InGRID

Supporting expertise in inclusive growth

www.inclusivegrowth.eu

Deliverable 7.9

POTENTIAL PATHWAYS FOR THE FUTURE DEVELOPMENT AND SUSTAINABILITY OF THE INGRID RESEARCH INFRASTRUCTURE

Karolien Lenaerts, Monique Ramioul,
Kenneth Nelson, Stephanie Steinmetz,
Albert Esteve, András Gábos, István György Tóth,
Janna Besamusca, Charlotte Articus,
Ralf Münnich & Natalie Shlomo

December 2021

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 730998

Abstract

This document lays out some potential pathways for the future development and sustainability of the InGRID research infrastructure. It is a summary and compilation of key documents prepared under the InGRID-2 project and its work plan, including three strategic notes on data, methods and policy.

This report constitutes Deliverable 7.9, for Work Package 7 of the InGRID-2 project.

December 2021

General contact: inclusive.growth@kuleuven.be
p.a. InGRID
HIVA - Research Institute for Work and Society
Parkstraat 47 box 5300, 3000 LEUVEN, Belgium

For more information karolien.lenaerts@kuleuven.be

Please refer to this publication as follows:
Lenaerts, K., Ramioul, M., Nelson, K., Steinmetz, S., Esteve, A., Gábos, A., György Tóth, I., Besamusca, J., Articus, C., Münnich, R., & Shlomo, N. (2021), Potential pathways for the future development and sustainability of the InGRID research infrastructure, Deliverable 7.9, Leuven, InGRID-2 project 730998 – H2020

This publication is also available via http://www.inclusivegrowth.eu

This publication is part of the InGRID-2 project, this project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 730998.

InGRID
Supporting expertise in inclusive growth

Contents

1. Introduction

This document presents Deliverable 7.9 'Roadmap of InGRID sustainability and innovation agenda' of WP7 on 'Strategic advancement of the Research Infrastructure (RI)' of H2020 project 'InGRID-2 – Inclusive Growth Research Infrastructure Diffusion: from data to policy'. The general objective of this work package is to organise coordinating and futuring activities for the further advancement, integration and sustainability of the research infrastructure. The specific objectives of WP7 are:

- to simulate the discussion and debate on deepening and enlarging the RI;
- to evaluate with the research infrastructure community the operational running;
- to understand the future ERA and ESS context and European policy background in which InGRID will operate;
- to define RI data gaps and barriers-of-integration in Southeast Europe and Central Europe;
- to outline the relevant policy drivers and financing possibilities over the next 5-10 years;
- to revise and update the sustainability plan for InGRID RI beyond the project.

This deliverable relates to both this general objective and the specific objectives guiding the research, and, in particular, is connected with the last two points. The deliverable explores future pathways of the InGRID-2 research infrastructure. It recalls, revises and updates the future plan prepared under the first InGRID project (Van Gyes *et al.*, 2017), which ran from 2013 until 2017, and complements it with the insights from several outputs prepared under the second InGRID project, which ran from 2017 until 2021, more specifically: reports on the data forums, special interest groups and stakeholder platform conferences (incl. the spotlight reports) in WP6, reports on research infrastructure gaps in South-East and Central Europe, strategic reviews of the Greek and Slovak research infrastructure roadmaps, a survey targeting InGRID-2 user communities on their future needs, and three strategic briefing notes, along with other deliverables. These insights are further enriched with discussions with the Advisory Board and Executive Committee. The coordination team also participated in a number of brokerage meetings with representatives from other (European) research infrastructures and European Research Infrastructure Consortiums (ERICs), which also helped shaped the thinking on the future of the InGRID-2 research infrastructure. Of importance here, are the agreements that have been concluded with leading networks and stakeholders and that help secure the future of some of InGRID's key infrastructures, such as the OECD (ICTWSS), the European Commission's Joint Research Centre (EUROMOD), the International Microsimulation Association, and others (see Deliverable 7.10 for more details). Where relevant or necessary, this deliverable refers to the InGRID-1 future plan, users' needs survey (Szekér & Van Gyes, 2015) and scoping exercises (a first exercise by Lenau *et al.* (2016) on future needs and challenges from the point of statistical methodology; and a second exercise by Hamon-Cholet *et al.* (2017) on the relationship and interaction the InGRID RI has with the Central and East-European countries). Especially the section on the positioning of the InGRID research infrastructure is largely kept from the previous future plan, as this information still holds for the most part.

The InGRID-2 research infrastructure integrates, innovates and opens the existing, but distributed European social sciences research infrastructure on 'poverty and living conditions' and 'working conditions and vulnerability'. This advanced infrastructure aims to provide the European research community with new and better opportunities to fulfil its key role in the development of evidence-

based European policies for inclusive growth. Inclusive growth is economic growth that creates opportunity to all segments of the population and distributes the dividends of increased prosperity, both in monetary and non-monetary terms (such as health, education and employment aspects), fairly across society (OECD, 2015). Improved well-being and less inequality are key factors in this regard. Inclusive growth is a top priority in the European Union's strategy, for several years already (Schmidt, 2014). The European Commission, for example, argues in its Europe 2020 strategy launched in 2010, that *'in a changing world, we want the EU to become a smart, sustainable and inclusive economy.'* As part of this strategy, and in the accompanying EU Employment Package, the EU wants to create more and better jobs and enhance social inclusion. Targets have been set to reach high employment levels and combat poverty. With the Employment Package, the European Commission wants to boost job creation and focus on job-rich growth (COM 2020 final; European Commission, 2013) as well as to ensure the quality of new jobs. In the proposed 2015 revision of this European Employment Strategy, titled 'moving towards more inclusive labour markets', it is stated:

> *'Job quality has a particularly important role to play, encompassing adequate earnings, training opportunities and access to lifelong learning, the possibility for career progression, measures to improve work-life balance, the quality of the work environment and safe transitions between jobs and back into work for those losing their job.'*

Today, still, as the InGRID-2 project has reached its conclusion, inclusive growth is an EU policy priority. As the current Commission President Ursula von der Leyen stated in her Political Guidelines for the Next European Commission 2019-2024:

> *'I want Europe to strive for more when it comes to social fairness and prosperity. This is our Union's founding promise. I am proud of our unique European social market economy. It is what allows our economies to grow – and what drives poverty and inequality to fall. It ensures that social fairness and welfare come first. Strengthening our social market economy is acutely important at a time when we are redesigning the way our industry and our economy work.'*

An action plan to support the full implement of the European Pillar of Social Rights was put forward, including key proposals on ensuring a fair minimum wage for all workers, improving the working conditions of digital platform workers, as well as related to the Child and Youth Guarantee.

Similar to the future plan developed in the first InGRID project, this deliverable does not present a long-term, full-vision of the future of the InGRID research infrastructure, but has a more explorative scope and a more limited time frame in mind. Most importantly, this is because, in contrast to the future plan developed under the first InGRID project - which could be read as a plan of a following integrating project, financed within the European Research Infrastructures sub-programme of the EU Horizon 2020 science programme - there currently is no call for proposals within the European Research Infrastructures of the Horizon Europe programme of the European Commission that fits with the thematic scope of the InGRID-2 research infrastructure, although the InGRID-2 team still hopes such a call may be launched in the future. As a result, there are several possibilities as to how the InGRID research infrastructure could be further developed and improved, and it is not yet clear which one(s) of these pathways will materialise. The timing, too, is less clear. The potential pathways forward are discussed in this deliverable. Building on the future plan elaborated in the first InGRID project, this deliverable will also lay out the improvements that were made in the InGRID-2 project, and highlight areas where further advancement is required.

2. Positioning the InGRID-2 research infrastructure

The InGRID-2 research infrastructure wants to support the involved social scientists as 'lead' users of the project to access, order, analyse, re-use data and help them translating their evidence-based knowledge to the practitioners' field of European policy innovation (end users). Key in the approach of this top-level, interdisciplinary SSH field is looking for and interpreting problematic trends in social situations and workplaces and to monitor or assess possible policy influences and innovations on these trends. The focus arenas of the InGRID-2 infrastructure are to provide and improve integrated and harmonised data, analytical facilities to link policy and practice, and indicator-building tools to translate this knowledge into benchmarks for policy innovation. Integration and access to a series of relevant data of different national and international providers (pre users) is the starting point.

This section was adapted from the research proposal and agenda underpinning the InGRID-2 project and the future plan developed under the first InGRID project (see Van Gyes *et al.*, 2017).

2.1 Concept of a European social science research infrastructure

A research infrastructure is a facility or platform that provides resources and services for the scientific community to conduct top-level research in their respective scientific fields and to foster innovation. To this end, research infrastructures enable researchers to access, order, analyse, store and reuse data and knowledge in ways that are otherwise impossible. Research infrastructures may be single-sited, distributed, or virtual, and include major scientific equipment and instruments; collections, archives or scientific data; and related systems and features.

A European approach for a research infrastructure is defended in science policy (European Research Area) based on the following arguments:[1]

- to organise the access transnational to all European researchers and not only the ones coming from an individual Member State;
- to look for cost-effective synergies by avoiding duplication of efforts; coordinating and rationalising the use; sharing investments; pooling resources;
- to trigger the exchange of best practice and develop interoperability of facilities and resources;
- to connect national research communities and increase the quality of research and innovation.

A European social science research infrastructure, such as the InGRID research infrastructure, aims to facilitate research on the key societal challenges that Europe is facing, by supporting the social sciences research community. InGRID thus offers 'helping hands' to the research community that it serves. InGRID also goes beyond supporting research, in that the project has an elaborate training programme. InGRID-2 is a distributed research infrastructure (see OECD, 2014 for the definition), which brings together 16 institutes with key expertise, equipment, instruments, collections and data based at different locations across Europe. Renschler *et al.* (2013) developed the following definition for research infrastructures in the social sciences: *'durable institutions, technical tools and platforms and/or services that are put into place for supporting and enhancing research as "public good" resources for the social science*

1 http://ec.europa.eu/research/infrastructures/index_en.cfm.

community' (p. 14). Renschler *et al.* (2013) define five essential features of research infrastructures that are intrinsically interlinked:

- *public good*: the research infrastructure provides services and resources that are non-exclusive, non-competitive, and available to all;
- *user-oriented*: the services correspond to the needs of researchers and can take on various forms, such as data, tools, education and training, and methodological expertise, all contributing to the advancement of a specific field of science;
- *durable and stable* on a long-term basis to avoid losing accumulated benefits. This kind of maintenance requires effective strategies of recognition and legitimation to anchor the research infrastructure in science policies;
- *adaptability*: to remain closely aligned with the needs of users and to gain continued stakeholders support innovation has to be a key feature;
- *method-focused*: infrastructures support the methodological approach of researchers by enhancing opportunities for analysis and replication. Transparent and open access to data and analytical tools and stimulating harmonisation and comparability are important elements in this support.

2.2 Rationale of InGRID as European distributed research infrastructure

Based on this general concept of a European research infrastructure that integrates distributed social sciences 'helping hands' in a specific field of science, the rationale of the InGRID RI can be explained. First, we delineate the community-of-research the InGRID research infrastructure wants to serve. In a second step, we give an overview of the different methodological and servicing 'institutions' that are gathered in the InGRID-2 research infrastructure for this purpose.

2.2.1 Help to tackling the key European social challenges

At the start of the second InGRID project in 2017, the European economy was still recovering from the financial-economic crisis starting in 2008. Despite the mild economic improvements and progress in the labour market around the time, many Europeans were still suffering from the consequences of the crisis. Many experienced decreases in their living standards during the crisis, resulting in poverty, but also growing inequalities between regions, age groups and household types. Over 122 million people - one in four Europeans - were at risk of poverty and social exclusion. Almost 25 million people registered as unemployed, and the new jobs that were available often were of poor(er) quality. Indeed, trend studies on job quality detect a move towards simpler and more intense forms of work organisation in Europe (Holman *et al.*, 2015).

In spite of the economic and social improvements made towards the end of the 2010s, the COVID-19 pandemic, which broke out in 2020 in Europe and is currently in its fourth wave, has once again challenged the European economies and labour markets. The impact of the COVID-19 crisis varies across sectors, companies and workers depending on a number of factors, for example, whether the activities were considered 'essential' or not, whether activities were affected by restrictions on travel and movement, whether activities were disrupted by supply chains issues, etc. Telework surged in the period 2020-21 and millions of workers found themselves in (temporary) unemployment. This crisis, too, highlighted and exacerbated existing inequalities in a number of areas, such as housing, poverty. The economic outlook published in the summer of 2021, however, already indicated some progress after the initial decline.

These developments, moreover, are set in a context of globalisation, demographic changes (notably related to aging and migration), financialisation, welfare-state retrenchment, changing power relations between labour and capital, technological transformations and skills mismatches in the labour market (OECD, 2015; Atkinson, 2015). While some of these trends have been around for decades, they are becoming increasingly intertwined and have accelerated in the past years. As a result, inequalities are

growing (Nolan *et al.*, 2014). In many areas of life, whether it be education, employment prospects or active ageing, chances are increasingly not equal. As the OECD stated: '*This takes a toll on the social fabric of communities, places a heavy economic cost on future growth and reduces trust in governments and institutions.*' (OECD, 2015).

The European policy level is defined more and more as a key lever to tackle these growing social problems (Salverda, 2015; Vandenbroucke *et al.*, 2015; Lechevalier & Wielgohs, 2015), but it is also more and more questioned as a policy actor that can deliver on these matters.

In his first speech towards the European parliament Jean-Claude Juncker, the former President of the European Commission, spoke about the fight against a new state in the Union: '*A 29th state is currently emerging within the borders of the European Union. It is the state where people without jobs live. A state in which young people became unemployed; a state in which we see people excluded, set back and left by the wayside.*' The 'Europe 2020' strategy is since 2010 the overarching policy initiative which joins all areas of EU competence and activity in order to prepare the EU societies for the future. It identifies three mutually reinforcing priorities. The priority of inclusive growth is aiming to raise labour market participation, fight poverty and strengthen social cohesion. Combating poverty and ensuring quality of work and employment are core elements in achieving this objective. More recently, and as indicated above, the current President of the European Commission Ursula von der Leyen again underlined the need to reconcile the social and the market in today's economy, and at the same time underscored the critical role of the (democratic) systems and institutions, including social dialogue and the social partners, in achieving the EU's ambitions in the social domain.

However, the arena of policymakers and decisionmakers is a searching community in this regard. Policy issues have to be clarified, evolving social and labour market situations clearly benchmarked, new policy solutions considered and assessed. Identifying the best policies for reducing inequality is a puzzle that is yet to be resolved.

2.2.2 Servicing a community specialising in European comparative evidence

InGRID serves the social sciences community that aspires to make an evidence-based contribution to this European policy challenge of inclusive growth. Evidence-based policymaking is a concept that is often put forward as the ideal and recommended strategy to do 'good and sustainable policymaking' (Sutcliffe & Court, 2005; UNESCO, 2010). 'Everyone is entitled to his own opinion, but not to his own facts' is in this regard a famous quote of the American senator Moynihan. The necessity to have this kind of common ground in politics and policy making seems essential in the current period of rising populist politics. This search for common ground can be defined as follows: 'Evidence-based decision making aims at assisting decision makers and practitioners to identify different policy options to solve a problem, and then to choose between them' (Unesco, 2010). In this context, the role of the research community is to provide this 'best' evidence from research, which will influence the view of policymakers and complements the opinion-based inspiration of policymaking.

The InGRID scientific community focuses on social in/exclusion, vulnerability-at-work and related social and labour market policies from a European comparative perspective. It is an interdisciplinary field of poverty research, labour studies, policy analysis and social statistics. Key tools in this social science research are all types of data: statistics on earnings, administrative social data, labour market data, surveys on quality of life or working conditions, and policy indicators.

Figure 1. Top-level research community served by InGRID

This European research community is, on the hand, involved in fundamental scientific activities of which proof can be find in the European Framework Programmes (e.g. RE-InVEST, IMPROVE, RESCUE, INSPIRES, GINI, NEUJOBS, UNTANGLED, ...). The community, however, also has links to policy innovation in networks, established and funded by the European Commission and the European agencies. The European Social Policy Network (ESPN) – an EU-funded social policy research network which covers 35 European countries and is managed by LISER - provides the Commission with independent information, analysis and expertise on social policies. In particular, the ESPN supports the Commission in monitoring progress towards the EU social protection and social inclusion objectives set out in the Europe 2020 Strategy. Other important networks in the field of poverty, living conditions and social policies are Net-SILC2 and Net-SILC3 (project-based networks linked to improving the EU-SILC data), and the more informal European Social Reporting Network. In the field of employment policies and vulnerability-at-work the European Employment Policy Observatory (EEPO) aims to improve European and national policymaking by providing information, analysis and insights on the design, implementation, monitoring and evaluation of policies. EurWORK, created by Eurofound, monitors - on a comparative basis - key developments in social dialogue and working conditions that affect work in all Member States and at the European level. It seeks to contribute to the development of evidence-based policymaking and practices that improve the quality of working life. A range of InGRID-2 partners are involved in one or more of these networks. They are in their turn linked to key circles of preparing and discussing European policy innovation: the Employment Committee (EMCO) and the Social Protection Committee (SPC) (notably the indicators subcommittees).

2.3 Integrated, but distributed resources and services

InGRID-2 serves and facilitates in an advanced way the European expertise on inclusive growth from data to policy. Integrating, opening, widening and innovating the InGRID research infrastructure-in-existence is the key goal of the InGRID-2 project. The information in this section is largely based on Van Gyes *et al.* (2017) but updated where necessary.

2.3.1 The InGRID-2 consortium

The following table presents the research infrastructures integrated in the InGRID-2 project (which included more partners than in the first project). The goal of InGRID-2 is to provide an optimised and quality-based access to existing data infrastructures and other research infrastructure allowing tor study inclusive growth in a comparative way. The consortium, therefore, includes datacentres that integrate data at the European level; competence centres with know-how on key official European data sources; institutes that invested in comparative policy databases and/or microsimulation models to investigate the impact of policies with comparative microdata; organisations providing standards of harmonisation and classification; new and innovative data collectors; statistical departments that can provide methodological help and sophisticated technical equipment.

Many of the teams involved in the InGRID consortium have extensive experience in large-scale social sciences projects (for example RECWOWE, EQUALSOC, WORKS, GINI, ImPRovE, NEUJOBS, WALQING, PIQUE, MEADOW, EurOccupations, Woliweb, SPReW, WorkCare, RISQ, AMELI). These transnational research activities resulted in new survey initiatives, in innovative use of existing European official statistics and surveys, in experiments with new ways of statistically analysing data.

A lot of energy has been dedicated to harmonisation of standards for questionnaires, classifications and statistical quality in these projects. The efforts of leading data centres such as LIS and CED to integrate and archive important data on a European level were recognised in particular. It is this distributed research infrastructure that InGRID has tried to stabilise and improve, and open up to a growing science community.

Table 1. The integrated research infrastructure of InGRID-2

Name and type	Specific data sources and/or infrastructure competences
General	
LIS Datacentre Integration of comparative data	Cross-national data archive and research centre, which fosters primary comparative research by providing access to household microdata. LIS collects and harmonises micro-data sets from upper- and middle-income countries that would otherwise be incomparable and places them on a secure infrastructure, and make them available to the Research Community. LIS houses 2 databases, the Luxembourg Income Study (LIS) Database, and the Luxembourg Wealth Study (LWS) database.
IECM – Integrated European Census Microdata (CED)	IECM was started in 2005 as a joint collaboration between the CED (Spain), the Minnesota Population Centre (USA) in partnership with 18 European statistical offices to coordinate, harmonise and disseminate integrated European Census microdata samples. The IECM database contains anonymised microdata samples encompassing as many as 58 censuses and in total about 115 million person records.
Poverty, living conditions and social policies	
TÁRKI Poverty and Living Conditions Data Centre (TÁRKI POLC)	Expertise on European statistical sources and surveys; EUROMOD; Integrated Poverty and Living Conditions Indicator System (IPOLIS) database; provides access to major data surveys (longitudinal and cross-sectional) of the Hungarian society.
IRISS (LISER) Expertise on official international data sources; microsimulation models	CEPS/INSTEAD (now LISER) was recognised by the EU in 1998 as one of only five 'Large Scale Facilities' in the social and economic sciences ('IRISS' infrastructure). It offers access to major international data sets together with scientific and methodological support for analysing these. Data hosted by IRISS allow for in-depth comparative analyses in a number of areas including income, employment and living conditions, wealth, values, ... Secondly, IRISS provides support for the development of microsimulation models, including advanced topics like cross-country comparison, behavioural analysis and long run (dynamic) concerns.

