SkiLMeeT

Data Management Plan

Authors: Suhendan Adıgüzel van Zoelen, all Partners

Reviewers: Sadegh Shahmohammadi, Joost van Genabeek, Linda Koopman

Grant agreement no. 101132581

Deliverable: D7.2

Due date: 31.12.2024

Funded by
the European Union

Document control sheet

Project Number:	**101132581**
Project Acronym:	SkiLMeeT
Work-Package:	7
Version:	Second version (M12)
Issue Date:	31.12.2024

Classification

Draft	
Final	X
Confidential	
Restricted	
Public	X

Legal notice

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under grant agreement No. 101132581. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.

Table of contents

List of abbreviations

Aldgate	ALDGATE STRATEGY GROUP SPOLKA Z OGRANICZONA ODPOWIEDZIALNOSCIA
SkiLMeeT	Skills for Labour Markets in the Green and Digital Transition
CEDEFOB	European Centre for the Development of Vocational Training
DMP	Data Management Plan
EC	European Commission
EU	European Union
FEEM	FONDAZIONE ENI ENRICO MATTEI
GA	General Assembly
GDPR	The EU General Data Protection Regulation
HE	Horizon Europe
IBS	FUNDACJA NAUKOWA INSTYTUT BADAN STRUKTURALNYCH
IPS	INSTITUTE OF PHILOSOPHY AND SOCIOLOGY
QAP	Quality Assurance Plan
LISER	LUXEMBOURG INSTITUTE OF SOCIO- ECONOMIC RESEARCH
MB	Management Board
NDA	Non-disclosure agreement
OJV	On-line Job Vacancy
REA	Research Executive Agency
RWI	RWI - LEIBNIZ-INSTITUT FUR WIRTSCHAFTSFORSCHUNG
EAB	External Advisory Board
TNO	THE NERHERLANDS ORGANISATION FOR APPLIED SCIENTIFIC RESEARCH
UU	UNIVERSITY OF UTRECHT
WP(s)	Work Package(s)

1. Executive Summary

SkiLMeeT project which stands for "**Ski**lls for **L**abour **M**arkets in the Gr**ee**n and Digital **T**ransition" is an interdisciplinary research project carried out by eight research institutions, exploring how to exploit the potential of the digital and green transition and help equip the European workforce with the skills required for the future of work.

We will analyse in detail:

- the extent of skills shortages and mismatches,
- the drivers of skills gaps,
- strategies to close the skills gaps, including strategies of re- and upskilling.

The project, started on January 1st, 2024 and will end on December 31st, 2026. It is a project in the Horizon Europe under Culture, Creativity and Inclusive Society work programme 2023-2024 on TRANSFORMATIONS-01-07 — Tackling European skills and labour shortages.

The **purpose of this Data Management Plan** (DMP) is to outline how the research data will be collected and will be handled during and after the project. It will serve to make SkiLMeeT data Findable, Accessible, Interoperable and Reusable (FAIR).

This document is the first version of the Data Management Plan (D7.2). The main topics are:

- what kind of data will be gathered?
- how the data will be gathered?
- how the data will be stored?
- how the privacy aspects of respondents are secured?
- overview or respondents and participants.

The Data Management Plan (DMP) will be updated as planned one more time (Month 24) over the course of the project and also whenever significant changes arise, such as (but not limited to): 1. new data, 2. changes in consortium policies (e.g. new innovation potential, decision to file for a patent), 3. changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving). The second version of the DMP (Month 12) is capturing ethical subjects especially regarding the interviews.

The DMP of SkiLMeeT is part of the pillars of the Open Research Data OpenAIRE of the European Commission enables open access of research data generated by Horizon Europe projects. SkiLMeeT will

follow the legal requirements for participating in the Open access to research data which are set out in the Grant Agreement.

TNO is lead beneficiary of the data management, but all other partners in the project have data management tasks and all partners will strictly follow the rules that have been described in this DMP. RWI supports TNO in research data management through the Research Data Centre Ruhr (FDZ Ruhr) at the RWI and with the experience of research data manager for the Untangled Horizon 2020 project.

The structure and contents of the plan were drafted according to 'Data Management Template (HE):V1.0 – 05.05.2021' provided at the participants portal grand management services.

2. Data Summary

2.1. The purpose of the data collection/generation and its relation to the objectives of the project

2.1.1. General overview

a. Various datasets are available on skills shortages and mismatches, the drivers of skills gaps, strategies to close the skills gaps and the re- and upskilling in digital and green transition. On the one hand, this concerns data from questionnaire research and databases and , big data from online platforms.

The purpose of the data collection/generation in SkiLMeeT is to integrate existing (quantitative) datasets and make them analysable. There is a need to analyse this existing data more thoroughly. Also, little attention has been paid to linking these data. By linking certain datasets and using big data from online platforms, we mean to enrich the information provided by existing data (see Table 1). This data is constructed to use the material coming from the project and to help the project team qualify the results (see table 2).

The term data is restricted to existing data, i.e. statistical data and contextual data (e.g. as derived from government reports, studies, etc.) collected by others, and new data, i.e. data collected in the SkiLMeeT project such as workshop/congresses data and focus group and/or individual interview data. In this DMP (M6) we will give an overview of this data and how it is managed.

a. We rely on secondary material already collected by statistical offices or research institutes to produce (WP2, 4) new results. We will use linked databases and micro-data to understand of transformations

and (causal) impacts requires (country) panel studies if it is applicable (Table 1). We also present the WPs.

b. SkiLMeeT project will collect new data that complement existing datasets and new qualitative data from workshops, focus group discussions, interviews and conference participation. In addition to LinkedIn's aggregated data, we will directly approach LinkedIn users to share their online profile with us. Methods of collecting this new data will be described in detail in coming version of the DMP. In WP5, we will conduct four occupational case studies in six partner countries (Bulgaria, Germany, Italy, Luxemburg, the Netherlands and Poland) with 15 in-depth qualitative interviews per occupation. In the mid-term and final congress, it is aimed to reach min. 180 persons.

