Research and Innovation Action

# Social Sciences & Humanities Open Cloud

Project Number: 823782 Start Date of Project: 01/01/2019 Duration: 40 months

## Deliverable 1.3 First Annual Progress & Activity Report

| | |
|---|---|
| Dissemination Level | PU |
| Due Date of Deliverable | 31/01/2020 (M13) |
| Actual Submission Date | 31/01/2020 |
| Work Package | WP1- Project Management and Administration |
| Task | T1.3 Project Reporting |
| Type | Report |
| Approval Status | |
| Version | V1.0 |
| Number of Pages | p.1 – p.70 |

**Abstract:**

This document is a periodic report presenting the project's progress in the first 12 months. It includes planned and completed activities and use of resources of the project partners (Beneficiaries and LTPs) and defines the current status of the SSHOC work plan.



SSHOC, "Social Sciences and Humanities Open Cloud", has received funding from the EU Horizon 2020 Research and Innovation Programme (2014-2020); H2020-INFRAEOSC-04-2018, under the agreement No. 823782

## History

| Version | Date | Reason | Revised by |
|---|---|---|---|
| 0.1 | 18/11/2019 | First suggestions on template by Martina Drascic | Ivana Ilijasic Versic |
| 0.2 | 09/12/2019 | Revision and agreement on final version of sections | PMB |
| 0.3 | 23/12/2019 | WP reports included in the template (WP1, WP2, WP3, WP4, WP5, WP6, WP7, WP9) | Martina Drascic, PMB |
| 0.4 | 28/01/2020 | WP8 report included, edited and revised | Ivana Ilijasic Versic |
| 0.5 | 29/01/2020 | Final revision, Risks, Conclusions, References sections edited | Ron Dekker |
| 0.6 | 31/01/2020 | Financial and Use of Resources sections added | Martina Drascic, Ivana Ilijasic Versic |
| 1.0 | 31/01/2020 | Final version for submission | Martina Drascic |

## Author List

| Organisation | Name | Contact Information |
|---|---|---|
| CESSDA ERIC | Martina Drascic | martina.drascic@cessda.eu |
| CESSDA ERIC | Ivana Ilijasic Versic | ivana.versic@cessda.eu |
| TRUST-IT | Marieke Willems | m.willems@trust-itservices.com |
| CLARIN ERIC | Daan Broeder | daan.broeder@di.huc.knaw.nl |
| CLARIN ERIC | Maria Eskevich | maria@clarin.eu |
| ESS ERIC | Gianmaria Bottoni | Gianmaria.Bottoni@city.ac.uk |
| UPF | Diana Zavala Rojas | diana.zavala@upf.edu |
| SHARE ERIC | Annette Scherpenzeel | Scherpenzeel@mea.mpisoc.mpg.de |
| SHARE ERIC | Johanna Bristle | Bristle@mea.mpisoc.mpg.de |
| LIBER | Vasso Kalaitzi | Vasso.Kalaitzi@kb.nl |
| DARIAH ERIC | Laure Barbot | laure.barbot@dariah.eu |
| DARIAH ERIC | Frank Fischer | frank.fischer@dariah.eu |
| CNR | Emiliano Degl'Innocenti | emiliano.deglinnocenti@cnr.it |
| UoN | Cees van der Eijk | Cees.Van_Der_Eijk@nottingham.ac.uk |

## Time Schedule before Delivery

| Next Action | Deadline | Care of |
|---|---|---|
| First suggestion of template and descriptions of planned text | 18/11/2019 | CESSDA |
| Revision and agreement on final version to be filled | 22/11/2019 | CESSDA |
| Presentation and agreement with PMB on the template and timeline | 25/11/2019 | CESSDA |
| Comments and suggestions by PMB | 09/12/2019 | PMB |
| Comments included in the template - beginning of work on content | 13/12/2019 | CESSDA |
| Collection of data by beneficiaries and LTPs | 30/01/2019 | CESSDA |
| All input by WP Leaders included in the template | 20/01/2019 | PMB and CO |
| Final editing and addition of Financial reports in the document (M1-6 & M7-12) | 31/01/2019 | CESSDA |
| Peer review | 31/01/2019 | PMB, and 1st Tier to issue recommendations |
| Final revision and submission | 31/01/2019 | CESSDA |

# Executive Summary

This deliverable is the first annual progress report for the SSHOC project, and as such, one of three planned progress reports to be delivered during the project lifetime. It states the progress and status of work in the first year of the SSHOC project (M1-M12). The information gathered here will be updated after six months, (and thus form the official EC periodic report); the next progress report will follow in 2021.

The report starts with a general plan for the SSHOC project in Year 1 (2019), and lists the main objectives set up to be achieved. The main part of this deliverable follows activities described in chapters which go into details about highlights and achievements on the level of each Work Package and on the level of Tasks, with notes on deviations from the DoA. The final chapter contains the summary on the delivery of project outputs (namely deliverables), achievement of project milestones, use of resources, and risks met in the Year 1 across all Work Packages.

