

Challenges in the Harmonisation of Survey Data on Ethnic and Migrant Minorities in Europe: the EthmigSurveyData Approach



Laura Morales (EthmigSurveyData COST Action Chair, Sciences Po, CEE & LIEPP, CNRS)



InGRID2 project Expert Workshop
28th & 29th January 2021

Presentation of the International Ethnic and Immigrant Minorities' Survey Data Network (COST Action 16111)

Surveys as essential sources of information on the inclusion of ethnic and migrant minorities (EMMs)

1

Surveys provide a wealth of information about lifestyles, behaviours, attitudes, social connections, opportunities, barriers, skills, etc.

2

Unlike official statistics, they allow to connect several of these aspects at the individual level: they provide a more complete picture about mechanisms

3

When designed with a longitudinal element, they are exceptional sources of information about interconnection of factors fostering/hindering life opportunities and integration

4

They allow sophisticated analyses of the multiple factors driving the inclusion/exclusion of migrants, and multiple comparisons across various groups

Objectives of COST Action ETHMIGSURVEYDATA

- (1) Produce a comprehensive compilation of the data that is available on EMMs across Europe and beyond
- (2) Articulate the mechanisms to archive and share the data and make them available to a wider public on open data platforms
- (3) Bring together the expertise that will allow a post-harmonization process to facilitate comparative analyses of data that currently stand alone
- (4) Coordinate efforts to obtain additional research funding to extend data sharing and to identify blank spots to be investigated further
- (5) Promote further comparative analysis and the dissemination of research results on EMMs' integration
- (6) Create the research capacity to promote similar survey data production in the European countries where they still do not exist
- (7) Train the future generations of European researchers in best practices for data production, sharing and analysis in this area of research; and
- (8) Provide a platform for the coordination of future data collection to maximize comparability

Four core tasks to implement our goals



Our goals focus on the study of EMMs' economic, social and political integration with survey data, and on the production, compilation, sharing and pooling of such data

1

Compilation of an **EMM Survey Metadata Registry**: an online search tool

2

Generation of an **EMM Survey Question Bank**: via the CESSDA European Question Bank (through involvement of Sciences Po team in H2020 SSHOC project), pilots starting now

3

Undertaking of **data post-harmonization** to generate pooled datasets on various dimensions of EMM's integration from suitable individual survey datasets (pilots being designed)

4

Design and implementation of an **EMM Data Playground** to dynamically analyze and visualize a selection of data from the post-harmonized pooled datasets (worked not started)



All these tasks will converge into components of the [EMM Survey Data Hub](#)

Presentation of the Ethnic and Migrant Minorities' Survey Metadata Registry

A FAIR tool for EMM survey data

Producing the EMM Survey Registry: a collective undertaking of the ethnic and migration studies data community

The **ethnic and migration studies data community** brings together **various stakeholders** of **quantitative survey research** undertaken with **EMM populations**. For the purposes of executing its scientific work, the data community is organised into three distinct groups:



ethmig
survey
data

ETHMIGSURVEYDATA

- An international network funded by the COST Association with more than 200 EMM-focused researchers from Europe and beyond
- Provides the intellectual impetus for the scientific work undertaken by the data community



Task 9.2 team of SSHOC

- Participates in and contributes to SSHOC on behalf of the data community
- Primarily coordinates and manages the scientific work undertaken by the data community



FAIRETHMIGQUANT

- An Open Science project funded by the French Agence Nationale de la Recherche (ANR)
- Ensures the inclusion of French surveys in the scientific work undertaken by the data community



Motivation and context: *An opportunity to add value using EMM survey data*

There **remains an unused potential in EMM survey data** because these data are often **difficult to find, access, and reuse**.

Researchers continue to **face distinct challenges when surveying EMM populations**. Four challenges stand out: (1) sampling, (2) (non)response, (3) mode of interviewing and (4) trust in researchers or data-collectors.