Table 1. **Main data sources to be used in SkiLMeeT**

Data source / Origin/Provider	Countries	Period	Main observation unit	N	WP	Description + data utility
European Labour Force Survey (EU-LFS)	EU + some non-EU countries	1995-present	Employee	As the original data will be aggregate to the sectoral or to the regional level, the final size will depend on which variables are exactly used to calculate the aggregates	2-5	• **Description:** This is a EUROSTAT dataset and its ad hoc modules (Eurostat) providing valuable information on accidents and work-related health problems, work organisation and working-time arrangements, and the reconciliation of work and family life. LFS is conducted yearly since 1983. The last available edition is 2019. • **Data utility:** The datasets prepared in WP2 will be used for WP 3, 4 and 5
European Working Conditions Survey (EWCS)	EU + some non-EU countries	2005, 2010, 2015, 2021	Persons aged 15 or older in employment	1500 per country; 50000	3-5	• **Description:** This is a EUROFOUND dataset and covering working conditions of employees and self-employed. Since 1990, it is conducted every 5 years. The last available edition is 2015. It provides information on employment status, working time duration and organisation, work organisation, learning and training, physical and psychosocial risk factors, health and safety, work-life balance, worker participation, earnings and financial security, as well as work and health. • **Data utility:** The datasets prepared in WP3 will be used for WP 3, 4 and 5
European Union Statistics on Income and Living Conditions (EU-SILC)	EU countries	2005,2020	Persons aged 16 and over within private households		3-5	• **Description:** This is a EUROSTAT dataset and is conducted yearly since 2003. The last available edition is 2019. It provides information on education; health, labour and income, employment status, activity status, occupation, hours worked; type of contract; job mobility.

						• **Data utility:** The datasets prepared in WP3 will be used for WP3 and WP5.
Programme for the International Assessment of Adult Competences (PIAAC)	31 countries (incl. 20 EU countries)	2012/13, 2022/23	Adults		2-4	• **Description:** The Survey measures adults' proficiency in key information-processing skills - literacy, numeracy and problem solving - and gathers information and data on how adults use their skills at home, at work and in the wider community. • **Data Utility:** The datasets prepared in WP2 will be used for WP 3 and WP4.
LinkedIn - Professional Network Data on Companies and Individuals	EU + about 200 countries worldwide. We focus on EU-countries	2015-present	Employees and job seekers		2-5	• **Description:** 1. Individual data from platform users: 2. Donation campaign platform users: LinkedIn, Meta and Google 3. Anonymized profile LinkedIn data at individual level (including skills and experiences) and Workforce Dynamics (related areas of expertise per company, including skills) 4. Aggregated data from LinkedIn on transitions towards green and digital jobs • **Data Utility:** The datasets prepared in WP2 will be used for WP 3, 4 and 5.
RevelioLabs Data						• **Description:** Revelio Labs provides workforce intelligence by analyzing resumes, and other publicly available employment data. The company uses natural language processing (NLP) and machine learning to extract insights from textual job market data, offering a comprehensive view of labour market trends, hiring practices, and workforce composition. Data would provide insights in occupational transitions to green and digital jobs.

						• **Data Utility:** The datasets prepared in WP2 will be used for WP 3, 4 and 5.
European Skills, Competences, Qualifications and Occupations (ESCO)	EU	2017-present			2-4	• **Description:** The European Skills, Competences and Occupations classification (ESCO) identifies and categorises skills, competences, qualifications and occupations relevant for the EU labour market and education and training. ESCO systematically shows which skills are essential or optional (context-dependent) for different green and digital occupations and makes clear which occupations are similar based on the required skills. ESCO provides also the lists of green and digital skills. • **Data Utility:** The datasets prepared in WP2 will be used for WP 3 and 4.
O*NET	US	1997-present			2, 5	• **Description:** O*Net is a commonly used taxonomy of occupations and skills. O*Net clarifies the importance of general concepts of knowledge, skills, abilities, work activities, values and interests for 923 occupations in the United States and relates them to relevant tasks and detailed work activities. • **Data Utility:** The dataset prepared in WP2 will be used for WP 5.
European Patent Office (EPO) Data	EU	1990-2018	Patents		2, 4	• **Description:** The text of the patents will be used to construct indicators of technological change with a particular focus on green and digital transitions in WP 2. • **Data Utility**: The dataset prepared in WP 2 will be used for WP 3 and 5.
US Patent Office (USPTO) Data	US	1990-2018	Patents		2	• **Description:** The text of the patents will be used to construct indicators of technological change with a particular focus on green and digital transitions in WP 2.

						• **Data Utility**: The dataset prepared in WP 2 will be used for WP 3 and 5.
Google Patents	US	1990-2018	Patents		2	• **Description:** The text of the patents will be used to construct indicators of technological change with a particular focus on green and digital transitions in WP 2. • **Data Utility**: The dataset prepared in WP 2 will be used for WP 3 and 5.
Further big data sources: PayScale (Professional Network Data on Companies and Individuals); Glassdoor (Professional Network Data on Companies and Individuals); Indeed (Professional Network Data on Companies and Individuals)	EU	Present			2-4	• **Description:** Big data sources form different platforms on companies and individuals. This data set is optional: it will possibly be used to supplement data from LinkedIn. • **Data Utility**: The dataset prepared in WP 2 will be used for WP 3 and 4.
European skills and jobs survey (ESJS)	EU	2014, 2021	Adult employees	Wave 2021, N=46,000	2	• **Description:** The ESJS provide information on sociodemographic characteristics, job characteristics, job-skill requirements (literacy, numeracy, digital, analytical, manual and interpersonal skills), skill mismatches (vertical; horizontal; mismatches in specific skills; skill gaps and deficits; skill mismatch transitions), initial and continuing

						vocational education and training participation, and labour market outcomes (wages, job insecurity, job satisfaction). • **Data Utility**: The dataset prepared in WP 2 will be used for WP 3.
European Structure of Earnings Survey (SES)	EU	Every 4 years, last 2018	Employees working in enterprises with at least 10 employees	N: country dependent, e.g. in 2018, in Luxembourg: 57.000; in Germany: 997.000	2-3	• **Description:** The survey provides comparable data across countries and over time on earnings. It is a large sample survey of enterprises on the relationships between the level of pay and individual characteristics of employees (sex, age, occupation, length of service, highest educational level attained, etc.) and those of their employer (economic activity, size, and location of the enterprise). • **Data Utility**: The dataset prepared in WP 2 will be used for WP 3.
German Qualification and Career Surveys (QCS)	Germany	6 waves between 1985 and 2018	Workers	20,000-35,000 per wave	3	• **Description:** Representative labour force cross-sections on qualification and working conditions in Germany.
German Sample of Integrated Labour Market Biographies (SIAB)	Germany	1975-2021	Employees	2% random sample of all employees	3-5	• **Description:** Administrative social security records of all employees liable to social security contributions in Germany. • **Data Utility**: The data will be used to study workers' labour market careers and skill (mis-)match in the context of vocational training and technological change in WP 3 and 5.
UN data	Worldwide	1995-2020, Every 5 years	Migrants	All migrants from regions of origin to destination countries	2-4	• **Description:** The United Nations Global migration database is a database on the number of international migrants by country of birth and citizenship, sex and age as enumerated by population censuses, population registers, nationally representative surveys and other official statistical sources from more than 200 countries and territories in the world. • **Data Utility**: The datasets prepared in WP2 will be used for WP4.

