will be improved in general, because relevant data – especially for the tertiary sector – are rarely available from public and commercial sources. Thus, existing datasets will be extended as well as new datasets generated. Furthermore, the survey data will allow the identification of promising energy applications and DSM potentials in the selected sector in different European countries. These data have a high potential for re-use after the end of the REFLEX project within further research projects. They are useful for other researchers and for stakeholders from industry and policy making.

1.6.3 DATA FOR EXPERIENCE CURVES

Endogenizing technological learning through experience curves allows for an enhanced assessment of the evaluation of impacts from policy measures or alternative incentive schemes on realizable future cost reduction. In addition, in view of current rapid and necessary changes in energy systems (driven partially by policies and partially by markets) and the ensuing need for flexibility, the endogenous modelling of the cost development of existing and new energy-related technologies in bottom-up models will become even more important.

However, the data and experience curves required to do so are not readily available. A comprehensive review of many energy supply (and some energy demand) technologies have been published by Junginger et al. (2010), but these require updating. Since then, recent studies have been published for some individual technologies (e. g. Bolinger and

Wiser 2012; Candelise et al. 2013 or Chen et al. 2012). However, an up-to-date overview is not available. Especially with regard to technologies needed for increasing the flexibility in energy systems (such as storage technologies or DSM-devices) little or no experience curves have been published. Thus, to advance the energy models included in REFLEX beyond the state-of-the-art by implementing these experience curves, data collection will be required to devise or update experience curves for existing technologies and to estimate experience curves for new technologies.

Furthermore, it will require smart and innovative incorporation and interlinkage of these experience curves in various sectoral energy models to comprehensively assess the effects of technological learning and the demand for increased flexibility in energy systems. The outcome – a state-of-the-art and up-to-date overview of experience curves and underlying database – could benefit other energy models outside the project, i.e. developed in the EU as well as worldwide) to meet the challenges of modelling our changing energy systems for the coming decades.

1.6.4 GENERATED INTERMEDIATE MODEL OUTPUT DATA

These data are only intermediate results of the EMS and will transfer between the models during the iteration process of the EMS. A relevant benefit of these data for further applications outside the framework of REFLEX is not expected.

1.6.5 GENERATED FINAL RESULT DATA OF THE EMS

The overall objective of REFLEX is to support the SET-Plan by strengthening the knowledge base for transition paths towards a low carbon energy system based on a cross-sectoral analysis for the entire energy system of the European Union. Due to the complexity of this system, it is obvious that the implementation of the SET-Plan requires in-depth knowledge on the interrelationship between the different energy sectors (electricity, heat and mobility), energy technologies but also on the interdependencies between energy and non-energy industries, environment (beyond greenhouse gas emissions) and society.

The result data of the EMS within REFLEX helps to understand and investigate the complex links, interactions and interdependencies between the different actors and technologies within the energy system as well as their impact on society and environment. Based on the EMS result data, recommendations for effective strategies for a transition of the European energy system to a low-carbon system will be derived. Policy makers at EU level as well as at regional level can use these findings when developing policy measures. Furthermore, the data of the REFLEX project can be used as a reference or starting point for further research work on the future design of the energy system of the European Union.

1.7 DATA PROTECTION AND EXPLOITATION STRATEGY

In order to ensure efficient implementation of dissemination and exploitation activities amongst the participants, a Consortium Agreement (CA) was signed by all partners. The CA is among other things dealing with the exact details on the participants' background data, the rights to, the protection of and the exploitation of data/results generated solely and/or jointly during the project. Moreover the CA sets up specific rules on how to deal with dissemination activities and to ensure open access to all peer-reviewed scientific publications. The model used is the DESCA template, version 1.0 (www.desca-2020.eu).

The following basic rules apply:

- All participants define their individual background data required for their successful participation in the project (own model input data and commercial model input data purchased before the start of the project). The rights to this background remain with the respective owner but royalty-free access to other participants is granted if not restricted by third parties and if it is required to enable other participants to carry out their research and development activities in the context of the project.