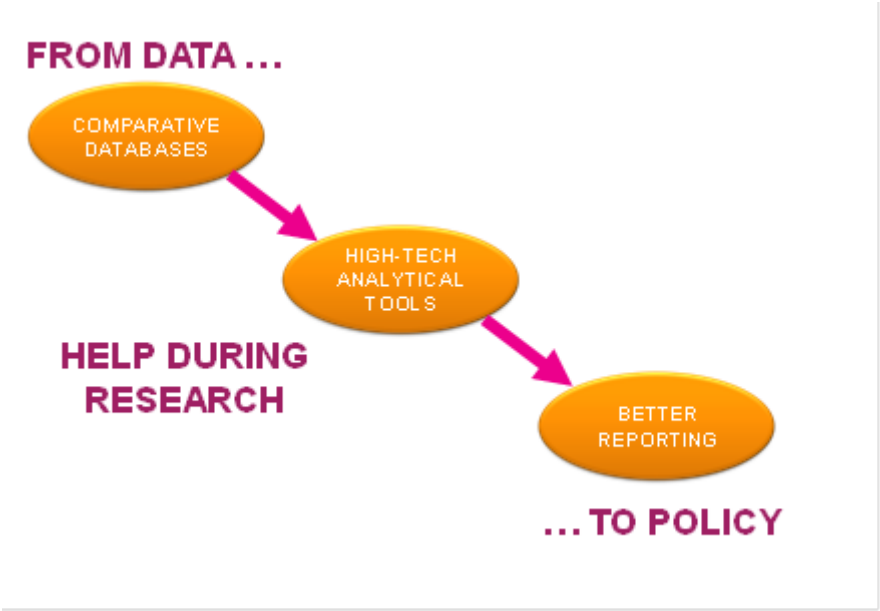
Name and type	Specific data sources and/or infrastructure competences
SOEP by DIW National best practice survey; innovative data	The German Socio-Economic Panel (SOEP) is an independent research-driven infrastructure unit that serves the international scientific community by providing nationally representative longitudinal data and related data sets on private households in Germany. The SOEP RI contains a variety of studies. The centrepiece is SOEP-Core, where every year (since 1984) about 15,000 households and more than 30,000 individuals are surveyed. As early as June 1990, the SOEP expanded to include the states of the former German Democratic Republic (GDR). In 2013 and 2015, SOEP added two major subsamples of about 3,000 immigrants, and in 2016, SOEP will include a random sample of about 2,000 refugees who applied for asylum in Germany between 2013 and 2016. Another important new addition is the Innovation Sample (SOEP-IS), which is designed to improve and develop survey methodologies for assessing the determinants of human behaviour. It offers the international research community a unique platform for cutting-edge research.
British household surveys by UEssex	The British Household Panel Survey (BHPS) was carried out by ISER-UEssex from 1991-2009 (Waves 1-18). The main objective of the survey was to further the understanding of social and economic change at the individual and household level in Britain (the UK from Wave 11 onwards), to identify, model and forecast such changes, their causes and consequences in relation to a range of socio-economic variables. From Wave 19, the BHPS became part of a new longitudinal study called Understanding Society, or the UK Household Longitudinal Study (UKHLS), conducted by ISER. Understanding Society is an innovative world-leading study about 21st century UK life and how it is changing. It captures important information about people's social and economic circumstances, attitudes, behaviours and health. The study is longitudinal in its design and of high quality.
SOFI by SU Comparative policy databases	SOFI has a strong tradition in comparative welfare state research of highest international standards. SOFI also hosts a number of unique comparative and institutional data sets on major social policy programmes. The Social Policy Indicators Database (SPIN) integrates a number of unique institutional and comparative data sets on public policies that are hosted by our institute. This includes the Child Benefit Data set (CBD), the Parental Leave Benefit Data set (PLB), the Social Assistance and Minimum Income Protection Interim Data set (SAMIP), the Social Citizenship Indicator Programme (SCIP), the Social Insurance Entitlements Data set (SIED), the Social Policy in East Asia Data set (SPEAD), and the Out-of-Work Benefits Data set (OUTWB). Together, these data sets provide a vast amount of quantified indicators on major social benefit and transfer programmes, covering up to 35 countries (including EU 27) and for some policy programmes with data going back to the 1930s.
CSB by UA Comparative policy databases; Expertise on official surveys, EUROMOD	The Herman Deleeck Centre for Social Policy (CSB) at the University of Antwerp organises on-site access to and transfers user knowledge of microsimulation models (including EUROMOD), large-scale survey data on income and living conditions (including EU-SILC and SHARE), and hypothetical household simulations of tax-benefit systems and the cost of essential goods and services (HHoT/ EUROMOD, CSB-MIPI and the CSB EU Reference Budgets Database). The CSB-MIPI data set, developed by CSB, provides data on the three main pillars of minimum income protection (minimum wages, social assistance for working age households and guaranteed pensions). It covers 25 EU member countries and 3 US States. CSB-MIPI contains cross-national and cross-temporary comparable model hypothetical household simulations for a large range of household types and income cases as well as long-term time series on gross benefit levels. In addition, it addresses the conditionality of social assistance benefits, the associated rights and in-kind benefits. In addition, visitors can get access to the CSB EU Reference Budgets Database, which brings together the results of two international projects on the out-of-pocket cost of essential goods and services for households. The CSB EU Reference Budgets Database covers the cost of food for 26 EU countries, and the cost of essential goods and services related to housing, primary health care and personal care for 8 European countries.
EUROMOD by UEssex Microsimulation tool	EUROMOD is a state-of-the-art policy analysis tool that allows studies on the functioning of, and interplays between, different types of tax and benefit programmes, thereby making in-depth distributional policy analysis possible. EUROMOD links microdata from household surveys and policy regulation (codified into analysable units) in a single user interface. EUROMOD is now expanded to cover all EU Member States and has been used as the software platform for many 'spin-off' country models outside the EU and is currently being used as the platform for a UNU-WIDER-funded project which will build tax-benefit microsimulation models for several African countries.
Working conditions, vulnerability and labour policies	
HIVA-KU Leuven	Organises on-site access to and transfers user knowledge on the EWCS and ECS, organised by Eurofound.

Name and type	Specific data sources and/or infrastructure competences
CEE Expertise on official surveys	The infrastructure provides expertise on data about work and employment systems at different geographic levels: French surveys: The French survey on working conditions (CT), a survey on Health and career path (SIP) and a survey where occupational physicians evaluate employees' exposure to occupational hazards (SUMER); linked employer/employee surveys: A survey on organisational change and computerisation (COI), a survey on working conditions and on psychosocial risks at work (CT and RPS), a survey on work and employment relations (REPONSE) and a survey on training and skills developments (DIFES and DEFIS). European surveys: EWCS (European Working Conditions Surveys) and ECS (European Company Surveys). The last edition of the ECS has been inspired by the MEADOW guidelines, the main output of the FP6 MEADOW project coordinated by CEE; ESENER (European Survey of Enterprises on New and Emerging Risks) is an employer survey looking at how safety and health risks are managed in European workplaces, organised by the European Agency for Safety and Health at Work. International surveys: PIAAC and TALIS (from OECD).
WageIndicator by CELSI Innovative data	WageIndicator databases are collected by the WageIndicator Foundation, based in Amsterdam. Data management is performed by CELSI. The WageIndicator databases cover four different databases. Core is the WageIndicator web survey on work and wages. It is a continuous, volunteer, anonymous, multilingual, multi-country comparable questionnaire, posted on the WageIndicator websites in 87 countries in all five continents, and the target population is the national labour force. The survey data is organised in annual releases, available from 2001 onwards. It has 720 variables concerning wages and benefits, workplace and firm characteristics, working conditions, attitudes, education, occupation, industry, household and personal characteristics. Sample sizes are large: in 2015 more than 72,000 individuals entered valid wage data. The websites exploit also a mini-survey with a limited set of questions of the web survey, with more than 283,000 individuals entering valid wage data in 2015. This database provides new, elsewhere not available data, allowing for in-depth analysis of wages across countries; accompanies with WageIndicator Minimum Wages database; Cost-of-Living database & Collective Agreements database
AIAS/AISSR by UvA Expertise on official data sources; comparative policy data; harmonisation standard	The Institutional Characteristics of Trade Unions, Wage Setting, State Intervention and Social Pacts (ICTWSS) includes nearly 200 variables and covers 51 countries and 55 years (1960-2014). The WISCO database links occupations to tasks, job profiles and educational requirements, consisting of an occupation database with 1,700 occupational titles, translated into the languages of more than 70 countries, and coded according to the international classification ISCO-08. Expertise on the EU microdata sets EU-SILC and EU-LFS.
Statistical quality management and methodological equipment	
S3RI by Southampton	The University of Southampton Statistical Sciences Research Institute (S3RI) has a world class reputation for research and development in statistical methodology, delivering and providing technical solutions for real applications and problems, and delivering and supporting training and knowledge transfer. Researchers are interested in visiting S3RI for working alongside experts in a wide range of statistics topics including survey sampling, experimental design, statistical modelling, biostatistics and Bayesian analysis. In addition to excellent pool of academics, S3RI offers high quality desk space and computing facilities that can be utilised by visitors to perform their analyses.
Economic and Social Statistics Department UNI-Trier	Large-scale simulation facilities and local expertise in the fields of survey sampling methods, small area estimation, synthetic data generation, Monte Carlo simulations for survey methods, microsimulations, estimation methods for producing Census outputs by combining survey and administrative data sources, and the use of the R software.
UNIPI-DEM	The unit UNIPI-DEM of the Pisa University is a statistician research team, which has developed high quality research skills on the estimation of poverty and living conditions indicators at small area level (e.g. unplanned areas or domains).
CMIST by UNIMAN	Social Statistics and CMIST can provide excellent facilities for visiting researchers to benefit from the expertise of its staff which include statisticians, demographers, medical sociologists, data science, socio-economics and in addition, gain access to data through the UK Data Service.
CEPS	Expertise on using new web-based methods to observes new skills and new jobs in the labour market; as leading think tank at the European level, extensive how-to-knowledge on science valorisation towards the European policy arena.

2.3.2 Range of data resources and analytical services

InGRID wants to support social scientists as 'lead' users of the project to access, order, analyse, re-use data and help them translating their evidence-based knowledge to the practitioners' field of European policy innovation (end users). Key in the approach of this top-level, interdisciplinary SSH field is looking for and interpreting problematic trends in social situations and workplaces and to monitor or assess possible policy influences and innovations on these trends. It is about analysing and interpreting comparative data; connecting this kind of scientific analysis to a greater or lesser extent to policy assessment and evaluation; translate these findings and observations in a concise and transparent way to policy makers. Indicator-building is an important element in this scientific process of providing policy evidence. A policy indicator is a quantitative or qualitative measure of how close one is to achieving a set goal (policy outcome). Indicators help to analyse and compare performance across population groups or geographic areas, and can be useful for determining policy priorities and are as such a much used instrument in European policymaking of inclusive growth.

Figure 2. Three steps of facilitating research by InGRID



The components of the InGRID-2 research infrastructure are hereby (a) providing and improving integrated/harmonised data, (b) analytical facilities to link policy and practice, and (c) indicator-building tools to translate this knowledge into benchmarks for policy innovation.

Figure 3. Facilities and services of InGRID

The data expertise and competence centres of InGRID-2 are providing in this regard to the defined research community the following types of social sciences services and facilities:

1. international expertise and competences on key data collections, mostly European surveys or data (EU-SILC, LFS, European Household Finance and Consumption Survey, European Working Conditions Survey, European Company Survey, ESENER survey), but also data from Member States (census data, national working conditions surveys and employer' surveys, socio-economic panels, ...). Taking data protection rules into consideration, researchers are guided and helped in the use and analysis of these data;
2. integrated data sets: multi-country, harmonised income (LIS) and census microdata (IECM), indicator databases built-from the mentioned surveys (EUROMOD, IPOLIS); the CSB-EU Reference Budgets database;
3. innovating data collection strategies beyond these 'official' data: web survey initiatives like the WageIndicator databases, MEADOW/InGRID protocols for organising/improving working conditions surveys, access to national best practices like the German SOEP (Socio-Economic Panel) (DIW), British Household panel and Understanding Society UK Household longitudinal study or French REPONSE (linked employer/employee survey);
4. tools for harmonised classification: WISCO or the world data base of occupations; translation and classification tables for national surveys and indicator-building based on textual policy databases;
5. comparative policy databases: Institutional characteristics of Trade Unions, Wage Setting, State Intervention and Social Pacts (ICTWSS); the Social Policy Indicators Database (SPIN); CSB-MIPI data set on minimum income protection;
6. microsimulation environments: tax-benefit microsimulation model for the European Union EUROMOD, LIAM2-based microsimulation models; UNI-Trier social statistics large simulation infrastructure (including AMELIA);
7. statistical quality standards: small area estimating techniques (UNIPI); S3RI centre for technical support on small area estimation, sampling and inference for Big Data; data linkage;
8. indicator-building techniques: high expertise on indicator-building: quality of work (KU Leuven), vulnerable groups (TARKI, UvA), social policies (SU, UA, LISER).
9. IPOLIS is a comprehensive, multidimensional, cross-country and over-time comparative database on the quality of life of vulnerable groups in the European Union.

2.3.3 Stakeholders of the InGRID-2 research infrastructure

Besides its lead users, the InGRID-2 research infrastructure aimed to serve two additional stakeholder groups, i.e. pre-users and end-users of the InGRID research infrastructure. End users are defined as policy innovators. Based on scientific evidence they strive for policy innovations and re-formulations. They belong to European or national politics, public administrations or civil society. They can be just simple or plain users of the scientific results, but also the ones that finance or commission the research. Pre-users are particular data providers or collectors. These include European agencies, official statistical bureaus and international or national data initiatives of European interest.

The InGRID-2 research infrastructure of integrated and comparative (EU-wide) data sources is fed with a series of data initiatives. As the inventory papers of InGRID-1 (see Gabos & Kopasz, 2014; Szekér & Van Gyes, 2015; Greenan & Seghir, 2015) and of InGRID-2 (see Besamusca & Steinmetz, 2019; Meylemans & Lenaerts, 2020; Kostolny *et al.*, 2021) show, data inflow come from European official survey initiatives like EU-SILC, LFS and European household finance and consumption survey; fragmented academic data gatherings (ESS, SHARE, ISSP); empirics coming from European and international agencies and international organisations; major national survey instruments such as census data, established working conditions surveys (mainly in the Nordic countries, in some cases matched employer-employee data) and country best-practices such as the German Socio-Economic Panel study. These initiatives are complemented with own, more first-mover data procurements based on web(crawling) data or policy indicator-building. In InGRID-2, in particular, the potential use of administrative data and of text databases for both analysis of poverty, vulnerability, and living and working conditions, and for indicator-building, were explored.

Particularly interesting stakeholders are in this regard international organisations. Both the OECD and ILO are involved in providing data and in conducting or commissioning applied policy research that focuses on policy innovation. OECD also supervises some international surveys (PIAAC; PISA). In the InGRID-2 project, the connections with these international organisations were strengthened through dedicated research visits from InGRID-2 researchers at their facilities, which involved close collaboration using their datasets for InGRID-2 outputs (e.g. ICTWSS database, research on how to measure the platform economy, research on education using PIAAC, etc.). The same observation applies to European agencies and institutions: European Foundation for the Improvement of Living and Working conditions (Eurofound): European Working Conditions Survey; European Company Survey; European Quality of Life Survey; European Agency for Safety and Health at Work: ESENER survey; European Centre for the Development of Vocational Training (Cedefop): Employer survey on skill needs; European institute for Gender Equality; European Union Agency for Fundamental Rights: surveys on discrimination; Eurostat: EU-SILC; Labour Force Survey; Structure of Earnings Survey; European Central Bank: Household Finance and Consumption Survey (HFCS), etc.

2.4 Different than the same, the same as different ones

The mission of inGRID-2 as a research infrastructure is thematically and policy driven: it is about serving a research community focusing on the EU 'policy strategy' of 'inclusive growth'. As such, the InGRID research infrastructure differs from other previous and current, recognised social sciences research infrastructures at the European level, e.g. European Social Survey (ESS), SHARE survey, Generations and Gender, and CESSDA (data archives).

- Especially the first three are one harmonised source of data collection. The InGRID research infrastructure has a more *distributed network* structure and is specialising in integrating different data sources. The research infrastructure deliberately does not start from the perspective of infrastructure integration by archiving, nor starting new data collections (as very often the case in European social science infrastructure initiatives), but by (a) keep integrating and improving the existing data, and (b) keep developing the inclusion of also the next - as important - research steps of analysis

tools and knowledge transfer practices; As such the infrastructure is more comparable with European research infrastructures in other sciences (Wernli *et al.*, 2013).

- The *distance in the value chain* between the technical resources of the InGRID research infrastructure and the academic collaboration, based on these instruments, is *very short* and guaranteed by the composition of the consortium. Many InGRID partners are not 'technical' providers (archives, data collection initiatives or IT installations) that have to 'sell' something to interested 'lead users', but 'lead users' that integrate, improve and open their cumulated RI to other users. In this opening and servicing, collaboration is established - thematically so to speak - with what can be defined as specialised technical providers. European data centres/archives as LIS (income household) and IECM (census data) are involved as partners. The data products from large-scale surveys, organised by official European and national agencies, are integrated and users are given expert advice and service on the use. National best practices of these kind of data products are given the opportunity to connect and to increase the use of their products to a European audience of social scientists (e.g. SOEP in Germany, UK Understanding Society Study).
- The infrastructure is within the social sciences servicing an *interdisciplinary* community of mainly sociologists, economists and political scientists.
- As InGRID pays deliberately attention to not only situational data, but also to integrating policy data and connect in analytic tools these policy indicators to social trends, a *strong link is made with innovation*, not so much business or commercial innovation, but political, public service or social innovation. Contributing to this kind of evidence-based European policy necessitates of course a continued integration/harmonisation of policy data and indicators, besides the development and improvement of methods to link these policy data to situational data in analysis;
- As the focus is not on one source of data - as the others mainly focus on surveys - the distributed infrastructure has also *a more open character*. Looking to the potentials of innovative forms of 'big' data, namely internet data and administrative data, which are new or currently only rarely used in the dedicated research community, is already a component of the RI and will be a point of attention.
- As the InGRID research infrastructure is built upon existing, distributed data collections and not on one data initiative that is managed by the RI itself, the excellence of the RI is of course more *depending on 'others'*: national or international data providers. Strong networking and stakeholders exchange should and is as a result an important strategic success factor for the RI performance.

3. Innovation agenda

3.1 Innovation agenda: moving from InGRID-1 to InGRID-2

The first part of this recalls the innovation agenda that was proposed and implemented when moving towards an advanced research infrastructure (from the first to the second InGRID project). For that reason, this section extensively builds on the future plan related to the InGRID-1 project developed by Van Gyes *et al.* (2017) and the work plan of the InGRID-2 project. We recall the main innovations in the second InGRID project with reference to the first InGRID project, as this helps to clarify the issues that remain to be addressed following InGRID-2. These are tackled in the following section.

In order to strengthen the integration of the activities and to secure the impact, the different services and activities of the InGRID research infrastructure are clustered into two thematic pillars, 'Poverty and living conditions' and 'Working conditions and vulnerability'. This two-pillar conceptualisation of the research infrastructure is linked to the dual flagship implementation of the European Inclusive Growth Strategy with (a) social policy initiatives focusing on tackling poverty and social exclusion, and (b) initiatives related to the employment challenge of new jobs and new skills.

Although the relationship between research and policy is not linear, Sutcliffe & Court (2005) specify that evidence informing policymaking should be accurate, of high quality and objective. This implies that the research evidence should be representative, relevant, free from bias and statistically correct, have a high level of credibility, and follow from rigorous and tested processes and methods. Whether the evidence is timely and topical, generalisable across specific cases and has policy implications will determine if the evidence is useful for policymaking. Finally, research evidence should be accessible, usable and understandable for policymakers. To meet these requirements, the current European RI on inclusive growth is facing three types of advanced challenges, which InGRID tackles as horizontal focus areas.

The first and main challenge is related to data, necessary to provide good evidence (accurate, credible and relevant). Procuring, managing and integrating data is key to the InGRID research community to provide their evidence. As such, the community is still confronted with several problems regarding data: problems with access to existing data, problems in the collection of new data, an extensive need for new data and the need to properly deal with quality limitations of existing data. Integration of comparative data and the improved harmonisation of data efforts is as a consequence the first focus area of InGRID RI innovation.

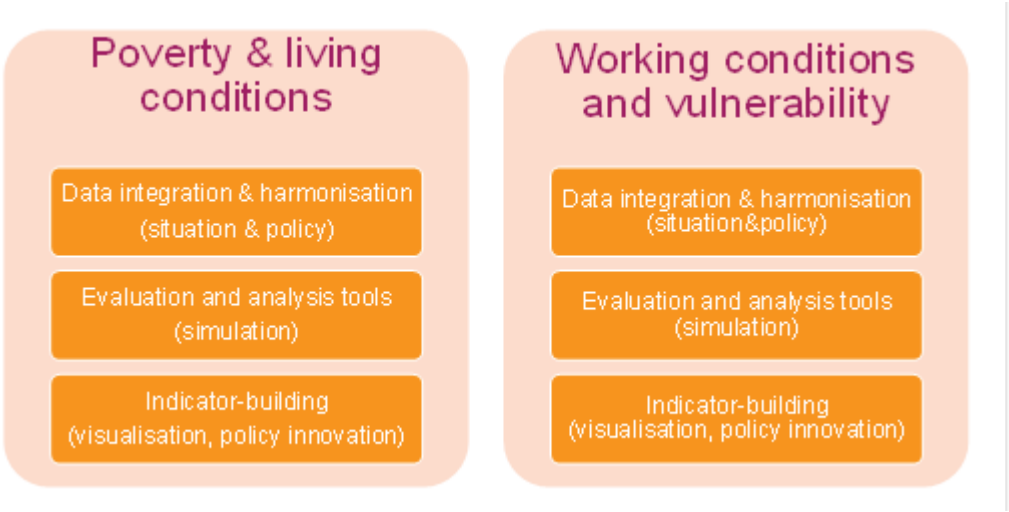
A second type of challenges relates to the methods within social sciences to support among others the credibility of the evidence. Interpreting and applying new or existing evidence requires high-quality tools. Besides statistical quality guidance and improvement, the InGRID infrastructure delivers here as services statistical expertise and simulation environments. Facilities to linking policy monitoring and evaluation to situational data is an important approach in this regard of the InGRID infrastructure.

The third challenge is the gap between policy and research, which can be related to the tensions that exist between the world of researchers and the world of policymakers. Show and tell is important in developing evidence-for-policy. At the European policy level indicator-building is an important process in this regard. European social and economic governance is framed in scoreboards (e.g. MIP scoreboard), prepared by indicators' sub-groups (e.g. sub-groups of Social Protection Committee-SPC and Employment Committee-EMCO) and driven by benchmarked policy targets (e.g. the 5th EU2020 goal of at least 20 million fewer people in or at risk of poverty and social exclusion). Indica-

tors are the road signs to know one is heading in the right direction to meet or exceed the policy targets. Knowledge exchange and improvement of European indicator-building belongs as a consequence to the core concept definition of the InGRID infrastructure.

Harmonised data integration, analytical tools linking policy and situation and indicator-building are thus the three focus areas of the InGRID-2 infrastructure to help and service the defined lead users.

Figure 4. Pillars and focus areas of the innovation agenda



The following table summarises the innovation agenda for the InGRID research infrastructure, as it transformed into an advanced infrastructure.

Table 2. Innovation agenda of InGRID (moving towards an advanced research infrastructure)

POVERTY & LIVING CONDITIONS	WORKING CONDITIONS & VULNERA-BILITY
Data integration/harmonisation	
Extension of Integrated Poverty and Living conditions Indicators System (IPOLIS) in scope and coverage	Harmonisation/integration data on collective bargaining and minimum wages
Exploration of harmonisation longitudinal data on educational careers	Harmonisation/integration working conditions data on vulnerable groups
Data linkages (and small area estimation) from statistical standards perspective	Integrated micro series of working conditions surveys and international surveys on employers' behaviour
Combining data tools for dynamic microsimulation	New types of (web)data and its use
	Historical data of the EU-LFS: feasible to compile?
Improvement of analytical tools	
Conceptualisation and measurement of out-of-work benefits	New methods to examine employers hiring practices and skill transferability
Extending EUROMOD (new policies and new tools)	New methods to measure new occupations and new forms of work
Integrating data on welfare services	New methods to measure working conditions using administrative databases: availability - feasibility
Small area estimation techniques and regional poverty measurement	Exploration of microsimulations approaches in comparative working conditions research
Valorisation/reporting tools and new indicators	
Household hypothetical tool and representing policy relevant indicators	Developing multidimensional vulnerable group indicators
Indicator protocols on migrants' social rights	Developing policy indicators on OSHA management
Demographic factors and poverty indicators	Developing indicators to assess progress in working conditions

3.1.1 Poverty, living conditions and social policies

3.1.1.1 Data integration and harmonisation

In recent years, major progress has been made in the establishment of social surveys that cover all, or nearly all, EU countries. EU-SILC and EU-LFS are two prominent and commonly used examples. However, the analysis of European survey data has for long been restricted due to lack of contextual data on social policy and other institutional arrangements relevant for poverty and social exclusion outcomes. By establishing more formal collaboration between different data providers and research infrastructures, InGRID **harmonises and integrates** different types of relevant RIs. By combining comparative contextual level data on social policies and comparative data on poverty and social exclusion outcomes in a single data shell, InGRID made it easier for researchers and policymakers to conduct institutionally informed analyses of social exclusion outcomes, which is at the core of evidence-based policymaking. The **IPOLIS** database established in the first InGRID project plays an important role in this process of integrating outcome-oriented and contextual level data in the areas of poverty, living conditions and social policy.