To facilitate access to and quality of knowledge on the inclusion of EMMs in the economic, social and political domains of public life across Europe, the ethnic and migration studies data community have created the [EMM Survey Registry](#), a free online tool that follows the FAIR principles and displays survey-level metadata about existing quantitative surveys to EMMs.

Adopting the FAIR principles

The [EMM Survey Registry](#) has been designed and developed **in line with the FAIR principles**, so that a wide range of users can untap the potential of quantitative surveys undertaken with EMM (sub)populations.



Findable

- Easier to locate surveys, as the EMM Survey Registry is like a 'census' of existing surveys
- Easy to navigate information (i.e. metadata) about each survey due to the user-friendly interfaces (front and back-end) and the systematic documentation of the information (via a metadata schema)



Accessible

- Metadata about each survey is made publicly available online
- Metadata about each survey specifies how to access the data set(s), technical documentation, questionnaire(s), and other relevant publications



Interoperable

- Metadata is set up so that they can be easily shared with humans
- Metadata is also set up so that they can easily be shared with machines (primarily other social sciences data archives and repositories)



Reusable

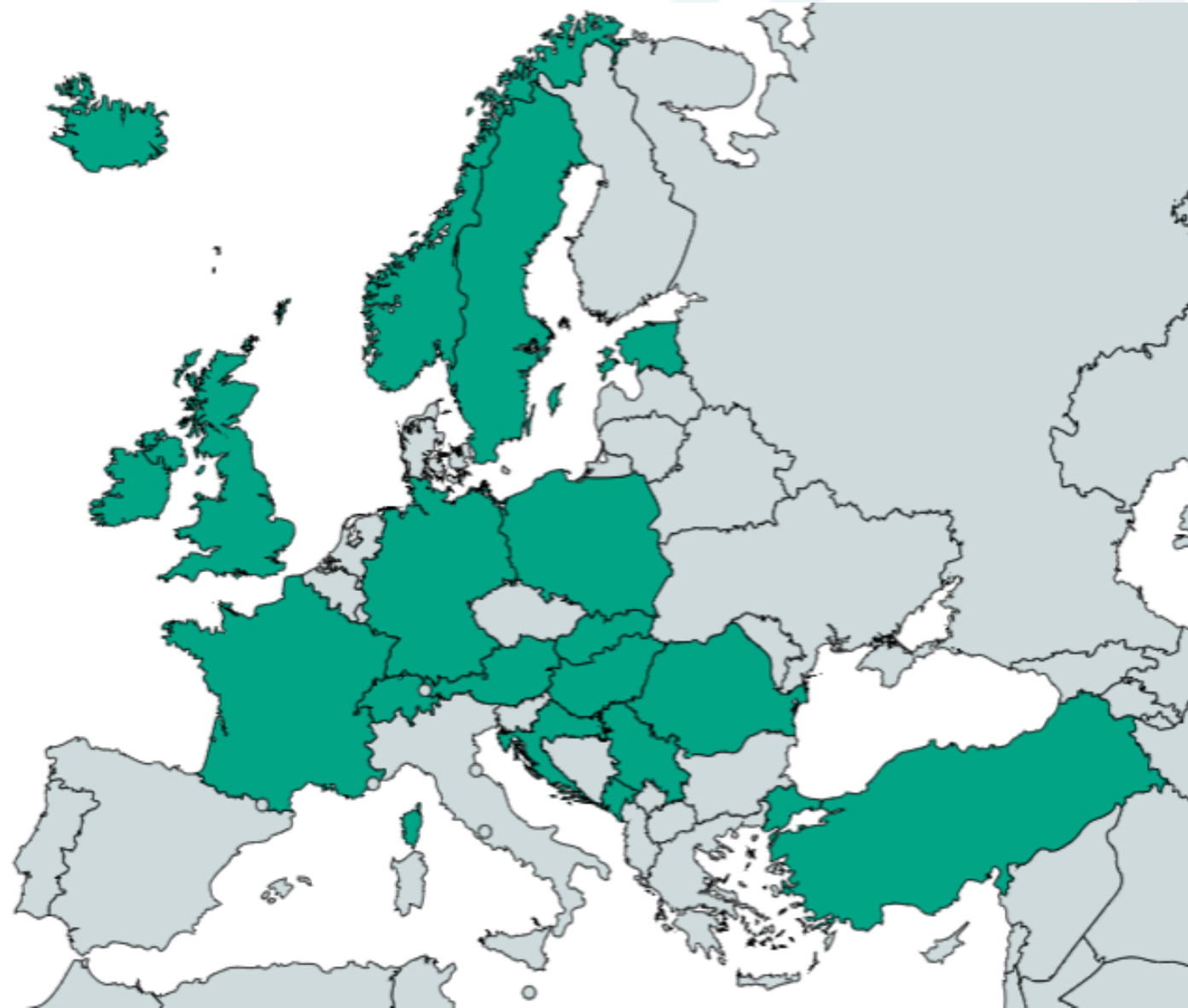
- Metadata is detailed, informative, organised, and structured; metadata itself can be used for research
- Reuse of the survey is promoted/facilitated since the metadata includes information about how to access the relevant original sources (e.g. data set(s), technical documentation)