- During the project: Background data that is acquired by individual partners during the project, e. g. in the context of projects on behalf of other clients run in parallel, will be treated as pre-existing background data.

- The property rights to data collected and data/results generated during the project belong to those involved in its collection and generation. When more than one consortium member is involved in the creation of results it will be jointly owned by the respective consortium member.

Dissemination and exploitation of data and results will be executed in accordance with EU laws and with respect to specific laws in the participating countries. Before any dissemination activity will take place, respective legal aspects will be examined and clarified. This is particularly the case for data from the DSM-survey and data purchased from commercial providers. The possibility for protection of generated results within the project (consortium) will be also examined before publication. All participants have departments specifically devoted to managing intellectual property. These departments will manage the relevant protection processes.

Within REFLEX the dissemination and exploitation of data will be coordinated by the Exploitation and Innovation Manager (EIM) regarding knowledge management and innovation activities. The EIM is responsible for:

- maintaining a registry of background data,
- maintaining a registry of data gathered and generated in the work packages during the project,
- assessing the opportunities for exploitation, for example by following political events in the energy sector or searches of other scientific databases for similar developments and
- proposing specific exploitation measures, e. g. policy briefs and events.

In REFLEX periodic analysis of transfer opportunities to adjust the exploitation strategies takes place. All consortium partners contribute to the exploitation plan of the project throughout its life span. Thereby, the EIM is in close contact and regularly informed about the exploitation plans of the partners to use synergies and to ensure the best and suitable use and exploitation of results. Furthermore, the EIM will regularly advise the consortium and individual partners about possible strategies.

The exploitation strategy is outlined below:

- First it will be decided whether to disseminate a dataset and in what way.
- Participants inform the EIM and other consortium members if they wish to publish or disseminate any datasets, whether in a direct way or indirectly.
- Before any dissemination activity may take place the participants must examine the possibility of protecting generated results.
- Upon (affirmative) dissemination decision the dataset will be made available (regarding different dissemination types see section 2.2)

2 FAIR DATA

All collected and generated data will be implemented in the DWH in a standardized way. The DWH includes several databases which are depending on the structure and contents of the different datasets. For managing the databases the database-management-tool “Mesap” will be used, which provides also several data preparing and identifier mapping functions. Mesap is developed and commercially provided by the Seven2one GmbH.

A selection of existing data as well as data collected and generated during the project will be made available to interested research groups and interested parties from policy and industry. The following section outlines how the data will be exploited and made accessible for verification and re-use and how data will be curated and preserved upon closure of the project.

2.1 MAKING DATA FINDABLE

For making data findable, a data catalogue will be prepared, which will be implemented in the REFLEX project website (www.reflex-project.eu). The catalogue gives an overview of all provided datasets and affords the metadata of them. The scope and design of the metadata will be oriented on the metadata structure of the “Open Power System Data” platform (www.data.open-power-system-data.org/) see Table 8.

Table 8: Scope and contents of the metadata for a provided dataset

Category	Content
Name	Name of dataset - a concise one (short but informative)
ID	Dataset identifier
Description/Notes	Short description of scope/contents of the dataset. Also specific remarks (e. g. restrictions, data gaps etc.)
Keywords	List of used keywords for the dataset
Version	Dataset version given as number and/or date. Also the information if it is the latest available version
Last changes	Short description of changes to previous dataset version
Timescale	If the dataset is a time series: • Values for which years e. g.: - 2010-2050 (yearly) - 2010-2030 (5-year-steps) - 2010, 2012, 2017, 2022 • Values structure in the course of year if applicable e. g.: - seasonal - hourly - quarter-hourly • Structure of type days if used
Spatial reference	Spacial reference of values with scope and level of differentiation/aggregation e. g.: - EU 28 (NUTS 0) - EU 28 + X (NUTS 3) - Poland, Germany,…
Sectoral reference	Sectoral reference of values, if applicable e. g.: - Households, Industy, Traffic… - Road Traffic, Rail Traffic, Aviation... - C24_C25 , D, E (Codes from Eurostat)
Sources	Used sources to prepare/provide the dataset, if possible with links to the primary data/original input data
Attribution	Recommended text for attribution
Contact	Contact information for questions/remarks
Access	Terms of data access/usage (licence, free of charge, XX € protective charge etc.)
Field documentation	List of used fields within the dataset with following subcategories: • field name: e. g. capacity_installed • type (format): e. g. number (float) • unit: e. g. MW • description: e. g. installed electrical capacity at the end of year