This progress report, and the two that will follow at the end of each project year, form the basis for more extended reporting to the European Commission and provide overview and further improvements to the ongoing work.

## Abbreviations and Acronyms

| | |
|---|---|
| API | Application Programming Interface |
| AUSSDA | The Austrian Social Science Data Archive |
| BBMRI | Biobanking and BioMolecular resources Research Infrastructure |
| CESSDA | Consortium of Social Sciences Data Archives |
| CentERdata | Stichting CentERdata (Foundation CentERdata) |
| CIDOC | International Committee of Documentation (Comité International pour la Documentation) |
| CIDOC-CRM | CIDOC Conceptual Reference Model |
| CLARIN | Common Language Resources and Technology Infrastructure |
| CNR | Consiglio Nazionale delle Ricerche (Italian Research Council) |
| CNRS | Centre national de la recherche scientifique (French National Centre for Scientific Research) |
| CO | Coordinator |
| CRMdig | CRM Digital; a model for provenance data |
| CRMsci | CRM Scientific Observation Model |
| CUNI | Univerzita Karlova (Charles University in Prague) |
| DAI | Deutsches Archäologisches Institut (German Archaeological Institute) |
| DARIAH | Digital Research Infrastructure for the Arts and Humanities |
| DBSS | Dried Blood Spot Samples |
| DDI | Data Documentation Initiative |
| DG RTD | Directorate-General for Research and Innovation |
| DNA | The Danish National Archives |
| DoA | Description of Action (Annex 1 of the GA) |
| EC | European Commission |
| EMM | ethnic and migrant minorities |
| EOSC | European Open Science Cloud |
| EQB | European Question Bank |
| ERIC | European Research Infrastructure Consortium |
| E-RIHS | European Research Infrastructure on Heritage Science |
| ESS | European Social Survey |

| | |
|---|---|
| ETL | Extract, Transform, Load |
| ETHMIGSURVEYDATA | The International Ethnic and Immigrant Minorities' Survey Data Network |
| FAIR | Findable, Accessible, Interoperable, Reusable (data) |
| FORTH | Foundation for Research and Technology – Hellas |
| GA | Grant Agreement |
| GDPR | Generic Data Protection Regulation |
| GESIS | GESIS - Leibniz-Institut für Sozialwissenschaften e.v. (Leibniz Institute for the Social Sciences) |
| ISIC | International Standard Industrial Classification of All Economic Activities |
| ISCO | International Standard Classification of Occupations |
| KNAW | Koninklijke Nederlandse Akademie van Wetenschappen (The Royal Netherlands Academy of Arts and Sciences, Netherlands) |
| LIBER | Association of European Research Libraries |
| LTP | Linked Third Party |
| NACE | Statistical Classification of Economic Activities in the European Community |
| NG | The National Gallery (London) |
| NSD | Norsk senter for forskningsdata (Norwegian center for research data) |
| OECD | Organisation for Economic Co-operation and Development |
| PARTHENOS | Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies |
| PMB | Project Management Board: CO + WP Leaders |
| ProPASS | Prospective Physical Activity, Sitting and Sleep Consortium |
| PSNC | Poznan Supercomputing and Networking Center |
| RDF | Resource Description Framework |
| REST | REpresentational State Transfer |
| Sciences PO | Fondation Nationale des Sciences Politiques (The Paris Institute of Political Studies) |
| SHARE | Survey of Health, Ageing and Retirement in Europe |
| SSH | Social Sciences and Humanities |
| SSHOC | Social Science and Humanities Open Cloud |
| SSHOCro | SSHOC Reference Ontology |

| | |
|---|---|
| SWC | Semantic Web Company GMBH |
| T | Task |
| TMT | Translation Management Tool |
| TRUST-IT | TRUST-IT SERVICES LIMITED |
| UCL | University College London |
| UGOE | Georg-August-Universität Göttingen Stiftung öffentlichen Rechts (The University of Göttingen) |
| UKDS | UK Data Service |
| UI | User Interface |
| UNOTT | The University of Nottingham |
| UPF | Universitat Pompeu Fabra (Pompeu Fabra University) |
| VCR | Virtual collection registry |
| WP | Work Package |

# Table of Contents

# 1. Introduction

The SSHOC project has three mandatory reporting periods covering M1-M18, M18-M36, and M36-M40. In addition to Periodic and Final reports to be submitted to the Project Officer (PO), internal reporting has been set up, in order to ensure that all project obligations (as stated in the DoA, Part B) are being fulfilled, and the progress of each Work Package is documented and on track. The practice of regular reporting and updating of all involved partners has been set up: partners deliver their reports continuously (input is collected at regular meetings or requested by partners in emails, and reported directly in the EC portal by Coordinator), or every two-weeks or monthly (at regular coordination meetings between Coordinator and WP leaders, used for planning within Work Packages and Tasks, management decisions and process control), and every 6 months (use of resources reports sent to Coordinator). Collected information will be used to prepare three annual progress & activity reports which have been set as deliverables covering 3 delivery periods (M1-M12, M13-M24, M25-36).

