

Research and Innovation Action

Social Sciences & Humanities Open Cloud

Project Number: 823782

Start Date of Project: 01/01/2019

Duration: 40 months

Deliverable 9.2 Midterm evaluation report

| | |
|-------------------------|--|
| Dissemination Level | PU |
| Due Date of Deliverable | 30/06/21, M31 |
| Actual Submission Date | 19/08/21 |
| Work Package | WP9 - Data Communities |
| Task | Task 9.1 Identifying shared and unique challenges for SSH data communities |
| Type | Report |
| Approval Status | Waiting EC approval |
| Version | V1.0 |
| Number of Pages | p.1 – p.18 |

Abstract:

This report provides an overview of services, tools, procedures, and data originating from WP9 (*Data Communities*), and of the extent to which they have become available in the data communities for which they were primarily intended. The report is linked to SSHOC D9.3, which will report in detail on the reception and usage of these products in the various user communities of WP9.

The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/ her sole risk and liability. This deliverable is licensed under a Creative Commons Attribution 4.0 International License.



History

| Version | Date | Reason | Revised by |
|---------|------------|-------------------------------------|-------------------|
| 0.0 | 07/07/2021 | Structure and Initial content | Cees van der Eijk |
| 0.1 | 13/07/2021 | Draft content for section 2 | Ami Saji |
| 0.2 | 29/07/2021 | Draft content for Section 4 | Carmen Di Meo |
| 1.0 | 12/07/2021 | Finalisation of content and editing | Cees van der Eijk |

Author List

| Organisation | Name | Contact Information |
|--------------|-------------------------|----------------------------------|
| UNOTT | Van der Eijk, Cees | cees.vandereijk@nottingham.ac.uk |
| Sciences Po | Ami Saji | amikatherine.saji@sciencespo.fr |
| CNR | Carmen Di meo | dimeo@ovi.cnr.it |
| CNR | Emiliano Degl'Innocenti | emiliano.deglinnocenti@cnr.it |

Executive Summary

Three of the Tasks in WP9 (Data Communities) are associated with particular user communities. These are Ethnic and Migration Studies (associated with T9.2), Electoral Studies (associated with T9.3), and Heritage Science (associated with T9.4). Each of these three tasks has as a major part of its remit the production of services, tools, procedures and data of direct relevance for their respective user communities. This report provides an overview of the state of development of these products, and of the extent to which they are available to the user communities in question. A detailed reporting of the extent of usage in these communities, and of the feedback thereby engendered will be presented in the report of Deliverable D9.3, at the end of 2021.

This report describes for each of these tasks the major activities undertaken in the construction of their respective services, tools, procedures and data, and the current state of development for each of these.

The major product that has been generated by Task T9.2 (Ethnic and Migration Studies) is the EMM Survey Registry. This Registry is by now fully developed in terms of its structure and functionality, and the currently available version is a publicly accessible beta version. It is also very well developed in its coverage of surveys, with information from more than 1200 surveys from almost 30 countries having been ingested, with many more being added in the near future. It has already been used by 300-400 unique researchers from different sectors (academe, industry, administration) and countries.

The second major activity of T9.2 consists of a feasibility study of developing an EMM Question Data Bank (QDB) as part of the CESSDA-led European Question Bank. It is thus quite possible that such an EMM-QDB will eventually be developed into a product available to the user community, but that is not a commitment in the current SSHOC project.

As far as T9.3 (Electoral Studies) is concerned, the main product to be developed is a Knowledge Graph (KG) in Electoral Studies. This product has been developed into an alpha version that is based on a large volume of ingested material (almost datasets and almost 4000 publications). The structure and functionality of the KG as specified in the design stage is operational, but subject to further improvement and refinement. A first stage of testing has been concluded with a small panel of expert users, which flagged up various issues to be implemented in an enhanced way before further testing and wider access in the user community can productively be undertaken. The implementation of these enhancements is currently under way.

For T9.4 (Heritage Science) the main product to be developed is a digital platform –named RESTORE– that allows the recovery, integration and reuse of existing digital materials of different kinds (e.g., archival, museal, textual, artefactual) and stored using a variety of standards. This platform is to be piloted in the context of a project centered on the 14th/15th century merchant Datini. The platform has been fully specified and developed into an alpha version.