Each dataset can be unambiguously identified via the combination of dataset name and the version label. Both will be included in the unique dataset ID. With regard to the publication and transparency of results of the project work (e. g. in journals), the preparation of tailored dataset packages for the publications is considered. A package contains together with a short content description of the package a compilation of:

- the metadata of the published results
- the metadata of the relevant datasets required to verify the results, as long as these can be made available.

It will be discussed within the REFLEX consortium, if for these dataset packages digital object identifiers (DOI) will be used, which could be easily included in publications as a reference.

2.2 MAKING DATA OPENLY ACCESSIBLE

Within REFLEX three different possibilities for data dissemination will be considered, as described in Table 9.

Both, the collected input data as well as generated data, will be made available, mostly as open access according to the guidelines of the EU. The specification of data which will be made openly available is still under discussion. It will be decided within the project consortium case by case and after that updated in this section of the DMP.

Table 9: Possibilities of data dissemination considered in REFLEX

Dissemination	Description
Open Access Publication	Owners will be granting royalty-free access of a meaningful selection of generated results to other participants and to the public, possibly restricted by appropriate embargo periods and/or respecting restrictions from editors of scientific journals and organizers of conferences.
Commercial Exploitation	Data suitable for commercial exploitation (e. g. for a commercial re-use by consulting companies) will be managed by the project partner ESA^2 which was funded after completion of the EU funded innovation project ESA2 explicitly with the purpose to exploit research results (including research data) related to (coupled) energy systems modelling.
Indirect	Parts of the generated data will be disseminated only indirectly as part of intermediate or final results of models and/or as qualitative outcome based on post-analysis of results.

The access to all provided data will be offered via the project website (www.reflex-project.eu/). The planned formats are:

- csv
- xlsx
- sql

In the case of an *Open Access Publication*, the dataset can be easily downloaded. The download links for different formats are given within the metadata in the category “Access”.

In case of *Commercial Exploitation* of a dataset, a registration procedure for all those interested in such datasets will be implemented. This includes the opportunity, to differ the conditions for access depending on type of the inquirer or planned re-usage (e. g. dataset is free of charge for public scientific institutions for scientific work, but with a charge in case of commercial re-use by a company). After registration of the request of a dataset, a time-limited download link will be provided via e-mail to the registered contact together with the terms of usage and as the case may be with the invoice. The requisition will be implemented in the metadata, which are available free of charge in any case, in the category “Access”.

2.3 MAKING DATA INTEROPERABLE

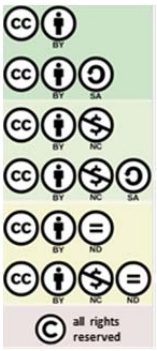
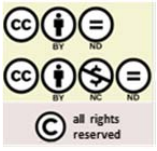
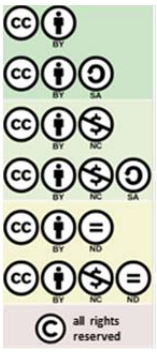
To increase the interoperability of provided data, commonly used vocabularies for the metadata contents as well as for the identifiers and the contents of the identifiers within the datasets will be applied. These include standardized name conventions and codes used in official statistics (e. g. for countries, regions etc.). Furthermore, specific energy system topics related name conventions will be orientated on the “Open Power System Data” platform (www.data.open-power-system-data.org/). An additional mapping procedure or the provision of mapping tools for data users is not envisaged.