These reports will include the project plan and objectives relating to the period, progress across WPs and Tasks, risks met and contingency plans, planned versus actual deliverable submission, planned versus actual milestones achievement, and use of resources status relating to the period provided by internal financial reports collected by the Coordinator every six months.

In more detail, this report, D1.3 provides an overview of the work performed from M1 to M12 of the SSHOC project and all challenges faced in the process. If applicable, it also specifies deviations from the planned timeline and contingency plans outlining corrective actions to be taken.

**Plan for Year 1 of SSHOC project**

In 2019, year one of the project, the Consortium planned to review the DoA, set up appropriate teams, hire employees to be working on tasks in the following years, discuss inter-relations and dependencies between the Work Packages in order to work better together and reduce the amount of overlaps, achieve the first milestones in each task and deliver the work promised in the GA for the first year of the project. The entire Consortium planned to meet twice in 2019, at two Consortium meetings, and set up regular virtual calls to ensure progress of each Task and Work Package. For year one, SSHOC Project planned to deliver in total **22 Deliverables** and achieve **16 milestones**, as presented in the timeline (in Figure 1) below.

Due to a significant number of Deliverables being scheduled for M12 (which coincided with the holiday period as the timing of the EC approval of SSHOC project generated its start to be January 2019), some of the processes were slower than expected. This resulted in the project not delivering all the planned deliverables in Year 1. In total, SSHOC Consortium managed to achieve 13 Milestones and successfully submit 13 Deliverables to the EC. Three of those were already approved in Year 1. In addition, 9 deliverables are either in preparation or review process, and all delays were communicated with the SSHOC PO. The following chapters detail the work progress in each Work Package and Task, explain the activities conducted in Year 1 and, where applicable, deviations from the initial plan.

| Year 1 2019 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| WP1 Project Management and Administration | | | | | | | | | | | |
| | | | D1.1 | | D1.6 | | D1.2 | | | | |
| | | | MS 1 | | | | | | | | |
| WP2 Communication and Dissemination | | | | | | | | | | | |
| | | D2.1 | | | | | | | | | |
| MS 2 | | MS3 | | | MS 4 | | | | | | MS 5 |
| WP3 Lifting Technologies and Services into the SSH Cloud | | | | | | | | | | | |
| | | | | | D3.1 | | MS 9 | | | | |
| | | | | | D3.2 | | MS10 | | | | |
| | | | | | MS 8 | | MS 11 | | MS 12 | | |
| WP4 Innovations in Data Production | | | | | | | | | | | |
| | MS 16 | D4.1 | | | D4.16 | | | | | | D4.3 |
| | | | | | MS 19 | | | | | | D4.12 |
| | | | | | | | | | | | D4.14 |
| | | | | | | | | | | | D4.18 |
| WP5 Innovations in Data Access | | | | | | | | | | | |
| | | | MS 25 | | MS 21 | | | | D5.1 | | |
| | | | | | MS 24 | | | | D5.17 | | |
| WP6 Fostering Communities, Empowering Users, & Building Expertise | | | | | | | | | | | |
| | | | | | D6.1 | MS 38 | | | | | D6.2 |
| | | | | | | | | | | | D6.7 |
| WP7 Creating the SSH Open Marketplace | | | | | | | | | | | |
| | | | | | | | | D7.1 | | | |
| WP8 Governance/ Sustainability/ Quality Assurance | | | | | | | | | | | |
| | | | | | | | | | | | D8.2 |
| WP9 Data Communities | | | | | | | | | | | |
| | | | D9.1 | | D9.6 | | | | | | D9.7 |
| WP10 Ethics requirements | | | | | | | | | | | |
| | | | | | | | | | | | |

***Figure 1.*** *Planned Deliverables and Milestones in Year 1 of SSHOC project.*

# 2. Progress and activity reports per WP

## 2.1. WP1 - Project Management and Administration

In the first year of the project, WP1 focused on establishing the bases for a proper and sound management of this large project. The coordination team worked on setting up financial and administrative procedures, tools for monitoring of the overall consortium's delivery of the work plan and establishing platforms and channels for efficient and collaborative scientific and administrative coordination (T1.1 and T1.3). Procedures for ensuring high quality and proper risk management were also set in this year (T1.2). As a result, partners achieved one milestone, and submitted two deliverables planned for Year 1.