Abbreviations and Acronyms

| | |
|-------------|---|
| ANR | Agence Nationale de la Recherche (France) |
| AUSSDA | Austrian Social Science Data Archive |
| CESSDA | Consortium of European Social Science Data Archives |
| CIDOC - crm | International Committee for Documentation - Conceptual Reference Model |
| CNR | Consiglio Nazionale delle Ricerche (Italy) |
| EMM | Ethnic and Migrant Minority |
| EOSC | European Open Science Cloud |
| EQB | European Question Bank |
| FAIR | Findability, Accessibility, Interoperability, and Reusability |
| GOGS | Name of an Open Source Distributed Version Control System |
| MVP | Minimum Viable Platform |
| SPARQL | standard query language and protocol for Linked Open Data and RDF databases |
| SSHOC | Social Sciences & Humanities Open Cloud |
| SSHOCro | SSHOC Reference Ontology |
| SWC | Semantic Web Company |

Table of Contents

| | |
|---|----|
| Introduction and purpose | 6 |
| T9.2: Ethnic and Migration Studies | 7 |
| Brief description of the Task and the user community | 7 |
| Major activities in T9.2 and production of services, tools, procedures and data | 7 |
| Development of the EMM Survey Registry | 7 |
| Feasibility testing of an EMM Question Data Bank (QDB) | 10 |
| T9.3: Electoral Studies | 10 |
| Brief description of the task and the user community | 10 |
| Major activities in T9.3 and production of services, tools, procedures and data | 11 |
| T9.4: Heritage Science | 12 |
| Brief description of the task and the user community | 12 |
| Dissemination Activities | 14 |
| EURHISFIRM and RESILIENCE | 15 |
| Conclusion | 16 |
| References | 17 |

1. Introduction and purpose

This report constitutes Deliverable D9.2, known in the Consortium Agreement as *Midterm Evaluation Report*. It centres on the services, tools, procedures and data originating from WP9 (*Data Communities*). Its purpose is, on the one hand, to provide an overview of these products, and on the other hand, to consider their reception and usage in the various user communities that are represented in WP9. In these respects, the aims of this Deliverable overlap to a considerable extent with those of Deliverable D9.3, which is referred to in the Consortium Agreement as a *Usability Evaluation Report*. To avoid wholesale duplication of the same material of these two Deliverables, it was agreed to consciously link them in such a way that D9.2 will function as an explicit precursor to D9.3. This implies that D9.2 focusses particularly on the description of the products emanating from WP9, while D9.3 will emphasise in as much detail as possible their usage in the various user communities and the feedback deriving from that usage. This will not fully eliminate any overlap in content between Deliverables D9.2 and D9.3, but it will keep it limited as much as possible.

At the start of the SSHOC project WP9 represents the following user communities:

- Ethic and Migration Studies (constituting within WP9 Task T9.2)
- Electoral Studies (constituting T9.3 in WP9)
- Heritage Science (constituting T9.4 in WP9)

At later moments, two other user communities have become part of WP9, each of which exists in the form of a separate Consortium that represents specific user communities:

- Since 2020, the *Eurhisfirm* Consortium, which represents the user community of economic historians
- Since 2021 the *Resilience* Consortium, which represents the user community of religion studies.

These two additional user communities are included in WP9 without having been tasked with separate responsibilities to produce services, tools, procedures and data in the context of WP9 and SSHOC. While they do make important contributions in terms of such products, this is defined and structured within their own Consortia, and reported in those contexts. In view of this, the current report focusses mainly on what has been produced in Tasks T9.2, T9.3 and T9.4 for their respective user communities.

The structure of this report reflects the qualitatively distinct character of what has been developed in the various tasks of WP9: each is reported upon in a separate section. Thus Section 2 reports on the work by and for Ethnic and Migration Studies, Section 3 does so for Electoral Studies, and Section 4 does so for Heritage Science. Section 5 comments on some of the work conducted in Eurhisfirm and Resilience. Section 6 concludes and links up to SSHOC Deliverable D9.3.

2. T9.2: Ethnic and Migration Studies

2.1 Brief description of the Task and the user community

As part of Work Package 9 Data Communities of the SSHOC project, Task 9.2 Ethnic and Migration Studies brings together the data community focused on ethnic and migration studies. Its overarching objective is to make quantitative survey data on ethnic and migrant minorities (EMMs) more accessible and (re)usable to a wide range of users (e.g. (non)academic researchers, policy makers) in Europe and beyond. As a means to achieve this objective, the Task 9.2 team has planned for two tangible outputs: first, launching a free online database and tool that will be FAIR (Findable, Accessible, Interoperable, Reusable)-compliant and will display compiled survey-level metadata for over 800 quantitative EMM surveys from over 30 different European countries, i.e. the **EMM Survey Registry**; and second, testing the feasibility of developing an **EMM component as part of the CESSDA-led European Question Bank (EQB)**, covering the EMM surveys identified via the development of the EMM Survey Registry.

The SSHOC project is thus providing the funding and technical support needed by the Task 9.2 team to successfully deliver these two outputs. Moreover, as Task 9.2 pertains to the data community for ethnic and migration studies, it involves the participation and support of two different initiatives also invested in making EMM survey data FAIR: COST Action 16111 – ETHMIGSURVEYDATA and FAIRETHMIGQUANT. The former is an international network funded by the COST Association with more than 200 EMM-focused researchers primarily based across 35 countries in Europe and neighbouring countries and it provides the intellectual impetus for the two outputs. The latter is an Open Science project funded by the French Agence Nationale de la Recherche (ANR) and is responsible for ensuring the inclusion of the French EMM surveys in the two outputs, as well as promoting the FAIR principles with the ethnic and migration studies data community in France.