Compiled metadata

The [EMM Survey Registry](#) currently captures **600+** surveys from **18 different countries**:

- **Austria** (n=69)
- **Croatia** (n=38)
- **Estonia** (n=12)
- **France** (n=12, more uploaded soon)
- **Germany** (n=78)
- **Hungary** (n=33)
- **Ireland** (n=21)
- **Iceland** (n=7)
- **Montenegro** (n=10)
- **Norway** (n=16)
- **Poland** (n=5)
- **Romania** (n=9)
- **Serbia** (n=12)
- **Slovakia** (n=11)
- **Sweden** (n=45)
- **Switzerland** (n=10)
- **Turkey** (n=32)
- **United Kingdom** (n=179)

More countries will be added in the next 1-2 months: B&H, Bulgaria, Cyprus, Finland, Italy, Luxembourg, and Spain.
Around 30 countries eventually.



The front-end of the EMM Survey Registry: The key functionalities

EMM Survey Registry

The EMM (Ethnic and Migrant Minorities) Survey Registry is a database of quantitative surveys that have been undertaken with EMM (sub)samples across Europe and beyond. Survey-level metadata is available for each of the surveys included in the EMM Survey Registry. Explore and learn about the different surveys and find specific surveys using the search, filtering (simple and advanced) and sorting features. Once the desired filtering, sorting and/or search parameters have been identified/selected, the list of surveys will be automatically updated.

Access to advanced filtering

Simple filtering (Advanced)

1.0. COUNTRY

UNITED KINGDOM

1.5. SCOPE OF SURVEY

National (48)

Subnational (113)

1.9. REPRESENTATIVE OF THE POPULATION

Yes (47)

No (1)

1.10. TYPE OF SURVEY

Single cross-section (4)

Repeated cross-section (multiple waves with different samples) (25)

Longitudinal/panel survey (multiple waves with the same or partially overlapping samples) (19)

Simple filtering

Free text search: search for country, keyword, institution, scope, etc.

Showing: 1 - 25 of 48 search results

Sort by: [country](#) | [scope](#) | [region](#) | [start date](#) | [end date](#) | [EMM target population](#) | [sample size](#)

[Clear all selections](#) [Advanced filtering](#)

1.5. Scope of survey: National

1.0. Country: United Kingdom

Access to advanced filtering

United Kingdom

Understanding Society: Wave 1

Understanding Society: Wave 1

Access to full record about a survey

Scope of survey: National

Name of region(s) Eng.:

If subnational, type of subnational area: -

Starting date of survey: 2009-01-08

End date of survey: 2011-03-07

Survey in development/not yet completed: No

Date when the survey first became a part of an international survey programme: Not applicable

EMM target population: All residents of ethnic minority identification in the city/region/country

Survey designed as a general population survey: Yes

Total net/achieved sample: 33178

Language(s) of technical survey documentation in ISO code: ENG

Institution/team responsible for data production: Institute for Social and Economic Research (ISER), University of Essex

Institution/team that owns the data: Economic and Social Research Council (ESRC)

Help

Keyword search

Sorting

Filter summary

XML

Abbreviated record about a survey





Demo of the front-end of the EMM Survey Registry:
<https://ethmigsurveydatahub.eu/emmregistry/>

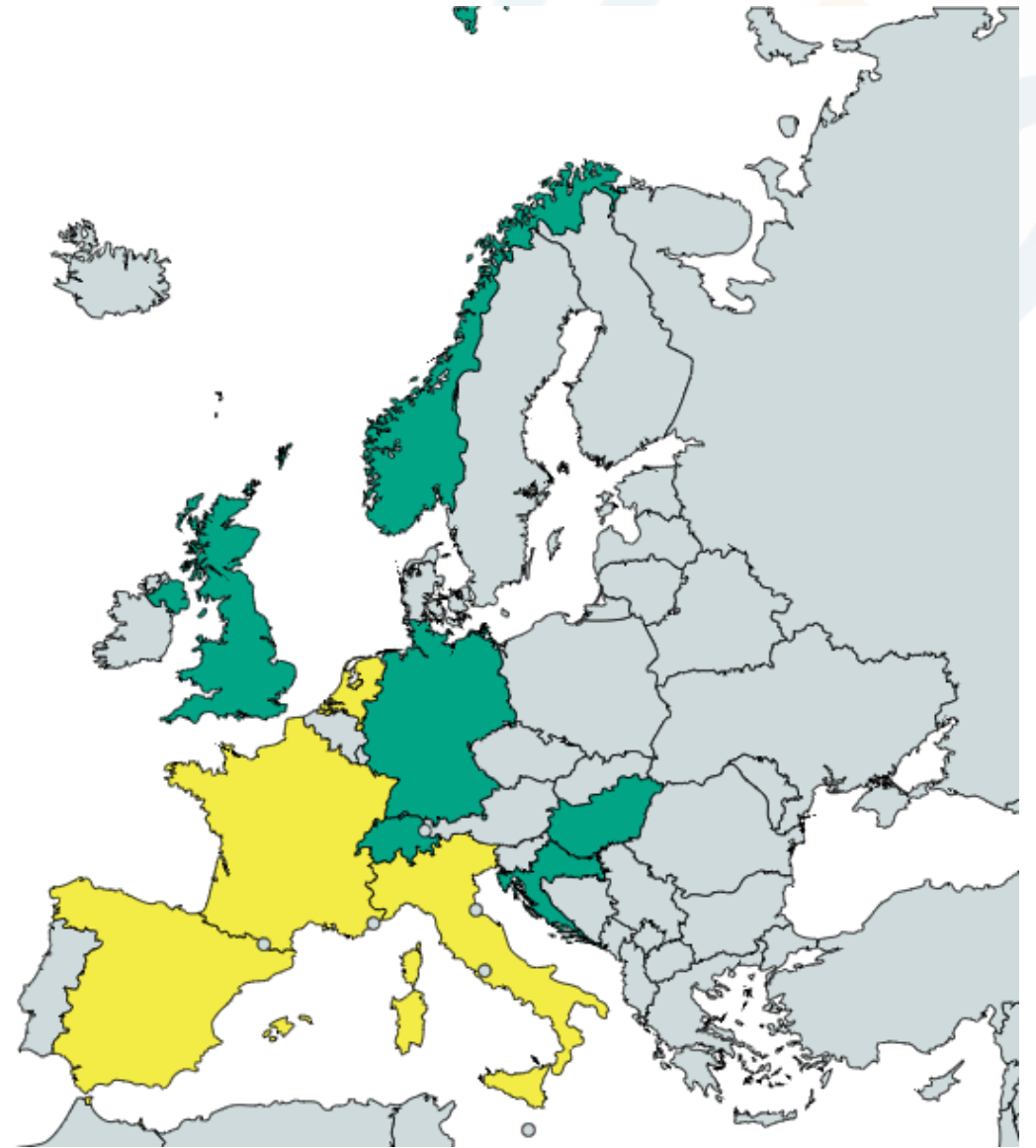


Presentation of the EMM Question Data Bank (QDB) pilots

A necessary step prior to post-harmonisation

Developing the QDB pilots:

- The QDB pilots **draw on the survey metadata** compiled first through the EMM Survey Registry
- The QDB pilots will only include **national and subnational** surveys from **10 of the 35 COST Action participating countries**.
- **6 of the 10 countries** are **already included** in the pilots (in **green**); the **remaining 4 countries** (in **yellow**) will **soon join the pilots** once their survey metadata files are validated and uploaded to the [EMM Survey Registry](#).



Developing the QDB pilots:

The QDB pilots will focus only on **three specific topics***:



Pilot 1

Civic life and political participation



Pilot 2

Stereotypes, prejudices and discrimination (intergroup relations/beliefs)



Pilot 3

Belonging and social identity, cultural and social norms

*The topics all correspond to specific variables from the [EMM Survey Registry metadata schema](#).

Launching the QDB pilots:

Initial stages implemented for Pilot 1 on Civic Life and Political Participation

1

Setting the **inclusion/exclusion criteria** for the questionnaires

2

Defining the **metadata schema** (based on the EQB metadata schema)

3

Selecting the CNRS's [MyCore](#) as the tool/platform for **storing** and **managing** the questionnaires

4

Establishing a detailed **workflow** for **collecting** the questionnaires

Implementing Pilot 1: *Collecting the questionnaires (step 1)*

The **first step** is to **collect** the **questionnaires** for **pilot 1** using the detailed workflow. To date questionnaires for **136 surveys** have been collected.



Identify all the questionnaires meeting the inclusion criteria by documenting them on an [online tracker](#)



Locate and download each questionnaire meeting the inclusion criteria



Upload and store the retrieved questionnaires to [MyCore](#)