### 2.1.1. WP1 progress

#### 2.1.1.1. Task 1.1 Administrative Project Management

This task´s aim in Year 1 was to set up the basis for an efficient management of the SSHOC project and provide tools and procedures that would ensure a project of this size to deliver its contractual obligations. It was focused both on (1) administrative procedures and tools, (2) organisational activities and management of the Consortium, and (3) strategic and scientific management of the project activities closely related to its collocation and adjustment to a broader EOSC landscape.

To allow the project to start, Coordinator fulfilled all the basic project needs, such as setting up the collaborative platform (Basecamp[1]), regular management calls, internal project document repository (Google Shared Drive[2]) and **distributed the 1st instalment** to the Beneficiaries in a timely manner.

The task was led by CESSDA ERIC as the project coordinator, in close cooperation with WP leaders coming mostly from the 1st Tier organisations (ESS ERIC, SHARE ERIC, CLARIN ERIC, DARIAH ERIC, E-RIHS, Trust-IT, LIBER, and the University of Nottingham in addition) forming the Project Management Board (PMB), responsible for ensuring day-to-day sound and cost-effective management and delivery of the project outputs. As planned, partners worked to prepare tools for reporting, cost and time management, procedures and guidelines for activity planning and updates, and submission of deliverables to the Commission. This work led the partners to achieve the first project milestone ***MS1 Integrated project management and reporting tools and templates***[3] which meant that all prepared tools and templates were agreed and accessible to all partners via the project document repository and Basecamp collaborative platform in April 2019 (M4 of the project). A large part of this work also resulted in a ***D1.1 Project***

1 Basecamp: Project Management & Team Communication Software: https://basecamp.com/

2 The internal SSHOC Project repository is stored at CESSDA Shared Drive and accessible to all project partners: https://drive.google.com/drive/shared-drives

3 Details on all the tools provided for the project are explained in the Milestone 1 report shared with partners in the SSHOC document repository

***Management Plan***[4], submitted to the EC on 26 June 2019, and approved 18 November 2019. This document explains the most important management procedures and sets the basis for proper project management and reporting for a project of this size. The plan contains information about governance, project calendar, reporting roles and responsibilities for all partners and defines the communication mechanisms for the project's partners, templates for reporting and formal communication of the project's activities and outcomes.

The Work Package leaders have managed and coordinated work in their respective WPs and their teams since month one, while regularly reporting to PMB. To ensure that, **regular PMB meetings** were held every two weeks. In total, **20 virtual calls** (via GoToMeeting platform), and two physical (face-to-face) meetings were held as part of the project's consortium meetings.

The team also organised two meetings to gather the entire Consortium: **SSHOC Kick-off** meeting was held on 11-12 March 2019 in Utrecht, the Netherlands, organised by the Coordinator, and the **2nd Consortium meeting** was held in Florence, Italy on 14-15 October 2019, with CNR as the local organiser. These events allowed participation of all relevant stakeholders, with the aim to discuss the planned activities, ongoing work, and achieved outputs between all the leading people in the project. In November 2019, the team started to plan the 3rd Consortium meeting, to be held in March 2020 in The Netherlands.

In this task, the 1st Tier partners also formed the **Scientific Board** with the main purpose to advise the Coordinator and the Consortium on strategic issues of the project. The Board held one virtual meeting, and two face-to-face meetings (co-located with Consortium meetings). The Board focused on broader and more strategic EOSC eco-system through engagement in relevant governance bodies such as EOSC Executive Board, EOSC Working Groups or by collaborating with EOSC Secretariat, and connecting with projects, EC departments and data communities of relevance for SSHOC (i.e. EURHISFIRM[5] project or meeting with EC DG RTD at the EOSC Conference in Budapest). Specific attention was given to close collaboration and alignment of efforts with other cluster projects: EOSC-Life[6], ESCAPE[7], ENVRI-FAIR[8], and PANOSC[9], financed through the same call (INFRAEOSC-04). All these additional activities are done in order to secure that developments in SSHOC follow the fast-changing environment and adjust accordingly.

#### 2.1.1.2. Task 1.2 Quality Assurance & Risk Assessment

Task 1.2 is led by the coordinator CESSDA ERIC and partners ESS ERIC, SHARE ERIC, DARIAH ERIC, CLARIN ERIC, LIBER, TRUST-IT, E-RIHS and the University of Nottingham as members of the Project Management Board and the 1st Tier Strategic Board. In Year 1, Task 1.2 partners managed its main purpose: to define and specify the appropriate mechanisms and processes to maintain a high-quality level in the whole

4 Amir Spahic, Ivana Ilijasic Versic, Franciska De Jong, & Cees van der Eijk. (2019). SSHOC D1.1 Project Management Plan (Version v1.3). DOI: 10.5281/zenodo.3595441

5 Project website of EURhisFIRM EC funded project: https://eurhisfirm.eu/

6 EOSC Life project website: https://www.eosc-life.eu

7 ESCAPE project website: https://projectescape.eu

8 ENVRI-FAIR website: https://envri.eu/about-envri-fair/

9 PANOSC project website: https://www.panosc.eu

project structure and outcomes and establish a basis for efficient risk management. This was done in several ways:

(1) creating tools and procedures for risk and quality management,
(2) creating internal platforms to make sure all the 40 SSHOC partners are being informed about the developments and procedures in a project of this size,
(3) developing means to keep up with the events surrounding SSHOC and making sure it is aligned with the EOSC related developments at all times.