2.2 Major activities in T9.2 and production of services, tools, procedures and data

2.2.1 Development of the EMM Survey Registry

The EMM Survey Registry—which has been co-created by COST Action 16111 - ETHMIGSURVEYDATA, the H2020 SSHOC project, and the ANR project FAIRETHMIGQUANT—is a free online tool for discovering and learning about existing quantitative surveys undertaken with EMM (sub) populations in Europe and beyond. This tool not only brings together previously scattered information about these surveys but also offers detailed and structured metadata for each of the surveys. Designed with the end-users in mind,

this tool can also easily be leveraged by anyone interested in doing research or policy-related work using quantitative surveys on EMMs' integration and inclusion.

The EMM Survey Registry is currently live and available in beta version. This version is fully functional and displays metadata for over 1200 surveys from nearly 30 different countries. The tool has already welcomed 300-400 unique users each month from all over the globe and from different disciplines, sectors, industries, etc. It aspires to showcase survey research on EMMs in 30 European and neighbouring countries. In the coming months, an additional 100-200 surveys will be added to the EMM Survey Registry covering additional countries. Survey producers will also be invited to contribute their own metadata (for free but with quality supervision) to help expand the metadata collection of the registry, as well as share their research with an international audience.

The fully operational beta version of the EMM Survey Registry can be accessed online¹.

More specifically, this development included the following aspects:

- Reporting on the EMM Survey Registry in the form of SSHOC Deliverable D9.4², submitted to and approved by the EC
- Inclusion in the SSHOC Service Catalogue³; NOTE: The “Type” and “Property” have been wrongly identified and will be updated in the autumn of 2021 to “Tools & Services” and “Sharing & Discovery” respectively
- In the pipeline: the integration of the EMM Survey Registry in the SSH Marketplace⁴
 - NOTE 1: This link provides a mock-up of the entry available in the development environment
 - NOTE 2: this mock-up contains 3 additional entries for the SSH Marketplace linked to the Registry. These additional entries will also be included in the SSH Marketplace as companions to the service/tool-specific entry
- The team applied to an Open Call for Thematic Service Providers by EOSC-Pillar⁵; application was accepted and EOSC-Pillar has offered to provide us in-kind support in 2 ways: advising on how to properly on-board the EMM Survey Registry to EOSC and referring relevant contacts within France (namely Huma-Num) who can help us explore options for transferring the hosting services of our EMM Survey Registry from a private to institutional provider. This in-kind support from EOSC-Pillar is currently being formalized and we anticipate starting this collaboration with EOSC in early fall 2021

¹ EMM Survey Registry: <https://ethmigsurveydatahub.eu/emmregistry/> (accessed Aug 2021).

² Ami Saji, & Laura Morales. (2020a). D9.4 Database with the metadata of surveys to EMMs across Europe (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.4558307> (accessed Aug 2021).

³ EMM Survey Registry in SSHOC Service Catalogue: <https://www.sshopencloud.eu/emm-survey-registry> (accessed Aug 2021).

⁴ SSHOC Marketplace: <https://sshoc-marketplace.acdh-dev.oeaw.ac.at/tool-or-service/h8UTUS> (accessed Aug 2021).

⁵ Link to the public call: <https://www.eosc-pillar.eu/open-call-thematic-service-providers> (accessed Aug 2021).

- The team is actively exploring options for developing an OAI-PMH compliant API for the EMM Survey Registry, so that its metadata can be made interoperable with a variety of social sciences platforms/archives
- The team is currently working to formalize support to be provided to prospective users
- The team conducted SSHOC webinar 1 - 26.10.2020⁶
- The team conducted SSHOC webinar 2 - 20.09.2021⁷
- A variety of additional important dissemination activities related to the EMM Survey Registry (presentations, social media and online outreach, publications) was implemented⁸
- List of relevant materials related to the development/management of the EMM Survey Registry funded through SSHOC and with the support of the data community:
 - Report (v1) analysing the metadata from the EMM Survey Registry⁹. NOTE: v2 is currently in the works and will be released sometime in fall 2021.
 - Quality check workflow document¹⁰.
 - Conditions of use of the EMM Survey Registry and its metadata¹¹.
 - User guide for contributing metadata directly to the EMM Survey Registry¹².
 - Video tutorials about how to contribute metadata directly to the EMM Survey Registry and Youtube channel playlist with important recordings/videos about the EMM Survey Registry¹³.
 - COVID-19 collection announcement, including links to ENG/FRA/SPA infosheets¹⁴.

⁶ SSHOC Webinar on Introducing the newly launched Ethnic and Migrant Minorities (EMM) Survey Registry: <https://www.sshopencloud.eu/events/sshoc-webinar-introducing-newly-launched-ethnic-and-migrant-minorities-emm-survey-registry> (accessed Aug 2021).

⁷ SSHOC Webinar on Showcasing a survey for free on the EMM Survey Registry:

<https://www.sshopencloud.eu/events/sshoc-webinar-showcase-your-survey-free-emm-survey-registry> (accessed Aug 2021).