2.4 INCREASE DATA RE-USE

The datasets will be made available to third parties as soon as they are generated, prepared and reviewed for publication/commercial exploitation and when the conditions of dissemination are decided and possible protections of dataset are clarified within the consortium. However, additional restrictions by setting appropriate embargo periods and/or respecting restrictions from editors of scientific journals and organizers of conferences are also possible. A generally valid statement regarding the embargo periods is not possible at the moment. It can differ from case to case.

The use of Creative-Commons-Licences (CC) for provided data will be discussed within the consortium. Table 10 gives an overview of conceivable licences for the different types of dissemination. Which licence is used will be decided dataset by dataset and after that updated in this section of the DMP.

Table 10: Conceivable licences for the different types of dissemination

Dissemination	Data group	Licences
Open Access Publication	Existing model input data	
	Collected and generated new model input data	
	Generated final result data of the EMS	
Commercial Exploitation	Existing model input data	
	Collected and generated new model input data	
	Generated final result data of the EMS	
Indirect		

Regarding the data quality assurance the processes given in Table 11 are implemented in the project work.

Table 11: Processes of data quality assurance

Data group	Processes
Existing model input data	- Harmonization of model input data to ensure a consistent analysis within the EMS and regarding the defined scenario storylines. For the same information have to be used the same dataset (values) in all models. (Consortium decision) - Harmonized data will be provided to all models before initializing the EMS run
Collected and generated new model input data	- Minimum two internal reviews of the generated new model input data - Additionally Peer-Reviews in case of publication in Journals - Harmonization of model input data to ensure a consistent analysis within the EMS and regarding the defined scenario storylines. For the same information have to be used the same dataset (values) in all models. (Consortium decision) - Harmonized data will be provided to all models before initializing the EMS run
Generated intermediate model output data	- Check of plausibility of intermediate output data of a model during EMS runs by the responsible modeller before implementation of the data in the DHW for data transfer to another model
Generated final result data of the EMS	- Minimum two internal reviews of the generated final result data - Additionally Peer-Reviews in case of publication in Journals

The consortium will continue to provide the data via the REFLEX project website for a limited period of time after the end of the REFLEX project. The project website will be maintained during this period. After this period, an appropriate reference/link to the final data repository will be integrated in the REFLEX project website, which is no longer being maintained after that.

The final repository for these data has not been chosen yet. The choice of repository will depend on:

- location of repository
- research domain
- costs
- open access options
- prospect of long-term preservation.

The following approaches for long term data provision (or a combination of them) are conceivable:

i) The data remain in the DWH of the project partner ESA². The data provision is transferred to the website of the ESA² Company (www.esa2.eu).

ii) Another repository could be ZENODO https://zenodo.org/. This is online, free of charge storage created through the European Commission's OpenAIREplus project and is hosted at CERN, Switzerland. It encourages open access deposition of any data format, but also allows deposits of content under restricted or embargoed access. Contents deposited under restricted access are protected against unauthorized access at all levels. Access to metadata and data files is provided over standard protocols such as HTTP and OAI-PMH. Data files are kept in multiple replicas in a distributed file system, which is backed up to tape every night. Data files are replicated in the online system of ZENODO.

   Data files have versions attached to them, whilst records are not versioned. Derivatives of data files are generated, but the original content is never modified. Records can be retracted from public view; however, the data files and records are preserved. The uploaded data is archived as a Submission Information Package in ZENODO. Files stored in ZENODO will have MD5 checksum of the file content, and it will be checked against their checksum to assure that a file content remains correct. Items in the ZENODO will be retained for the lifetime of the repository which is also the lifetime of the host laboratory CERN which currently has an experimental programme defined for the next 20 years. Each dataset can be referenced at least by a unique persistent identifier (DOI), in addition to other forms of identifications provided by ZENODO.

iii) A third option is provided by the Technische Universität Dresden, which is currently setting up an institutional, inter-disciplinary repository with long-term archive in the project OpARA. It will provide open access long-term storage of data, including metadata and will go into production in 2017.