The first resulted in the submission of the main deliverable of the task - ***D1.2 Quality Assurance & Risk Assessment Plan***[10] in December 2019. This deliverable details on the quality assurance processes by defining activities and instruments to be used for the regular quality monitoring and risk assessment of the project in the form of a handbook for project partners. A special focus was given to the quality assurance of deliverables, and other project outcomes. Supporting tools, such as templates, guidelines and Risk Register, were created to regularly keep track of the identified risks and maintain the overall quality of work.

The second approach, explained in detail in the abovementioned Deliverable, encompasses all created tools and internal platforms which will assure the Consortium has all the relevant information available at all times. Except for collaboration, communication and archive platforms, this includes the creation of **SSHOC Wiki page** - ***SSHOC Guidelines***[11], an internal web page available to all project partners containing updated information on SSHOC procedures, management, reporting, visibility, and updates.

To make sure that SSHOC is visible to and aware of the related events/events of interest, and to align the dissemination and other activities with the EOSC related developments at all times, task partners set up two types of engagement platforms:

(1) set up an EOSC-Hub/cluster projects technical alignment team to make sure all the technical updates are taken into account, and
(2) set up a close collaboration with other clusters' communication and dissemination teams to achieve EOSC outreach alignment.

[10] Martina Drascic, Ivana Ilijasic Versic, Carsten Thiel, Daan Broeder, Gianmaria Bottoni, Marieke Willems, … Vasso Kalaitzi. (2019). SSHOC D1.2 Quality Assurance & Risk Assessment Plan (Version V1.2). DOI: 10.5281/zenodo.3595453

[11] https://sites.google.com/cessda.eu/sshocwiki/ Internal Wiki page available to SSHOC partners who are members of the SSHOC document repository (CESSDA Google account on Google Sites (www.sites.google.com))

#### 2.1.1.3. TASK 1.3 PROJECT REPORTING

This task's main focus in Year 1 was to set up a regular communication with the EC and secure continuous reporting on the EC portal, to set up procedures and formats for project reporting within the Consortium, and to establish and conduct the first internal reporting.

During year 1 the Coordinator maintained regular **communication with the EC** Project Officer and Legal Officer, to inform and report any issues, delays and discuss the content of the 1st Amendment to be submitted in early 2020. The continuous and periodical reporting on the work progress, and financial and use of resources reporting to the coordinator every 6 months was set up and agreed already in the project. The procedures were set up and explained at the project Kick-off and later in Deliverable 1.1 and Deliverable 1.2, followed by provided templates for reporting available in SSHOC Document repository and listed in Milestone 1 report.

Continuous reporting was done during regular calls for Task, Work Package, and PMB or in direct communication with the Consortium, while periodic internal reports are conducted every 6 months. For this purpose, the Coordinator has set up the reporting calendar and conducted the **first internal financial/use of resources reporting for M1-M6** of the project. Gathered information was presented to the Consortium at the 2nd Consortium meeting in Florence. Also, in November 2019, due to the holiday season, the Coordinator announced and started the second internal reporting for M7-M12 of the project. Information on both of the periods combined is represented in this report (See section "*Use of Resources*"). As a result of the established monitoring and project reports, the first Deliverable of this task is this document, *D1.3 First Annual Progress & Activity Report,* to be submitted in January 2020.

### 2.1.2. Note on deviations from the plan and risk monitoring

The only deviations from the plan in WP1 was the delay in submitting the Deliverables. *D1.1 Project Management Plan* was submitted in June (M6), instead of April (M4), *D1.2 Quality Assurance & Risk Assessment Plan* in December (M12) instead of August (M8), and *D1.6 Data Management Plan* is still pending and expected to be delivered in early 2020. All the delays were due to the need for including several partners in the writing and review process, and the review period overlapped with several other deliverables preparations, which were given priority. There are no negative consequences of these delays and all were communicated to the EC Project Officer (PO).