⁸ see Saji Ami & Morales Laura. (2020a). D9.4 Database with the metadata of surveys to EMMs across Europe (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.4558307> (accessed Aug 2021).

⁹ see Saji, Ami, & Morales, Laura. (2020b). Reusing and sharing metadata: A case study of quantitative survey research on the integration of ethnic and migrant minorities (EMMs) across Europe. 12th Annual European DDI User Conference (EDDI20). Zenodo. <https://doi.org/10.5281/zenodo.4305692> (accessed Aug 2021).

¹⁰ Saji, Ami, & Morales, Laura. (2021). Guidelines document.

<https://ethmigsurveydatahub.eu/emmregistry/guidelines-document-metadata-template-version-control/> (accessed Aug 2021).

¹¹ Conditions of use: <https://ethmigsurveydatahub.eu/conditions-of-use-of-the-emm-survey-registry-and-its-metadata/> (accessed Aug 2021).

¹² Saji, Ami, Barjan, Isabella, Douchet, Léa. (2021). EMM SURVEY REGISTRY Back-end user guide

<https://ethmigsurveydatahub.eu/emmregistry/user-guide-beta-version-emm-survey-registry/> (accessed Aug 2021).

¹³ Video tutorial series: <https://www.youtube.com/playlist?list=PLo7wmLY6eAWKFFAg8rT2BIRmUuTJRt1li> and <https://www.youtube.com/playlist?list=PLo7wmLY6eAWlwMy8uoGCVyYcC4uQSRrZh> (accessed Aug 2021).

¹⁴ <https://ethmigsurveydatahub.eu/help-us-to-create-a-covid-19-collection-for-the-emm-survey-registry/> (accessed Aug 2021).

- Zenodo community¹⁵ NOTE: Many of the items referenced in point B are already part of or will be added to this Zenodo community.

2.2.2 Feasibility testing of an EMM Question Data Bank (QDB)

COST Action 16111 - ETHMIGSURVEYDATA, the ANR project FAIRETHMIGQUANT, and the H2020 SSHOC project are jointly testing the feasibility of setting up, as part of the CESSDA-led European Question Bank (EQB), a component dedicated to EMM surveys identified via the development of the EMM Survey Registry (i.e. the EMM Question Data Bank). If such a stream is set up, users will be able to use this tool to explore existing survey questionnaires and retrieve specific questions to be used in their own surveys. This work will lead to a feasibility report in the form of Deliverable D9.5, by December 2021.

Feasibility testing is undertaken via three thematic pilots covering surveys from 10 different countries. More details about the on-going work plan/strategy for the EMM QDB can be given upon request.

3. T9.3: Electoral Studies

3.1 Brief description of the task and the user community

The Team of the Data Community project Electoral Studies (Task 9.3 of WP9) consists of Cees van der Eijk (University of Nottingham), Martin Kaltenboeck, Albin Ahmeti and Sotirios Karampatakis (Semantic Web Company), Sylvia Kritzinger, Hajo Boomgaarden Johann Gruendl and Julia Partheymueller (University of Vienna), and Veronika Heider (AUSSDA). The work that is undertaken in this Task aims to serve the user community of electoral studies, which is a broad, variegated and international universe of scholars and practitioners in all phenomena related to elections, portrayed in more detail in Deliverable D9.6 (*Demarcation Report of Electoral Studies User Community*)¹⁶. Within this broad user community, the work of T9.3 focusses more specifically on the scholarly sub-community of analysts of the electoral behaviour of citizens, for reasons explained in detail in Deliverable D9.7 (*Design and Planning of Knowledge Graph in Electoral Studies*, particularly section 2.1¹⁷). The main component of the work planned in T9.3 is the

¹⁵ Zenodo community: <https://zenodo.org/communities/ethmigsurveydata/?page=1&size=20> (accessed Aug 2021).

¹⁶ Van der Eijk, Cees. (2020a). SSHOC D 9.6 Demarcation Report of Electoral Studies User Community. Zenodo. <https://doi.org/10.5281/zenodo.4558308> (accessed Aug 2021).

¹⁷ Van der Eijk, Cees, Kritzinger Sylvia, Partheymüller Julia, Kaltenböck Martin, Ahmeti Albin, & Karampatakis Sotirios. (2021). SSHOC D9.7 Design and Planning of Knowledge Graph in Electoral Studies (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.4558309> (accessed Aug 2021).

construction of a Knowledge Graph in the field of Election Studies and an associated Election Studies Analytics Dashboard.

