The Ethnic and Migrant Minorities' (EMMs') Survey Registry: A tool to help make EMM survey data FAIR

The Social Sciences and Humanities Open Cloud adopts the RDA Metadata Standards Directory Working Group recommendation

Ethnic and Migration Studies is a component of the Social Sciences and Humanities Open Cloud (SSHOC) H2020 project. It is dedicated to improving the access, (re)use, and dissemination of quantitative survey data on ethnic and migrant minorities' (EMMs') integration. The work is led by Prof Laura Morales of Sciences Po and is supported by a small research team, comprised of 1 FTE with Ami Saji and short-term research assistants. Given the focus of the Ethnic and Migration Studies task, it works closely with the COST Action 16111 - Ethmigsurveydata, a research network launched in 2017 that brings together EMM-focused researchers from all sectors (academic, think tanks, government, civil society organizations, private companies), as well as the ANR flash-funded project, FAIRETHMIGQUANT.

When the SSHOC project concludes in 2022, it will have:

1. developed the EMM Survey Registry, a free online tool that will display survey-level metadata from over 800 existing quantitative surveys on EMMs' integration from 30+ countries (primarily in Europe),
2. implemented a rigorous pilot to test the feasibility of setting up, as part of the CESSDA-led Euro Question Bank, a component dedicated to questionnaires from quantitative EMM surveys.

The challenge

With persistent demands to understand and study the integration of EMMs through data-driven approaches, there has been an abundance of quantitative survey research being conducted and generated by the ethnic and migration studies research community. Unfortunately, much of this data is still fragmented and hidden, making it difficult to leverage and realise their full potential in stimulating within and cross-disciplinary collaboration and innovation. The SSHOC Ethnic and Migration Studies team, the Ethmigsurveydata COST Action and the FAIRETHMIGQUANT project are therefore developing the EMM Survey Registry as a means to tackle this very scientific challenge. Specifically, as the EMM Survey Registry will serve as a "one-stop shop" to learn about and explore existing quantitative surveys to EMMs (via the compiled survey-level metadata), it will facilitate access to and quality of knowledge on the inclusion and integration situation of EMMs in the various economic, social and political domains of public life across Europe and its neighbouring countries.

The RDA outputs adopted

For our work, we consulted the recommendations developed by the Metadata Standards Directory Working Group. Below we have outlined the specific way in which we adopted this working groups recommendation:

The EMM Survey Registry (currently in beta version) brings together previously scattered information about quantitative EMM surveys to a single access point. In order to offer the compiled information in a coherent and systematic fashion, the EMM Survey Registry has developed a rich metadata schema (tested and reviewed rigorously by EMM-targeted researchers) that will be used to describe each of the identified surveys. Moreover, this metadata schema has been designed to be compatible with DDI Codebook, so that survey-level metadata can be harvested, re-used, interoperable, etc. This decision to adopt DDI for the EMM Survey Registry metadata schema stems from the Metadata Standards Directory Working Group Recommendation of using the established metadata standard for your respective discipline (in this case, DDI since we are working with survey data from the social sciences). providing Scholix-compliant information, without having to do additional conversion/workflow tasks which can have time and therefore cost associated with them.

The easier it is for publishers to share this data, the more these links can be used to shore up connections between articles and data, help with reproducibility and reporting, and ensure that researchers get credit for sharing their data.

Find out more at:

www.rd-alliance.org/recommendations-outputs

Visit

rd-alliance.org   enquiries@rd-alliance.org

or write us at

Benefits of adoption and impact

The EMM Survey Registry has been designed to be a free online tool that will provide meaningful and detailed information about existing quantitative surveys on EMMs' integration through a user-friendly interface. Moreover, it has been conceptualized to conform with a metadata standard (i.e. DDI) that is widely used within the social sciences and survey research community. For these reasons, the EMM Survey Registry will not only offer improved access to and (re)use of quantitative survey data on EMMs' integration, but also provide a concrete and cogent example of how and why to make data from a multi-disciplinary field FAIR (findable, accessible, interoperable, reusable).

The adoption process

The EMM Survey Registry is currently developing an API that will offer the DDI Codebook compliant survey-level metadata. To date, the SSHOC Ethnic and Migration Studies team (with Ethmigsurveydata and FAIRETHMIGQUANT) has completed the following steps (in terms of setting up the aforementioned API):

1. defining a metadata schema for the EMM Survey Registry that will provide useful and meaningful information to users of quantitative EMM surveys,

2. collaborating with DDI experts (i.e. the CDSP at Sciences Po) to match metadata variables of the EMM Survey Registry to DDI Codebook components, and

3. setting up a mechanism on the EMM Survey Registry that allows a DDI Codebook compliant XML file to be requested and generated for each survey already being described and displayed on the EMM Survey Registry. In the coming months, the SSHOC Ethnic and Migration Studies team will work with DDI and data science experts (primarily from the SSHOC project community) to transform the current mechanism that produces DDI Codebook compliant XML files into an API that conforms with OAI-PMH. Moreover, this API will be made available on the EMM Survey Registry site.

Lessons learned

The SSHOC Ethnic and Migration Studies team and the researchers linked to Ethmigsurveydata and FAIRETHMIGQUANT had limited prior experience and engagement with survey metadata documentation. While DDI and RDA provide a wide-range of free resources to learn about metadata standards for documenting research data, we sometimes had a difficult time understanding how to actually adopt and use the recommended metadata standard for our discipline, i.e. DDI. As such, for researchers that are fledglings or new to metadata standards, it might be helpful to provide more opportunities to obtain tailored and targeted support (e.g. coaching, training).
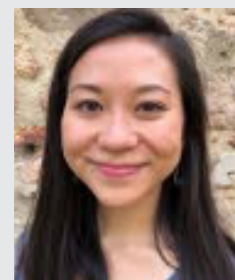
About SSHOC

Social Sciences & Humanities Open Cloud (SSHOC) is a project funded by the EU framework programme Horizon 2020 and unites 20 partner organisations and their 27 associates in developing the social sciences and humanities area of the European Open Science Cloud (EOSC). SSHOC partners include both developing and fully established European Research Infrastructures from the social sciences and humanities, and the association of European research libraries (LIBER). Between them, they have expertise across the entire data cycle from creation and curation to optimal re-use, training and advocacy.

During the 40-month lifespan of the project, (from January 2019 to April 2022) SSHOC will transform the current social sciences & humanities data landscape with its disciplinary silos and separate facilities into an integrated, cloud-based network of interconnected data infrastructures.

Contacts:
Laura Morales - SciencesPo
Ethnic and Migration Studies SSHOC task lead
Ami Saji - SciencesPo
Ethnic and Migration Studies Researcher