French 'Panel Tous salariés'	France	1976-present	Employees	Random sample of all employees	5	• **Description:** It is a database tracking a sample of employees since 1976. Each observation corresponds to an individual's employment in a company over the course of a year (or a period of unemployment benefit for an individual in a given year). • **Data Utility**: The datasets prepared in WP2 will be used for WP5.
Online Job Vacancies (OJV)	Germany	2017 - present			2; 4	
Online Job Vacancies (OJV)	EU27 + United Kingdom + Norway + Switzerland	2018 - present	Job ad		2; 4	• **Description:** This dataset, provided from Lightcast, contains online job ad data, with a set of information attached, such as the EU region it was posted in, the description of the ad and salary.
Community Production (PRODCOM)	EU	1995-2017 / 2018 - present	Manufacturing goods	On average 4288 products per year.	4	• **Description:** this is a EUROSTAT dataset. For what is relevant to SkiLmeeT, it collects data on sold (€) manufacturing production at the product level.
Comunicazioni obbligatorie (COB). Italian linked employer employee data.	Italy	2012 - present	Anonymised employees	Roughly 2.5 million	4	• **Description:** this dataset contains worker flows from one job to another. It provides information, besides worker flows, on education, tenure, local labour market, age and other workers characteristics.
Italian Labour Force Survey	Italy	2009 - 2014	Anonymised employees	Roughly 150'000 observations for each trimester	4	• **Description:** the dataset contains employee level data on salary, sector, occupation, contract type and demographics.
Italian occupational database including tasks and skills	Italy	2012 - present	Occupations	796 distinct occupations	4	• **Description:** this dataset contains, for each occupation, the skills necessary and the level needed of the skill to carry out that occupation.

French establishment survey on energy consumption and CO2 emissions	France	1997 – 2019	Firm	Between 15000 and 1000 unique firms	5	• **Description:** this dataset provides, besides balance sheet data, a survey on consumption and expenditure for energy products
French occupation data	France	2003 – 2019	Firm-occupation	411 distinct occupations	5	• **Description:** This dataset contains occupational employment (in FTE) and wages for French establishments.
European Company Survey (ECS)	EU	2004/05, 2009, 2013, 2019, 2020	Representative sample survey of business establishments with at least 10 employees	Representative sample survey of business establishments with at least 10 employees	2-4	• **Description:** The dataset covers workplace organisation, workplace innovation, employee participation and social dialogue and also includes questions on skills use, skills strategies and digitalisation. • **Data utility:** The datasets prepared in WP2 will be used for WP3 and 4.

Table 2. **EU-wide indicators for labour and skills shortages and skills mismatch in SkiLMeeT**

Indicator type	Construction of specific indicators	Data sets used	Analysis level	Time period	Related Tasks	Partners Involved
Supply of skills	Levels of literacy, numeracy, and problem-solving skills based on cognitive assessment	OECD PIAAC	Countries, sectors, occupations	2012, 2023		LISER
Supply of green and digital skills	Levels of green and digital skills based on self-reported assessment	LinkedIn	Countries, Sectors, Occupations	From 2015	3.4, 4.8, 5.1, 5.2	TNO, LISER
Demand for skills	Worker-level indices for non-routine cognitive analytical, non-routine cognitive interpersonal, routine cognitive, and manual tasks, and a synthetic measure of routine task intensity (RTI).	OECD PIAAC	Countries, sectors, occupations	2012, 2023		LISER
		OJV, LinkedIn	Countries,			

Demand for green and digital skills	Levels of green and digital skills based on job postings		sectors, occupations	From 2015	3.2, 4.7, 5.4, 5.6	UU, LISER, FEEM
Skills mismatch	Estimation of matching efficiency using number of vacancies and unemployed workers	Vacancies: Cedefop data; Unemployed: EU-LFS	Countries, occupations	From 2020		LISER
Skills mismatch	Workers' self-assessed skill mismatch	OECD PIAAC	Countries, sectors, occupations	2012, 2023		LISER
Labour shortage	Occupational shortage ratio: number of new recruitment and number of unemployed, skill content of occupations Companies' difficulties to recruit Based on OECD Skills Shortage index: Model-based forecasting of future occupations in shortage	EU-LFS, OJV, ECS, BCS	Countries, sectors, occupations	2014 – 2023, Forecasts for 2028, 2033		LISER
Skill shortage	Based on OECD Skills Shortage index: (1st hourly wage growth, employment growth and growth in hours worked by occupation; 2nd indicator is translated into a skills index): Model-based forecasting of future skills in shortage	EU-LFS, EU-SILC, OJV, O*NET, OECD PIAAC	Countries, occupations	2005-2021, Forecasts for 2028, 2033		LISER
Occupational mobility	Similarity index for skill profiles between occupations, relative size of worker flows between occupations	ESCO, QCS (skill profiles), EU-SILC, LinkedIn (worker flows) , SIAB	Countries, occupations	2005-2020, 1975-2021	3.4	LISER, TNO
Occupational mobility	Importance of green or digital skills in an occupation, correlation of Balassa index	OJV, LinkedIn	Countries, occupations	From 2015	3.2, 3.4	LISER, TNO
Skills of brown jobs	Description of skills of brown jobs	Italian O*NET	Italy, sectors, occupations	From 2012	4.5	UU, LISER. FEEM

Green and brown skill specificity	skill distance between brown and green occupations	OJV, Italian O*NET	Countries, sectors, occupations	From 2014	4.5	FEEM
Green production and technology	measure of manufacturing production that is green and of green technologies	PRODCOM,PATSTAT	Countries, sectors, occupations	From 2018	2.4, 4.3, 4.4	FEEM
Carbon content of jobs	Carbon content of jobs	EACEI, DADS	France sectors, occupations	from 2003	5.3	FEEM

2.1.2. Overview of the analytical objectives of each of the WPs, the kind of data collected/generated and the main variables

In this section, we give a detailed overview at the level of the Work Packages of the analytical objectives, the kind of data collected, the main variables used, and the type of analysis conducted.

Work package 1: Conceptual framework and joint infrastructure

Task 1.1: Overview of literature and initiatives

Analytical objectives and tasks: This task prepares big data using online job vacancies (OJV) from Lightcast for European countries. The main task is to prepare a list of keywords regarding different skills.

Responsible: RWI

Variables: List of keywords regarding different skills, e.g. green, digital, routine skills

Method and analysis: A key challenge is how to extract skill requirements from the job postings and classify them into skill clusters (e.g. green, digital, routine/non-routine skills). We use Natural Language Processing (NLP) techniques to prepare the job ads; we next compile a list of keywords describing the skills we want to extract (we rely on the description of green skills, digital skills, routine/non-routine skills provided by, e.g., O*Net, ESCO or OECD PIAAC). To match job ads to the keywords, we use semantic matching to find job ads with a high similarity to some of the keywords defining skills, and ensure the quality of the results with manual annotation.