3.2 Major activities in T9.3 and production of services, tools, procedures and data

As stated in the previous section, the main product that is envisaged to be generated by T9.3 is a Knowledge Graph in the field of election studies. As described in Deliverable D9.7 (see footnote 17) a Knowledge Graph is “a formal and structured representation of a domain of knowledge that offers stakeholders an interface for navigating the accumulated knowledge in the domain, and for discovering relationships between information from otherwise separate and unconnected databases within or relating to that domain” (p.11; for a more detailed discussion, see the report of Deliverable D9.7, and particularly section 2.2). It is generally envisaged that KGs will be invaluable for the further development of social sciences and for a better utilisation of existing data. Yet, their development in the domain of the social sciences and humanities is still in its infancy. The aspirations of T9.3 with respect to the construction of a KG should therefore be regarded as a pilot or feasibility study (as emphasised in the report of D9.7). The realisation of this pilot involved so far two major deliverables: D9.7 and D9.9.

D9.7 specified the design and planning of the KG to be constructed, which included the precise demarcation of the substantive domain; the specification of the desired functionality and of the universes of datasets and publications to be covered; the approach to existing ontologies and further ontology and taxonomy development; and technical specifications including choice of software platform. In this way D9.7 provided the basis for the actual construction of a pilot-KG.

In Deliverable D9.9 (*Delivery of User-Validated Knowledge Graph and Election Studies Analytics Dashboard*¹⁸) the choices specified in D9.7 have been implemented in an application based upon the Pool Party software platform¹⁹. The first version of the pilot-KG can be accessed online²⁰, while the D9.9 report provides (in section 4.3) a concise User Guide. This application covers close to 4000 publications, and almost 1100 datasets.

An important aspect of KGs is that they should evolve in at least two ways. One of these relates to the ingestion of material (in this case: publications and datasets). Only by regular updating can a KG become and remain relevant for its intended user community. In the remainder of the SSHOC project, such updating will be undertaken, partly by the development team, and partly by involving members of the

¹⁸ Karampatakis Sotirios, Ahmeti Albin, Kaltenböck Martin, Gründl Johann, Partheymüller Julia, & Van der Eijk Cees. (2021). SSHOC D9.9 Delivery of user-validated Knowledge Graph, and Election Studies Analytics dashboard (V1.0). Zenodo. <https://doi.org/10.5281/zenodo.4700170> (accessed Aug 2021).

¹⁹ <https://www.poolparty.biz/> (accessed Aug 2021).

²⁰ First version of the pilot-KG: <https://sshoc.poolparty.biz/GraphSearch/> (accessed Aug 2021).

user community in ways described in Deliverable D9.8 (*User-community involvement plan*²¹). The other way in which KGs should evolve is by testing (as described in the report of D9.7, section 7) and improving upon imperfections flagged up in actual usage. Here too, the aspiration is to involve members of the user community (as described in D9.7 and D9.8). At the moment of writing of the current report such testing has only yet been conducted using a small panel of five evaluators who are experts in the substantive field covered by the KG and who are also familiar with the architecture and structure of the KG as implemented in the PoolParty software platform. This testing --in April/May 2021-- resulted in the identification of a limited number of aspects in the user-interface that were deemed to require amendment before wider user-community involvement will be appropriate. These amendments are currently being implemented, and when completed will allow for the kind of end-user involvement in further testing that was envisaged in D9.8.

4. T9.4: Heritage Science

4.1 Brief description of the task and the user community

The Task 9.4 “Heritage Science and Humanities” data pilot is led by the CNR (Consiglio Nazionale delle Ricerche, Italy) and consists of Task leader Emiliano Degl’Innocenti, CNR - OVI SSHOC research fellows Carmen Di Meo and Francesco Coradeschi, and CNR - OVI SSHOC member Maurizio Sanesi.

The T9.4 team aims to realise a MVP (Minimum Viable Platform) named “RESTORE” that focuses on the recovery, integration, accessibility and reuse of digital resources, either provided by GLAM (Galleries, Libraries, Archives, Museums) institutions or Heritage science research institutes. In the cultural heritage domain digitised materials are collected, mapped and stored using a variety of standards. This implies that the realization of the RESTORE platform requires these standards to be made interoperable.

RESTORE (smaRt accESs TO digital heRitage and mEmory) is a digital environment (i.e.: a platform) where the complete workflow and the tools selected by the SSHOC T9.4 are made available for testing²². The CNR - OVI Institute also coordinates a project, co-funded by the Regione Toscana²³, with the scope of reconstructing the history of the city of Prato by means of the integration of archival, museum, and

²¹ van der Eijk, Cees. (2020b). D9.8 User-community involvement plan. Zenodo. <https://doi.org/10.5281/zenodo.4558312> (accessed Aug 2021).

²² RESTORE platform - Homepage: <http://restore.ovi.cnr.it/> (accessed July 2021).

²³ RESTORE is a project coordinated by the CNR - OVI Institute and co-funded by the Regione Toscana, in the context of the call POR-FSE 2014-20 - Programma operativo regionale (Por) del Fondo sociale europeo (Fse): <https://www.regione.toscana.it/por-fse-2014-2020> (accessed July 2021) - ending in 2022.

textual sources first converted into datasets, centered on the figure of the famous merchant Francesco di Marco Datini (ca. 1335 - 1410). The “Datini” pilot is the outcome of the RESTORE project. The approach, the software components and the technological solutions developed for the RESTORE project are to be seen as achievements and by-products provided to the above project by SSHOC T9.4, which developed the digital infrastructure needed. From the RESTORE project, an in-kind contribution to the platform development comes from the CNR - OVI research fellows Alessia Spadi and Federica Spinelli, who are in charge of - respectively - frontend and backend services of RESTORE.