Despite the achievements of the EU and the research community in setting up comparative micro-level data sets in areas of poverty and social exclusion, several questions are still hard to address based on existing survey data. One obstacle that plagues many European social surveys, such as EU-SILC and EU-LFS, is lack of inclusion of **hard-to-reach-groups**, such as homeless people, migrants and institutionalised people. In order to assist countries in setting national targets for inclusion of hard-to-reach groups, more and better frameworks for cross-country data collection and usage are needed. Also in this context of hard-to-reach groups, IPOLIS serves a special purpose. In the InGRID-2

project, the IPOLIS database was extended to vulnerable and hard-to-reach groups of special relevance for European social integration, such as Roma and disabled persons, but also migrants and refugees. Several of these vulnerable and hard-to-reach groups often require clearly targeted, coordinated and comprehensive policy responses.

Another problem in European research and policymaking concerns availability of comparative **longitudinal data** on poverty and social exclusion outcomes. Many comparative socioeconomic surveys only include data for relatively short time spans. For example, cross-sectional EU-SILC data is at best (depending on country coverage) only dating back a decade or so (waves are available from 2003 on). Longitudinal data is essential to go beyond descriptions and get closer to causal analysis, and therefore more efforts must target data restrictions for causal analysis. In terms of data integration and harmonisation IPOLIS was expanded with **historical data and now includes more countries than in the first InGRID project**.

Data restrictions also become apparent when the units of analyses are shifted **from nations to regional or local levels**, where sample sizes often are insufficient to yield accurate and reliable estimates. Since the regional, or local, differentiation of poverty and social exclusion is substantial in some European countries, the InGRID-2 project explored new possibilities in statistical analysis to improve analyses of the spatial dimensions of European social integration. Better integration of census data, social survey data and administrative records is central in this process of facilitating policy relevant research on social inclusion outcomes, particularly in small area estimation of social inclusion outcomes.

Analyses of poverty and social inclusion outcomes should not only focus on changes in the lower segments of society. In terms of the political economy of redistribution, also changes in the upper segments of society are important, not the least for cross-class interest organisation and political mobilisation (Korpi & Palme, 1998). In this perspective, greater possibilities for joint analyses of income and wealth distributions would considerably enhance the policy scope of existing RIs. The integration of wealth data from the Household Finance and Consumption Survey (HFCS) into **EUROMOD** in InGRID-2 substantially increased our possibilities to simulate how wealth taxes are distributed and are contributing to social inequality, both directly through redistribution of income and capital, but also indirectly by means of cross-class political coalition making. Information about wealth also taps into poverty analyses by its interaction with many minimum income benefits, where wealth (in various forms) often is considered part of the means-test.

EUROMOD provides excellent opportunities for static microsimulations. However, considering longer-term perspectives in an ageing society**, dynamic microsimulation techniques** also play an important role in InGRID-2. Dynamic microsimulation infrastructures were developed and made accessible to a larger audience than before, through specific training sessions and reinforced exchange between European data and infrastructure developers (including the International Microsimulation Association). The development of microsimulation models is a demanding process, which may induce developers to build on existing tools in order to save time and reduce risks of errors. New possibilities to link static and dynamic microsimulation models by better integration of two important microsimulation platforms (EUROMOD and LIAM2) were explored. LIAM2 is an open source and user-friendly modelling and simulation framework mainly set up for the development of discrete-time dynamic microsimulation models with cross-sectional dynamic ageing (presently used in at least 5 European countries).

3.1.1.2 Improvement of analytical methodologies and tools

The European welfare states are in the process of rapid and profound **transformation**. A tendency in many European countries is that cost efficiency has come to the forefront of policy reform (Nelson, 2011), while social inequalities have been on the rise (OECD, 2008, 2011; Perrons & Plomien, 2010; Nolan *et al.*, 2014). Cash benefit programmes that were built up in the 20th century to protect citizens from economic vagaries in the labour market have increasingly been criticised as

counter-productive and have in several European countries been scaled back (Korpi & Palme, 2003; Montanari *et al.*, 2008; Nelson, 2013). Particular concerns have been raised in relation to labour force participation, and in many countries changes to welfare states have been introduced to create more employment friendly societies. This development coincides with an expansion of public commitment in areas of social investments, such as education, training and care arrangements of various kinds, thus areas other than cash benefits.

The **reorganisation** of the European welfare states is expected to have far-reaching implications for the lives of citizens, and indeed this is the premise on which many reforms were based. However, it is evident that countries have re-organised their welfare states differently, something that has provided new impetus to ongoing discussions about the causes and consequences of welfare states (Esping-Andersen, 2002; Ferrera & Hemerijck 2003; Morel *et al.*, 2011). The balance between income protection and investments to support people to participate fully in society is at the centre of this debate. Although many European countries need adjustment to better support the development of human capital from early childhood and onwards, thereby preventing social exclusion at the individual level and encouraging economic growth at the societal level (Esping-Andersen 1996), it has become increasingly clear that employment growth alone is insufficient for social inclusion. In the immediate decades preceding the 2008 global financial crisis, several European countries have experienced substantial increases in employment, but stagnant or even increased poverty rates (Marx & Nelson, 2012). It is therefore likely that both elaborate forms of income redistribution and employment-friendly reforms are necessary for the continued success of the European welfare states in generating high levels of social equality (Cantillon, 2011; Nelson, 2012).

Due to the nature of welfare state institutional change and the shift in public commitments from income protection to stimulating labour supply and employment, InGRID has devoted particular attention to activities that encourage research on the interplay between cash benefit programmes and public services. This integration of 'cash' and 'care' is often absent in comparative research (Daly & Lewis, 2000). Welfare states intervene in market principles in various ways and shape social inequality through multiple and often complex pathways. To account for these intricate processes and facilitate policy inferences, it is necessary to untangle how the different types of policies interact, rather than conceptualising income protection and social investment strategies as two separate and competing areas. To do so, more and better comparative data on public service provisions, for example, is needed. InGRID-2 therefore defined **new tools and indicators for the comparative analysis of cash benefit schemes**, and developed **new protocols and methodologies for the establishment of more and better data on public services**.

Policy evaluations and recommendations on best practices require high quality micro-level data as well as reliable information on institutional arrangements that are systematically codified into quantitative indicators that are comparable across countries and over time. Typically, institutional data of this kind are not ready made out there. Instead a serious amount of infrastructure research is often needed. The relevant dimensions of social policy programmes need to be defined and protocols for their measurement need to be established, even before any data collection can be initiated. Due to the problems involved in collecting comparative data on the institutional structure of social policies, most comparative welfare state research is still based on less accurate expenditure data (Gilbert, 2009; Kuhner, 2007; Doctrinal *et al.*, 2015). Perhaps the most serious obstacle here is the close association between social spending and needs, and the often poor linkage between expenditures and levels of provision that citizens, in various circumstances, may receive from the welfare state. Comparative research and effective policymaking require more precise and valid indicators on the institutional design of policies in areas that we expect are crucial for social development.

In order to improve methodologies and tools for European research on poverty, living conditions and social policy, a new framework for the **conceptualisation and measurement** of benefit coverage was developed in InGRID-2, including covering the associated topic of non-take up. Whereas conceptualisation and measurement of income replacement in social benefits programmes have been

extensively discussed in the literature (Ferrarini *et al.*, 2014), considerate confusion exists in terms of the inclusiveness of social policies, and how to appropriately define this central dimension of the European welfare states. By identifying pitfalls and possibilities in conceptualisation and measurement of benefit coverage and take-up, possibilities for more effective impact assessment in comparative research are substantially improved.

In order to facilitate further development of RIs that can assist policymakers and researchers to realise the social objectives outlined in the EU 2020 Growth Strategy, hypothetical household tax-benefit analyses and microsimulations were developed in InGRID-2 and extended to new policy areas. Hypothetical household analysis is commonly used when researchers translate social policy legislation into standardised indicators that can be readily used in quantitative research (Korpi, 1989; Bradshaw et al, 1993). Eligibility criteria and entitlement levels are calculated for a set of representative hypothetical households where the underlying assumptions are transparent and easily accessible for reasons of validity. Typical examples include social insurance replacement rates and social assistance adequacy rates. One specific challenge tackled in InGRID concerns the adaption of model family techniques into EUROMOD. The establishment of the Household Hypothetical Tool (HHoT) in EUROMOD is a major innovation developed in InGRID, which bridges research on hypothetical households and microsimulation. Possibilities for comparative analysis are substantially improved by further developments of HHoT in EUROMOD for expert analyses on social policy institutional structures.

The policy analyses offered by EUROMOD itself fruitfully complements hypothetical household-based analyses of social policy. Limitations of EUROMOD are often defined by the underlying structure of EU-SILC, which is used as input data for microsimulations. Hitherto, it has been problematic to simulate contributory benefits in EUROMOD, essentially because EU-SILC lacks necessary information about the employment history of respondents. Although the extension of the HHoT tool in EUROMOD opens up new possibilities to analyse contributory unemployment benefits based on hypothetical household techniques, additional possibilities for research would be created if contributory programmes were to be included in the basic structure of EUROMOD. Parental leave benefits could be used as a pilot case in this regard. Due to the role of parental leave benefits as an effective social investment, for example, in terms of increasing fertility in an aging Europe, stimulating female labour force participation, and reducing poverty, this extension of EUROMOD constitutes a substantial improvement of European microsimulation RIs.

Small area estimation and analyses of the **spatial dimensions** of European social integration are not only improved by better integration and harmonisation of existing RIs. Improvements also require methodological innovations. InGRID has substantially improved possibilities for research on poverty and social inclusion outcomes at regional and local levels. New methods for poverty and inequality estimates in small geographical areas were developed by using **auxiliary information** from LFS rather than censuses and by developing new techniques for small area estimation that are more appropriately designed for policy impact assessment. Considering the latter, **constrained small area estimation** is one fruitful alternative that InGRID has to explore. Furthermore, in poverty analysis, the usual baseline is set at the national level. However, when poverty is disaggregated to smaller geographical units it is often meaningful to define baselines more narrowly, something that brings up issues of regional and local differentiation in cost of living and price levels. Intra-country comparisons of regional and local poverty indicators therefore require some degree of price adjustment. InGRID developed appropriate strategies to estimate sub-national purchasing power parities, which can be used to **adjust regional and local poverty estimates** to geographical differences in price levels and cost of living.

3.1.1.3 Indicator-building and valorisation towards European policy innovation

By extending core RIs with new protocols and tools, new data will be generated in areas that currently are difficult to research, or areas that need to be studied in more depth. In addition, usability of data

will increase, and in this way increase the value of existing RIs. Besides the ambition to integrate state-of-the-art RIs and develop new methodologies, InGRID thus paid considerable attention to improve the **scope and accessibility** of data.

In terms of the scope of existing European data infrastructures in the area of poverty, living conditions and social policy, at least three shortcomings were identified under InGRID-1 that relate to particular challenges that currently are facing the European welfare states. The first challenge concerns increased migration into Europe, particularly of refugees from the Middle East and North Africa. The ways in which the European countries have responded to recent flows of immigrants is unclear, as well as the extent to which immigrants are included in the welfare state of the country of destination. Studies on welfare states and migration are lacking and there are significant gaps in research, particularly in the area of minimum income protection (Sainsbury, 2012). InGRID RIs thus developed new protocols for data collection and develop new indicators on the social rights of immigrants. One important source of input information for this exercise is the country reviews published by the European Social Policy Network (ESPN), which is an important stakeholder of the InGRID RI.

Patterns of family formation and dissolution are changing in Europe, causing new challenges for the European welfare states to reconcile family and work (Lewis & Giullari, 2005). Particularly our understanding of how new patterns of family formation affect social inclusions outcomes relevant to the social targets in the Europe 2020 Growth Strategy needed to be improved, both at the national and regional (local) levels. By combining two well-known data sources, EU-SILC and EU-LFS, **a new indicators system** was created to improve infrastructures for analysing the role of family formation for joblessness, relative income poverty and material deprivation in European countries and its regions. Propositions were made in InGRID-2 on a new set of indicators for analyses on **regional social inclusion outcomes** that are robust to the complex hierarchical structure of survey data, the presence of influential observations in analysing socio-economic data at sub-national level, and if necessary, the temporal distribution of data.

The value of existing RIs for policymakers, researchers and stakeholders ultimately depends on their usability, i.e. that people have access to data and necessarily know how in terms of analysing data. In InGRID-2, we improved European infrastructures in the area of poverty, living conditions and poverty in several ways. Besides networking activities and training activities, the **functionality** and **user friendliness** of tools was improved. Examples are the HHoT tool in EUROMOD, the IPOLIS web interface or the visualisation of the Social Policy Indicators database (SPIN).

3.1.2 Working conditions, vulnerability and job quality

3.1.2.1 Data integration and harmonisation

By establishing more formal collaboration between data providers and infrastructures, InGRID can **harmonise and integrate** different types of relevant research infrastructures, thereby creating new possibilities for European and global research on working conditions, vulnerability and job quality. The main areas which needed to be extended and improved, are the harmonisation and integration of collective bargaining and minimum wage data for 55 countries, present in the ICTWSS database of industrial relations from 1960 onwards, and beyond and of existing European microdata sets with respect to vulnerable groups and their intersections on the labour market. InGRID explored the possibilities to develop an integrated European working conditions microdata series following the IPUMS-International methods (IPUMS harmonises census data from all over the world) and worked on the harmonisation and integration of employee and employer level data from international surveys on work organisation, job quality and performance. By combining **various and novel comparative sources of data** on working conditions, vulnerabilities and job quality in one framework, InGRID-2

made it easier for researchers and policymakers to conduct institutionally informed analyses of the labour market developments, which forms the core of evidence-based policymaking.

In the past years, information about **bargaining coverage and trade union membership** in the ICTWSS database has been collected by national correspondents. This data has been widely used for country-level variables in studies related to industrial relations and working conditions. In addition, also micro level data has become available through data-archives challenging data harmonisation concerning these variables. While the characteristics of minimum wage setting is coded for 55 ICTWSS countries and recently also for Eurofound countries, this information is **only available as text for other countries**. For global comparisons, however, **uncoded text** can be used only if **coded into a data set that is harmonised with existing databases**. In addition, information about the minimum wage levels will facilitate the global analysis of minimum wage setting in relation to real wage levels and working conditions.

The IPOLIS database allows to identify vulnerable groups in the labour market from a poverty perspective through single categories, e.g. women, migrants. The InGRID-2 project **expanded the identification of vulnerable groups** by an inventory of several social categories (such as gender, ethnicity, migrants, religion, nationality, disability and sexuality). The focus was on the identification of **important intersections based on the harmonisation and inventarisation of European microdata sets** with respect to labour market vulnerability in terms of participation rates, working conditions and other relevant labour market outcomes.

The inventory work conducted under InGRID-1 was followed by a **scoping exercise regarding the development of an integrated working conditions microdata series of national working conditions surveys** following the methods of IPUMS-International.

In InGRID-1 existing international surveys covering the topic of work organisations, job quality and performance have been identified. They are either conducted at the employee (e.g. ESS, PIAAC) or at the employer (e.g. ECS, CIS) level. This exercise was **expanded in InGRID-2 by building at intermediate- or meso level data sets based on microdata.** Such data sets are more easily available for researchers and political stakeholders than the microdata sets as there would not be any strong legal constraints on access issues. As these new data sets are designed by the research community, they **are not restricted by administrative categories**, such as the meso data sets produced by the OECD or Eurostat. The underlying hypothesis is that, when smartly defined and analysed, **mesodata sets can bring approximately as much information as micro ones**. Even though the observation unit is not an individual, mesodata sets have to be considered as data sets rather than collections of indicators, and these data sets can be treated as panel data (comparable to the methods used in Hurley *et al.*, 2015).

Non-random samples are increasingly being used in the area of working conditions and job quality research. These new data sources commonly include **web-based surveys, social network data and Big Data**. Yet, with non-probability sampling being relevant for social science research, **new challenges to sampling and inference occur**. Classical statistical inference is done following an inductive logic by linking the (not entirely known) population of interest to observed sample data in probabilistic terms. In non-probability sampling, there are major issues with the assumptions underlying this logic (Steinmetz *et al.* 2014). Examples include the lack of a list frame, unknown selection processes, the potential difficulty of person identification, and the uncertain coverage of the target population, which might result in selectivity. During the implementation of the first InGRID project, there was **no common framework for sampling and inference with regard to non-probability sampling**. Therefore, InGRID-2 investigated **different ways of compensating for the selectivity of the sample** for the purpose of carrying out statistical inference, e.g. selection modelling, propensity score matching, and likelihood based approaches.

3.1.2.2 Improvement of analytical methodologies and tools

To increase the understanding of working conditions, vulnerability and job quality and to create new possibilities for European research in these areas, InGRID-2 has devoted time and resources to the **enhancement of existing methodologies and the development of new analytical tools** in these research domains. This was done by focusing in particular on new methods and approaches to analyse **employers hiring practices and skill transferability**, **newly arising occupations and types of work** and to explore the **usability of administrative data** for comparative research on **job quality** in Europe. As data collection and availability are closely intertwined with the methodological step, the issues of working conditions, vulnerability and job quality were tackled from several different angles and on the basis of a range of data sources.

A first task dealt with **employer practices** and **skill transferability.** Under InGRID-1, a research gap was identified in the area of employers approach towards selection and recruitment, especially in a cross-border setting. These issues are particularly relevant given the EU's efforts to improve labour mobility and in light of the refugee crisis (related to skill transferability issues) and **called for an examination of the tools and data sources available**. A second observation in this regard is that employers, as well as job seekers, are increasingly turning to social media to find a match (Lenaerts *et al.*, 2016). This is reflected by employers' strong presence on these platforms and the fact that a growing number of vacancies is posted online (job boards and social networks). While the **potential of social media as a data source for labour market analysis** is clear, it has been largely overlooked. Yet, work that uses other forms of web-based data has shown that these new types of data carry a huge potential for labour market research. For that reason, the InGRID-2 project had several tasks that were dedicated to this issue. Another dimension on which relatively little research exists is employer practices and skill transferability. Although several efforts have been done at the European level, little is known about the **information that employers use to make hiring decisions** and the degree to which skills are transferable across countries. Moreover, recent international surveys have tried to capture skills from the demand and supply sides of the labour market: for instance, PIAAC (household survey) has been carried out in 24 countries by the OECD in 2012 (OECD, 2013). The Cedefop pilot employer survey on skill needs has been undertaken in 2012 in 9 EU Member States. Also the World Bank has collected household and employer data (Skills towards employment and productivity) in 2012 and 2013 to provide relevant information for understanding skill acquisition and job skills requirements. Due to these various and new types of data sources this calls for a deeper inquiry into the potentials and pitfalls for labour market analyses. InGRID-2 aimed to fill this gap.

A second strand of innovations was devoted to **new occupations and new forms of work**. While the identification of new and emerging occupations is important for policymakers, education institutes and career centres, the traditional identification process is rather cumbersome and based on data sources that typically are not updated regularly. **Web-based data**, which are available in real-time and which can be extracted easily are a solution to these issues. Within InGRID, a new methodology that uses data extracted from online job portals was developed to alleviate these issues, but this methodology needs further testing, fine-tuning and dissemination. This work was done in InGRID-2, with the aim to **provide a tool to track new and emerging occupations**. While some new occupations emerge in traditional sectors, others may arise in **newly developing sectors, such as the sharing economy**. These **new types of labour can bring vulnerabilities** with them (e.g. can one earn a living wage, what are the working conditions). As these new types of work only recently started to emerge and only a few studies exist on this sector, it is unclear what the future will bring. Therefore InGRID-2 had several activities aiming to **shed more light on the sharing economy,** by **investigating the data and statistics available** and by **conducting methodological case studies** with the aim to support and stimulate further research on the topic.

In the past decades several attempts have been made to measure job quality and decent work based on different surveys and data sources. In this regard, the quality of the microdata stemming from different surveys and administrative sources has been largely improved. InGRID-2 explored the avail-

ability of administrative data for a selected number of EU countries for comparative research on quality of work in Europe. Furthermore, collective bargaining is a key instrument in producing social norms on working conditions issues as it enables employers and workers to define the rules governing their relationships. In recent years social partners at the European and national level have accepted that the development of enterprise-level bargaining is, to a certain extent, useful in times of crisis as it allows for new forms of internal flexibility. This is an ongoing process which sometimes described as 'organised decentralisation', and is mirrored in higher-level agreements setting collective bargaining at the enterprise-level. However, depending on the covered topics during negotiations, this decentralisation offers possibilities to derogate from the legal rules or creates incentives to multiply spaces for dialogue on working condition issues. In addition, collective bargaining incentives or obligations are more and more a lever for public policies targeted on working conditions. Taking stock on the knowledge about national industrial systems produced by ILO, Eurofound and OECD, in InGRID-2 new ways of capturing heterogeneity in negotiated responses were developed through the analysis and coding of working conditions clauses in collective agreements.

A final innovation could be to explore the potential of **microsimulations** in the area of working conditions. Microsimulations are an important tool to investigate the impact of policies on the situations of people. Facilitated by EUROMOD, microsimulations are rather successful and more and more used to compare and investigate benefit policies in Europe. In working conditions research this approach is currently almost completely absent at the European level. Therefore, InGRID looked into t**he potential of microsimulation models in the area of working conditions** based on **the job-type** approach of quality of work.

3.1.2.3 Indicator-building and valorisation towards European policy innovation

In policymaking, indicators are important for monitoring the status-quo and for measuring progress towards policy objectives. By extending the core research infrastructures with new protocols and tools, InGRID will develop indicators and valorisation tools for various RIs and thereby create new possibilities for European research on working conditions, vulnerability and job quality. In particular, the development of **labour market indicators for vulnerable groups**, **OSH policy indicators** and **working condition indicators** are central.

When it comes to the creation of **labour market indicators for vulnerable groups** a core challenge is that most of the common social surveys in Europe are limited in sample size per country. This makes the identification of and the comparison between different vulnerable groups across countries difficult. A possibility to overcome this, is to use the European microdata sets (EU-LFS and EU-SILC) providing larger samples sizes per country and allowing to identify in a more reliable way vulnerable groups. In the framework of InGRID **the identification of vulnerable groups and their intersections** (e.g. gender*migrants) was extended in order to **create rankings and indicators** of them for a **range of labour market outcomes** (such as participation and working conditions). As wages are not satisfactory measured in the EU-LFS, additional data sets such as the EU-SILC and the WageIndicator survey ought to be used.