Information on risk monitoring in WP1 for Year 1 is presented in the table below:

***Table 1.*** *Risk monitoring in WP1 - Year 1*

| Risks monitoring WP1 | | | | |
|---|---|---|---|---|
| No of risk (from the DoA) | Description | Have risk-mitigation measures been applied? | Did the risk materialise so far? | Comments |
| New risk | Underestimated organisational efforts needed to coordinate a big scale project with considerable amount of Beneficiaries and Linked Third Parties, causing delays in the project delivery | **Yes** | **Yes** | Mitigation measures applied: Adding an extra employee in the coordination team, focusing on communicating delays to the PO frequently, and setting up internal procedures, timelines and delay registry to coordinate and eliminate future delays. |

## 2.2. WP2 - Communication, Dissemination and Impact

The main objective of WP2 is to widely promote the innovation and benefits of SSHOC to targeted stakeholders, maximize visibility, activate end-user communities in strong collaboration with WP6, disseminate results, and demonstrate impact. For the first year, WP2 focused on defining the Overall Communication and Outreach Strategy (D2.1) building the SSHOC community, leveraging on the existing and extensive network of European Research Infrastructure Consortia and Landmarks in the social sciences and humanities and the network of the Association of European Research Libraries. The WP has set the basic channels, tools and mechanisms in place for a coordinated and aligned communication, with tailored messages to the identified target stakeholders on SSHOC results and activities.

In the first year of the project the **website** was launched with basic information on the project, the Consortium, expected impact and a space for news and events. During the first year, and in the second iteration of the website information was expanded on SSHOC in the wider EOSC, the marketplace and the training network. At the end of the first year WP2 started the preparation for the third iteration of the website in which the marketplace page will include a service catalogue to raise awareness on which services will be delivered during the project lifetime.

Collaborations were set in place with WP6 on raising awareness of the SSHOC and its activities in stakeholder engagement activities and the targeted promotion of the training network. With WP7, alignment activities took place on the evolving branding of the marketplace as well as its visibility on the SSHOC website in every iteration.

### 2.2.1. WP2 progress

#### 2.2.1.1. Task 2.1 SSHOC Cluster communication & Dissemination strategy

The main goal of T2.1 is to define the project's communication and branding approach, and thereby set the basis for impactful SSHOC-related communication of the different ERICs, clusters and landmarks involved in the project. Since the main activities of T2.1 were of strategic character, they predominantly were set in the first six months of SSHOC. By providing orientation for communication and dissemination, T2.1 set the basis for T2.2 – the actual outreach and communication activities – and T2.3 – the SSHOC web presence.

Within T2.1 the first Deliverable of the SSHOC project was submitted on 5 April 2019 and approved by the EC on 18 November 2019: ***D2.1 SSHOC Overall Communication and Outreach Plan***. The document set out a strategy for communication, and additionally for dissemination. It provided guidance to the projects' partners in order to make the most of their activities by outlining the main objectives, target groups, tools, and activities for communication and dissemination of the SSHOC.

Furthermore, within T2.1 and as described in D2.1 several **guidelines and templates** were produced that aimed at supporting the distributed and coherent communication activities of the project partners.

Guidelines and templates produced explained and supported (1) the production of online content for the SSHOC web presence, (2) the organizing, reporting and evaluating of SSHOC-related events and (3) the drafting of SSHOC-branded documents, e.g. slideshow presentations and report documents.

#### 2.2.1.2. Task 2.2 Maximising Impact: SSH Communication and outreach as part of EOSC

Task 2.2 encompassed the production and dissemination of content following the strategy set out in T2.1 and defined in D2.1. This included web content for the SSHOC web presence, for social media, press releases and other outlets. Stakeholders groups identified in D2.1 were captured in the profiled ***MS2 Community database creation & management*** in M1. This database was set up and to identify and collect SSHOC stakeholder connections from partners. In year 1 WP2 and WP6 started populating this database with stakeholder connections, no personal contacts were collected for data privacy reasons. The database will be used to identify who should be contacted on a one-to-one basis for engagement activities as well as a two-step communication via the partners who have these direct connections.

SSHOC uses various communication channels leveraging the project partner networks with the SSHOC website as the central channel. Under T2.2 regular updates were published on the website.

In the first year:

- **22 news pieces were published on the SSHOC website**
- **38 events were promoted via the SSHOC website** relevant for the project stakeholders of which **21 were SSHOC activities**

In addition to these a number of **news-items deliverables were digested**[12] and published under project results. These updates were disseminated in a wider network via SSHOC and partner channels. In total **4 newsletters and direct messages**[13] were sent, on the latest SSHOC achievements and activities to **167 subscribers**.

Publications in journals and external channels were made in the EOSC Hub magazine[14], VOB journal[15], SERISS newsletter.