Task 1.2: Conceptual framework

Analytical objectives and tasks: The aim of task 1.2 is to create a conceptual framework and glossary of key terms and concepts related to the topics covered by the project.

Responsible: RWI

Variables: NA

Method and analysis: The conceptual framework considers available data, indicators, measurement, and definitions (e.g., of shortage or mismatch). This framework thus provides a reference point for all analytical work packages (WPs 2-5) and is also used in WP6 in the context of the engagement of stakeholders (Task 6.2), the research community (Task 6.3) policy and policymakers. The conceptual framework must guarantee that all SkiLMeeT's tasks are performed in a coherent way, and that the project's procedures and output can be clearly communicated within and outside the project. This is

crucial for the SkiLMeeT project because the research on data and indicators, particularly in WP2 and WP3, need clear and transparent definitions.

Task 1.3: Protocols for research activities

Analytical objectives and tasks: The aim of task 1.3 is to ensure an efficient implementation of the project and high-quality outcomes, considering ethics, gender, privacy and GDPR, achieve the project outcomes.

Responsible: TNO

Variables: NA

Method and analysis: Protocols for the qualitative and quantitative research tasks will be developed that provide:

1. Information on the research procedures: data collection, data analysis, guides to identify, contact and interview individuals, reporting;
2. Quality control mechanisms: reviewing the quality of the outputs;
3. Documents to support partners in their project work (e.g. letter of consent, reporting template, summary sheets).

The protocols and supporting documents will be made publicly available via the project website.

Task 1.4: Joint infrastructure

Analytical objectives and tasks: The aim of the task is to ensure a coherent approach among project partners, to facilitate communication with researchers, policymakers and stakeholders not directly involved with the project, and to facilitate making available data (including indicators) and methods to outside researchers.

Responsible: RWI

Variables: Not applicable

Method and analysis:

Work package 2: Data on labour and skills shortages, skills mismatches, and their drivers

Task 2.1: Preparing online job vacancy data (OJV) to extract skills needs

Analytical objectives and tasks: This task prepares big data on online job vacancies (OJV) from Lightcast for several European countries. There are several steps: preparing the raw text (including potential translation in English); extracting the skills from the text; and classifying the skills extracted to a taxonomy (e.g., O*Net or ESCO).

Responsible: RWI, LISER

Variables: skills in high demand, content of occupations and changes therein over time

Method and analysis: Using LLM to extract and classify skills (Mixtral, ChatGPT)

Task 2.2: Preparing big data from online platforms to analyse the supply and demand of skills in different EU countries

Analytical objectives and tasks: Preparing big data from online platform LinkedIn to analyse the supply and demand of skills across EU-countries. The supply will be derived from user profiles on LinkedIn and the demand will be extracted form OJV-data and job vacancies at LinkedIn.

Responsible: TNO

Variables: labour mobility, skills similarities of occupations, over- and under-use of skills

Method and analysis: We will use datasets provided by 1) Revelio Labs, 2) a donation campaign among platform users (LinkedIn, Meta and Google) and 3) LinkedIn (aggregated data on transitions towards green and digital jobs). We will apply extraction methods and natural language processing to analyse the user profile data and job vacancy data.

Task 2.3: Data generation for labour and skills shortages in the EU

Analytical objectives and tasks: Generate the data needed to calculate indicators on labour and skills shortages in the EU in WP3 tasks. The task involves two main steps: first, access and store existing data sources collected by Cedefop, Eurofound, European Commission, Eurostat, OECD, LinkedIn and Lightcast. Second, at the sector level and/or occupation level, merge the data needed to examine past/present/future labour and skills shortages.

Responsible: LISER

Variables: Labour and skills shortages in the EU

Method and analysis: access and store existing data sources, merge of data

Task 2.4: Patent data on advances in digital and green technologies

Analytical objectives and tasks: Generate exposure measures to advances in digital and green technologies. The task involves several steps: preparing the patent data (title and abstract); collect a list of keywords for the different technologies; conduct keyword matching using NLP; defining a walkover from IPC/CPC codes to industry of use.

Responsible: LISER

Variables: exposure to advances in digital (robotics, AI) and green technologies

Method and analysis: NLP

Task 2.5: Collecting data on vocational training curricula and linked technological change embedded in patents

Analytical objectives and tasks: Collecting data on vocational training curricula and linked technological change embedded in patents

Responsible: Utrecht University

Variables: Technology exposure, curricula change, content of curricula change (skills)

Method and analysis: Natural language processing methods for deriving these variables from text documents

    - Identify technology via breakthrough patents (Kelly-Papanikolaou-Seru-Taddy '21): discontinuous change in the innovation space – impactful and novel patents
    - Linking patents to training curricula using NLP methods. Idea: rely on textual similarity following the method by Kogan et al. (2019)
- Outcome: training curricula linked to technological change embedded in patents at the level of occupations
- Will be used in Tasks 4.7, 5.4, and 5.6 (jointly with RWI)
- Discussion: The need to share methodologies between, Ludivine, Ulrich, and Felix Grimm was suggested. They are working together to explore methodologies for extracting relevant information from data patent.

Work package 3: Indicators on labour and skills shortages and mismatches

Task 3.1: Heterogeneity of skill levels and worker tasks in Europe

Analytical objectives and tasks: Task 3.1 aims to establish stylised facts about:

the variation in basic cognitive skills, skills mismatch, and worker tasks across countries, occupations, and sectors;

changes in the skill levels and worker tasks between the early 2010s and the early 2020s

Responsible: IBS

Variables: OECD PIAAC 2012 and 2023 surveys.

Method and analysis: Task 3.1 will make use of the data sets for skills and worker tasks, prepared in task 2.3. These data sets cover 21 EU countries. The indicators for skills supply include levels of literacy, numeracy, and problem-solving skills. The data analysis will focus on determining indicators for skills demand include worker tasks measures of non-routine cognitive analytical, non-routine cognitive interpersonal, routine cognitive, and manual tasks, and a synthetic measure of routine task intensity (RTI). The skills mismatch indicators will be workers' self-assessed overskilling and underskilling indicators.

Task 3.2: Indicators for green and digital skills and specialized versus diversified skill sets

Analytical objectives and tasks: generate (a) indicators on the distribution and growth of green and digital skills across EU countries; and (b) measures for how specialized or diversified skill sets are to assess reskilling needs.

Responsible: FEEM

Variables: OJV, skill supply and demand data (task 2,3), share of green or digital skills in an occupation, distribution across sectors, countries and several years

Method and analysis: NLP, principal component analysis.

Task 3.3: Indicators for past, current and forecasted skill shortages and mismatches

Analytical objectives and tasks: (a) generate indicators of past and current labour and skills shortages and mismatches at sectoral and/or occupational levels based on data sources merged in WP2; and (b) forecast future labour and skills shortages.