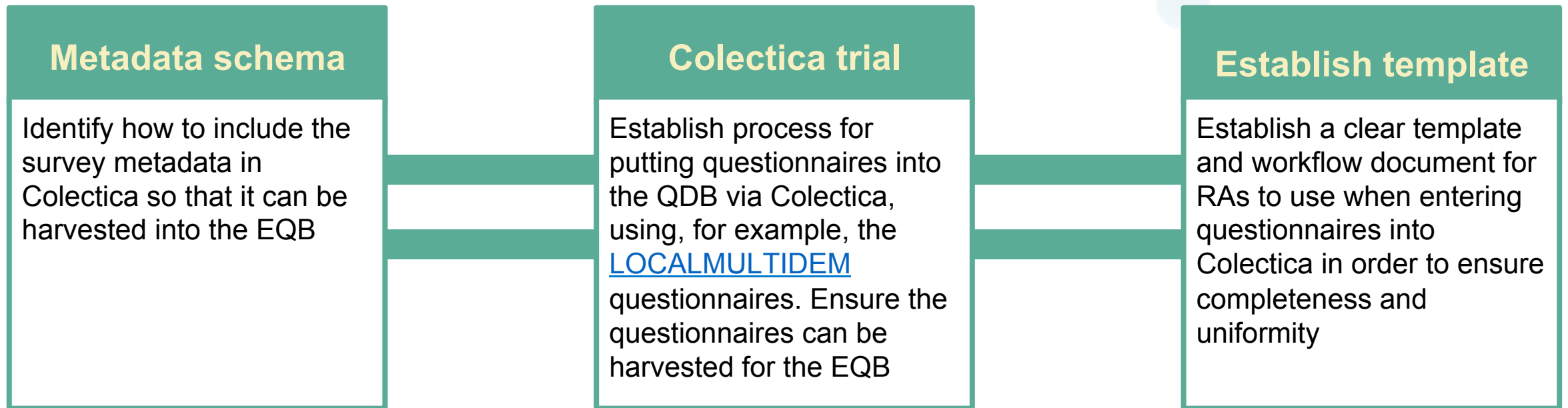


Record on the [online tracker](#) when and how the questionnaire was collected



Implementing Pilot 1: *Editing q'aire metadata on Colectica (step 2)*

The **second step** is to **use** Colectica, a software for **recording questionnaires** and their **metadata**. This step is currently **underway**, in parallel with step 1.



Implementing Pilot 1: *Next steps*

1

Link individual questions in questionnaires of Pilot 1 to hierarchy of concepts identified by WG3

2

Recruiting RAs to add questionnaires to the system

3

Integrate the questionnaires into the EQB

Linking questionnaire items to concepts

Another necessary step prior to post-harmonisation

Preparatory work on dimensions of integration and concept hierarchy:

In order to post-harmonise survey items, we need to work with a hierarchy of concepts: we first link integration dimensions to the topics covered by the survey questionnaire as revealed by the EMM Survey Registry survey-level metadata

Integration Dimensions	Topic(s) Covered by the Survey Identified by EMM Survey Registry
1. Labour Market	1.13.15.Labour market integration
2. Health	1.13.10.Health/health access
3. Education	1.13.6.Educational attainment/trajectory, human capital, skills
4. Housing/Residential Profile	1.13.9.Housing/housing access 1.13.25.Space use, spatial consequences
5. Income and Poverty	1.13.13.Income-related (and/or poverty)
6. Leisure	1.13.3.Consumption and/or leisure 1.13.26.Time use
7. Language	1.13.16.Language skills/training
8. Religion	1.13.29.Religion
9. Cultural and Social Norms (Pilot 3bis)	1.13.8.Gender relations, gender identity, sexuality 1.13.20.Political inclusion and participation, and social/political attitudes

Preparatory work on dimensions of integration and concept hierarchy:

Integration Dimensions	Topic(s) Covered by the Survey Identified by EMM Survey Registry
10. Legal Status and Rights	1.13.1.Asylum seekers and refugee issues 1.13.2.Citizenship and naturalization 1.13.11.Human smuggling and trafficking 1.13.18.Legal status/administrative situation
11. Belonging and Social Identity (Pilot 3)	1.13.12.Identity (ethnic, national, racial, religious) and belonging 1.13.20.Political inclusion and participation, and social/political attitudes
12. Civic Life and Political Participation (Pilot 1)	1.13.24.Social cohesion and/or civic engagement and/or networks 1.13.27.Transnational patterns (e.g. remittances, travel, engagement with 'home' country politics, etc.) and diasporas
13. Stereotypes, Prejudices and Discrimination (Intergroup Relations/Beliefs) (Pilot 2)	1.13.5.Discrimination, racism and/or xenophobia 1.13.14.Interethnic contact & conflict 1.13.21.Public attitudes about migration and migrants
14. Demographic flows, patterns, characteristics and behaviours	1.13.4.Demographic characteristics/behaviours 1.13.7.Family reunification, marriage, family relations 1.13.19.Migration trajectory (past/future) 1.13.22. Reasons for migration/migration drivers 1.13.23.Return migration

Preparatory work on dimensions of integration and concept hierarchy:

From dimensions of integration to building a hierarchy of sub-dimensions, concepts and sub-concepts: an illustration