A key point in the data pilot project is that of treating the data to make them readable and linkable in terms of a *PERSON - EVENT - OBJECT*-related logic. The release of the “data pilot” project contributes to the implementation of interoperability and sustainability of data management and storage in the Humanities and the Heritage sectors, and to the implementation of tools for the Heritage scientists and stakeholders, proposing a virtuous model for data integration. These achievements are realized by means of a full selection of instruments, together with a documentation on their development, that work with each standardized and validated procedure in use within the different fields of knowledge, referring to the SSHOCro (reference ontology), CIDOC - crm, and making use of open access sources, following FAIR protocols.

In terms of user communities, the task pilot involves people and datasets coming from the following institutions:

- Archives (Prato State Archives)
- Museums (Museo di Palazzo Pretorio, Prato)
- Superintendencies (The Archival and Bibliographic Superintendency)
- Coordination Agencies (Central Institute for the Union Catalogue of Italian Libraries, ICCU)
- Multiple research institutes in the field of Heritage Science and SSH, both within academic departments and research infrastructures active at a national level.

The platform developed will elaborate a sustainable workflow together with a toolkit supporting data management and integration of heterogeneous resources - from ingestion, to publication, through processing. In fact, the pipeline is currently clearly designed and articulated, supported by a set of open access tools, consisting of the following sequence:

- Data acquisition from GLAM partners, storage in a trusted datastore (i.e.: CKAN) ⇒
- Knowledge extraction through parsing of original data ⇒
- Thesauri design and creation (archival, museum, textual entities expressed in the mapping) ⇒
- Modeling of the entities extracted with CIDOC - CRM ⇒
- Data Transformation in RDF (through custom Python parsers and other tools, including 3M) ⇒
- Code storage and documentation (JupyterLab) ⇒
- Code control, versioning and documentation (GOGS) ⇒
- Data import in the triplestore (OpenLink Virtuoso) ⇒
- Visualization of semantic data with different interfaces (i.e.: SPARQL endpoint, Faceted browser, Graph Browser) and data dissemination.

Such a digital ecosystem will allow each user - and especially those coming from research institutes - to get open access to a unique data universe, where both the original resources and their triples (RDF), could be visualized and, if needed, downloaded. Furthermore, using the RESTORE tools makes it possible for other data providers to contribute to the project dataspace with their own datasets.

The RESTORE approach and the technological solutions developed, can be extended to other projects and reused in similar contexts by other institutions within the GLAMs, the Digital Humanities and Heritage Science research communities. Furthermore, the project's development will benefit from the collaboration with DARIAH-ERIC²⁴ (ESFRI Landmark for the humanities and social sciences) and E-RIHS²⁵ (ESFRI project for the heritage science), as well as the European Open Science Cloud (EOSC)²⁶. In this respect, the development of the RESTORE digital platform is aligned with the SSHOC²⁷ and the IPERION-HS²⁸ projects, with the aim to deliver a toolkit for heterogeneous data modeling and integration, as well as to establish a set of good practices for the research data management and storage.

4.2 Dissemination Activities

After the first iteration, results will be evaluated by a focus group of domain experts, using smart impact indicators, such as "user stories", and "usability scale surveys", according to the protocol validated by the Italian Ministry for Public Affairs called eGlu (2018)²⁹. The whole procedure will be refined, optimized and extended until the end of the project to better match the partners' needs;

- Report on MS49 "Heritage Science and Humanities Pilot alpha release"³⁰

²⁴ Digital Research Infrastructure for the Arts and Humanities - DARIAH-ERIC. ESFRI Landmark for the Digital Humanities, cfr. <https://www.dariah.eu/> (accessed Aug 2021).

²⁵ European Research Infrastructure for Heritage Science - E-RIHS. ERIHS is the only RI in the field of Social and Cultural Innovation in the ESFRI Roadmap, cfr. <http://e-rihs.eu> (accessed Aug 2021).

²⁶ European Open Science Cloud - EOSC: "The European Open Science Cloud is an environment for hosting and processing research data to support EU science. The process to create the EOSC was initiated by the Commission in 2015. It aimed to develop a trusted, virtual, federated environment that cuts across borders and scientific disciplines to store, share, process and re-use research digital objects (like publications, data, and software) following FAIR principles." Cfr. https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/european-open-science-cloud-eosc_en (accessed Aug 2021).

²⁷ Social Sciences and Humanities Open Cloud - SSHOC, the reference Social and Cultural Innovation thematic cluster for the EOSC, funded by EC under the H2020-INFRAEOSC-2018-2 topic: <http://sshopencloud.eu> (accessed Aug 2021).