InGRID inventories of comparative policy databases (in preparation) are showing that comparative text bases exist on OSH policies. However, the **development of OSH policy indicators**, although a key EU policy area in the field of working conditions, can be considered a **'black box'**. Existing **international text databases** (ILO country profiles, OSH national systems descriptions by the European Agency for Occupational Safety and Health) could be used **to develop OSH policy indicators** which also can be linked to existing working conditions databases. These efforts were done in the InGRID-2 project.

In InGRID several methods based on **multidimensional indicators** have already been tested and compared to create **measures for Quality of Work and employee vulnerability**. To deepening this perspective and to enhance the understanding of economic circumstances and employer practices associated with job quality and favourable employee outcomes InGRID investigated the possibilities

of a **norm-based indicator on quality of work**, based on existing, predominantly EU-wide surveys, while assessing and benchmarking the available methods to determine a risk indicator for quality of work dimensions. Moreover, to enhance our understanding of economic circumstances and employer practices that are associated with job quality and favourable employee outcomes, and to identify socio-organisational contexts that foster economic performance, creativity and innovation, InGRID proposed **a scoreboard** linking indicators from employer level surveys and indicators from employee or household level surveys that are relevant from a policy perspective.

3.2 Innovation beyond the InGRID-2 project

Similar to the InGRID-1 project, a survey was organised among the InGRID-2 community in order to gather the insights and opinions of experts within research domains related to inclusive growth in Europe (Szekér & Van Gyes, 2016; Szekér *et al.*, 2020). Using this user needs survey as an input, and coupling its results with the insights from the roundtable discussions, the strategic roadmaps as well as of three strategic notes, some thematic priorities as well as challenges and opportunities related to data, methods and policy or indicator-building are laid out.

3.2.1 Thematic priorities

In the InGRID-2 user survey, of which the results are described by Szekér *et al.* (2020), respondents were asked to indicate the key topics in their field of expertise that are expected to gain in relevance for policymaking and research in Europe in the coming years. In the area of poverty, inclusive growth and social policy, the most common themes that were mentioned related to migration and integration, (health) inequalities, ageing, poverty research and measurement (indicators, estimations), importance of closing gaps between old and newer EU Member States (convergence), digital health, distribution of resources, housing, social exclusion and multidimensional poverty. Many of these topics are already in the scope of the InGRID research infrastructure, but some have received somewhat less attention and other topics have not yet been covered at all (e.g. digital health). Regarding working conditions, vulnerability and labour policy, the following topics were put forward often by survey respondents: digitalisation and its impact on work, minimum wages, skill needs and mismatches, lifelong learning, ageing, gender gaps, unemployment, new forms of employment and the social protection of these workers, work-life balance and job quality and working conditions. In this area, too, most topics are already tackled by the InGRID research infrastructure, but there still are some gaps.

After identifying key topics that are likely to become more important in the future, in both research and policy, survey respondents were asked if the currently available data, indicators, methods or tools suffice to analyse these issues. Strikingly, only 36% of the respondents find that this is the case (Szekér *et al.*, 2020) (46% of policymakers vs. 35% of academics). A majority of the respondents, especially academics in EU countries, thus signals that the currently available ways for analysis are not sufficient to keep up with new or upcoming research topics and policy questions. Below, we elaborate on the challenges in terms data, methods and indicators and explore what innovation could be envisaged following the InGRID-2 project.

3.2.2 Data challenges

This section was taken from Esteve (2021).

3.2.2.1 Data availability

Thematic coverage: The primary mission of surveys and official statistics is none other than to collect and disseminate representative data on issues of social concern. These data should also contribute to improving our understanding of social problems and to generating evidence for public

intervention. Social challenges are diverse in nature and vary in importance depending on the political agenda. Some of them are structural in nature (ageing, economic inequality, working conditions, etc.) while others are specific to certain periods (the refugee crisis, COVID-19, etc.). Population ageing is challenging the pillars of the welfare states and affecting its long-term economic sustainability. Job uncertainty is affecting transitions to adulthood among the youth. The EU has enough statistics to monitor the long-term implications of these transformations thanks to their data collection efforts (surveys) harmonised at the EU-Level (e.g. EU-LFS, EU-SILC, SHARE).

However, EU statistics are less adapted to capture short-term impacts that may have also profound implications for our society. The COVID-19 is a clear example of that. The monitoring of COVID-19 development and implications on persons and families would have required constant updates of data on several issues. The ability of EU statistics to provide rapid and flexible data for emerging issues is hampered by the difficulties that European agencies such Eurostat have to make rapid decisions, due to its bureaucratic procedures. The design and implementation of surveys addressing emerging issues often involves a long and difficult negotiation processes that require the presence of an intermediate institutions such national research councils with connections with official statistical offices. The EU programmes on data infrastructures have not promoted the development of new surveys, instead they have focused on improving existing data infrastructures and providing better and more widespread access to users from all over Europe. This could be a key area to further develop within the InGRID research infrastructures, e.g. on how to better capture short-term impacts or to report on issue in 'real time'.

Comparability: The development of European-wide studies requires programmes that promote statistical comparability across data sources from multiple countries. Insufficient comparability affects the process of decision-making across Europe and comparative research. The availability of and access to official statistics varies greatly between countries. Prior to Eurostat's work on the harmonisation of surveys, national versions of many standard surveys existed in different countries, as in the case of the Labour Force Survey (LFS). Before the expansion of the European Union, harmonised LFS questionnaires were already available for the different countries. This helped Eurostat to harmonise the LFS data on the basis of standard questionnaires. However, there are still differences between the contents of Eurostat LFS microdata and the contents available directly from countries. Comparability has not been fully achieved in one of the most important surveys that exist in Europe to monitor trends in the labour force market and the economically active population. Countries are faced with the 9 demands of Eurostat and the historical consistency of their statistical series. Historical samples of LFS have been left behind and are not all part of the Eurostat's collection. Within Eurostat, EU-SILC data is perhaps the most comparable source across European countries. Despite that, comparability issues also arise due to different modes of collection and sample design.

Out of Eurostat, the availability of statistical data varies greatly from country to country. Despite many countries having similar statistical products, coordination within Europe is less frequent and, therefore, disparities grow. Some countries are almost entirely dependent on registers. Other countries have to schedule specific surveys to obtain information on topics of societal interest. If having a harmonised statistical area is a priority for Europe, Eurostat should exercise a stronger leadership in its construction. The current approach based on the preparation of guidelines is not enough. National statistical offices do not have the economic and human resources to develop new products as they feel the pressure to satisfy the internal demand. European institutions are not funding data collection efforts for long periods of time. National statistical offices should be responsible for it, but they do not feel the need to satisfy demands at the European level while receiving lots of requests from the national authorities. All these institutional problems are real barriers to achieving full comparability across data sources in Europe both in terms of coverage, collection methods and contents.

Sub-population coverage: An inclusive society implies having information on each and every one of the social groups that make it up, especially those that are statistically invisible. The lack of visibility of hard-to reach groups is mainly due to their small numbers but also because they are not present in population registers. Because of their small numbers, such groups are hardly represented in general representative surveys, and often they are not even present in the population frame (e.g. homeless people). Collecting data from such groups requires alternative strategies. Efforts are being made to generate representative samples of hard-to-reach populations using techniques like snow-balling sample designs. International migrants also belong to this category due to their high mobility patterns. Their official place of residence does not always correspond to their actual place of residence. Those international migrants who are included are not necessarily representative of their community. During the field work, survey and census takers gather a lot of information regarding which populations are effectively targeted and which are not. This information is rarely shared with the research community despite its value. Oversampling is a strategy to overcome the lack of visibility of some groups, but it is not an option for the hard-to-reach population that do not even have a known probability of being selected for sampling.

In any case, European statistics today are far from having an optimal solution for collecting information on sub-populations. Again, the answer depends on the possibilities of each country. One example is the refugee crisis resulting from the conflict in Syria and the Middle East. A number of studies have been carried out about these populations, most of them qualitative, with the exception of those carried out in Western and Nordic countries that have good records. Refugee studies at the European level would be limited by the availability of data. The low and insufficient coverage of some population groups ranked number two among the limitations expressed regarding data availability in Europe (in the InGRID-2 user survey, see Szekér *et al.*, 2020).

Country coverage: The lack of universal statistical coverage covering all countries is another obstacle to be faced by comparative research at European level. With the exception of the sources coordinated by Eurostat, the availability of statistics and the practices of access to information are very heterogeneous. This means that academic development or statistical production on some topics is systematically concentrated in a few countries, usually the most advanced and rich and therefore the most homogeneous in terms of standard of living and population characteristics. For instance, the statistical coverage of 10 Eastern and Southern European countries is lower than that of Scandinavian or Western countries. Moreover, the available statistics are neither comparable, nor are they based on similar questionnaires, nor they have the same frequency of publication or comparable sample sizes. Access to information is really diverse, not only in terms of legal limitations on access, but also in terms of the very existence of samples for scientific use. Researchers therefore complain that it is difficult to carry out European wide or comparative studies beyond the possibilities offered by Eurostat sources. For example, individual follow-up panel data are available for a limited number of countries (e.g. the United Kingdom, the Netherlands, Germany and the Scandinavian countries). The number of comparative studies focusing on these three countries is impressive in some fields of sociology and demography.

Lack of sample: The increasing complexity of research questions requires the design of questionnaires with great conceptual detail to analyse the deep causes and consequence of social change. More conceptual details usually mean more questions, thus lengthy questionnaires, that require in person interviews or a close follow up of the respondent. This usually has a negative impact of the sample size because budgets are limited, and such surveys are more expensive. So, one common problem of these sophisticated surveys is statistical significance. Even in surveys with large sample sizes (e.g. over 10K cases), the analytical power of the sample clashes with the number of cases. This problem compromises the likelihood of drawing valid conclusions about countries and of making statistically significant relationships. The surveys that are best equipped in terms of questions have the greatest

constraints in terms of sample size. If data are available for more countries (such as in the case of EU-SILC or SHARE), a common strategy to gain statistical significance is to combine data from multiple rounds and countries. Lack of sample size also affects the visibility of minority groups in these type of surveys. Increasing sample size is, however, costly. If surveys are longitudinal, attrition becomes a real challenge as following individuals over time is more and more challenging.

Linking records: The possibility of linking records from different sources on the same individual is perceived as an opportunity to increase the research value of existing data. This brings advantages, such as avoiding the design of lengthy questionnaires, minimises the respondents' burden and build sophisticated data sets that allow for more complex analytical frameworks. However, legal and technical impediments stand in the way. Legislation on personal data protection and statistical confidentiality are major obstacles. The conditions of access to the register data are defined at the country level. In general, only persons residing or affiliated to national research institutes and universities can attend. However, an increasing number of statistical offices are using record linkages to generate census-like statistics and other type of products. Door-to-door traditional population censuses are being replaced by new products consisting in the harmonisation and combination of data across data sources. For many countries, this is a real challenge because person ids are not consistent across all data sources, but efforts are being made across European countries to overcome these limitations. The future of census-like statistics will very much depend on the success in combining data from multiple sources. This could open the doors to more sophisticated statistical products for researchers.

As a hybrid formula and potentially feasible in some countries, there is the possibility of having individual registers that can eventually be linked to ad hoc surveys. However, this option, present in countries such as the United Kingdom or the Netherlands, requires an excellent coordination between research councils and statistical institutes, which is not always the case in all countries. From a technical point of view, the integration of registers is seriously compromised by the effective possibilities of linking registers, as not all countries have common and sufficiently implemented individual indicators in the registers to facilitate the link. Fragmentation of registers and forms of capture are obstacles to data harmonisation. Efforts to reconcile identifiers come up against the limitations of the statistical agencies themselves: the inability to impose these changes on the different administrations in charge of collecting the information, the decentralisation of registers, the existence of different data cultures. Eurostat could contribute and provide guidance to the countries willing to move towards a stronger combination of register data.

3.2.2.2 Access to data

Difficulties in accessing microdata: The existence of appropriate data for research does not necessarily mean that they can actually be used for this purpose. Beyond the problems of coverage and representativeness, access to information is another important limitation. The conditions of access are determined by different factors: Firstly, by the obligation to preserve the privacy of individuals. Secondly, by the reluctance of statistical agencies to make strategic information public. And, thirdly, by the limited capacity that the offices in charge of preparing the data have to respond to all requests.

Individual privacy is a right to be protected and the dissemination of individual microdata might put this right at risk. The balance between privacy and scientific use of information is complicated and allows for multiple solutions, which makes it even more difficult to reach broad consensus, as evidenced by the responses to the InGRID-2 survey. In the INGRID-2 survey, the ranking of priorities varies according to the research interests of the participants. Thus, some prefer to achieve greater geographic detail than conceptual detail and vice versa. Finding the perfect balance for all requests it is a challenge. Ideally, requests should be customised to the needs of researchers, but this will increase the workload of data providers, usually facing shortages of personnel for these particular matters. Having a European Research Data Council could certainly contribute to unify access conditions

across European countries. The role of such an institution would be different than Eurostat. As the latter is primarily concerned with collecting and harmonising data for Europe to inform policy makers. Social scientists are not the first priority for Eurostat.

Beyond data privacy concerns, there are additional limitations to the dissemination of data when these data might reveal some patterns that are to be protected in order to avoid discrimination or stigmatisation of certain social groups. For instance, educational authorities might be reluctant to disseminate information about school performance of certain schools to avoid stigmatisation. Social scientists might feel the same and decide not to show some particular results.

Attempts at the European level to harmonise the conditions of access to information through the creation of a single entry portal have always come up against the reluctance of countries to unify criteria and formats for accessibility to their micro and metadata. Reconciling all documentation is a key challenge, as harmonisation affects not only documentation but conditions of access and dissemination. In addition, harmonisation is complicated because of the heterogeneous and uneven implementation of national data archives in the EU countries. Some countries do not have national data archives and statistical officers are in direct contact with social scientists.

Costs of access to information: The costs of accessing microdata, in terms of financial resources and time, are another important constraint to consider. The conditions of accessibility to microdata vary from country to country and they also vary depending whether you are a national researcher or not. In some cases, access to third-country data is conditioned to the existence of a formal collaboration with a local researcher. There are statistical agencies that require a formal application to obtain the data while others have their data sets available on the internet for download. The economic costs of accessing the data or requesting specific data vary widely. Understanding access conditions might be complicated. It is not always obvious which data are available and which data can be requested. Previous knowledge about the system of the collaboration with national researchers is a must. This situation generates all sort of inequalities regarding data accesses. Not all scientists have the same connections and institutional links to get access to data. Some are more international than others and have more resources from their institutions to cover the costs of data access. For this particular reason, projects like InGRID-2 are important at the EU level because contribute to universalise and democratise access to data infrastructures across European countries. Any follow-up project should bear this in mind.

Eurostat has set its own conditions regarding access to data, which are obviously the same for all members of the European Union. For instance, the requirements for the consultation of EU-SILC and EU-LFS are subject to the preparation of a research project, the presentation of a working team and the approval by all countries involved of the application. The procedure is lengthy and not flexible enough. On the other hand, once access has been authorised, the data provided is often more than what the researcher needs. There are more effective platforms and data extraction systems in the market from which Eurostat will largely benefit (e.g. IPUMS).

Levels of access: The level of access to information refers to the conceptual and geographical detail in which the information is presented, the characteristics of sample and the periodicity of the data. National statistical agencies maintain different protocols regarding access to data, which varies depending on the type of data to be disseminated and the work involved in the preparation of the data sets. Some agencies have developed scientific use microdata files that can be used for research purposes. Aggregated data is often publicly available (without restrictions) but their value for research is comparatively lower than for microdata. Within Eurostat, there are different levels of access. Because European countries have different approaches to data confidentiality, Eurostat has adopted the most restrictive approach. The harmonisation of policies regarding data access would promote more comparative work at the European level.

Confidentiality and privacy limitations regarding access: The disseminations of information collected at the individual level is many times constrained by privacy and confidentiality limitations to protect the identity of individuals. These limitations have significant implications regarding data access. A standard procedure to protect people's identity in social surveys and administrative data is to anonymise the personal records and suppress information that could potentially lead to the identification of the respondent. When this happens, the statistical and substantive relevance of a particular data set might decrease. As an alternative, countries provide access to full counts and complete data set through secure data access sites and under sever measures of control. Within the field of statistics, and other related disciplines, there is permanent research on which methods are more efficient in keeping the relevance of the data while protecting individual privacy at the highest standards. Limiting the amount of conceptual and geographic detail available in the data is a standard response to privacy concerns (e.g. suppressing geographic detail). However, permutation and record swapping are other alternatives, including the construction of synthetic data sets.

In the years to come, the increasing pressure for confidentiality, which both affects and justifies the conditions of access to and recodification of information, may further limit the availability of information. Such trend runs against the idea of fostering the use of administrative data for scientific use purposes. Many of these data are presented at the aggregate level, not allowing researchers to access individual microdata, which, of course, reduces the analytical value of the data.

The metadata: Metadata comes last in the ranking of limitations, but it is nevertheless a key aspect for the harmonisation and dissemination of data at European level. Metadata accompany the publication of data by providing information on the characteristics and contents of the samples and information on the data collection process. This is key information for researchers, as it not only facilitates the manipulation and processing of the data, but also provides insight and knowledge about the potential biases of the data. Providing details on the sample structure, the allocation of weights, the collection of information or the imputation methods are essential to analyse and interpret results correctly. However, metadata standards across European countries and surveys are heterogeneous and this conditions the harmonisation of metadata at the European level. The e-portal developed under the InGRID-2 project presents an important step forward in this regard. The e-portal will be maintained for five years after the InGRID-2 project end, and could be further improved upon and expanded in a future project.

3.2.3 Methodological challenges

This section was taken from Articus et al. (2021).

Data availability often determines the scope and speed of methodological advances and evolution of new problem- or case-oriented approaches. Within and beyond the research infrastructure InGRID, two main areas can be seen where innovations in statistical methodologies have taken place: On the one hand, statistical information and indicator values are increasingly needed for local geographical areas; often these are administrative units referred to as NUTS 3 or LAU. This has generated much research on Small Area Estimation. On the other hand, due to many reasons including reduction of response burden, new and alternative data sources are being considered in data collection and survey processes. This includes the simultaneous use of very different data sources and combinations of classical survey data, registers, administrative data and new data sources.

What are the lessons learned from the InGRID research infrastructure? Various methodological advances have taken place in the areas of poverty mapping and working conditions. The integration and cooperation between applied researchers and statistical methodologists have shown to be an important asset to combine theory with empirical foundations and furnishes developing the use of modern data sources with advanced statistical methods to provide a sound basis for policy support.

However, the next decade has to face important changes within Europe and between European regions. The measurement of regional disparities via statistical information and indicators strongly urges building European infrastructures on data, data provision, and appropriate (survey) statistical methodologies. Important developments have to take place in the combined use of data sources. This certainly necessitates the development of adequate statistical methods, often referred to as multi-source estimation. Special attention has to be laid on the granularity of the data and the information needed. Additionally, and in combination with Big Data, also algorithmic advances for numerical solutions of new statistical methodologies have to be developed to allow for statistical inference on large and combined data sources. Finally, as a very important but challenging asset, European-wide data platforms need to be further supported. This will be an essential feature for open and reproducible research and evidence-based policy. The synthetic but realistic data platform, as a basis for European microsimulations, could be a first important step to achieve these goals.

3.2.3.1 Geographical precision

Because most societal phenomena under study show clear regional patterns, there is a strong interest in small scale statistical results at low levels of geography. While the relevance of regional statistics seems indisputable, their provision raises two interconnected questions: The first is the question of how regional units for statistical analysis should be defined. Secondly, a fundamental question of all disaggregated analysis is whether results of adequate precision can be obtained based on the available data. We will address both of these questions below.

Regarding the definition of regional units, there are the standard regional entities for statistical purposes in European Official Statistics, i.e. NUTS 1 to 3 and LAU. They establish a harmonised system of regional entities, generally (i.e. when possible) based on administrative subdivisions, there with largely facilitating the provision of comparable results in the European Union in a standardised system. The lowest level is Local Administrative Units (LAU), i.e. low-level administrative entities such as municipalities, communes, or parishes. These frequently vary strongly in size, both with respect to area and population.

Besides this clearly relevant definition of regional units, there is a growing interest in more flexible spatial disaggregations to complement it. The rising availability of geo-coded data allows to also exploit the potential of grid-level analyses or of defining entities specifically targeted for the analysis at hand. These might or might not be oriented on administrative boundaries. As such, very low-scale divisions specifically targeted at the phenomenon under study are possible. This flexibility, however, comes at a price of limits in comparability and availability of covariates in model-based studies.

Addressing the second question, it is obvious that the strong disaggregation of the available sample data in the context of a regional analysis routinely results in small sample sizes for most or even all of the regions under study. Consequently, traditional direct estimates lack accuracy. Small Area Estimation techniques, specifically designed to obtain reliable estimates in this case, are a solution. This has been one of the key methodological research fields in InGRID.

3.2.3.2 New data sources

Due to the need to obtain more granularity in data for policy and empirical research, alternative data sources are being considered for use in statistical processes. Here we consider three main alternative data sources and discuss their opportunities as well as their limitations.

Administrative (Register) Data: Statistical agencies around the world are directing resources into advancing the use of administrative data in official statistics and national statistical systems. Administrative data are defined as secondary data sources since they are produced by other agencies as a result of an event or transaction relating to administrative procedures of organisations, public administrations, and government agencies. Nevertheless, they have the potential to become important data sources for the production of official statistics by significantly reducing the cost and burden of

response and improving the efficiency of statistical systems for government, businesses, and citizens. There are many examples of successful applications of incorporating administrative data in statistical systems, particularly in the area of business statistics.

Transforming administrative data into statistical data for official statistics systems is not without costs. Often, administrative data need to be adapted and processed to make it suitable for statistical systems. Some examples are transforming population definitions to meet statistical populations, transforming entities (households, businesses) into statistical units, and transforming administrative variables into statistical variables. Zhang (2012) was one of the first to conceptualise a Total Administrative Data Error Framework. Errors are defined according to representation (objects) and measurement (variables). Errors for objects (frame errors, selection errors, and missing redundancy) are classified broadly as coverage errors. Coverage errors occur if either the accessible data include objects that are not in the target data, or if we are unable to access data that we would ideally want, but are not available in the accessible data. Statistical agencies are becoming more involved with processes to improve the quality of administrative data and efforts are made to carry out statistical data editing procedures to also satisfy the quality assurance framework of the European Statistical systems: relevance, accuracy and reliability, timeliness, coherence, and accessibility.

Whilst individual administrative data is generally not available to researchers, one can use aggregated statistics as auxiliary information for obtaining more detailed granularity in statistical estimates through model-based Small Area Estimation techniques, one of the key research themes within InGRID. In addition, statistical agencies may release administrative data that have been linked to survey data for research purposes through the usual protocols of releasing confidential data, i.e. through data archives and safe data enclaves.