Responsible: LISER

Variables: (a) For past and current labour shortages, we use variables regarding new hires and unemployed at sectoral and/or occupational level. (b) For past and current skills shortages, we use variables regarding difficulties of companies filling vacant positions at sectoral and/or occupational level. (c) For past and current skills mismatches, we use data regarding skills or education, wages,

working hours, and employment growth and compare the composition of job vacancies with the composition of employment at sectoral and/or occupational level. (d) For future labour and skills shortages, we develop a theoretical model that endogenizes skills requirements across occupations, accounting for equilibrium skill prices and resulting worker allocation across occupations, sectors, and EU regions.

Method and analysis: Using data sources merged in WP2, we use different statistics to measure past and current labour and skills shortages and mismatches at sectoral and/or occupational levels and we develop a theoretical model to forecast future labour and skills shortages.

Task 3.4: Mapping the similarities and distances between occupations

Analytical objectives and tasks: Generate a similarity index and measure of distance in skill requirements between occupations.

Responsible: LISER

Variables: similarity index, skill distance

Method and analysis: Using data on the skill content of occupations (based on survey data or online job vacancies), we use different statistics of similarity (e.g., cosine similarity) to generate a matrix of skill similarity between each pair of occupations.

Task 3.5: Indicators of matching efficiency in European countries

Analytical objectives and tasks: This task develops matching functions that allow us to create indicators on countries' ability to match individuals to jobs over time and contributing factors in our sample of countries, thus quantifying the degree of mismatch by occupation.

Responsible: RWI

Variables: Unemployment, vacancies.

Method and analysis: We use data on the unemployed and on vacancies to estimate matching functions. We compute Beveridge curves at the occupation-level by country to create indicators of labour market tightness for each EU27 country. The matching functions, in turn, allow us to compare matching efficiency across countries. Differences between green and digital jobs are of particular interest in this context.

Task 3.6: Synthesis of indicators

Analytical objectives and tasks: This task provides an overview on the indicators constructed in WP3 and compares them to existing indicators, also stressing the contributions of the SkiLMeeT indicators. The task highlights similarities and differences regarding (a) definitions of skills and labour shortages and mismatch, and of (2) methodologies, e.g., with respect to the combination of different data sources measuring supply and demand of labour and skills. It furthermore highlights how our indicators can be used to (1) quantify shortages and mismatch, (2) analyse potential drivers of shortages and mismatch, and (3) evaluate pathways to reduce shortages and mismatch. This synthesis is an important input for creating feedback loops with stakeholders (Task 6.2), and synergy effects with and dissemination to the research community and research initiatives (Task 6.3).

Responsible: RWI

Variables: All variables created in WP3.

Method and analysis: Synthesis of indicators.

Work package 4: Drivers of skills shortages and mismatches

Task 4.1: Digitalisation, globalisation and supply factors as drivers of skill-use at work

Analytical objectives and tasks: This task analyses to what extent technology, skills supply, and globalisation drive the evolution of routine and non-routine work across countries.

Responsible: IBS

Variables: OECD PIAAC surveys (2012 and 2023), EPO Patent measures of digital technologies, UN and WTO data for trends in globalisation (prepared in WP2).

Method and analysis: The analysis of task 4.1 extends the static analysis of Lewandowski et al. (2022) by studying changes in technology, supply of skills and globalisation over time and their relationship with shifts in the structure of job tasks and skill use at work. Heterogeneities across workers by age, education and gender will be explored.

Task 4.2: Drivers of skill needs and shortages in the labour market

Analytical objectives and tasks: Identify the drivers of skills needs and shortages related to digitalisation, demographics and globalisation. The task consists of four main steps: first, we map skill needs and requirements across occupations, regions and countries over time as calculated in the previous tasks. Second, we measure potential skills gaps derived from the skills mismatch indicator calculated in Task

3.3. Third, we examine the drivers of skills needs and potential skills gaps in terms of digitalisation, demographics (mainly ageing) and globalisation (mainly migration). Fourth, we examine the heterogeneity of these effects across age, education and gender.

Responsible: LISER

Variables: skills gaps, skills needs and requirements; digitalisation, demographics and globalization drivers

Method and analysis: Construction of a skills needs and requirements map and a measure of potential skills gaps. Econometric analysis of drivers and heterogeneous effects.

Task 4.3: Skills needs and shortages in the context of the green transformation

Analytical objectives and tasks: identify the skills needs of firms engaged in green production and innovation activities and potential skills shortages in regional labour markets. The goal is to inform EU policymakers on the reskilling investments required to boost green industrialisation in Europe

Responsible: FEEM

Variables: OJV-derived green skills, EPO patents, PRODCOM, SBS

Method and analysis: 1) We estimate the relationship between the green job vacancies and measures of green production and innovation, conditional on other determinants of skill demand. 2) We analyse wage offers in green job vacancies relative to other jobs to reveal potential skill shortages across occupations and regions. 3) We use the time dimension to identify the causal effect of expanding green production and inventing green patents on skills demand in EU regions using a shift-share identification strategy that exploits variation at the product-region-year level.

Task 4.4: The green transition and reskilling needs

Analytical objectives and tasks: conduct an analysis of the skills requirements and reskilling needs across EU sectors and countries by jointly analysing skills supply and demand.

Responsible: FEEM

Variables: OECD PIAAC, OJV, EPO patents, PRODCOM.

Method and analysis: We test, via regression analysis, whether an increase in green production and patenting activity is associated with an increase in the demand for certain skills (e.g. STEM and technical skills) at the sector-by-country level. We then compare the difference in the skill set required in sectors

producing green goods and using green patents with the skill set of polluting sectors using measures of skill proximity (from task 3.4) that are relevant to inform policymakers on the reskilling requirements associated with the energy transition.

Task 4.5: Skills loss and worker outcomes in the green transition

Analytical objectives and tasks: estimate the transitional costs on the labour market associated with the decarbonisation of the economy. evaluate both the loss of skills post-displacement for brown occupations and the differential effects of skills losses on outcomes such as wages, employment and tenure.

Responsible: FEEM

Variables: Italian linked employer-employee data on work-flows. Italian occupational database on skills and tasks.

Method and analysis: map via a transition matrix, green and brown jobs together with other dimensions such as educational attainment, gender and the geographical distribution of such jobs, and estimates the skill distance using both factor analysis and an angular separator distance.

Task 4.6: Energy price shocks and skills demand

Analytical objectives and tasks: This task explores how the sharp increase in energy prices in 2022 affected the skills needs, employment, wages and the greening of the economy. The analysis will compare Germany, which experienced a larger shock because of its energy mix, with France, where the impact was more modest as it relies more on other energy sources.

Responsible: RWI

Variables: The analysis combines information on energy prices by source and detailed job vacancy data with administrative data from France and Germany.