Stakeholder Engagement on social media helps to build a strong SSH community, where proper engagement ensures effective impact. Social Media channels set-up in SSHOC are **Twitter**[16], **LinkedIn**[17],

12 SSHOC Website publication of Deliverables produced: https://www.sshopencloud.eu/publications/deliverables
13 SSHOC newsletter: https://www.sshopencloud.eu/newsletter/sshoc-newsletter
14 Publication in EOSC-Hub magazine: https://www.eosc-hub.eu/sites/default/files/Magazine%233_Web_version%20Final-compressed.pdf
15 Ilijašić Veršić, I. und Ausserhofer, J. (2019) „Die Sozial- und Geisteswissenschaften und ihre Interoperabilität mit der European Open Science Cloud: Was ist SSHOC?“, Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare, 72(2), S. 383-391. doi: 10.31263/voebm.v72i2.3216.
16 SSHOC Twitter: https://twitter.com/SSHOpenCloud
17 SSHOC LinkedIn: https://www.linkedin.com/company/18997546/

**Flickr**[18], **YouTube**[19] and from M1-M11 Slides were uploaded to the **SlideShare channel**[20]. On twitter SSHOC has a community of 773 followers. From M1-M6 SSHOC had a LinkedIn account set-up in personal format, allowing publication of long articles. In M6 of the project, due to updated LinkedIn regulations, SSHOC moved to a professional type account. The new account has grown to 63 connections in M12.

Task 2.2 aims at covering the implementation and maintenance of a "cluster" **project identity and branding** where each partner, ERICs and landmarks initiatives are represented and enhanced toward existing communities and new communities. To this end, for the consortium to use, a branding package was delivered with a **Logo Pack**[21], **Branding guidelines**[22], and a template package for ppt presentation, deliverables, meeting agenda and milestones[23] prepared in collaboration with WP1.

Close alignment was established with WP7 on the branding of the Marketplace and with WP6 on the training community. In this context, the dynamically evolving and stakeholder tailored SSHOC **Communications Toolbox**[24] was designed and published on the project website. In Year 1 the following communication collaterals were delivered for the toolbox:

- Video's:
  - 1 general project video[25]
  - Video's at the Kick-Off Meeting[26]
  - Video of the launch of the training community[27]
- General presentation
- General Flyer[28]
- Roll-Up Banner[29]
- Post-It training community[30]

18 SSHOC Flicker: https://www.flickr.com/photos/170844433@N02/
19 SSHOC Youtube channel: https://www.youtube.com/channel/UCw-mY8v84yeHW2z4KG3ZLtA
20 SSHOC Slideshare: https://www.slideshare.net/SSHOCInitiative
21 SSHOC logo pack: https://www.sshopencloud.eu/logo-pack
22 SSHOC brand guidelines: https://www.sshopencloud.eu/sites/default/files/SSHOC%20Guideline_booklet.pdf
23 Available at the SSHOC Shared Drive document repository and the project Wiki under visibility guidelines
24 SSHOC communications kit: https://www.sshopencloud.eu/communications-kit
25 https://www.youtube.com/watch?v=UP7dBiV3-yQ
26 Interview with Ron Dekker, CESSDA Director: https://www.youtube.com/watch?v=gUWEragkniw
Interview with Franciska de Jong, CLARIN-ERIC executive director: https://www.youtube.com/watch?v=cgJbOgfXvzY
Interview with Astrid Verheusen, LIBER executive director: https://www.youtube.com/watch?v=eqoCi1Cu2yY
Interview with Frank Fischer, DARIAH co-director: https://www.youtube.com/watch?v=xfWuho7tVfA
27 Launch of the SSHOC Training Community - Interview with Ellen Leenarts - DANS-KNAW: https://www.youtube.com/watch?v=DWuHgQGKhdA
28 SSHOC flyer: https://www.sshopencloud.eu/sites/default/files/SSHOC_flyer_A5_October2019.pdf
29 SSHOC banner: https://www.sshopencloud.eu/sites/default/files/sshoc_banner_0.pdf
30 https://www.sshopencloud.eu/sites/default/files/SSHOC%20Post-IT_76x76_August2019_V0.2_0.pdf

- 2 press releases[31]
- Posters are tailored to the audience attending the specific conference - 5 posters were designed and presented:
    - EOSC Hub 2019
    - LIBER 2019
    - OS FAIR 2019
    - CLARIN 2019
    - EOSC Symposium 2019

In M1-12 SSHOC (co)-organised **9 events** and attended **35 events** to raise awareness on the project and its activities[32]. SSHOC organised events and presence at external events has been promoted via the website, SSHOC and partner social media channels and newsletters. For these events WP2 liaised with partners on collaterals, social media promotion and post event reporting and publication of presentations. WP2 has been working liaising with the **EOSC ecosystem** for communication and community engagement purposes via the following events:

- EOSC Hub week 2019[33]
- "EOSC services, collaborations and the RDA"[34] RDA Plenary 14 side-event, resulted in joint report.
- EOSCatRDA[35] where SSHOC joined the EOSC secretariat kiosk
- EOSC Symposium 2019[36].

The communication and outreach strategy includes the regular monitoring of impact. For this aim, T2.2 has provided **monthly progress** reports from M4, reporting on activities carried out, news and events promoted, status of compliance with KPIs set. In addition, the **SSHOC Dashboard**[37] has been set up to allow easy access for all partners to monitor the SSHOC website and social media metrics and to shape our strategy accordingly.