Big Data: In the American Association for Public Opinion Research (AAPOR) report on Big Data (Japec, *et al.*, 2015), sources of Big Data are defined as follows: Social media data; Personal data (e.g. data from tracking devices); Sensor data; Transactional data; Administrative data. The first four sources of data are organic and follow the general principle of Big Data: large volumes of data at high velocity and in varying formats. In the context of researching social phenomena, we can usefully classify Big Data into two classes: whether the data records are identifiable, that they can be associated with a single physical unit in space or time, or not. It is this classification which informs how we can use Big Data for research purposes and what can be achieved.

If the elements in a dataset can be meaningfully associated with a unit at a given place and time, such as an individual, institution, product, or geographical location, then Big Data can be made fit for the purpose of statistical inference. Examples of this include satellite imaging for agricultural surveys and censuses, product bar-codes from stores to collect data for constructing price indices, and traffic loop counters for counting the number of vehicles crossing a specific intersection. Big Data may not cover the target population exactly or there may be selective missing patterns in the data that cannot be treated as random, which complicates the statistical modelling and its interpretation. In addition, measurement errors need to be considered when combining sources of data or when the data available may only be proxies for the data needed. Although these problems are very challenging, researchers have developed techniques that are designed to compensate for measurement errors in their statistical modelling. For example, one way of dealing with measurement errors is to ensure that there are high-quality, random samples available with overlapping variables. Such random samples can help to compensate for poor coverage, for example, through capture-recapture techniques, or for measurement errors and selectivity when data values are missing.

Other sources of Big Data, such as Twitter feeds, other forms of social media, and google searches require new forms of analytics as well as visualisation. This in itself is an important area of research and requires new skills and algorithms. However, if the data cannot be made identifiable at some level then it is of limited use for statistical inference.

Some examples of using Big Data in official statistics is reported in Daas, *et al.* (2015) where they used identifiable Big Data of traffic loop detection records to measure traffic intensity at known intersections. Traffic loops can count the number of vehicles per minute that pass at a specific location as well as measure the speed and length of the vehicles. When analysing the data, it was clear that it suffered from selective missing data problems due to some computers failing to submit data. In the second example, they used non-identifiable Twitter messages according to pre-specified 'buzz' words and showed a correlation between Twitter sentiment and the official Dutch Consumer Confidence Survey. This application, however, is not replicable research. Other examples of the use of Big Data in official statistics are: (1) The Australian Bureau of Statistics is using satellite data consisting of crop areas at specific time points for their agricultural official statistics production (Tam & Clarke, 2015). (2) In Italy, researchers are investigating the use of Big Data as covariates for Small Area Estimation models (Marchetti, *et al.*, 2015). All these examples indicate that when Big Data can be made identifiable, there are potentially large benefits to incorporating the data into statistical systems.

Non-probability samples: Non-probability samples, such as quota sampling, snowballing, and convenience sampling, have typically been designed to survey hard-to-capture populations in sociological research. However, given increasing nonresponse rates in our random surveys and the popularity of conducting web surveys due to their low costs, this type of sampling has become more prevalent for more mainstream data collections.

At first sight, web surveys seem to be an excellent replacement for traditional offline data collection methods. Indeed, using the Internet for data collections are common amongst both academic and commercial researchers nowadays. It is also worth mentioning that this strategy is particularly appealing amid the COVID pandemic when conditions for face-to-face surveys are and will likely continue to be restricted in many parts of the world. However, a line of inquiry that deserves scholarly attention is the reliability of Internet-based data as it is not grounded in probabilistic methods of sampling. Opponents to the use of web survey data have long been questioning its quality, especially when the aim is to analyse and provide statistical inference to the general population.

One of the key issues associated with online data collection is the phenomenon of self-selection. Self-selection occurs when data collectors are not in control of the selection process and it is largely or completely left to individual targets to select themselves for the survey. It is determined by factors such as computer literacy, Internet penetration, and interest to participate, which are rarely evenly distributed in the population. The underlying logic of self-selection is thus inherently different from the probability sampling paradigm, which follows random sample selection. Consequently, while probability sampling enables statistical analyses to produce accurate and unbiased estimates of the general population, self-selection can lead to biased estimates, and therefore wrong conclusions are drawn from the collected data. Another challenge for Internet-based data is participation in online surveys. When data is collected online, individuals without access to the Internet or the link to the survey platform are naturally excluded. Since specifying the size of the population being able to take part in any type of web survey is almost impossible, it is extremely hard to capture response rate or structure of the sample relative to a total sample that is aware of the survey. Finally, self-administered surveys are shown to be more likely than interviews to suffer problems associated with satisficing. Respondents tend to provide answers requiring less effort (e.g. rounding up responses and providing inaccurate numbers) when fatigued by the survey. Unfortunately, notwithstanding their advantages and popularity, most web surveys are not immune to the shortcomings mentioned above. It is thus important for researchers to address these shortcomings and develop innovative approaches to improve analyses using data from web surveys.

The problem with nonprobability samples is that it is not possible to generalise to a population, calculate estimates and confidence intervals, or carry out statistical inference. There are generally two approaches to compensate for selection bias in nonprobability samples: A model-based approach and a quasi-randomisation approach that integrates the nonprobability sample with a probability reference

sample. Quasi-randomisation approaches are common and include two main approaches: Sample matching and post-hoc adjustments through propensity score matching. Both require the use of a probability reference sample. In sample matching, a nonprobability sample is drawn with similar characteristics to a target probability-based sample. Units are matched in the nonprobability sample to the probability reference sample based on a set of variables that explain both participation and the outcome variable so that the covariates are balanced. Then, inference is carried out using the survey weights of the probability reference sample. Propensity scores matching stacks the nonprobability sample and the probability reference sample and estimates the probability of participation in the nonprobability sample using, for example, a logistic regression model where explanatory variables in the model explain both participation and the outcome variable. The predicted probabilities are then used in the calculation of initial participation weights for the nonprobability sample, for example through propensity score stratification or through the inverse probability. This step is then followed by post-stratification in which auxiliary population information is used to adjust and benchmark the initial participation weights. The benchmarking is often implemented using a raking adjustment to known population aggregates. This step further reduces the impact of selection bias and other potential coverage errors in the nonprobability sample.

One of the most successful and prominent examples of a nonprobability web survey program is the WageIndicator (WI) data. It was initiated in The Netherlands in 2001 by Pauline Osse and Kea Tijdens as a platform for employees and employers looking for information about income and is supported by InGRID. As of 2020, the WI organisation is operating in over 80 countries worldwide. The WI web survey is designed by experienced researchers and supported by world leading academic partners such as University of Amsterdam and Harvard Law School. According to Wageindicator.org, the WI web survey identifies the labour force as its target population. The respondents of these multilingual web-surveys are volunteers recruited through the national WI websites and a wide range of websites of WI partners. In Huang and Shlomo (2021), an application is shown using the 2016 nonprobability WI dataset of the Netherlands to measure the gender pay gap.

In summary, whilst there is still a need for random probability samples to adjust for statistical biases in new types of data sources, combining these data sources with new forms of data allows for a richer environment of available data with more granularity and geographical detail for policy and empirical research purposes. Clearly, there is a major challenge to adjust and analyse multiple data sources. Another particular challenge is the need to develop optimal record linkage techniques for combining multiple sources of data. This is, however, offset by the requirement to respect the confidentiality of data subjects by reducing the probability of disclosure of sensitive information on individuals or institutions. This involves a balance between reducing disclosure risk by 'degrading' data in various ways, whilst retaining, or being able to recover, sufficient information so that the data are still suitable for linkage and efficient statistical analysis. It is clear that the future of social research is evolving with the emergence of different forms of data, both organic and collected using structured methods, and the need to incorporate and combine multiple sources of data. Software and algorithms are being developed and the involvement of statisticians in these is essential to ensure that the data that become available retain their integrity and thus usability for statistical analysis and inference whilst preserving confidentiality.

3.2.3.3 Challenges within Europe

When taking a European perspective on important societal topics, the comparability of data products from different member states comes into focus. Ensuring full comparability, however, is demanding: It requires an alignment of the scope of a statistic in terms of covered variables, of concepts, definitions, and methods, and ideally an alignment of the data gathering process. Further, the spatial resolution of results should be the same in all countries under study. In many cases, there is a conflict of comparability with national path-dependencies and legal, regulatory, or cultural peculiarities.

An example from the experiences obtained in the InGRID-2 project are the tremendous difficulties encountered in the attempt of creating a harmonised data set on labour and work conditions. Based on a comprehensive overview of related data products in 17 member states (Szekér & Van Gyes, 2015), Desiere and Lenaerts (2020) have investigated the possibilities to integrate and harmonise these national surveys in order to create a common dataset for European analyses. They conclude that this seems not feasible in the short and medium term because the aim and scope of the considered national surveys differ strongly. This also results in different definitions of the target populations. Further, the conditions of data access vary strongly. For this specific field, they conclude that comparability of data is very limited, harmonisation is difficult, and the use of European data products is a more rewarding strategy. They do, however, point out that national data products are frequently based on considerably larger sample sizes, allowing for a stronger disaggregation with regard to spatial units or e.g. demographic groups.

Bertarelli *et al.* (2020) come to a similar conclusion: In an attempt to employ Small Area Estimation (SAE) to obtain small-scale results for the At-risk-of-poverty results in a European border region, they conclude that the availability of comparable covariates at the target resolution level for all countries is a major obstacle to this model-based approach. They, therefore, test the usability of remote sensing data products as covariates in model-based SAE.

From this experience we can take three complementary strategies or perspectives to further build up comparable data within Europe in the future: A first strategy is the further strengthening of European surveys, particularly with respect to the realised sample sizes, in order to broaden their potential for regional analysis. Secondly, the preconditions for a harmonisation of available national data products could further be strengthened. Finally, alternative data sources, which frequently are generic 'cross-border' data, could fill gaps and further widen the scope of possible analyses.

3.2.3.4 Modern methods for microsimulations

The aim of the research infrastructure InGRID was to elaborate European-wide empirical research focusing on poverty and working conditions. As part of the research, microsimulation methods were considered. Microsimulation methods date back to Orcutt (1957) who pointed out that heterogeneity and complexity of individuals and their decisions can hardly be observed or modelled using aggregate data. Indeed, Harding *et al.* (2010) define Microsimulation is a technique used to model complex real life events by simulating the actions of and/or impact of policy change on the individual units (micro units) that make up the system where the events occur which perfectly meets the aims of InGRID-2 research.

Within Europe as well as within InGRID, the well-known EUROMOD microsimulation model is applied (see https://euromod-web.jrc.ec.europa.eu/). EUROMOD can be classified as a static microsimulation, where scenarios can be considered by directly changing policies such as tax systems (Sutherland & Figari, 2013). In general, EUROMOD microsimulations are then based on the European Union Statistics on Income and Living Conditions (EU-SILC). The demands for European-wide microsimulations, of course, often exceed capabilities of the EU-SILC data, by contents, variables, or geographical scaling.

In addition to static microsimulations, dynamic models offer great opportunities for research in the social and economic sciences as they allow for a longitudinal evolution of the population. Dynamic models offer further possibilities for the analysis of changing populations (Burgard *et al.* 2020a). A comprehensive overview of different types and methodological differentiations of dynamic microsimulations is provided by Li and O'Donoghue (2013).

According to Li and O'Donoghue (2013), dynamic microsimulation methods cover two tasks. Of course, those outlined above and policy-oriented aims of answering scenarios or what-if-questions to provide a sound basis for evidence-based policies. However, they also highlight a second task as microsimulation is a tool to generate synthetic micro-unit-based data. Indeed, the strength of the

latter has to be pointed out as an important data and methods strategy aiming to provide the necessary basis for a European-wide data and simulation lab to foster evidence-based policy research.

What are the methodological challenges within this data strategy? Before answering this question, one shall put together possible different aims within this data strategy. European policies are less and less based on the country-level and more focus is on area-level. Survey data often do not provide accurate insights from NUTS 3 and downwards. Additionally, in the era of Big Data, an increasing demand for geo-coded data is observed. And finally, there is a strong demand for comparative studies across countries in Europe. The earlier research infrastructure Data Without Boundaries (see http://www.dwbproject.org/) was aiming to provide European-wide microdata access. Country-specific disclosure limitations, however, did you allow to provide the necessary basis for cross-border access to microdata. An alternative strategy is to build synthetic but realistic geo-coded data. The German MikroSim research unit of the German Research Foundation (see Münnich *et al.*, 2021) is currently researching statistical methodologies to provide the basis for building a German data lab. The focus is creating a synthetic but very realistic German universe based on the Census 2011 results with the entire population using Open Street Map to provide fully geo-coded data.

To come back to the above question, the construction of these kinds of data sets makes use of many different data sources, mainly survey data, but also other sources including administrative data as far as they are available. The first step is to construct a cross-sectional data set with many variables. In order to provide also a dynamic microsimulation environment, population forecasts and other developments have to be considered. The variety of methods under consideration cover all about survey statistics, and, hence, also the statistical methodologies presented within the InGRID research infrastructure. Special emphasis, of course, has to be laid on cross-sectional and longitudinal predictive modelling methods, imputation methods, and calibration while keeping observed variations and heterogeneities of the data. In this context, one has to point out that simply applying predictive modelling has to consider inferential properties of the original data. This often can be achieved by modern machine learning algorithms. Future research certainly has to focus on the impact of various predictive modelling strategies on the construction of datasets. However, measuring synthetic but realistic data still needs further research in terms of possible limitations using these datasets, also considering regional disparities on a micro level (Burgard *et al.*, 2021). First results within the MikroSim research, however, show that policy scenarios provide very interesting and useful insights (Dräger *et al.*, 2021).

Nevertheless, two aspects still have to be mentioned in developing such a database. Firstly, information on the lowest level, i.e. the geo-coded household information, is beyond statistics. Sophisticated mathematical optimisation strategies have to be further developed to allocate households to living space appropriately while respecting very different margins within hierarchies of statistical information. In a multi-stage procedure, all persons could first be located in census grid cells (1000 x 1000 and 100 x 100) and then within the grids of real addresses (Burgard *et al.*, 2019, Burgard *et al.*, 2020b). Additionally, before providing these data as open data, new disclosure control methods have to be developed and tested to prove that the identification of real units is not possible. First attempts in this context are provided in Ahmed *et al.* (2021).

In any case, constructing a European-wide synthetic but realistic microsimulation data set, based on street maps with geo-coded data provides a rich and important source for open and reproducible research and enables the European research community to improve evidence-based policy as well as (statistical) methodological research

3.2.4 Challenges related to indicators and policymaking

This section was taken from Gabos et al. (2021).

Bridging the gap between policymaking and research is a major challenge for policy-oriented research and evidence-based policymaking. When asked in the InGRID-2 user survey the about priorities and highly important issues, academics and policymakers seemed to be in general agreement regarding the top five challenges: both groups identify the 'limited cooperation between researchers and policymakers' as the main challenge, followed by 'the lack awareness of available research outcomes and their usability among policymakers' and 'the mismatch between policy and research cycles and thematic priorities'. Also 'communication issues and lack of mutual understanding' and 'limited cooperation among researchers' were frequently mentioned as challenges by both groups. Under the general agreements though, some nuances in weights given to the various issues appeared. Policymakers more than academics indicated that issues such as 'lack of trust and openness', 'lack of resources to conduct state-of-the art policy-oriented research' and 'challenging political climate' are important challenges concerning this issue.

When respondents were asked to explain how InGRID-2 can best help to bridge the gap between data and policy, it was highlighted that InGRID-2 could increase the interaction between both parties through events such as round tables, seminars, workshops, and conferences. Also, it was emphasised that InGRID-2 could improve communication by 'advertising' research results by more easily accessible ways of communication like short policy-oriented articles, podcasts, presentations, graphs, newsletters, visualisations, and infographics, in addition to long and complex research reports and peer-reviewed journal articles. It was also emphasised that universities, governments and other financing bodies can contribute to bridging the gap by putting more value to non-academic output.

Finally, it was mentioned that joint trainings and more frequent networks between policymakers and researchers could help in better proliferating between academia and policy.

It is among the very aims of InGRID to serve the social sciences community that wants to make an evidence-based contribution to a European policy strategy of inclusive growth. In addition, it is also the aim of the project to create sustainable infrastructures that can support European policymaking with innovative tools to access relevant analyses and data produced by the European research community, living and working in very diverse corners of Europe. In what follows, we present two cases for this: the examples of the Integrated Poverty and Living Conditions Indicator System (IPOLIS) on the one hand and the Institutional Characteristics of Trade Unions, Wage Setting, State Intervention and Social Pacts (ICTWSS) database on the other. Both of these are intended to build bridges between academia and policy both on a European level and on national levels.

3.2.4.1 The Integrated Poverty and Living Conditions Indicator System (IPOLIS)

The process of building up IPOLIS and its main characteristics

IPOLIS was the core outcome of a work package on innovative tools and protocols for poverty and living conditions research of the first InGRID project. Within the InGRID-2 project, the extension and further development of IPOLIS under InGRID-2 was carried out within the frame of WP8 'Data harmonisation and integration regarding poverty and living conditions'. IPOLIS fits within the frame defined by the overall InGRID project objectives according to several respects.

- IPOLIS is related to all three focus areas of the InGRID-2 project: (1) it relies on and integrates harmonised data; (2) it links data, specifically quality of life outcomes with policies and (3) promotes indicator development.
- Material living conditions in general and poverty and social exclusion in particular (also as defined by the Europe 2020 strategy target), stays at the core of the integrated indicator system.
- IPOLIS is conceived to be an innovative tool by including interactive data visualisation.
- It will allow not only researchers, but also the broader stakeholder community to follow the situation of most vulnerable groups.
- IPOLIS builds mainly on the European statistical system, while other data sources are also considered as inputs.

The aim of the work regarding IPOLIS was to build a platform to improve infrastructure for monitoring, analysing and evaluating the situation of the most vulnerable groups (Gábos & Kopasz, 2014). Nine specific vulnerable groups were identified at the beginning of our work in InGRID:

1. easy-to-reach groups: (a) children (0-17 years), (b) young people (15-30 years) and (c) older people (65+ years);
2. hard-to-identify groups: (d) migrants and people with migrant background, (e) Roma, (f) travellers;
3. hard-to-reach groups: (g) institutionalised people, (h) undocumented immigrants and (i) homeless people.

The building up of IPOLIS can be divided into two separate phases. In the first phase (InGRID-1, 2013-2017), IPOLIS was produced for the easy-to-reach, age-specific vulnerable groups: children, young people and older people. The selection of these three vulnerable social groups was supported by the following considerations:

- the risk of poverty and of social exclusion is higher than population average for children, young adults and older people in almost all countries, when examined by age (e.g. Eurostat, 2010);
- age easily identifies groups both in administrative and survey type data collections, which is not the case with other attributes;
- important prior efforts to monitor poverty, living conditions, quality of life and well-being exist for these age groups, especially for children.

After the delivery of the first phase of IPOLIS to the European Commission in February 2016 (database) and in February 2017 (visualisation tool). For the setup of the indicators of the monitoring system and for carefully analysing myriads of topical methodological problems and decisions, a number of research papers/deliverables were drafted, discussed and published (Gábos & Kopasz, 2014, 2015; Schäfer, Zentarra & Groh-Samberg, 2015; Kopasz, 2015; Limani, 2017).

The second phase of the IPOLIS project focused on extending the indicator system database with additional vulnerable groups, once they can be coherently identified in a large data infrastructure and robust indicators can be produced. The details on the extension of IPOLIS to include additional groups were set out in additional working papers (Bernát & Messing, 2016; Schepers, Juchtmans & Nicaise, 2017).

As a result, the IPOLIS database was further developed to better facilitate new research on poverty, living conditions and social policy, as well as to extend it with additional vulnerable groups. Four groups were considered, such as (i.) disabled people, (ii.) migrant people and people with migrant background, (iii.) Roma people and (iv.) people living in institutions.

After careful evaluation of the underlying data infrastructure (see Gábos & Kopasz, 2018, also involving knowledge from the participants of the expert workshop on 'Methods and data infrastructure to measure the quality of life of various vulnerable groups: extending IPOLIS', held on 25-27 April 2018 in Budapest), the selection of disabled people and migrants has been decided for the extension of IPOLIS under the InGRID-2 project.[2] Nevertheless, a methodological and data infrastructure report on these four vulnerable groups was prepared under InGRID-2 (Gábos *et al.* 2020).

The value added of IPOLIS for academia and policy

The overall aim of IPOLIS did not change though the development process:

2 While the data infrastructure on the Roma is improving (see Bernát & Messing, 2016), it cannot yet provide statistically robust and timely data for most of the countries where the share of Roma population is considerably high. As the institutionalised population is concerned, the data infrastructure is poorly developed and far from being able to provide indicators for IPOLIS or for any similar initiative.

- to improve infrastructure for analysing and monitoring the situation of most vulnerable groups in general;
- to monitor the situation of children, young people, older people, people with migrant background and people with disabilities in particular, in the fields of poverty, living conditions and quality of life; and
- to explore relationships between indicators and to detect cross-country patterns according to selected measures.

It was always conceived to serve as a resource for various user groups (researchers, policymakers at different levels, NGO experts, journalists, students, etc.).

The value added of IPOLIS lies in its integrated system, in its broad definition of quality of life domains and in the visualisation tool that makes analysis easy and user friendly.
- The importance of integration comes from the fact that while data and indicators in the field of poverty, living conditions and quality of life are widely available, it is rarely possible to analysed in a harmonised database the interlinkages across various domains of the quality of life. IPOLIS provides at the same time an integrated, multidimensional frame for analyses and interpretation, a selection of most relevant measures.
- The easy-to-handle and flexible visualisation tool helps users to explore the database and to detect cross-country and cross-time patterns and correlations in a flexible, user-oriented way. The data visualisation tool attached to the IPOLIS is embedded in an online platform (www.ipolis.tarki.hu), having a mutual direct linkage with the InGRID-2 project website (link here).
- The online platform has a simple design and structure, focusing on the online tool that allows visual analytics in IPOLIS. Users can easily navigate between vulnerable group modules and find information about both the project and IPOLIS, including useful links to outputs and events related to them.

Technical aspects of IPOLIS

The actual version of the IPOLIS data visualisation tool provides users with the following options to explore the indicator database:
- cross-country comparative analysis (column charts, spider web charts);
- time-series analysis (line charts);
- bivariate correlations (scatterplots);
- paired cross-module analysis.

In addition, several in-built features assist users to prepare visual outputs that are most convenient for their purposes. The most important features are as follows:
- tutorial;
- indicator selector, including three levels: domain, indicator and breakdown;
- country selector, including options for individual country and group selection;
- year selector;
- scale fixing option;
- value displaying option;
- benchmark selector (EU-28, EU-27, EU-15, EU-12, overall/adult population where available);
- ranking and highlighting;
- direct download options as either .cvs files or images;
- share through the social media (Facebook, Twitter).

Some features are selective for specific analytical options. The public version of the IPOLIS platform was launched in February 2017 and the updated/extended version is to be opened in August 2021.