Method and analysis: The analysis uses a shift-share design combining the local industry structure (region x industry) and the evolution of the energy prices to identify whether the energy price shock (a) reduced employment in energy-intensive local industries, (b) increased or reduced the demand for green jobs and (c) had any influence on wages.

Task 4.7: Effect of innovation on skills mismatches

Analytical objectives and tasks: analyse the role of technological change for skills acquired through vocational training and for skill mismatch

Responsible: Utrecht University

Variables: technology exposure, skill content, skill mismatch

Method and analysis: regressions, instrumental variables

Task 4.8: Drivers of labour-market mismatch

Analytical objectives and tasks: This task analyses the sources of labour-market mismatch as identified in Task 3.5 in which we estimate Beveridge curves and matching efficiencies to explore the role of search frictions and country-specific regulations.

Responsible: RWI

Variables: We combine data from EU-LFS on job search-related information with data from e.g., OECD statistics on country-specific institutional factors (e.g., minimum wages). Moreover, Cedefop has information on the prevalence of Working From Home (WFH). This data allows us to study if WFH can provide remedy against skills shortages by broadening the search scope. An advantage of Cedefop data is its broad coverage across countries and occupations.

Method and analysis: The main specification uses standard regression techniques such as OLS. To study whether firms offering WFH attract more skilled workers in relation to comparable firms that do not offer WFH, we use COVID-19 as a quasi-experiment to perform a DiD estimation and supplement this approach with matching methods to find suitable control groups.

Work package 5: Pathways to reduce skills shortages and mismatches

Task 5.1: Working conditions and well-being for workers facing a job-to-job transition

Analytical objectives and tasks: This task aims to (a) explore workers' job-to-job transitions in Europe, (b) how they are related to skills transitions and (c) examine the role of working conditions in this context. The focus is on identifying and comparing the differences in working conditions between workers' previous and current jobs at an occupational level, covering a broad range of factors including wages, working hours, occupational safety and health (OSH), as well as overall well-being at work.

Responsible: RWI

Variables: We utilise various sources of data such as EWCS and SES. We use EU-SILC and LinkedIn data as well as potentially other sources of big data such as Payscale (prepared in WP2) across European countries to track the job-to-job transitions of workers across different occupations and business sectors based on their profile and perform in-depth analyses of representative data coming from country-

specific administrative and survey data, e.g., the German social security records and the French 'Panel Tous salariés'.

Method and analysis: Standard regression techniques such as logit and OLS models.

Task 5.2: Skills transferability across jobs and the quantity of worker flows

Analytical objectives and tasks: Measure skill loss and transferability of skills across occupations and how this correlates with the amount of worker mobility between occupations

Responsible: LISER

Variables: Transferability of skills across occupations

Method and analysis: Based on the skill similarity index, we analyze how much skill similarity between one occupation and another encourages mobility (following the logic in Gathmann and Schoenberg, 2010), A high skill similarity but low mobility points to other factors like entry barriers (e.g., occupational licensing or certification requirements) inhibiting job mobility.

Task 5.3: Carbon intensive occupations, skills transferability and the quality of worker flows

Analytical objectives and tasks: evaluates the effect of a climate policy, the EU emission trading scheme, on skills losses, exploring whether wage and skills losses are more pronounced in carbon intensive occupations that are expected to be more vulnerable, and which workers manage to avoid wage and skills losses, thus identifying pathways to beneficial labour market outcomes in the green transition.

Responsible: FEEM

Variables: French establishment survey on energy consumption and CO2 emissions. French occupational data. O*NET.

Method and analysis: Regression analysis. Shift-share.

Task 5.4: Skills mismatches between supply and demand, and worker mobility

Analytical objectives and tasks: assess the role of skill gaps for workers' labour market careers

Responsible: Utrecht University

Variables: skill gaps, wages, occupational mobility, technology exposure

Method and analysis: regressions, instrumental variables

Task 5.5: The effect of automation technologies on the educational choices of youth

Analytical objectives and tasks: study the role of training for workers' resilience to technology shocks

Responsible: IBS

Variables: technology exposure, skill requirements, skill supply via vocational training, wages, occupational mobility

Method and analysis: regressions, instrumental variables

Task 5.6: Role of training for workers' resilience to technology shocks

Analytical objectives and tasks: The aim of task 5.6 is to study the role of training for workers' resilience to technology shocks.

Responsible: UU

Variables: skill gaps, wages, occupational mobility, technology exposure

Method and analysis: regressions, instrumental variables

Task 5.7: Migration and cross-border commuters as a way to alleviate skill shortages

Analytical objectives and tasks: Determine how much international mobility by immigrants and cross-border workers help to alleviate skill gaps in a country

Responsible: LISER

Variables: Skills supplied by migrants, cross-border workers and resident population, skill needs of the local economy

Method and analysis: Measure the skills that immigrants and cross-border commuters bring to a country (here: Luxembourg); compare it to the skills of the resident population and to the skill needs. Quantify how much migrants and cross-border workers reduce the skill gaps in the local economy.

Task 5.8 : Occupational case studies on the impact of upskilling and reskilling on skill transitions

Analytical objectives and tasks: Examine the impact of upskilling and reskilling on skill transition and a better skills match.

Responsible: IPS

Variables: Upskilling, reskilling, skill transition, skills match.

Method and analysis: Four occupational case studies will be conducted in six partner countries (Bulgaria, Germany, Italy, Luxemburg, the Netherlands and Poland) with 15 in-depth qualitative interviews per occupation. A standardized interview protocol will be constructed by IPS, TNO, IBS, LISER and RWI. Data will be transcribed and a thematic content analysis will be done using the approach as described in Braun & Clarke, 2006.

Work package 6: Engagement, dissemination and communication

Task 6.1: Dissemination, exploitation and communication strategy

Analytical objectives and tasks: SKILMEET sessions have a dissemination purpose but will also offer opportunities to gather feedback about the project's results and the directions of the policy work.

Responsible: Aldgate

Variables: NA

Method: Webinars, mid-term and final conferences.

Task 6.2: Stakeholders' engagement

Analytical objectives and tasks: Stakeholders' engagement - This task gathers contacts and builds a community of stakeholders

Responsible: IPS, Aldgate

Variables: NA

Method and analysis: Action research

Task 6.3: Engagement with research community

Analytical objectives and tasks: This task gathers contacts and builds a community of researchers.

Responsible: RWI

Method and analysis:

- present and explore final results with the researchers and international partners.
- discuss the main elements of the Strategy Papers
- facilitate the exchange of experiences, knowledge and insights.
- celebrate, scale and diffuse SkiLMeeT's project results.

2.2. What types and formats of data will the project generate/collect?

It is provided above an overview in section 2.1.2 of the analytical objectives of each of the research WPs, the data that need to be collected/generated in each of the WPs and the variables which are important.