²⁸ Integrated Platform for the European Research Infrastructure ON Heritage Science - IPERION-HS, project funded by the European Commission, H2020-INFRAIA-2019-1, cfr. <https://www.iperionhs.eu/> (accessed Aug 2021).

²⁹ Eglu Protocol Guidelines: <https://docs.italia.it/italia/designers-italia/design-linee-guida-docs/it/stabile/doc/user-research/usabilita.html#protocollo-eglu-ig-pr-la-realizzazione-di-test-di-usabilita> (accessed July 2021).

³⁰ Degl'Innocenti Emiliano, Di Meo Carmen, Coradeschi Francesco, Sanesi Maurizio, Brunoni Elisa, Kritsotaki Athina, & Tsoulouha Eleni. (2021). MS49 Heritage Science and Humanities Pilot alpha release (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.5217113> (accessed Aug 2021).

- PUBLICATION: E. Degl'Innocenti et al. (2021) C. Di Meo, A. Spadi, F. Spinelli (2021) RESTORE: SMART ACCESS TO DIGITAL HERITAGE AND MEMORY in: Electronic Imaging & the Visual Arts: EVA 2021 Florence: 14 June 2021 / edited by Vito Cappellini – Firenze: LEONARDO LIBRI srl, ISBN 978-88-596-2159-1;³¹
- PUBLICATION: Spadi, Alessia, F. Spinelli and Emiliano Degl'Innocenti. "RESTORE: Smart Access to Digital Heritage and Memory." in: "Proceedings of the 17th Italian Research Conference on Digital Libraries, Padua, Italy (virtual event due to the Covid-19 pandemic), February 18-19, 2021." IRCDL (2021);³²
- PUBLICATION: RESTORE: Degl'Innocenti Emiliano, Alessia Spadi, Federica Spinelli, RESTORE - smaRt accESs TO digital heRitage and mEmory, in: Federico Boschetti, Angelo Mario Del Grosso, Enrica Salvatori (edd.) AIUCD 2021 - DHs for society: e-quality, participation, rights and values in the Digital Age. Book of extended abstracts of the 10th national conference, ISBN: 9788894253559, 3-11.³³
- Kleemola et al. (2020). SSHOC D8.2 Certification plan for SSHOC repositories³⁴.

5. EURHISFIRM and RESILIENCE

As stated in Section 1 of this report, the consortia EURHISFIRM (economic history of firms) and RESILIENCE (religion studies) are included in WP9, but without having been tasked with distinct responsibilities to produce services, tools, procedures and data in the context of WP9 and SSHOC.

These consortia do, however, generate a variety of services, tools, procedures and data for their respective user communities in ways that are defined and structured by their own executives, steering committees and participating members. In the context of the current report it is therefore relevant to refer to these products by way of linking to how they have been reported within these consortia.

For the EURHISFIRM consortium we refer for a list of generated products³⁵, which contains a complete overview of deliverables and reports about all the products generated by the consortium. For RESILIENCE we refer to the main webpage of the consortium³⁶, and particularly the subpages about services, results, and news and events.

³¹ accessible here: <http://digital.casalini.it/9788859621591> (accessed July 2021).

³² available here: <https://www.semanticscholar.org/paper/RESTORE%3A-Smart-Access-to-Digital-Heritage-and-Spadi-Spinelli/028367f0dbb73091fed601d48aa722787b47393f#paper-header> (accessed Aug 2021).

³³ available here: http://amsacta.unibo.it/6712/1/AIUCD2021_BOA-versione3A.pdf

³⁴ Kleemola, Mari, Alaterä, Tuomas J., Koski, Niko, Ala-Lahti, Henri, Jerlehag, Birger, L'Hours, Hervé, De Jong, Franciska, Van Uytvanck, Dieter, Parkola, Tomasz, Degl'Innocenti, Emiliano, Giacomi, Roberta, Sanesi, Maurizio, & Van Horik, René. (2020). SSHOC D8.2 Certification plan for SSHOC repositories (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.4558303> (accessed Aug 2021).

³⁵ available here: <https://eurhisfirm.eu/index.php/publications/> (accessed Aug 2021).

³⁶ available here: <https://www.resilience-ri.eu/> (accessed Aug 2021).

6. Conclusion

This report aimed to provide an overview of the products --services, tools, procedures and data--generated by the three major tasks of WP9 of SSHOC: ethnic and migration studies (T9.2), electoral studies (T9.3), and heritage science (T9.4). It also aimed to provide a perspective on their usage and reception in the user communities associated with these three tasks. Having reviewed the work of these three tasks, we therefore arrive at the following conclusion about products and their usage:

- The EMM Survey Registry (T9.2) is fully functional as a beta version, and is publicly available to the intended user community. This version is fully functional and displays metadata for over 1200 surveys from nearly 30 different countries. The tool has already welcomed 300-400 unique users each month from all over the globe and from different disciplines, sectors, industries, etc. Additional surveys are added to the Registry.
- The EMM Question Data Bank (T9.2) is in development as part of the CESSDA-led European Question Bank. This project aims to be able to arrive at a conclusion about its feasibility by the end of 2021.
- The Knowledge Graph in electoral studies (T9.3) is available in an alpha-tested version. It is under revision to accommodate feedback from a highly-qualified evaluator panel so that the next (beta) version can be made publicly available in the intended user community.
- The RESTORE platform (T9.4) is developed in alpha version and will be evaluated by a focus group of domain experts, upon which the product will be refined, optimized and extended until the end of the SSHOC project to optimally match the needs of the intended user community.