Coherence and interlinkages across domains and target groups

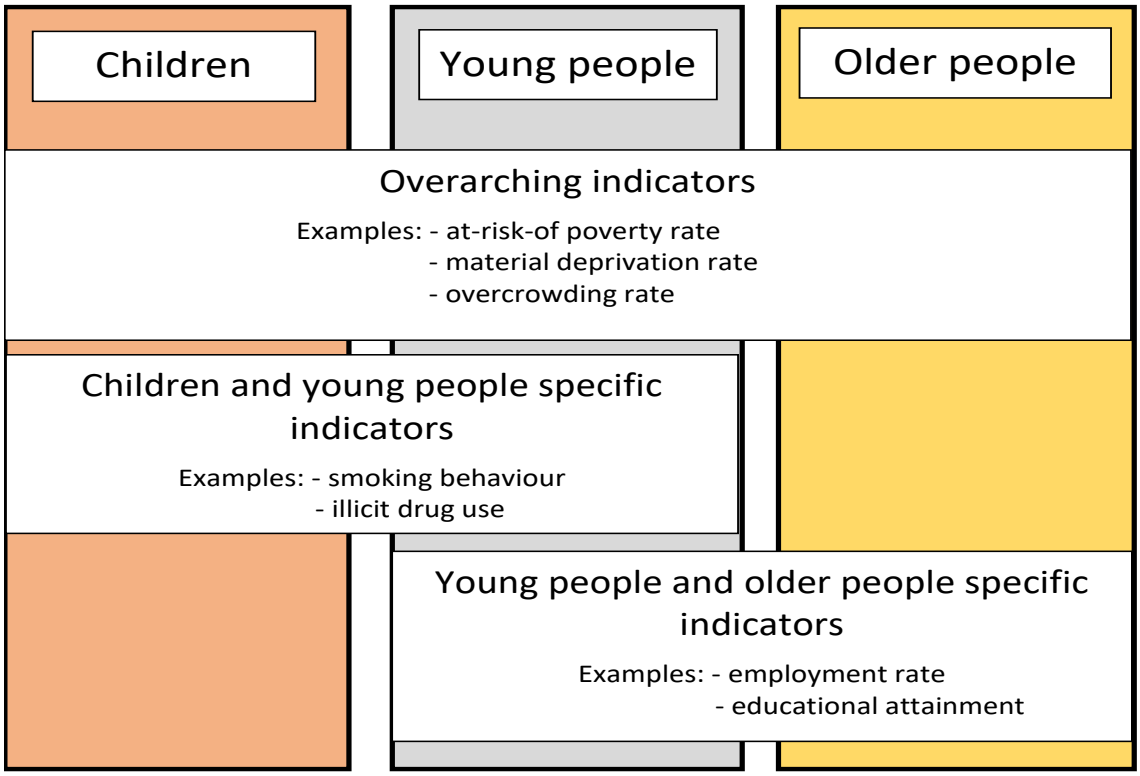
A major dilemma faced when setting up a complex database is that with very few exceptions, prior indicator system initiatives relate either to one specific vulnerable group (e.g. children, older people, etc.) or to a certain domain for the population as a whole. When setting up IPOLIS the situation was no different (Gábos & Kopasz, 2014, 2018).

For an integrated database that is carefully designed to cover special target groups and a complex set of special domains, it is very important to:

- ensure the coherence of the indicator system structure at the level of domains, components and subcomponents;
- set up direct linkages at indicator level between groups to allow for a comparative assessment of their relative positions - primarily according to the dimensions of poverty and material living conditions;
- consider that each stage of life cycle has its own characteristics and thus we need to pay special attention to age-group specific problems.

Figure 5 shows in a simplified way how the linkages between vulnerable groups like children, young people and older people were established in the first phase. Each portfolio of indicators belonging to a specific vulnerable group is represented in the figure by a differently coloured vertical rectangle. A set of indicators, referred here to as overarching indicators, characterises all three groups. These measures should have the same definition, preferably should be produced on the same data source and based on the same methodology. The application of these criteria was facilitated by the fact that vulnerable groups in the first version of IPOLIS were defined by age, but - depending on identification and data robustness - the situation was different for disabled people and migrants and people with migrant background. Further, while for household level indicators, like household income and material living conditions, the solution was relatively straightforward but for variables like perceived general health or physical activity (all relevant for all three age groups), finding a sufficiently comparable data source proved to be a difficult task. In addition, some of the potential indicators can be relevant for not only one, but two vulnerable groups. For example, this is the case with risk behaviour indicators, which are relevant for both children and young people, or with employment rate which is an important indicator for both young and older people.

Figure 5. Linkages across vulnerable groups in the first version of IPOLIS

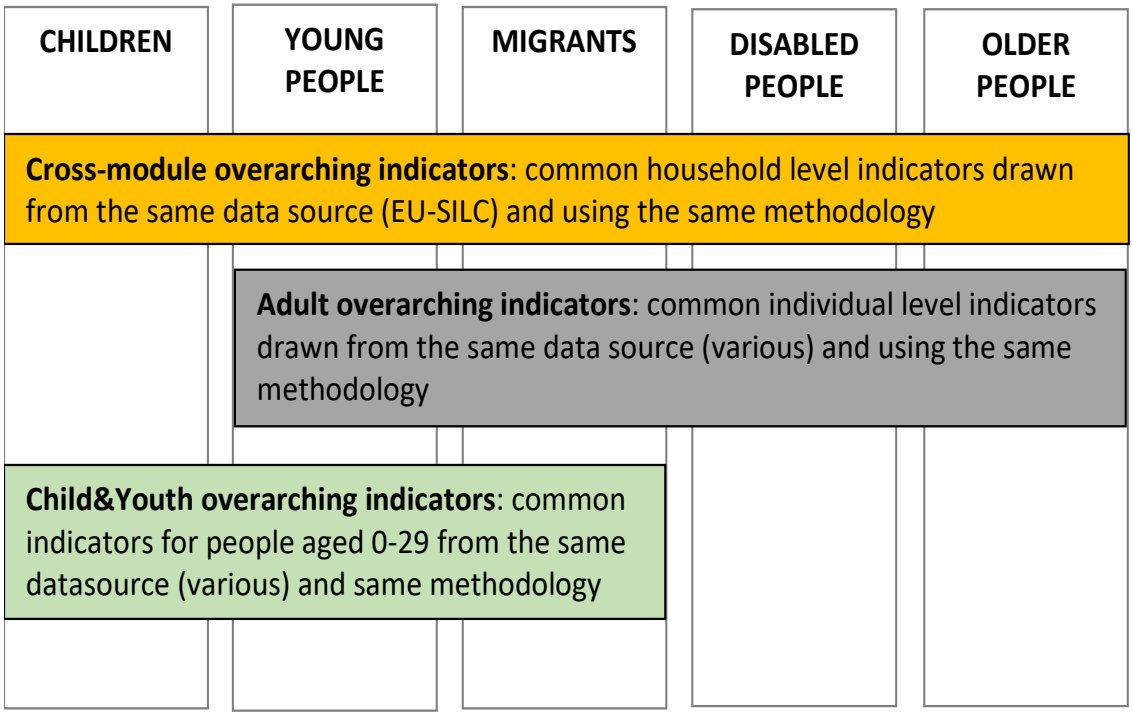


Source Gábos and Kopasz (2014: Fig. 1)

Figure 6 shows how the inclusion of two additional vulnerable groups has changed the picture. Similarly to what has been previously discussed, mainly household level indicators were suitable to make this link across all groups and specific variables from specific data sources had to be included for the additional target groups in the database. The groups of overarching indicators in the extended new structure were then similar to those in the simpler version of IPOLIS:

- overarching indicators: embrace all the five modules and include household level indicators, like household income and material living conditions;
- adult overarching indicators: involve all modules but children, and include individual level indicators on several aspects of quality of life;
- child and youth overarching indicators: some of the individual level indicators at young age can be relevant for not only one, but three vulnerable groups: children, young people and migrants. These indicators include, for example, risk behaviours, educational outcomes, social relations.

Figure 6. Linkages across vulnerable groups in the extended IPOLIS



Source Based on Gábos and Kopasz (2014: Fig. 1)

The types of analysis supported by IPOLIS

As described above, the first version of the IPOLIS data visualisation tool included three analytical options (cross-country comparison, time-series analysis and bivariate scatterplots) and two cross-module analysis opportunity for each module. The extension of IPOLIS towards disabled people and migrants and the new design of the analytical features in the data visualisation tool is presented in Figure 7. While all the three analytical options (cross-country-comparison, time series analysis and bivariate scatterplots) remained, cross-modules options were set in accordance with the incorporations of the new modules and the groups of overarching indicators previously presented in Figure 6. This analytical framework provides largest flexibility to all users to explore the underlying database.

As an example for the usefulness IPOLIS indicators for analysis with the aim to analyse and evaluate the situation of the children (as one of the most vulnerable groups in the society), as well as to inform policymaking efforts and initiatives, is provided by Limani *et al.* (2020). The report, published as an InGRID-2 working paper, aimed to contribute to such efforts by providing detailed information and analysis on children's material living conditions since the economic crisis. The main objective of the paper was to descriptively analyse child poverty and the relationship between parental background and material living conditions among children, across and within European Union (EU) member states, while also capturing trajectories in the period of and following the Great Recession. The report provides descriptive statistics on main trends, while the focus of the analysis was on the changes in indicators over time and by parental education. It covers the EU-28 Member States, plus the three additional EEA member countries (Iceland, Norway and Switzerland). As a next step, the results of the analysis were linked more closely to public policies in the field, gain in a cross-European perspective (Gábos, Kopasz & Limani, 2020). This would not have been possible without the complexity and detail of IPOLIS.

Figure 7. Analytical features across vulnerable group modules

CHILDREN	YOUNG PEOPLE	MIGRANTS	DISABLED PEOPLE	ELDERLY PEOPLE
Cross-country analysis	Cross-country analysis	Cross-country analysis	Cross-country analysis	Cross-country analysis
Time series analysis	Time series analysis	Time series analysis	Time series analysis	Time series analysis
Bivariate scatterplots	Bivariate scatterplots	Bivariate scatterplots	Bivariate scatterplots	Bivariate scatterplots
Cross-module analysis	Cross-module analysis	Cross-module analysis	Cross-module analysis	Cross-module analysis
Child & Youth analysis	Child & Youth analysis	Child & Youth analysis	Youth & Adult analysis	Youth & Adult analysis
	Youth & Adult analysis	Youth & Adult analysis		

Challenges and tasks for further extensions

Policy analysis as such can be based on various outcome indicators on the one hand but also, for a detailed and careful analysis some genuine policy variables are needed to be incorporated. To expand on this, a specific project output (Eneroth *et al.*, 2019) was dedicated to revise and enrich the existing set of policy indicators (Limani, 2017). The aim of this task was also to tailor the set of indicators that can help the users to assess cross-country variation in quality of life outcomes against policy and context information to the vulnerable group structure of IPOLIS. The proposal tried to strike a fruitful balance between some core indicators based on legislative frameworks, and indicators on distributive impacts derived from socio-economic surveys. Considering the former, the focus was put on major cash benefit schemes of outmost importance for economic wellbeing of the vulnerable groups identified in IPOLIS. More specifically, minimum income protection (minimum wages, social assistance, and minimum pensions), unemployment benefits, sickness benefits, child benefits, and standard pensions were proposed. All these indicators are based on model family analyses, where incomes for a pre-defined set of families are calculated based on social policy legislation (Eneroth *et al.*, 2019).

As detailed above, the second phase of the IPOLIS development aimed the integration of four additional vulnerable groups. Two of these (the Roma and institutionalised people) need further work in order to be ready for a full consideration.

The Roma

The main obstacle of including the Roma within the frame of the IPOLIS is the lack of a solid data infrastructure. This means that existing surveys are not adequate to provide a complete picture on the quality of life of the Roma on a regular basis across Europe.

The existing surveys that collect data on the Roma population in Europe, face several problems (Bernát & Messing, 2016).

- There is a lack of a baseline against which representativeness can be defined. In several countries, even the census does not include a category for 'Roma' or 'Gypsy'. Even when this is the case, censuses typically underestimate the share of Roma population, and are therefore generally imper-

fect sources for sampling. Depending on the method used to overcome this problem, the surveys may cover very different population segments (Messing, 2014).
- The definition of who is considered 'Roma' depends on how surveys operationalise the category of 'Roma', and they may end up with very different results in terms of basic indicators, such as employment rate, educational level, housing conditions, etc.
- The sensitivity of information on ethnicity, which may be very differently handled by members states – again, affecting validity and comparability.

However, quality of life indicators for Roma people would potentially be available if large-scale European-wide mainstream surveys (e.g. EU-LFS, EU-SILC, European Social Survey - ESS or the PISA) included data on the ethnic background of each respondent (Bernát & Messing, 2016: 13). This would provide a very good and comparative source for social inclusion indicators – not only across European countries, but also between the Roma and non-Roma populations in individual countries. It is important to note that there is a good practice in this respect: Hungary has already adopted this approach and the Central Statistical Office introduced a dual identification question on ethnicity to all non-compulsory surveys from 2014 (after piloting the on the LFS in 2013).[3] Until now, this is an isolated practice in the EU.

Institutionalised people

EU population surveys are usually restricted to private households. The exclusion of those living in collective households poses important questions.
- First, the size of the institutionalised population is not negligible, and is very likely to increase due to the growing number and percentage of older people across Europe.
- Second, a meta-analysis of surveys in the institutionalised population indicates that persons in this group differ in the distribution by age, gender, medical condition, economic activity, housing, social networks, etc.

Thus, we can assume that the inclusion of people living in institutions in general population surveys will change estimates of indicators related to these areas or indicators requested by the OMC (Eurostat, 2011; Schanze, 2017).

There are some initiatives that address those challenges. For example, the SERISS (Synergies for Europe's research infrastructures in the social sciences) project aims to better represent the European population, including important target groups for policymakers such as young unemployed, older persons in institutions and migrants; and to strengthen cross-national harmonisation of data and methodology across Europe. More specifically, the project examines the feasibility to include the institutional population into cross-national population surveys in the EU (Schanze, 2017). Also, two Eurostat coordinated cross-national surveys (EHIS, EU-LFS) allow for the inclusion of people living in institutions (Gábos & Kopasz, 2019).

3.2.4.2 The Institutional Characteristics of Trade Unions, Wage Setting, State Intervention and Social Pacts (ICTWSS)

The process of building up ICTWSS and its main characteristics

The update and extension of the ICTWSS database (https://www.oecd.org/employment/ictwss-database.htm), was the one of the core outcome of the work package 11 on data integration and harmonisation for working conditions and vulnerability pillar of the InGRID-2 project. The main aims were on the one hand to (a) ensure a continuation of the dataset and (b) to generally up-date and extend the information available in the data set with respect to year, covered countries and

3 For more information on the HCSO pratcice, please visit the relevant persentation from the expert workshop 'Methods and data infrastructure to measure the quality of life of various vulnerable groups: extending IPOLIS', Budapest, 25-27/4/2018. http://www.inclusivegrowth.eu/expert-workshops/call-6-expert-workshop-tarki#program.

covered variables (in particular on bargaining coverage, trade union memberships and minimum wages).

Originally, the database on Institutional Characteristics of Trade Unions, Wage Setting, State Intervention and Social Pacts (ICTWSS) was developed by Prof. Jelle Visser at the University of Amsterdam. It was first released in May 2007. In its initial form, the ICTWSS database combined data from various sources and projects with a main focus on trade union in EU and OECD countries (Visser & Ebbinghaus, 2000; Visser, 1991; Visser, 2006), collective bargaining and employment relations in Europe (European Commission, 2004), and social pacts (Avdagic, Rhodes & Visser, 2011). After its first release, the database has been updated every second or third year and more variables and countries have been added.

In 2021, in the framework of the InGRID-2 project, the ICTWSS database has been rebranded as the OECD/AIAS ICTWSS database. This new name reflects the joint effort by the OECD and AIAS-HSI to ensure the continuation of the database after Prof. Visser's retirement. The OECD/AIAS ICTWSS database develops and consolidate earlier versions of the ICTWSS database, notably in providing more detailed information on minimum wage settings in OECD (Besamusca, 2019; Besamusca, Garnero & Korinth, 2021) and expanding geographical coverage to Western Balkan countries. The first version of the OECD/AIAS ICTWSS database has been released in February 2021 and has been produced with the financial assistance of the European Union Programme for Employment and Social Innovation 'EaSI' (2014-2020), VS/2019/0185.

The value added of ICTWSS for academia and policy

The ICTWSS aims to provide researchers and policymakers with a set of country-level data in two key areas of modern political economies (see for more detail, Visser 2021):

1. the organisation and coordination of collective bargaining, wage setting, and social pacts; and
2. the organisation of employers and representation of employees in trade unions and works councils.

The database presents annual data between 1960 and 2019 for 56 countries: all current OECD and EU members: Australia, Austria, Belgium, Bulgaria, Canada, Chile, Colombia, Costa Rica, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Germany, Greece, Finland, France, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Malta, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Spain, Slovakia, Slovenia, Sweden, Switzerland, Turkey, the United Kingdom and the United States of America – with some additional data for the G20 countries: Argentina, Brazil, China, India, Indonesia, Russia and South Africa; and EU accession and Balkan countries: Albania, Bosnia and Herzegovina, Kosovo, Montenegro, Serbia, North Macedonia.

The database contains more than 100 variables, organised in 7 groups:

1. rights of association, collective bargaining and strikes, private and public sector;
2. wage setting: state intervention, coordination, centralisation, conflict resolution, indexation, extension, derogation, etc.;
3. social pacts, central agreements and social dialogue;
4. works councils and employee representation in the enterprise;
5. employer organisations;
6. trade union membership and union density;
7. collective bargaining coverage.

The types of analysis supported by ICTWSS

The ICTWSS lends itself to analyses in the broad fields of industrial relations, welfare states, labour economics and labour market institutions and stratification. The indicators presented in the ICTWSS

facilitate policy comparisons, for example in regard to the centralisation of wage bargaining, the analysis of social pacts or the organisation of employers and workers' representatives. During the InGRID-2 project, particular efforts were made to deepen the available data on minimum wage policy, which has (re-)gained the attention of academics and policymakers in recent years (Besamusca, 2019). First, politically, statutory minimum wages are increasingly seen as an instrument to fight in-work poverty in line with predistribution and social investment policy ideals, as evidenced by the draft directive on adequate minimum wages in the EU. The draft directive has sparked interest in minimum wage policies among thinktanks and policy institutions, who can use the ICTWSS to compare minimum wage policy models across countries. Academically, the ICTWSS data on minimum wages can be used to answer three main open questions: (1) the relation between minimum wage fixing regimes and policy goals like curbing poverty and inequality; (2) the involvement and power relations of social partners in minimum wage fixing; and (3) the consequences of differentiated minimum wage rates. The revamped ICTWSS database fills a gap in the EU data infrastructure, where information on the level of earned wages and minimum wages is readily available while information on the characteristics of minimum wage policies were missing.

Furthermore, the ICTWSS offers a range of indicators that can be employed to measure institutional characteristics in country comparative studies of economic or labour market performance. Most prominently, the variables of collective bargaining coverage and trade union density, both of which have been updated as well as harmonised, are used in many academic studies as a proxy for trade union strength and the institutionalisation of collective bargaining. Thanks to the large efforts of the OECD team, the revamped ICTWSS contains higher quality time series for both variables, enabling more longitudinal research.

Technical aspects of ICTWSS

The ICTWSS provides users on the one hand with a downloadable dataset containing information for all countries, years and included indicators. On the other hand, users are also able to customise their request via a newly developed user-interface (www.ictwss.org). Here, users can download, based on their needs, tables and visualisations (such as bar and line charts) based on a selection of variables, countries and years.

The most important features assisting users in their request are:

1. indicator selector (including breakdowns);
2. country selector (including individual and group of country selection);
3. year selector; as well as the option
4. to download those customised request either as cvs files or images.

An additional feature, of the new data interface is that country experts can enter data on the website backside.

The actual version of the ICTWSS provides users with the following options to explore the database:

- cross-country comparative analysis (tables and column charts);
- time-series analysis (tables and line charts).

Challenges and tasks for further extensions

Over the course of the InGRID-2 project and with the support of the European Union Programme for Employment and Social Innovation 'EaSI' (2014-2020, VS/2019/0185), the OECD/AIAS ICTWSS presents a consolidated version of the ICTWSS with high quality harmonised data, a longer time series, 56 countries. Crucially, its continuation in future years has been secured. Building on these achievements, a number of challenges to improve the data and relevance of the ICTWSS remain.

First, in the development from the ICTWSS to the OECD/AIAS ICTWSS, the scope of the database was reduced from 11 to 7 substantive groups due to data quality. Based on the available data on the four remaining groups, comprehensive harmonisation and data quality checks were considered infeasible within the time frame and resources of the InGRID2 project. This concerns the topics of (1) sectoral institutions and employer organisations, (2) number and membership of trade unions and union confederations, (3) membership shares, conflicts and divisions between and within trade union confederations, and (4) break downs of union membership and union density rates, for example by gender, occupation or sector. Since no comprehensive data infrastructures exist that can offer alternatives, the first challenge of the ICTWSS is to preserve and harmonise this part of the database for the future, preferably in cooperation with the ILO, Eurofound and social partners.

Secondly, data on employer density is much more scarcely available than data on trade union density. Since scholarship on collective bargaining and industrial relations indicates that employer behaviour and attitudes are crucial to maintain high levels of (sector level) collective bargaining and centralisation of wage setting, this represents an important data gap.

Thirdly, the methodology for the classification of minimum wage policies and minimum wage fixing regimes, which was developed in the context of the InGRID-2 project, will need to gain be tested and gain prominence in academic research. Furthermore, for longitudinal analyses, this group of variables will need to be back coded to previous years

4. Structural and organisational challenges

To advance and improve the InGRID research infrastructure, several organisational and structural issues have to be resolved. These challenges are briefly presented in this section of the report. This section refers back to some of the core issues identified by Van Gyes *et al.* (2017), explains how these have been tackled in InGRID-2, as well as what challenges remain to be addressed.

4.1 Synergies and cross-fertilisation

Successful integrating of a distributed RI includes of course the look for synergies (Van Gyes *et al.*, 2017). Within the two-pillar structure of the InGRID-2 project, looking for synergies meant investigating and developing the expansion or use of data, tools or methods of the one pillar in the other. In the innovation agenda on moving from the first to the second InGRID project, elements of cross-fertilisation related to the extension of IPOLIS in time, with information on new vulnerable groups, labour market variables and with policy indicators; the extension and enforcement of the EUROMOD approach; joint methodological exploration of the advantages of data linkages (in small area estimation of poverty, improving the robustness of non-probability web data and web surveys), the use of microsimulation in labour studies, and the use of web data for monitoring new jobs and new skills; proposals for joint harmonisation efforts to be continued or newly started on longitudinal educational careers data, methods to calculate intra-country comparisons of indicators, protocols for social rights indicators of migrants, methods for calculating multidimensional indicators (e.g. vulnerable groups, quality of work, OSHA policies). In the continuation of the InGRID research infrastructure, further efforts to strengthen the connection between the two pillars to foster synergies is required, as well as synergies with other projects and infrastructures.

