PROJECT	
Project number:	101188188
Project acronym:	SoGreen
Project name:	Social Aspects of the Green Transition (SoGreen)

Work Package	10
Deliverable Number	D10.1
Deliverable Name	Updated Data Management Plan (DMP)
Deliverable Type	DMP
Dissemination Status	PU
Lead Beneficiary	SHARE-ERIC
Authors	Stefan Gruber
Date Submitted	24 Aug 2026

SoGreen – Social Aspects of the Green Transition

Data Management Plan

Document history

1.0	23.05.2025	• Initial version.
2.0	17.08.2026	• Updated data collection status across all RIs • Paragraph on software and file formats refined and moved to section 2.2 on accessibility • Section 2.3 on interoperability refined and restructured • Section 6 added

1. Data summary	• The SoGreen project is the first cooperative project between the four largest European social science surveys: the *Survey of Health, Ageing and Retirement in Europe* (SHARE), the *European Social Survey* (ESS), the *Generations and Gender Programme* (GGP) and *Growing Up in Digital Europe* (GUIDE). • The goal is to enhance the capacities of each research infrastructure (RI) to generate valuable insights aimed at facilitating the study and evaluation of the social aspects of the green transition for different generations and socioeconomic groups across Europe, using a life-course perspective. • Central to the project's approach is the collaborative development of a new tool, the *Green Transition Questionnaire Module* (approximately 30 items — 23 core and 7 extended — developed in WP3). This module is developed to ensure coherence in addressing environmental themes which will then be fielded within the different surveys. • In addition to the survey data that will be collected in the participating RIs, contextual data on climate, air quality and OpenStreetMap data will be used in the SoGreen project – on e.g. the availability of public transport and green spaces in the area of residence. At a later stage of the project, the contextual data will be linked to the survey data using geographic identifiers of the respondents. **Short description of each RI** • **SHARE** is a unique dataset for studying the effects of health, social, economic and environmental policies over the life-course of European citizens aged 50 plus. Covering all continental EU Member States as well as Switzerland and Israel, SHARE can be considered a natural laboratory for assessing how policies and external shocks – like the COVID-19 pandemic – affect individuals' economic circumstances, physical and mental health, health behaviours, and social networks and activities. SHARE fields the SoGreen module in the self-completion questionnaire

	(SCQ) between its 10th and 11th regular wave of data collection, at the end of 2025, in Germany, Greece, Italy, the Netherlands, Poland and Sweden. • **ESS** is a cross-national survey conducted across Europe since 2001, measuring attitudes, beliefs and behaviour in more than thirty nations. ESS provides its cross-national web-interviewing infrastructure (CRONOS) to field modules of items focused on the green transition, selected/adapted in WP3. Eleven ESS countries take part in this panel study: Austria, Belgium, Czechia, Finland, France, Hungary, Iceland, Poland, Portugal, Slovenia and the UK. The first wave was conducted at the end of 2025, and the data were published in April 2026. The second wave was fielded in spring 2026, with data publication scheduled for October 2026. The third wave goes into the field in October 2026, and these data will be published in April 2027. • **GGP** is an interdisciplinary RI on population and family dynamics, collecting cross-nationally comparable longitudinal data on young adults, families, generational exchanges and life courses. Its core is the Generations and Gender Survey (GGS), whose archive includes individual-level data from more than 31 countries with over 300,000 respondents. The SoGreen module is fielded via the GGP In-Between Waves Survey (IBWS) in Croatia and Poland in 2026 and 2027. • **GUIDE** will be Europe's first comparative birth cohort study of children's and young people's wellbeing, tracking wellbeing and psychosocial development together with indicators of children's homes, neighbourhoods and schools across Europe. The survey items of the SoGreen module will be adapted to be accessible to the youngest populations for future data collection that will start after the SoGreen project. Cognitive interviewing takes place with approximately 20 children and young people recruited from the GUIDE CIPS in Ireland, to test the feasibility of the module for a younger age group.
2. FAIR Data **2.1 Making data findable, including provisions for metadata**	• The survey data will be provided through the research data centre (RDC) of each RI. • The SoGreen project website contains an overview of the project goals and results, and links to the RDC of each participating RI where the data can be downloaded. This ensures findability. • ESS and SHARE use Digital Object Identifiers (DOI) to make datasets permanently identifiable and locatable. The data registry links every DOI to a set of metadata (title, author, publication date, copyright etc.). GGP uses citation guidelines per country dataset and is implementing DOIs for its datasets by the end of 2026. GUIDE does not yet have data available, but there are plans to use DOIs as well. • Some parts of the metadata and survey documentation (such as questionnaires or show cards) are provided on the SoGreen website; the main part of the metadata is provided via the website of each RI.

	• Additional findability aids include the SoGreen newsletter and the datasets subpage of the project website, codebook and processing guidelines, and a cross-study guide. • Metadata are provided in machine-readable DDI format, so that datasets are not only findable to humans but also to search services and harvesters. • Datasets are versioned according to each RI's conventions, so that every release remains uniquely and persistently citable.