The report of Deliverable 9.3 will focus in more detail on usage of these products in their respective user communities, and on the feedback generated by that usage.

7. References

Degl'Innocenti, Emiliano, Di Meo, Carmen, Coradeschi, Francesco, Sanesi, Maurizio, Brunoni, Elisa, Kritsotaki, Athina, & Tsoulouha Eleni. (2021). MS49 Heritage Science and Humanities Pilot alpha release (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.5217113> (accessed Aug 2021).

Degl'Innocenti, Emiliano, Di Meo, Carmen, Spadi, Alessia & Spinelli, Federica (2021). RESTORE: SMART ACCESS TO DIGITAL HERITAGE AND MEMORY. In: Vito Cappellini (ed.) Electronic Imaging & the Visual Arts: EVA 2021 Florence: 14 June 2021 Firenze: LEONARDO LIBRI srl, pp.133-137. ISBN 978-88-596-2159-1.

Degl'Innocenti, Emiliano, Spadi, Alessia & Spinelli, Federica (2021). RESTORE - smaRt accESs TO digital heRitage and mEmory, in: Federico Boschetti, Angelo Mario Del Grosso, Enrica Salvatori (edd.), AIUCD 2021 - DHs for society: e-quality, participation, rights and values in the Digital Age. Book of extended abstracts of the 10th national conference, Pisa, 2021. ISBN: 9788894253559, 3-11.

Karampatakis, Sotirios, Ahmeti, Albin, Kaltenböck, Martin, Gründl, Johann, Partheymüller, Julia, & Van der Eijk, Cees. (2021). SSHOC D9.9 Delivery of user-validated Knowledge Graph, and Election Studies Analytics dashboard (V1.0). Zenodo. <https://doi.org/10.5281/zenodo.4700170> (accessed Aug 2021).

Kleemola, Mari, Alaterä, Tuomas J., Koski, Niko, Ala-Lahti, Henri, Jerlehag, Birger, L'Hours, Hervé, De Jong, Franciska, Van Uytvanck, Dieter, Parkola, Tomasz, Degl'Innocenti, Emiliano, Giacomini, Roberta, Sanesi, Maurizio, & Van Horik, René. (2020). SSHOC D8.2 Certification plan for SSHOC repositories (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.4558303> (accessed Aug 2021).

Saji, Ami, Barjan, Isabella & Douchet, Léa. (2021). EMM SURVEY REGISTRY Back-end user guide <https://ethmigsurveydatahub.eu/emmregistry/user-guide-beta-version-emm-survey-registry/> (accessed Aug 2021).

Saji, Ami & Morales, Laura. (2020a). D9.4 Database with the metadata of surveys to EMMs across Europe (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.4558307> (accessed Aug 2021).

Saji, Ami, & Morales, Laura. (2020b). Reusing and sharing metadata: A case study of quantitative survey research on the integration of ethnic and migrant minorities (EMMs) across Europe. 12th Annual European DDI User Conference (EDDI20). Zenodo. <https://doi.org/10.5281/zenodo.4305692> (accessed Aug 2021).

Saji, Ami, & Morales, Laura. (2021). Guidelines document. https://ethmigsurveydatahub.eu/emmregistry/guidelines-document_-metadata-template-version-control/ (accessed Aug 2021).

Spadi, Alessia, Spinelli, Federica & Degl'Innocenti, Emiliano. "RESTORE: Smart Access to Digital Heritage and Memory." in: "Proceedings of the 17th Italian Research Conference on Digital Libraries, Padua, Italy (virtual event due to the Covid-19 pandemic), February 18-19, 2021." IRCDL (2021).

Van der Eijk, Cees. (2020a). SSHOC D 9.6 Demarcation Report of Electoral Studies User Community. Zenodo. <https://doi.org/10.5281/zenodo.4558308> (accessed Aug 2021).

Van der Eijk, Cees. (2020b). D9.8 User-community involvement plan. Zenodo. <https://doi.org/10.5281/zenodo.4558312> (accessed Aug 2021).

Van der Eijk, Cees, Kritzinger, Sylvia, Partheymüller, Julia, Kaltenböck, Martin, Ahmeti, Albin, & Karampatakis, Sotirios. (2021). SSHOC D9.7 Design and Planning of Knowledge Graph in Electoral Studies (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.4558309> (accessed Aug 2021).