4.2 Knowledge exchange and innovation networking and stakeholdering

In the advancement of InGRID as European integrating research infrastructure, efforts to the further promotion, information-sharing and dissemination were very important (Van Gyes *et al.*, 2017). The networking activities in InGRID-2 concentrated on the outreach to the broader communities of users and the dissemination of the RI service access. Training events were organised for early stage researchers to stimulate the knowledge exchange and transfer. This included the recording of video tutorials inspired by MOOCs, which proved to be successful. Expert workshops were organised to encourage the discussion on RI improvements.

Besides these two types of knowledge exchanging events, the InGRID-2 project also included data forums, special interest group meetings and stakeholder platform conferences targeting different groups of experts and stakeholders, which proved helpful in making progress on specific data and methodological issues (e.g. reference budgets, dynamic microsimulation, etc.). As described above, the InGRID-2 research infrastructure of integrated and comparative data sources is fed with a series of data initiatives, including European official surveys such as EU-SILC, E-LFS, European household finance and consumption survey; fragmented academic data gatherings (ESS, SHARE, ISSP); data from European and international agencies; major national survey instruments like census data, established working conditions surveys (mainly in the Nordic countries) and country best-practices. These initiatives are complimented with own, more first-mover data procurements based on

web(crawling) data or policy indicator-building. Organising stakeholder involvement of these (external) data initiatives is crucial for the strategic advancement of the InGRID RI. Stronger networking with organisations such as the ILO and OECD looks an important issue to tackle in this strategic area. As described above, InGRID-2 has made good progress in this area and these efforts should certainly be continued and further extended in a follow-up project.

Indeed, when the InGRID-2 user community was asked about the issues they encounter when using different methods and tools, difficulties when combining different methods/tools emerged as the most or second most important issue (over 50% of academics and policymakers) (Szekér *et al.*, 2020). The lack of skills and training to work with these methods and the need for new and better methods are placed second and third. Other issues that were added by respondents were problems related to the use of specific software and the time to learn and experiment with new data. This thus indicates that efforts to exchange knowledge through training events are critical. Another suggestion made by the Advisory Board is to foresee training specifically targeting policymakers, e.g. on how to use or interpret data, how to compile basic descriptive statistics, etc. Short video tutorials could be helpful in this regard. Indeed, 87% of the respondents who completed the user survey indicate they would make use of online training resources if they were available (Szekér *et al.*, 2020). This share is larger for policymakers than academics, and larger for respondents from non-European countries.

The user survey also revealed a clear interest in an increased interaction between policymakers and academic through events such as round tables, seminars, workshops, and conferences (Szekér *et al.*, 2020). Events should embody investments in actual interaction and cooperation, as was the case for example in the special interest group meetings run in InGRID-2. Second, according to the user community, communication about research results could be more accessible and understandable for policymakers, e.g. by preparing short articles or focusing on visualisations (in English as well as other languages). One suggestion that also emerged from this consultation of the InGRID-2 community is to set up an online portal through which policymakers and academics can interact and exchange knowledge and information (see Szekér *et al.*, 2020).

4.3 Financial sustainability

InGRID started as an integrating research infrastructure at the transnational European level in 2013, funded under the EC FP7 programme, and continued as an advanced research infrastructure in 2017, funded by the EC H2020 programme. The European Commission funding was decisive to innovate and stimulate the integration and improved of decentralised and distributed research infrastructures (also see Van Gyes *et al.*, 2017).

These research infrastructures are funded by national sources or project-based European money. The InGRID-2 consortium includes, first of all, European data centres specialising in integrating data from various European and national data sources in social sciences. Some of them exist already for a long time (e.g. LIS), others are continuously looking for (EU) support (e.g. EUROMOD, IECM at CED) or drive on (mainly) national funding (e.g. SPIN, CSB-MIPI). Secondly, new data initiatives are continued by ad-hoc project-based initiatives (e.g. the WageIndicator survey) or by convincing national actors (e.g. MEADOW survey protocol of linked employer/employee surveys used in the Nordic countries). Third, the expert services in statistics, classifications, protocols and official comparative data sources (e.g. EU-SILC, EWCS, …) are built-up throughout involvement in EU research projects and by InGRID made available for distribution. A fourth group of 'best practice' national surveys base their financing on national funding.

The financial sustainability of these national research infrastructures remains a critical issue, as was also pointed out during the roundtable discussions and strategic gap analyses for South-Eastern and Central Europe, and the roadmaps for Greece and Slovakia. A critical factor in this regard is certainly the increased budgetary pressures that are noticed in relation to some of the key (national) data sources, compiled by pre-users (e.g. the funding of European agencies conducting surveys; cuts in

Ministries of Labour affecting national working conditions surveys, …). Further integration is by itself an impact factor in this regard. Past and on-going investments in social sciences equipment in a series of European FP large-scale projects, networks of excellence, etc. are in this way valorised and opened to others due to this integration. For a part of the main research infrastructures developed or maintained under the InGRID-2 project, agreements have been made and their long run sustainability appears ensured, e.g. ICTWSS database described above.

4.4 Coverage of the research infrastructure: CEE countries

As in many other research infrastructures of the European research area, integrating all countries of the European Union is a big challenge for the InGRID RI (see Van Gyes *et al.*, 2017). The former communist Central and East-European countries (CEE), that joined the EU the last decade, are in this regard as in many other sectors 'catching-up'. As Kovacs & Kutsar (2012) rightly state in their introduction social sciences had to be built up almost from scratch after the collapse of the communist system in the early 90s. In addition, the integration is also a necessary task taking the massive challenges of social inclusion and employment vulnerability into account of these regions. From an inclusive growth perspective these countries are a key-world-of-interest in the EU, on the one hand due to still worrying social situations, but also on the other hand the sometimes radical policy innovations implemented and given the sometimes large difference between these countries in social performance and innovation.

Nevertheless, data expert centres engaging in the kind of integrated and comparative InGRID data activities are more difficult to find in CEE countries. In the first InGRID project, only one partner was from CEE, namely TARKI from Hungary. In the second InGRID project, additional partners were included, such as CELSI in Slovakia and CIOP in Poland. Also a Greek partner PANTEION joined the consortium. Also in the comparative data sources that form the main pillar of the InGRID research infrastructure, the CEE countries are not under-represented. Additional efforts were also done to attract researchers and policymakers from South-Eastern and Central Europe to participate in InGRID-2 events, use the research infrastructure or take up a visiting grant. While these efforts were successful, further work is needed in this area.

4.5 Open and ethical

Promoted as part of the innovation-oriented European policy frameworks, the incorporation of emerging open science principles into research workflows increases transparency, reproducibility, dissemination, re-use and the transfer of knowledge - both within the research community as towards society (also see Van Gyes *et al.*, 2017). It maximises the potential for uptake, involvement by unforeseen collaborators, serendipitous discovery of knowledge, and the emergence of unplanned spaces of innovation. Depending on the type of research object, the principles are applied under the rubric of 'open access' (publications, working papers, presentation, educational resources), 'open data' (primary and derived data) or 'open source' (tools and code). As research infrastructure, all these three aspects of 'open science' deserve attention and action.

The same argument goes for promoting and guaranteeing research ethics. Familiarising new users with the InGRID RI and extending the use of the RI has to incorporate the necessary attention to research ethics. First-of-all, the principles and guidelines of the *European Charter for Access to Research Infrastructures* (ECARI), which includes and expands on the principles of the *European Code of Conduct for Research Integrity* (ECCRI) can act here as ethical framework. The ECCRI code of conduct sets out 8 principles, i.e. honesty in communication; reliability in performing research; objectivity; impartiality and independence; openness and accessibility; duty of care; fairness in providing references and giving credit; and responsibility for the scientists and researchers of the future. ECARI adds RI-specific

principles such as legal conformity to national and international law when designing access and use of RIs, transparency, non-discrimination when providing access, and encouragement of open access.

Taking these ethical guidelines and 'open science' principles into account, adherence to transparency and non-discrimination shall have to be ensured and verified in all steps of the process of providing access to the RI (transnational access, virtual access, etc.). This will be realised by (1) providing a clear and centralised point of contact for requests and information dissemination, (2) centrally verifying transparent and correct selection criteria before events or calls are announced, and (3) storing information on (non-)selection to verify complaints post-hoc if needed. At the same time, transparency of the results of this access to RI is embedded in the project, by (1) strongly encouraging the public dissemination of these results in terms of open access to data ('open data'), research findings ('open access') and research tools ('open source'), and (2) providing the required e-infrastructure to do so.

The open and ethical approach has to include also a strong policy on the protection of personal data. Personal data may be defined as any data permitting to identify the person involved. InGRID is about integrating existing research infrastructures. A series of these infrastructures of course collect and store data and this in accordance with national and EU legislation. In this regard the infrastructures follow most of the time specific procedures of access. Especially the infrastructures specialising in integrating microdata in comparative datasets (LIS, CED and UEssex) follow in this regard restricted procedures, which have been set when these infrastructures obtain the data from the actual data collectors. If they would not have this kind of approval in relation to protection of personnel data, they would not obtain the data from the (national) statistical offices. In practice, it means that they have procedures to avoid that data can be used to identify persons. Although the RI integration is only indirectly and at a secondary level involved in this issue of personal data protection of the research infrastructures, activities and quality control procedures have to set-up to guarantee this protection of personnel data by the research infrastructures.

A data management plan can act as a base document in these procedures of promoting 'open science' and guaranteeing research ethics. The InGRID-2 data management plan lays out in detail the procedures applied during the project and was updated regularly. It could serve as a basis for future research infrastructure projects.

5. Potential future pathways

The following potential pathways are currently being explored by the InGRID-2 consortium in order to continue the work carried out within the project and maintain and improve the infrastructure. For some of these options, it is already clear today that they are not feasible, or unlikely to materialise, in the short- to medium-run.

Option 1: Establishing a European Research Infrastructure Consortium (ERIC)

As described on the European Commission's website[4], ERIC is a specific legal form that facilitates the establishment and operation of research infrastructures with European interest, both existing and new research infrastructures (on a non-economic basis). There is, however, a consensus in both the InGRID-2 Executive Committee and the Advisory Board that it is too early to establish an ERIC for InGRID for several reasons. First, the research infrastructure is insufficiently developed to meet this goal: an infrastructure built on indicators and methods, rather than on original microdata, may not provide a sufficient strong basis. However, data collection is currently not allowed in the InGRID-2 project, which is an important issue in this regard. Another complication is that only Member States, associated countries, third countries (non-associated countries) and intergovernmental organisations can be members of an ERIC. Even though a Member State or country may be represented by one or more public entities or private entities with a public-service mission, e.g. research organisations or research councils, to exercise specified rights or fulfil specified obligations on its behalf. This means that establishing an ERIC based on the InGRID research infrastructure is a lengthy process requiring close collaboration with national authorities. Advantages of an ERIC are the legal capacity recognised in all EU countries, flexibility to adapt to specific requirements of each infrastructure, a faster process than creating an international organisation, exemptions from VAT and excise duty. An ERIC would also be an important step forward as regards the long-run sustainability of the InGRID infrastructure.

Option 2: Launching a third InGRID project under the EU's Horizon Europe programme

Another option that is being explored is to launch a third InGRID project under the Horizon Europe work programme. Horizon Europe is the EU's key funding programme for research and innovation with a budget of 95.5 billion euro. The programme facilitates collaboration and strengthens the impact of research and innovation in developing, supporting and implementing EU policies while tackling global challenges. It supports creating and better dispersing of excellent knowledge and technologies.

The preferred approach of the InGRID consortium would be to prepare and submit a proposal under the Horizon Europe work programme on research infrastructures. This would help ensure that the infrastructure can continue working in its current form, focusing on a further integration of the two thematic pillars, expanding knowledge exchange and innovation networking activities, and widening the transnational and virtual access to the infrastructure. However, as there currently is no call open or forthcoming under the research infrastructures programme that fits with the focus and scope of the InGRID-2 project and the research infrastructures it is composed of an alternative option would be to submit one or more projects to call under the thematic calls and the related missions. Both options are currently being explored by the InGRID-2 consortium. While a third InGRID project

4 https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/european-research-infrastructures/eric_en.

would help to ensure the sustainability of the infrastructure in the short- to medium-run, a solution for the longer run still needs to be developed.

Option 3: Setting up an Innovative Training Network (ITN)

A third option is to set up an Innovative Training Network, funded under the MSCA programme. Innovative Training Networks support competitively selected joint research training and/or doctoral programmes, implemented by European partnerships of universities, research institutions, and non-academic organisations. The research training programmes provide experience outside academia, hence developing innovation and employability skills. An ITN would be another type 'product' or 'output' of the InGRID research infrastructure and the two projects, which could serve as a lever to launching a new project under the research infrastructure programme or another programme funded by the European Union. One idea could be to launch an InGRID-3 project and an ITN, in parallel, in order to draw on the synergies that are thus created. This synergy lies in several aspects:

- an ITN can support InGRID-3 by creating bridges between the 'Poverty, living conditions and social policies' and 'Working conditions, vulnerability and labour policies' pillars;
- InGRID is an established advanced data infrastructure and solid research community beneficiary for ESR and innovative EU research training;
- strong foundation and pre-conditions for success and for innovativeness in research and training programme;
- build on experiences with InGRID TNA, summer school and expert network activities to train next generation employable ESR;
- supports key scientific objectives of InGRID in providing innovative data and methods for evidence-based policymaking; each PhD can tackle specific identified scientific and methodological gaps.

The starting point for the ITN would be to look at dynamics of inequality and the distribution of income in Europe, which is determined by redistribution policies as well as pre-distributive policies. The ITN will invest drivers of change and societal challenges, such as ageing, household composition, migration, technological transformations, … and it will model the impact of these changes using microsimulation tools and techniques. One key advantage is that such an ITN could be used as a way to attract researchers from CEE countries, allowing to strengthen the research infrastructure there. This is an important step forward into filling existing gaps. A disadvantage of the ITN is that it can only involve a small core group of the InGRID consortium.

Option 4: Broadening the network via a European Cooperation in Science and Technology (COST) Action

A fourth option under consideration is to launch a COST action. Such action help to connect research initiatives across Europe and beyond and enable researchers and innovators to grow their ideas in any science and technology field by sharing them with their peers. COST actions involve bottom-up networks with a duration of four years that boost research, innovation and careers. Although COST actions only receive funding to cover the expenses of networking activities rather than research, this could be an interesting way to keep the InGRID partners together, while broadening its community of researchers, policymakers and other stakeholders, for example in those countries where there are currently gaps (e.g. CEE countries).

Bibliography

Ahmed, A., Articus, C., Ertz, F., Münnich, R., Schmaus, S., Shlomo, N., & Weymeirsch, J. (2021). Overview report on the generation of synthetic universes for microsimulations. Utility and Confidentiality of Synthetically Generated Data Sets. Deliverable D12.6, Leuven, InGRID-2 project 730998 – H2020.

Articus, C., Münnich, R., & Shlomo, N. (2021). Strategic briefing note on methodological challenges, Deliverable 7.7, Leuven, InGRID-2 project 730998 – H2020

Atkinson, A. (2015). Inequality. What can be done? Harvard University Press, Cambridge.

Avdagic, S., Rhodes, M., & Visser, J. (2011). Social Pacts in Europe. Oxford University Press.

Bernát, A. & Messing, V. (2016). Methodological and data infrastructure report on the Roma population in the EU.

Bertarelli, G., Biggeri, L., Giusti, C., Marchetti, S., Pratesi, M., Articus, C., Caratiola, C., & Münnich, R. (2020). Methodological paper on intra-country comparisons of poverty rates, Deliverable 9.9, Leuven, H2020 InGRID-2 project.

Besamusca, J. & Steinmetz, S. (2019). Vulnerable groups in European data sets: an inventory of the measurement of vulnerable groups in 27 data sets covering the European area. Deliverable 11.3, Leuven, InGRID-2 project 730998 – H2020

Besamusca, J. (2019). Towards a global classification of minimum wage fixing machinery: an inventory of clas-sification schemes for minimum wage setting machinery, InGRID-2 Milestone 105, Leuven, InGRID-2 project 730998 – H2020.

Besamusca, J., Garnero, A., & Korinth, H. (2021). Detailed note on minimum wage setting mechanisms in the OECD/AIAS ICTWSS Database. OECD Publishing, Paris.

Blalock, H.M. (1984). Contextual Effects Models: Theoretical and Methodological Issues. *Annual Review of Sociology, 10*, 353-372.

Bradshaw, J., Ditch, J., Holmes, H. & Whiteford, P. (1993). A comparative study of child support in fifteen countries. *Journal of European Social Policy, 3*, 255-271.

Burgard, J. P., Dieckmann, H., Krause, J., Merkle, H., Münnich, R., Neufang, K. M., & Schmaus, S. (2020a). A generic business process model for conducting microsimulation studies, *Statistics in Transition New Series, 21*(4), 191-211.

Burgard, J. P., Krause, J., & Schmaus, S. (2021). Estimation of regional transition probabilities for spatial dynamic microsimulations from survey data lacking in regional detail. Computational Statistics & Data Analysis, 154, 107048.

Burgard, J. P., Krause, J., Merkle, H., Münnich, R., & Schmaus, S. (2019). Conducting a dynamic microsimulation for care research: Data generation, transition probabilities and sensitivity analysis. In: Workshop on Stochastic Models, Statistics and their Application (pp. 269-290), Springer, Cham.

Burgard, J. P., Krause, J., Merkle, H., Münnich, R., & Schmaus, S. (2020b). Dynamische Mikrosimulationen zur Analyse und Planung regionaler Versorgungsstrukturen in der Pflege. In: Mikrosimulationen (pp. 283-313), Springer VS, Wiesbaden.

Cantillon, B. & Vandenbroucke, F. (2014). Reconciling work and poverty reduction: how successful are European welfare states? Oxford University Press, Oxford.

Cantillon, B. (2011). The paradox of the social investment state: growth, employment and poverty in the Lisbon era. *Journal of European Social Policy, 21*, 432-449.

Cantillon, B., Marchal, S. & Luigjes, C. (2015). Decent incomes for the poor: which role for Europe? ImPROVE Discussion Paper No. 15/20, November 2015. Antwerp. University of Antwerp.

Cheli, B. and Lemmi, A. (1995). A Totally Fuzzy and Relative Approach to the Multidimensional Analysis of Poverty. *Economic Notes, 24*(1), 115-134.

Daas, P. J. H, Puts, M. J., Buelens, B., & Van den Hurk, P. A. M. (2015). Big Data as a Source for Official Statistics. *Journal of Official Statistics, 31*(2), 249–262.

Daly, M. & Lewis, J. (2000). The Concept of Social Care and the Analysis of Contemporary Welfare States. *British Journal of Sociology 51*(2), 281-98

Davies (2004). Is evidence-based government possible? Jerry Lee Lecture, presented at the 4th Annual Camp-bell Collaboration Colloquium, Washington DC; in Sutcliffe, S., & Court, J. (2005, November). Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries? ODI, Overseas Develop-ment Institute. Retrieved from http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/3683.pdf.

Desiere, S., & Lenaerts, K. (2020). Is it feasible to integrate and harmonise national surveys on working conditions? A scoping paper, Deliverable 11.5, Leuven, H2020 InGRID-2 project, https://www.inclusivegrowth.eu/project-output.

Docrinal, L., Nelson, K., Sirén, S. (2015). InGRID Deliverable: Comprehensive indicators for the analysis of out-of-work benefits: introducing the out-of-work benefits dataset. Stockholm: Swedish Institute for Social Research.

Dräger, S., Kopp, J., Münnich, R. T., & Schmaus, S. (2021). Analyse der Grundschulversorgung mit Hilfe klein-räumiger Mikrosimulationsmodelle. Trier University, Research Papers in Economics 1/21, https://www.uni-trier.de/fileadmin/fb4/prof/VWL/EWF/Research_Papers/2021-01.pdf.

Dupriez, O., & Boyko, E. (2010). Dissemination of Microdata Files: Principles Procedures and Practices. Interna-tional Household Survey Network.

Durlauf, S. (2001). A Framework for the Study of Individual Behavior and Social Interactions. *Sociological Meth-odology, 31,* 47-87.

Dusa, A. *et al.* (eds.) (2014). Facing the Future: European Research Infrastructures for the Humanities and Social Sciences. Scivero Verlag, Berlin.

Dusa, A., Oeller, C. & Wolff, S. (2014). A Common Agenda for the European Research Infrastructures in the Social Sciences and Humanities. In: Dusa, A. *et al.* (eds.) (2014). Facing the Future: European Research Infra-structures for the Humanities and Social Sciences, Scivero Verlag, Berlin, 209-232.

Ebbinghaus, B., & Visser, J. (2000). The Societies of Europe. Trade Unions in Western Europe since 1945. London: Palgrave Macmillan.

Eneroth, M., Gábos, A., Kneeshaw, J., Marchal, S., Nelson, K., & Nieuwenhuis, R. (2019). Towards an IPOLIS social policy model, Leuven, H2020 InGRID-2 project.

ESCO (2010). Overview of ESCO - a European classification of Skills, Competences and Occupations. Brussels: European Commission, DG Employment/DG Education and Culture.

Esping-Andersen, G. (1996). Welfare States in Transition: National Adaptations in Global Economics. London: Sage/UNAISD.

Esping-Andersen, G. (2002). Why We Need a New Welfare State. Oxford: Oxford University Press.

ESSC (2014). ESS Vision 2020: Building the future of European statistics, 21st Meeting of the European Statistical System Committee. Luxembourg.

ESSC (2015). Item 7 of the agenda, Draft regulation on statistics relating to persons and households, IESS: Inte-grated European Social Statistics, Work Programme Objective 7. In: 27th Meeting of the European Statis-tical System Committee Luxembourg.

Esteve, A. (2021). Strategic briefing note on data challenges, Deliverable 7.6, Leuven, InGRID-2 project 730998 – H2020.

Eurostat (2021). Statistical regions in the European Union and partner countries - NUTS and statistical regions 2021, https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-20-092.

Eurofound (2014). Policy lessons from the fifth EWCS: The pursuit of more and better jobs. Publications Office of the European Union, Luxembourg.

Eurofound (2015). First findings: Sixth European Working Conditions Survey - Résumé. Eurofound, Dublin.

European Commission (2009). Communication from the commission to the European Parliament and the council on the production method of EU statistics: a vision for the next decade. Retrieved from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0404:FIN:EN:PDF.

European Commission (2004). Industrial Relations in Europe 2004, Office for Official Publications of the European Communities, Luxembourg.

European Commission (2016). Employment and social developments in Europe 2015. European Commission. Directorate-General for Employment, Social Affairs and Inclusion, Brussels.

European Commission. Directorate-General for Employment, Social Affairs and Inclusion Directorate A (2016). Employment and Social Developments in Europe 2015. European Union, Luxembourg.