WPs 2, 5 and 6 are the WPs that SkiLMeeT collects data through CVs of the volunteers, interviews, focus groups and conference participation. This data collection concerns qualitative interviews and discussion with relevant stakeholder groups and policy makers.

WP2

For our study on the patterns and backgrounds of occupational transitions, particularly into green and digital industries, we need specific data related to job profiles and historical employment data of individuals.

Our primary focus is to analyses occupational histories and the type of transitions between different occupations. We aim to answer the following questions:

- What types of occupational transitions are occurring (or have occurred)?
- What are the demographic and professional backgrounds of the individuals making these transitions, such as Skills, education, age, previous occupations, and regional location?

Specifically, for transitions into or out of **the digital and green sector**, we seek to understand the profiles of these candidates in terms of their past occupations, age, regions, and other relevant data. To collect data from individuals, we are designing a data donation platform, where individuals would share their CVs with us. Participants can choose to either upload their CV or to fill in an online form. We are going to run a campaign for this data collection.

WP5. In order to identify promising pathways to reduce skills shortages and mismatches, quantitative data will be analysed (e.g. EU-SILC, EWCS, SES, OJV, LinkedIn, administrative data) and qualitative data will be collected (through occupational case studies with interviews and focus groups).

2.3. Will SkiLMeeT re-use any existing data and how?

SkiLMeeT will do research with re-used data (i.e. data collected by others). Table 1 provides an overview of the owners of the current data. We will not re-list this data. In section 3.4 , we have clarified how we re-use the data.

3. FAIR data

3.1. Making data Findable, including provisions for metadata

We are using existing data that is already Findable. For as far as the datasets are available to other users (see Table 1), then for this data already metadata exists. We are also creating new data so we do need to deal with protecting personal data when storing and analysing it.

The work relies on existing quantitative data sources as well as new data sets that are collected within the project through workshops, interviews (WP5 and WP6), web scraping and direct consultation of platform users (WP2). The data will be different, according to the different activities conducted in the project.

These materials will be kept available at TNO in a safe and findable way. Use of the data will be made available by indicating on the SkiLMeeT-website how to contact TNO, and data will only be provided guaranteeing the anonymity of respondents/interviewees. In the informed consent letters we will ask participants of workshops/focus group discussions whether or not they agree with any re-using of their anonymised data or not. After conclusion of the project, we will use the commonly used procedure used by the EU to maintain the information links to previous projects.

3.2. Making data openly Accessible

The project wants to combine and integrate existing European (and other international) datasets, new data and methodology to find the solutions and recommend the policy options on labour and skills shortages and mismatch, their drivers, and viable ways to tackle these shortages and mismatch, facilitated through feedback loops with stakeholders and the research community.

The project wants to build on EU-cross country and multilevel databases, using existing EU-wide surveys and online platforms. The project will provide open access to all aggregated datasets, methods, and tools (considering privacy aspects) to researchers, policymakers, and the general public. In all these aggregated datasets nothing can be traced back to individuals. However, the project will also build on national datasets. For legal reasons or for reasons of privacy, access to these datasets can be restricted to the research project by the respective institutions that collect and manage these data. Therefore, these data cannot be shared between the partners and cannot be made available after the project.

The workshop data will be made available, for as far as participants have allowed to access their data. This data will not be identifiable at the individual level. The data about the workshops will be made identifiable by metadata and can be accessed on a public database. Persons wanting access to this data will need to register on the TNO-website.

For the new qualitative data collected in workshops, focus group discussions, the datasets will be available as described under 2.1, but any information that makes the participants identifiable, will be removed. These data will be accessible as Excel or SPSS-datasets. These data types are commonly used, so no extra documentation will be provided to use the data.

When using the new qualitative data, we will collect the information about the identity of the person accessing the data. We will collect all relevant information about the identity. Because the information is only available for researchers, we will ascertain if the request comes from an existing and regulated research institute.

3.3. Making data Interoperable

Because we are using existing data, we are working on interoperable data. We make this data even more interoperable by aggregating EU-level data at the sector and the regional level. We will provide the rules for this aggregation. More cannot be done from our perspective.

The workshop, interview and new generated data cannot be connected to other data sources.

3.4. Increase data Re-use (through clarifying licenses)

Existing data cannot be licensed to permit the widest re-use possible. This data already exists and are managed by their current owners.

For the data collected during the project:

Licenses for Data Sharing: We will apply the Creative Commons Attribution 4.0 International (CC BY 4.0) license to our datasets. This license allows others to share, use, and adapt the data, provided that appropriate credit is given, a link to the license is provided, and any changes made are indicated. This ensures maximum flexibility and usability of the data for various scientific research endeavors.

Legal and Ethical Considerations: All data collected will be anonymized to ensure compliance with the General Data Protection Regulation (GDPR) and to protect the privacy and confidentiality of the

participants. Anonymization procedures will be documented and strictly followed to ensure no personal identifiers are included in the shared datasets.

Re-use Conditions: The data will be freely accessible for scientific research purposes. Users must provide appropriate attribution as specified by the CC BY 4.0 license. There are no restrictions on commercial use or adaptations, encouraging wide-ranging applications of the data in various research contexts.

Data Documentation and Metadata: Comprehensive documentation and metadata will accompany the datasets to ensure that other researchers can fully understand and effectively use the data. We will adhere to [relevant metadata standards, e.g., Dublin Core, DataCite] to provide detailed descriptions of the data collection methods, variables, and any processing steps undertaken.

Access and Availability: The anonymized datasets will be deposited in [name of repository], a reputable and accessible data repository. The data will be available for download without any embargo period, ensuring immediate access for the scientific community.

Support for Data Users: To support data re-use, we will provide a contact email for queries related to the dataset. Researchers can reach out for clarification on the data, documentation, or any other relevant aspects to ensure effective utilization of the data.

By providing clear licensing and thorough documentation, we aim to enhance the re-usability of our data, thereby contributing valuable resources to the broader scientific community and fostering further research and innovation.

4. Allocation of resources

The costs for making data FAIR in our project are limited to the costs pertain to keeping the data accessible. These costs will be borne by TNO. As far as is needed, we will depose our data at DANS (Data Archiving and Networked Services, The Netherlands), and will comply to the requirements of OpenAIRE (www.openaire.eu). This only deals with the workshops, interviews related data and the data collected via the data donation platform.

TNO decides on the long-term preservation (costs and potential value, how long and how what data will be kept). This means that TNO needs to take care of data security issues. TNO keeps data available on a protected TNO SharePoint site. This means that access can be regulated, and backup-copies are made continuously.

No associated costs will be involved with the data sharing.

The data preparation for making data FAIR is not costly since we will use mostly the existing data for our research.