Other institutional and thematic repositories will be considered and evaluated in the next months. The procedure will be discussed and decided at the end of the project lifespan. In any case the data will be available for third parties after the end of the project. The length of time for which the data will remain re-usable is not restricted.

3 ALLOCATION OF RESOURCES

The EIM is responsible for data management within the REFLEX project (see section 1.7). The estimated costs for making REFLEX data FAIR are 50.000 Euro.

The costs include:

- the clarification of data protection and licences,
- the final preparation of data by each project partner for publishing (without effort/costs for data collection/purchasing/generation etc.),
- the processes for assurance of data quality,
- the development and implementation of the data catalogue in the project website,
- the implementation of the registration procedure for access to commercially exploited datasets,
- the data hosting and backup for security,
- the data updating and maintenance of the data and of the data provision

during the project lifespan. These costs are covered by the project funds, mainly by the budgeted personnel costs.

The costs for long term preservation after the end of the project are difficult to estimate at the moment. They depend mainly on the level of convenience of data providing (with/without comprehensive search routines and/or additional consulting) but also on the scope and size of the collected and generated datasets. The permanent costs of preserving datasets on the ZENODO repository will be free of charge as long as the single dataset storage is no greater than the maximum 2GB of data. The permanent costs of preserving datasets on the OpARA repository are planned to be free of charge for TUD members. But the final decision on costs has not been taken.

The costs for long term preservation shall be covered by the collected charges from the commercial exploitation of datasets during the project lifespan and after.

4 DATA SECURITY

Most of the data handled in the REFLEX project are not sensitive regarding the laws governing data protection and data security. An exception represents the data from the DSM-survey. A provision/publication of these data is only possible in an anonymous form.

The DWH as well as the data provision via websites will be implemented on servers with regular backup and data recovery procedures.

5 ETHICAL ASPECTS

The data collection, data storage, data usage, data generation and data dissemination in this project do not affect to ethical issues.

6 OTHER

No other national/funder/sectorial/departmental procedures for data management will be used.

REFERENCES

Bolinger, M., Wiser, R. (2012): Understanding wind turbine price trends in the U.S. over the past decade. Energy Policy; 42, 628-641.

Candelise, C. Winskel, M. Gross, R. J. K. (2013): The dynamics of solar PV costs and prices as a challenge for technology forecasting. Renewable and Sustainable Energy Reviews; 26, 96-107.

Capros P. et al. (2016): EU Reference Scenario 2016, Energy, transport and GHG emissions Trends to 2050. Luxembourg: Publications Office of the European Union. https://ec.europa.eu/energy/en/data-analysis/energy-modelling.

Chen, X., Khanna, M., Yeh, S. (2012): Stimualtng learning-by-doing in advanced biofuels: effectiveness of alternative policies. Environmental Research Letters; 7, 045907.

DIW (edt.), Schröder, A.; Kunz, F.; Meiss, J.; Mendelevitch, R.; von Hirschhausen, C. (2013): Current and Prospective Costs of Electricity Generation until 2050, Berlin: Deutsches Institut für Wirtschaftsforschung.

IEA, NEA & OECD (2010): Projected Costs of Generating Electricity, Paris: International Energy Agency, Nuclear Energy Agency, Organisation for Economic Cooperation and Development.

Junginger, M., van Sark, W., Faaij, A. (eds.) (2010): Technological Learning In The Energy Sector. Lessons for Policy, Industry and Science. Edward Elgar Publishing Limited. Cheltenham, UK, ISBN 978 2 84844 834 6, p. 332.

Traber, T.; Kemfert, C. (2011): Gone with the Wind? – Electricity Market Prices and Incentives to Invest in Thermal Power Plants under Increasing Wind Energy Supply. Energy Economics 33 (2011), pp.249–256.

UBA (edt.), Icha, P. (2014): Entwicklung der spezifischen Kohlendioxid-Emissionen des deutschen Strommix in den Jahren 1990 bis 2013, Dessau-Roßlau: Umweltbundesamt.

VGB PowerTech (2011): Investment and Operation Cost Figures – Generation Portfolio, Essen: VGB PowerTech.