European commission. Directorate-general for research and innovation directorate A – policy development and coordination unit A1 – internal and external communication (2016). Open innovation, open science, open to the world: a vision for Europe.

European Commission. Directorate-General for Research Directorate B – European Research Area: Research Programmes and Capacity Unit B.3 Research Infrastructures (2010). A vision for strengthening world-class research infrastructures in the ERA. Report of the Expert Group on Research Infrastructures. European Union, Brussels.

European commission. Directorates-general for research and innovation (rtd) andcommunications networks, content and technology (connect) (2014). Background document public consultation'science 2.0': science in transition. http://ec.europa.eu/research/consultations/science-2.0/background.pdf.

Eurostat (2011). European Health Interview survey Task Force IV - report on institutionalized people, Document Tech-HIS/11_10. https://ec.europa.eu/eurostat/cros/system/files/TF%20HIS%20Report%20of%20the%20Task%20Force%20on%20institutionalised%20people.pdf.

Farago, P. (2014). Understanding how research infrastructures shape the social sciences: impact, challenges and outlook. Dusa, A. *et al.*, Facing the future: European research infrastructures for the Humanities and Social Sciences, Berlin, Scivero verlag, p. 21-34.

Ferrarini, T., Nelson, K. & Sjöberg, O. (2014). Unemployment insurance and deteriorating self-rated health in 23 European countries. *Journal of Epidemiology and Community Health, 68*, 657-662.

Ferrarini, T., Nelson, K., Palme, J. & Korpi, W. (2013). Social citizenship rights and social insurance replacement rate validity: pitfalls and possibilities, *Journal of European Public Policy, 20*(9), 1251-1266.

Ferrera, M. & A. Hemerijck, A. (2003). Recalibrating Europe's welfare regimes. In J. Zeitlin and D. Trubek (eds) *Governing Work and Welfare in a New Economy*. Oxford University Press, Oxford.

Figari, F., Paulus, A., & Sutherland, H. (2015). Microsimulation and Policy Analysis. In B. A. Anthony & B. François (Eds.), *Handbook of Income Distribution, 2*, 2141-2221. Elsevier, Oxford.

Freeman, R. (1984). Strategic Management: A stakeholder approach. Pitman, Boston.

Gábos, A. & Kopasz, M. (2014). Deliverable 20.1: Concept paper - For an integrated poverty and living conditions indicator system (IPOLIS) database. FP7 InGRID project

Gábos, A. & Kopasz, M. (2015). Methodological and data infrastructure report on children, Working paper. Leuven, InGRID project, M20.6.

Gábos, A. (2020). Upgraded and extended IPOLIS web interface, Deliverable 1.5, Leuven, InGRID-2 project 730998 – H2020.

Gábos, A., György Tóth, I., Besamusca, J., & Steinmetz, S. (2021). The need for and role of good indicators and valorisation, Strategic briefing on linking with policy innovation, Deliverable 7.8, Leuven, InGRID-2 project 730998 – H2020

Gábos, A., Kopasz , M., & Limani, D. (2020). Child poverty in Europe: the role of parental background and of family policies after the Great Recession (in Hungarian). In: T. Kolosi, I. Szelényi and I. Gy. Tóth (eds.): Társadalmi Riport 2020. Budapest: TÁRKI, 309-331.

Gábos, A., Kopasz, M., Baroma, R.S., Bernát, A., & Messing, V. (2020). Methodological ad data infrastructure report on additional vulnerable groups indicators: migrants, Roma, institutionalised people and disabled persons, Deliverable 8.3, Leuven, H2020 InGRID-2 project.

Gábos, A., & Kopasz, M. (2014). Conceptual report for the Integrated Poverty and Living Conditions Indicator System (IPOLIS).

Gábos, A., & Kopasz, M. (2015). Methodological and data infrastructure report on children.

Gábos, A., & Kopasz, M. (2019). Concept paper on the extension and further development of IPOLIS, Leuven, H2020 InGRID-2 project.

Garfinkel, I., Manski, C.F., & Michalopoulos, C. (1992). Micro experiments and macro effects. In Manski C. F. and I. Garfinkel (eds.) Evaluating Welfare and Training Programs. Harvard University Press, Cambridge.

Gautié, J., & Schmitt, J. (2010). Low-wage work in the wealthy world. Russell Sage Foundation, New York.

Gilbert, N. (2009). The least generous welfare state? A case of blind empiricism. *Journal of Comparative Policy Analysis: Research and Practice, 11*(3), 355-67.

Greenan N. & Seghir M. (2015). Inventory of linked employer-employee surveys on working conditions and occupational health and safety issues, Working paper, Leuven, InGRID project, M.21.10.

Hamon-Cholet S., Salibekyan Z. & Mottier L. (2016). Variations in Inclusive Growth Research Infrastructures within the European Union: barriers and drivers. The case of the Eastern European countries. Futuring briefing note 3, Deliverable 2.3.

Harding, A., Keegan, M., & Kelly, S. (2010). Validating a dynamic population microsimulation model: Recent experience in Australia. *International Journal of microsimulation,* 6(2), 46-64.

Hedström, P. & Swedberg, R. (1998). Social Mechanisms: An Analytical Approach to Social Theory. Cambridge University Press, Cambridge.

Holman, D., Holtgrewe, U., Ramioul, M. & Vandekerckhove, S. (2015). Convergence and divergence of job quality in Europe 1995-2010: a report based on the European Working Conditions Survey. European Foundation for the Improvement of Living and Working Conditions, Dublin.

Huang, Y., & Shlomo, N. (2021). Methods for Sampling and Inference with Non-probability samples: Improving the Reliability of Web Survey: An Empirical Evaluation of Weighting Adjustment, Imputation, and Sampling Matching Strategies to Measure Gender Pay Gap, Deliverable D11.8 (Chapter 3.4), Leuven, InGRID-2 project 730998 – H2020.

Hurley, J. *et al.* (2015). Upgrading or polarisation? Long-term and global shifts in the employment structure: European Jobs Monitor 2015. Eurofound, Dublin.

Japec, L., Kreuter, F., Berg, M., Biemer, P., Decker, P., Lampe, C., Lane, J., …, Usher, A., (2015). American Association for Public Opinion Research: Task Force Report on Big Data. https://www.aapor.org/AAPORKentico/AAPOR_Main/media/Task-ForceReports/BigDataTaskForceReport_FINAL_2_12_15_b.pdf.

Klasen, S. (2010). Measuring and Monitoring Inclusive Growth: Multiple Definitions, Open Questions and Some Constructive Proposals. ADB Sustainable Development Working Paper Series 12.] Retrieved from: http://www.adb.org/sites/default/files/pub/2010/ADB-WP12-measuring-inclusive-growth.pdf.

Kleiner, B. *et al.*, (eds.) (2013). Understanding research infrastructures in the social sciences. Seismo Press: Zurich.

Kopasz, M. (2015). Methodological and data infrastructure report for the elderly, Working paper, Leuven, InGRID project, M20.8.

Kopasz, M. (2015). Methodological and data infrastructure report on the elderly.

Korpi, W. & Palme, J. (1998). The Paradox of Redistribution and Strategies of Equality: Welfare State Institutions, Inequality and Poverty in the Western Countries. *American Sociological Review, 63*(5), 661-87.

Korpi, W. & Palme, J. (2003). New Politics and Class Politics in the Context of Austerity and Globalization: Welfare State Regress in 18 Countries 1975-1995. *American Political Science Review, 97*(3), 1-22.

Korpi, W. (1989). Power, politics and state autonomy in the development of social citizenship: Social rights during sickness in 18 OECD-countries since 1930. *American Sociological Review, 54*(3), 309-28.

Kostolny, J., Szudi, G., Guzi, M. & Kahanec, M. (2021). Vulnerable groups in European datasets. An inventory of the measurement of vulnerable group outcomes in 27 datasets covering the European area (Deliverable D11.4), Leuven, InGRID-2 project 730998 – H2020

Kostolný, J., Szudi, G., Kahanec, M., Gábos, A., Pawlowska, Z., & Pecillo, M. (2020). Research infrastructure gaps report Central and Eastern Europe, Deliverable 7.2, Leuven, InGRID-2 project 730998 – H2020.

Kovács I. & Kutsar D. (2012). Internationalisation of Social Sciences in Central and Eastern Europe: The 'Catching Up'–A Myth Or a Strategy ? Routledge, New York.

Kühner, S. (2007). Country-level comparisons of welfare state change measures: another facet of the dependent variable problem within the comparative analysis of the welfare state? *Journal of European Social Policy 17*, 5-18.

Lechevalier, A. & Wielgohs, J. (eds.) (2015). Social Europe: A Dead End. What the eurozone crisis is doing to Europe's social dimension. Djoef publishing, Kopenhagen,.

Lenaerts, K.; Beblavý, M. and Fabo, B. (2016). Prospects for utilization of non-vacancy Internet data in labor market analysis-an overview. *IZA Journal of Labor Economics, 5*(1).

Lenau S., Liégeois P., Marlier E., Münnich R. & Van Gyes G. (2016). Statistical research agenda for an EU inclusive growth research infrastructure: some key issues. Futuring briefing note 2, Deliverable 2.2.

Lenau S., Münnich R., Hermans M. & Van Gyes G. (2017). From knowledge management system to portal: scoping the e-infrastructure of the InGRID Research Infrastructure. Deliverable 5.12.

Lewis, J. & Giullari, S. (2005). The adult worker model family, gender equality and care: the search for new policy principles and the possibilities and problems of capabilities approach. *Economy and Society, 34*, 76-104.

Li, J., & O'Donoghue, C. (2013). A survey of dynamic microsimulation models: uses, model structure and methodology, *International Journal of microsimulation,* 6(2), 3-55.

Lieberson, S. (1985). Making It Count: The Improvement of Social Research and Theory. Berkeley: University of California Press.

Ligon, E. & Schechter, L. (2003). Measuring vulnerability. *Economic Journal, 113*(486), C95-C102.

Limani, D. (2017). Methodological and data infrastructure report on the policy and context indicators.

Limani, D., Gábos, A., Baroma, R.S., & Kopasz, M. (2020). Social Status Inequalities in Child Poverty in the Euro-pean Union. Working Paper, Leuven, H2020 InGRID-2 project.

Mahieu, B. *et al.* (2014). Evaluation of Pertinence and Impact of Research Infrastructure Activity in FP7 – EPIRIA. European Commission DG Communications Networks, Content & Technology, Brussels. Retrieved from: http://ec.europa.eu/smart-regulation/evaluation/search/download.do;jsessionid=nfJOpPvHBbGQ9bpzZdE73kALP8qztuLd3PdtVvo-5N1L0GrkBPAu!-1694519822?documentId=16972411.

Marchetti, S., Giusti, C., Pratesi, M., Salvati, N., Giannotti, F., Pedreschi, D., Rinzivillo, S., Pappalardo, L. & Gabrielli, L. (2015). Small Area Model-Based Estimators Using Big Data Sources. *Journal of Official Statistics, 31*(2), 263–281.

Marx, I. & Nelson, K. (2012). Minimum Income Protection in Flux. HoundMills: Palgrave MacMillan.

Meylemans L., & Lenaerts, K. (2020). Inventory of available administrative data for comparative research on quality of employment, Data availability in selected EU Member States, Deliverable 12.4, Leuven, H2020 InGRID-2 project.

Montanari, I., Nelson, K., Palme, J. (2008). 'Towards a European Social Model? Trends in Social Insurance among EU countries 1980-2000'. European Societies 10(5): 787-810.

Morel, N., Palier, B. & Palme, J. (2011). Towards a Social Investment Welfare State? Ideas, Policies, Challenges. Policy Press, Bristol.

Mueller-Langer, F., & Andreoli Versbach, P. (2014). Open access to research data: Strategic delay and the ambiguous welfare effects of mandatory data disclosure. Max Planck Institute for Innovation & Compe-tition Research Paper, (14-09). Retrieved from http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2458362.

Münnich, R., Schnell, R., Brenzel, H., Dieckmann, H., Dräger, S., Emmenegger, J., Höcker, P., … Stein, P. (2021). A Population Based Regional Dynamic Microsimulation of Germany: The MikroSim Model. *Methods, data, analyses, 15*(2), 241-264.

Munoz de Bustillo, R., Macías, F-E., Anton, J-I. & Esteve, F. (2011). Measuring more than money. Edward Elgar, Cheltenham.

Nelson, K. (2011). Synthesis Report on Improving the efficiency of social protection. The Peer Review in Social Protection and Social Inclusion and Assessment in Social Inclusion. European Commission, Brussels.

Nelson, K. (2012). Counteracting material deprivation: the role of social assistance in Europe. *Journal of Euro-pean Social Policy 22*(2), 148-163.

Nelson, K. (2013). Social Assistance and EU Poverty Thresholds 1990–2008. Are European Welfare Systems Provid-ing Just and Fair Protection Against Low Income?, *European Sociological Review, 29*(2), 386-401.

Nelson, K., & Nieuwenhuis, R. (2018). List of policy indicators to feed into IPOLIS Technical note, Leuven, H2020 InGRID-2 project.

Nolan, B., *et al.* (eds.) (2014). Changing Inequalities and Societal Impacts in Rich Countries: Thirty Countries' Experiences. Oxford University Press, Oxford.

OECD (2008). Growing Unequal. OECD, Paris.

OECD (2011). Divided We Stand: Why Inequality Keeps Rising. OECD, Paris.

OECD (2013). OECD Skills Outlook 2013: First Results from the Survey of Adult Skills. OECD Publishing. http://dx.doi.org/10.1787/9789264204256-en.

OECD (2013). New data for understanding the human condition. International perspectives. OECD Global science forum report on Data and research infrastructure for the social sciences. OECD Publishing.

OECD (2014). International distributed research infrastructues: issues and options. OECD Global Science Forum. OECD, Paris.

OECD (2015). In It Together. Why less inequality benefits all. Paris.

OECD (2015). All on Board: Making Inclusive Growth Happen. OECD Publishing, Paris.

Orcutt, G. H. (1957). A new type of socio-economic systems, *The Review of Economics and Statistics, 39*, 116-123.

Perrons, D. & Plomien, A. (2010). Why socioeconomic inequalities increase? Facts and policy responses in Europe. EU Report commissioned by DG Research. Brussels. European Commission.

Przeworski, A. (1974). Contextual models of political behavior. *Political Methodology 1*(1):27-61.

Qizilbash, M. (2003). 'Vague language and precise measurement: The case of poverty', Journal of Economic Methodology, Vol. 10, 41-58.

Renschler, I., Kleiner, B. & Wernli, B. (2013). Concepts and key features for understanding social science research infrastructure. In: B. Kleiner *et al.*, Understanding research infrastructures in the social sciences, Seismo press, Zurich, p. 11-18.

Sainsbury, D. (2012). Welfare States and Immigrant Rights: The Politics of Inclusion and Exclusion. Oxford: Oxford University Press.

Salverda, W. (2015). EU Policy Making and Growing Inequalities. European Economy discussion paper 8. European Commission. Directorate-General for Economic and Financial Affairs, Brussels.

Schäfer, A., Zentarra, A. & Groh-Samberg, O. (2015). Methodological and data infrastructure report on youth, Working paper, Leuven, InGRID project, M20.7.

Schanze, J. (2017). Report on sampling practices for the institutionalized population in social survey. SERISS. https://seriss.eu/wp-content/uploads/2017/02/SERISS_WP2_D2.16_Sampling-the-institutionalized-population.pdf.

Schepers, W., Juchtmans, G., & Nicaise, I. (2017). Reaching out hard-to-survey groups among the poor. Survey protocols, statistical issues and research design.

Steinmetz, S., Bianchi, A., Tijdens, K.G. & Biffignandi, S. (2014). Improving web survey quality – Potentials and constraints of propensity score adjustments. Callegaro, M., Baker *et al.*. (eds.) Online Panel Research: A Data Quality Perspective, pp 273-298. Series in Survey Methodology, Wiley.

Sutcliffe, S., & Court, J. (2005). Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries? ODI, Overseas Development Institute. Retrieved from http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/3683.pdf.

Sutherland, H., & Figari, F. (2013). EUROMOD: the European Union tax-benefit microsimulation model. *International journal of microsimulation,* 6(1), 4-26.

Szekér, L. & Van Gyes G. (2015). Inventory of working conditions and occupational safety and health surveys, Working paper, Leuven, InGRID project, M21.9.

Szeker, L. & Van Gyes, G. (2016). The future policy research agenda of an EU inclusive growth research infrastructure. InGRID Futuring briefing note 1. KULeuven-HIVA, Leuven.

Szekér, L., & van Gyes, G. (2015). Inventory of working conditions and occupational safety and health surveys, Milestone 21.9, H2020 InGRID-2 project.

Szekér, L., Smits, I., & Lenaerts, K. (2020). Survey report on the future needs, Leuven, InGRID-2 project 730998 – H2020.

Tam, S.M., & Clarke, F. (2015). Big Data, Official Statistics and Some Initiatives of the Australian Bureau of Statistics. *International Statistical Review* 83(3) , 436-448.

TÁRKI (2021). Codebook for extended IPOLIS: Disabled persons and migrants, Deliverable 8.4. Leuven, H2020 InGRID-2 project.

Tijdens, K.G. (2014). Reviewing the measurement and comparison of occupations across Europe. Working paper, Leuven, InGRID project, M21.2

Tijdens, K.G.; De Ruijter, E. & De Ruijter, J. (2012). Measuring work activities and skill requirements of occupations: experiences from a European pilot study with a web survey. European *Journal of Training and Development, 36*(7), 751-763.

UK Data Forum (2013). UK Strategy for data resources for social and economic research 2013-2018. UK Data Forum, Swindon. Retrieved from: http://www.esrc.ac.uk/files/research/uk-strategy-for-data-resources-for-social-and-economic-research/.

UNCTAD (2012). Macroeconomic Stability, Inclusive Growth and Employment. UN System Task Team on the post-2015 UN Development Agenda, Thematic Think Piece. Retrieved from: http://www.un.org/millenniumgoals/pdf/Think%20Pieces/12_macroeconomics.pdf.

UNESCO (Ed.). (2010). World social science report: Knowledge Divides. Unesco, Paris.

Van den Bosch, K. & Cantillon, B. (2006). Policy Impact. In M. Moran, M. Rein & R. E. Goodin (Eds.), The Oxford Handbook of Public Policy (pp. 296-318). Oxford: Oxford University Press.

Van Gyes G., Nelson K., Steinmetz S. & Hermans M. (2017). Short-term future agenda of the InGRID research infrastructure: servicing European research from data to policy on Inclusive Growth. Leuven, KULeuven-HIVA.

Vandenbroucke, F. *et al.* (2015). Unequal Europe: Recommendations for a more caring EU. Final report of the High-Level Group on 'Social Union'. Friends of Europe, Paris.

Visser, J. (1991). Trends in Union Membership. Chapter 4 in OECD Employment Outlook 1991, pp. 97-134.

Visser, J. (2006). Union Membership Statistics in 24 Countries, *Monthly Labor Review, 129*, 38-49.

Visser, J. (2019). The Position of National Union Confederations in Advanced Induszrialized Countries. Research paper submitted to the AFL-CIO. Washington DC.

Visser, J. (2021). OECD/AIAS ICTWSS Database. Detailed note on definitions, measurement and sources. Available at: https://www.oecd.org/els/emp/DetailedMethodnote-OECD-AIAS_DB.pdf.

Weber, M. (1978). Economy and Society: Au Outline of Interpretative Sociology. The University of California Press, Los Angeles.

Wernli, B., *et al.* (2013). Main components of social science research infrastructures. In: B. Kleiner *et al.*, Understanding research infrastructures in the social sciences, Seismo press, Zurich, p. 31-38.

Wissenschaftsrat (2011). Recommendations on research infrastructures in humanities and social sciences. Wissenschatsrat, Berlin. Retrieved from: http://www.wissenschaftsrat.de/download/archiv/10465-11_engl.pdf.

Zhang, L. (2012). Topics of Statistical Theory for Register-based Statistics and Data Integration. *Statistica Neerlandica,* 66, 41-63.

InGRID-2
Integrating Research Infrastructure for European expertise on Inclusive Growth from data to policy

Referring to the increasingly challenging EU2020-ambitions of Inclusive Growth, the objectives of the InGRID-2 project are to advance the integration and innovation of distributed social sciences research infrastructures (RI) on 'poverty, living conditions and social policies' as well as on 'working conditions, vulnerability and labour policies'. InGRID-2 will extend transnational on-site and virtual access, organise mutual learning and discussions of innovations, and improve data services and facilities of comparative research. The focus areas are (a) integrated and harmonised data, (b) links between policy and practice, and (c) indicator-building tools.
Lead users are social scientist involved in comparative research to provide new evidence for European policy innovations. Key science actors and their stakeholders are coupled in the consortium to provide expert services to users of comparative research infrastructures by investing in collaborative efforts to better integrate microdata, identify new ways of collecting data, establish and improve harmonised classification tools, extend available policy databases, optimise statistical quality, and set-up microsimulation environments and indicator-building tools as important means of valorisation. Helping scientists to enhance their expertise from data to policy is the advanced mission of InGRID-2. A new research portal will be the gateway to this European science infrastructure.

This project is supported by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 730998.

More detailed information is available on the website: www.inclusivegrowth.eu

Co-ordinator
Monique Ramioul

KU LEUVEN HIVA
RESEARCH INSTITUTE FOR WORK AND SOCIETY

Partners

TÁRKI Social Research Institute Inc. (HU)
Amsterdam Institute for Advanced Labour Studies – AIAS, University of Amsterdam (NL)
Swedish Institute for Social Research - SOFI, Stockholm University (SE)
Economic and Social Statistics Department, Trier University (DE)
Centre for Demographic Studies – CED, University Autonoma of Barcelona (ES)
Luxembourg Institute of Socio-Economic Research – LISER (LU)
Herman Deleeck Centre for Social Policy – CSB, University of Antwerp (BE)
Institute for Social and Economic Research - ISER, University of Essex (UK)
German Institute for Economic Research – DIW (DE)
Centre for Employment and Work Studies – CEET, National Conservatory of Arts and Crafts (FR)
Centre for European Policy Studies – CEPS (BE)
Department of Economics and Management, University of Pisa (IT)
Department of Social Statistics and Demography – SOTON, University of Southampton (UK)
Luxembourg Income Study – LIS, asbl (LU)
School of Social Sciences, University of Manchester (UK)
Central European Labour Studies Institute – CELSI (SK)
Panteion University of Social and Political Sciences (GR)
Central Institute for Labour Protection – CIOP, National Research Institute (PL)

InGRID-2

Integrating Research Infrastructure for European expertise on Inclusive Growth from data to policy Contract N° 730998

For further information about the InGRID-2 project, please contact
inclusive.growth@kuleuven.be
www.inclusivegrowth.eu
p/a HIVA – Research Institute for Work and Society
Parkstraat 47 box 5300
3000 Leuven
Belgium