The costs to gather new data FAIR will be covered as part of data gathering in each of the relevant WPs. Costs related to open access to research data are eligible as part of our Horizon 2020 grant and budgeted in the project

The Project Coordinator and Project Management Team are responsible for ensuring that data are handled accordingly to the DMP by all Consortium Members. MB will follow all the data management activities.

As stated in the GA, the 'enriched' data as well as new data to be collected in the project will be stored in the cloud storage facilities (SharePoint of TNO) or computer systems hosted by RWI, LISER, FEEM, UU or TNO to the extent permitted by data use contracts, available for other consortium partners. Both will provide the privacy and security enablers to protect the personal data and provide secure storage and access. All data generated will be centrally stored on the coordinator's central computers using the TNO's standard archiving data control procedures and routinely (daily) backed up onto secure areas on a central server, according to standard policies and procedures.

Existing data cannot be licensed to permit the widest re-use. This data is managed by their current owners. The new data that we create is not suitable to re-use. It is mainly available to ascertain reproducibility of our conclusions. Patents etc. are not the objective of the project. Open access costs for 3-4 peer-reviewed articles have been already budgeted. No associated costs will be involved with the data sharing.

The Project Coordinator (TNO) and Project Management Team are responsible for ensuring that data are handled accordingly to the DMP by all Consortium Members. The management board will follow all the data management activities.

TNO decides on the long-term preservation (costs and potential value, how long and how what data will be kept). This means that TNO needs to take care of data security issues. TNO keeps data available on a protected SharePoint site. This means that access can be regulated, and backup-copies are made continuously. Data will be stored in the RWI??? secured computer systems as well. RONALD Original data is not accessible, so any errors in use of open data can be corrected.

SkiLMeeT will make use of TNO regulations that are based both on national and EC ethics requirements .

5. Data security

For the protection of data during the project, we will apply organisational and technical measures to ensure that data is properly stored and made available. The organisational measures pertain to who may have access to data; in principle, only for research purposes. The technical measures are focused on pseudonymisation/anonymisation of any new research data, separating identifying and non-identifying data in separate files, use of secure servers and secure cloud servers. Our main data storage system is SharePoint Online. Information is stored in Microsoft's SharePoint Online service. The data is located within the European Union (Ireland and the Netherlands). Access to the information is possible worldwide based on a TNO account or a TNO partner account, but always on the basis of Multi Factor Authentication. Microsoft offers standard features to ensure the confidentiality, integrity and availability of the information. In addition, TNO makes a daily back-up of the information to a service in the Amazon S3 Cloud, also within the European Union. Both for SharePoint Online and for the backup use is made of encryption 'in transit' and 'at rest'.

Partners are advocated to use the TNO-data storage facility. They may have additional ways to store data. How they do this must be agreed upon with TNO.

6. Ethical aspects

Our study aims only to collect already available, existing information on the implementation of digital and green technology, use of skills, employment opportunities. SkiLMeeT research will mainly include secondary data analysis at different levels (micro and macro). For the research WPs (3, 4, 5 and 6), we rely on secondary material already collected by statistical offices or research institutes. We cannot share this data.

On the other hand, we will have workshops, interviews and focus group discussions in WP5 and WP6. We will also collect CVs from the job seekers or workers. For the recruitment of the participants of the workshops, interviews, discussions and collecting CVs, we rely on voluntary cooperation of these collaborating subjects. Participants of the workshops, webinars and scientific conferences can indicate their preparedness to share data. Participants will be asked to give informed consent for taking part. The following basic principles as "informed consent procedures" will be applied to the implementation of all research during SkiLMeeT in order to protect research participants (see D1.2 Research protocols):

› Avoidance of exploitation;

› Just distribution of benefits and burden;

› Respect for persons:

• Participants are treated as autonomous agents;

• Participants with diminished autonomy are entitled to protection;

› Respect for human dignity;

› Scientific validity;

› Scientific, social and/or educational relevance;

› Respect for rights and specific interests of (specific groups of) research participants, and/or the community/society

Furthermore, during the process of obtaining informed consent from participants, the researcher(s) of SkiLMeeT will provide information that is comprehensible for the target population, and made available beforehand as much as possible (so the subject can make a well thought decision) regarding the:

› voluntariness of participation;

› nature, purpose and duration of the research;

› procedures, including the expected duration and the extend of strain for participants;

› reasonably foreseeable factors that may be expected to influence participants' willingness to participate, such as potential risks, discomfort, adverse effects and benefits;

› right to decline to participate and withdraw from the research once participation has begun, without any negative consequences, and without providing any explanation;

› recording of voices and images, where applicable;

› confidentiality protection and limitations;

› procedures for incidental findings;

› applicable insurance guarantees;

› period of time to which the consent applies;

› re-use of specified data in the current, future or other research, where applicable;

› Incentives for participation;

› participants should be informed on the fact that/told that data will be stored and encrypted for a certain period of time;

› names and details of the responsible researcher and contact person(s) for questions about the research and rights of research participants;

Researchers of SkiLMeeT will keep adequate records of when, how and from whom informed consent was obtained, unless this could or proves to be detrimental to participants and/or where the formal registration of the informed consent has a negative effect on the execution of the study.

Furthermore, as the host institution, TNO has an appointed Data Protection Officer (DPO) under the General Data Protection Regulation 2016/679, and TNO has a well-defined a data protection policy, as well. We confirm that SkiLMeeT will made available the contact details of the DPO to all data subjects involved in all researches of the project.

In the informed consent forms and Information sheets (see D1.3 Research protocols) the below statement or similar information will be provided to the research participants:

"*Under the prevailing privacy legislation, you have various rights you can exercise to keep control of your personal data. For example, you have a right of access and correction to your personal data. More information on how to exercise your privacy rights can be found in the Privacy Statement on our website (link: Privacy statement | TNO). There you will also find the contact details of our Data Protection Officer.*"

SkiLMeeT will keep the data stored for 5 years after closing the project.

7. Other issues

SkiLMeeT will make use of TNO regulations that are based both on national and EC ethics requirements (see D1.2 Research protocols).

8. Further support in developing our DMP

We will look into other tools that can help us in our DMP. For our publications, we will use Zenodo, our website (if possible) and OpenAIRE.

SkiLMeeT

Follow us:

linkedin.com/company/SkiLMeeT

bsky.app/profile/skilmeet.bsky.social

twitter.com/SkiLMeeT

www.skilmeet.eu

Funded by
the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under grant agreement No 101132581.

SkiLMeeT Partners:

TNO innovation for life

RWI

LISER Luxembourg Institute of Socio-Economic Research

Utrecht University

ibs institute for structural research

Fondazione Eni Enrico Mattei

Institute of Philosophy and Sociology, Bulgarian Academy of Sciences

ALDGATE STRATEGY GROUP