Dimension	Sub-dimensions of integration			Broader concept (concept class)	Narrower sub-concept (for types of indicators)	Demographic-related considerations	Relevant comments	Example source question
	Context	Input	Outcome					
Identity		x	x	Self-identification	Self identification with ethnic group	May be asked to only specific groups or to general populations	Integration input or output depending on whether question asked before/shortly after migration (for 1st generation migrants) or later on (for 1st generation).	Localmultidem, Q13
Identity		x	x	Self-identification	Self identification with people from same religion	May be asked to only specific groups or to general populations	Integration input or output depending on whether question asked before/shortly after migration (for 1st generation migrants) or later on (for 1st generation).	Localmultidem, Q13
Identity	x		x	Self-identification	Self identification with (HOST COUNTRY) people	May be asked to only specific groups or to general populations	Integration context when asked of general population, integration output when asked to EMMs	Localmultidem, Q13
Identity			x	Self-identification	Self identification with people in neighbourhood			Localmultidem, Q13
Identity			x	Self-identification	Self identification with people with the same sex/gender			Localmultidem, Q13
Identity			x	Self-identification	Self identification with (HOST REGION) people			Localmultidem, Q13
Identity			x	Self-identification	Self identification with (CITY)			Localmultidem, Q13
Identity			x	Self-identification	Self identification with people of the same age			Localmultidem, Q13
Identity			x	Self-identification	Self identification with a particular social class			Localmultidem, Q13
Identity			x	Self-identification	Self identification with European people)			Localmultidem, Q13
Identity			x	Self-identification	Self identification with the world, humanity as a whole?			Localmultidem, Q13
Identity			x	Self-identification	Self identification with (HOME COUNTRY)			Localmultidem, Q13
Social attitudes	x		x	Trust	Trust in people (GENERAL)		Integration context when asked of general population, integration output when asked to EMMs	Localmultidem Q34
Social attitudes			x	Trust	Trust in [RESPONDENT 's ETHNIC GROUP] people			Localmultidem Q35
Political attitudes	x		x	Political confidence	Trust in different institutions		Integration context when asked of general population, integration output when asked to EMMs	Localmultidem Q36

From questionnaire items & concepts to post-harmonised datasets

Our plans for post-harmonisation

Also through thematic pilots: *Starting with Civic Life & Participation*

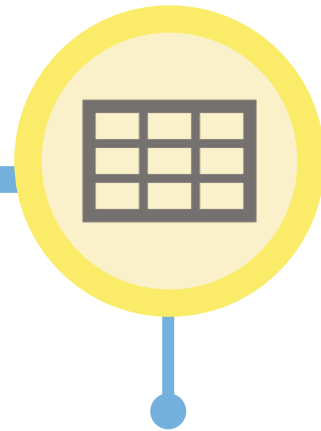
The **first step** is to **collect** the **questionnaires** for **pilot 1** using the detailed workflow. To date questionnaires for **136 surveys** have been collected.



We will select only surveys for which the datasets are freely available within the EMM Survey Registry collection



Using the EMM QDB online tool, we will identify the main concepts that can be post-harmonised across comparable surveys



Using the [GESIS Charmstats software](#), we will post-harmonise to common comparable categories



We will produce post-harmonised sub-sets of comparable cross-national surveys (e.g. local level vs regional level vs national level)



Ways to Collaborate with EthmigSurveyData

Are you (or your organisation) working with (or producing) surveys to EMM populations?

You can add the metadata to the EMM Survey Registry

You can contribute to one of the 3 QDB pilots or lead a new (topical) one if you have human resources to contribute

You can contribute to the conceptual hierarchy work

Any questions or suggestions?

Thank you for your attention!

Questions? Comments?

Join our community



<https://ethmigsurveydatahub.eu>



<https://ethmigsurveydatahub.eu/emmregistry/>



sshoc.project@sciencespo.fr
ethmigsurveydata@sciencespo.fr



<https://zenodo.org/communities/ethmigsurveydata/>