2.2 Making data accessible	• SoGreen follows the principle "as open as possible, as closed as necessary". For SHARE and GGP in particular, data are released only after registration and acceptance of the conditions of use, reflecting GDPR, personality-rights and disclosure-risk constraints.
	• Access to the SoGreen data is provided free of charge for non-commercial use globally, subject to European Union and national data-protection laws as well as the Conditions of Use of each RI.
	• Data access is given on an individual basis after a standardised registration procedure described for each RI at:
	  ○ https://share-eric.eu/data/data-access (SHARE)
	  ○ https://www.europeansocialsurvey.org/data-portal (ESS)
	  ○ https://www.ggp-i.org/form/accounts/register (GGP)
	• A link to the GUIDE registration procedure is not yet available, as the first full wave of data collection will not take place until 2027 at the earliest.
	• Access services are rendered through public data archives: for SHARE, the Centerdata Archive at Tilburg University; ESS data are archived and distributed via the ESS Data Portal, managed by SIKT (the Norwegian Agency for Shared Services in Education and Research); for GGP, access services are provided in-house through the GGP Data Portal, and in the long term the data are archived at Data Archiving and Networked Services (DANS) in the Netherlands.
	• Software and file formats: SHARE and GGP provide their data sets in Stata and SPSS format. ESS data are distributed via the ESS Data Portal in SPSS, Stata and CSV format and can, in addition, be retrieved programmatically via the ESS API (https://ess.sikt.no/en/api), which serves files in Stata, CSV and Parquet format and can be called from Python, R, Stata or curl. Independently of the formats offered, users working in other statistical software (such as R, SAS or Python) can use the data as well, since conversion between software packages is a standard procedure.

2.3 Making data interoperable	• Metadata standards: All participating RIs document their data with DDI, though at different levels of granularity – SHARE's Data&Documentation Tool (DDT) is based on DDI 2, ESS uses DDI Lifecycle, and GGP prepares metadata according to DDI Lifecycle 3.3. The documentation additionally draws on the controlled vocabularies maintained by the DDI Alliance and published through the CESSDA Vocabulary Service, as well as on general standards such as ISO 3166 (countries), ISO 639

	(languages) and ISO 8601 (dates). The metadata are therefore machine-readable and comparable across the four RIs, and comprehensively enough documented for researchers from other disciplines to interpret the data and combine them with other sources. Full semantic interoperability is not reached, however: variables are not linked to ontologies in RDF and the metadata are not modelled according to cross-domain frameworks such as CDIF or DDI-CDI, which would require variable-level semantic annotation beyond the scope of SoGreen. ● Multidisciplinarity and cross-nationality are key elements of SoGreen; the consortium's geographical coverage encompasses individuals, families and generations across all EU countries, the UK and Israel, often through multiple surveys. ● Interoperability is also increased via the SoGreen Knowledge Mobilisation Lab, a collaborative, transdisciplinary platform for knowledge exchange on the social aspects of the green transition.
2.4 Increase data re-use	● The data are made available to the scientific research community as soon as possible, i.e. as soon as they are cleaned and checked for data-protection / GDPR compliance. ● SoGreen data may be used for scientific research, subject to European Union and national data-protection laws, and are not restricted except for commercial use. ● Archiving the data together with detailed documentation of the (meta)data makes them re-usable without time restrictions. ● Timing: data are released as soon as they are cleaned and GDPR-checked. The first ESS collection (CRONOS-3 W6) was released as open access in April 2026. ● Preservation: data are archived long-term in the RIs' established repositories.
3. Allocation of resources	● The project is structured into 12 work packages, of which six are led by SHARE, two by ESS, two by GGP and two by GUIDE. ● SHARE-ERIC is the coordinator of the project, responsible for coordination — including monitoring that the action is implemented properly — and acting as the intermediary for all communications. ● Professor David Richter (SHARE-ERIC & SHARE BERLIN Institute) oversees the effective implementation of the project and the use of resources under the Consortium Agreement, supported by Deputy Coordinator Professor Agnieszka Chłoń-Domińczak (SHARE-ERIC & SGH Warsaw School of Economics) and by staff at SBI. ● SHARE-ERIC reviews required documents and information, verifies completeness and correctness, submits deliverables and reports, and distributes payments to the other beneficiaries. ● Costs of making data FAIR: data curation, documentation, DOI registration, disclosure control and archiving are covered by the project budget and by the RIs' core funding for long-term preservation.

	• Responsibility for data management: each RI's Research Data Centre is responsible for curating, documenting, publishing and preserving its own data; WP10 coordinates data management and maintains this DMP. • Long-term preservation is funded and carried out by the RIs' established archives (SHARE-ERIC; ESS via SIKT; GGP via DANS) as part of their ongoing infrastructure operation, rather than by the project alone.
4. Data security	• All RIs implement robust data-protection measures to ensure the confidentiality and privacy of their respondents. In panel surveys this applies also across waves, which poses both a challenge and a vital consideration in the generation and use of datasets. • Detailed technical and organisational measures are in place to ensure the security of data processing at each step, in compliance with the GDPR and national data-protection laws. • SoGreen data are released in pseudonymised format (as "scientific-use files" protected by de facto anonymity). Only the released datasets may be used for subsequent scientific research. • All names and other personal information that could identify individuals or households are removed from these datasets. • The secondary data sources, including air quality and OpenStreetMap data, are also provided to the scientific community and linked to the primary data. The linkage is conducted at a level that ensures full compliance with the GDPR and ethical standards. • During the project, data are held on secure, access-controlled servers at each RI, with regular backups and role-based access.
5. Ethical aspects	• Each RI obtains ethical approvals for the data collections and data linkages included in the SoGreen project. This ensures the project abides by the GDPR, common principles for conducting research, ethical considerations, and the FAIR principles. • Participation is voluntary and based on informed consent covering data processing, linkage to contextual data, and archiving and sharing for scientific re-use, in line with the GDPR and national law. • Children (GUIDE): the cognitive interviews with approximately 20 children and young people (GUIDE CIPS, Ireland) are conducted with guardian consent and age-appropriate child assent, with child-safeguarding measures and enhanced data minimisation; the material is used only for instrument development and is not publicly released. • Data linkage: linkage of survey data to air-quality and OpenStreetMap data is carried out at a level ensuring GDPR compliance and non-identifiability.
6. Other aspects	• This DMP complements, and does not override, each participating RI's own data-management policies and data-protection frameworks; national data-protection laws. Each RI's consortium rules continue to apply.