31 Launch SSHOC : https://www.sshopencloud.eu/news/"sshoc"-social-sciences-humanities-open-cloud-acronym-remember

32 In the two Appendices to this Deliverable tables listing all the events can be found

33 https://sshopencloud.eu/news/my-cloud-your-cloud-our-cloud

34https://www.sshopencloud.eu/news/eosc-esfri-cluster-projects-rda%C2%A0connecting-commonalities-and-collaborative-solutions-community

35 https://sshopencloud.eu/rda-14th-plenary-data-makes-difference

36 https://sshopencloud.eu/eosc-symposium-2019

37 SSHOC Dashboard: https://datastudio.google.com/reporting/1JTtWjrE7ee81DbpkWZc6596J22VNQncR

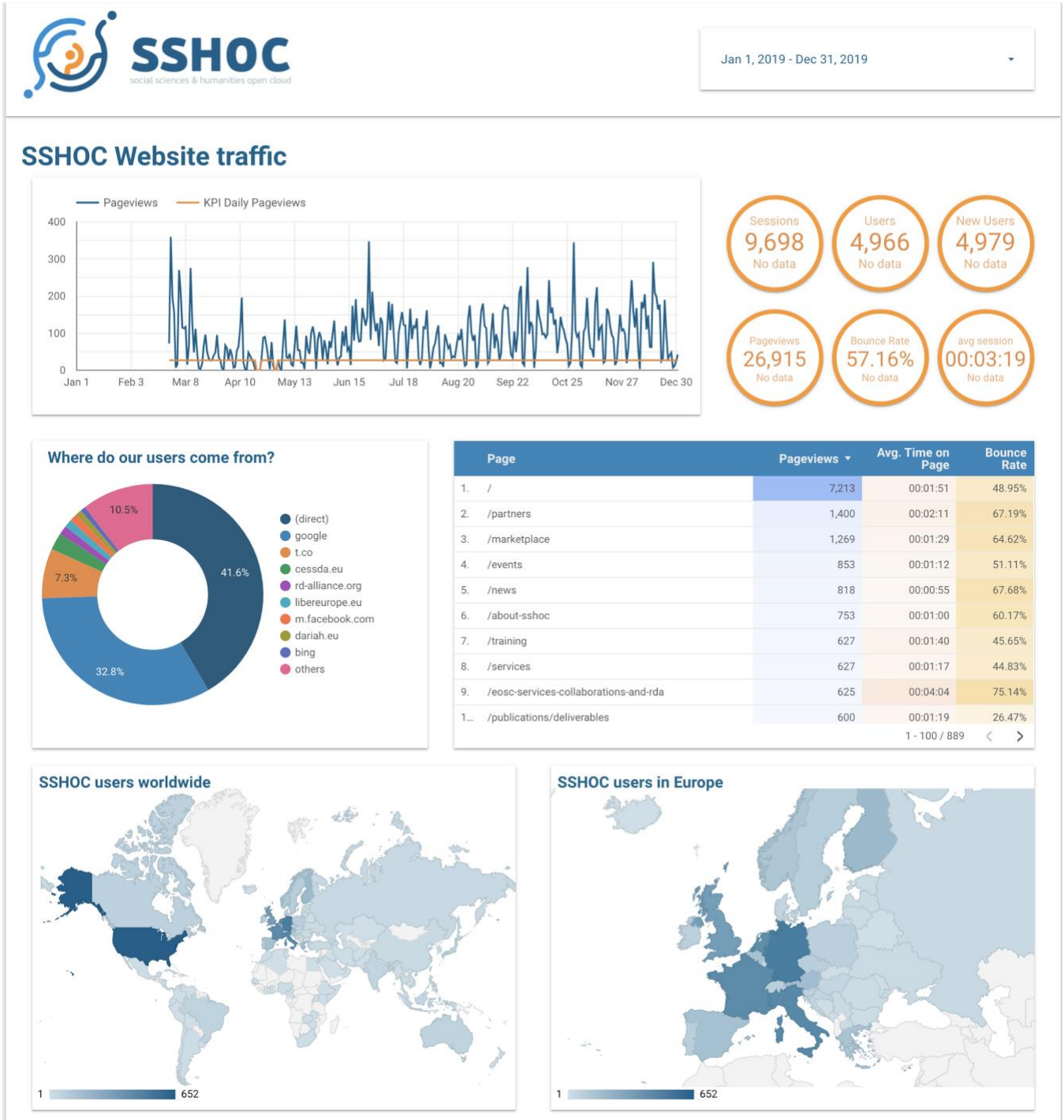


***Figure 2.*** *A part of SSHOC Dashboard presenting M1-M12 traffic*

In alignment with the FAIR principles, all project **reports, publications and presentations** are shared via the website. SSHOC WP1 and 2 decided that deliverables, reports and presentations should be made available via **ZENODO** to improve the sharing of the projects results. The SSHOC ZENODO account was set

up and all deliverables submitted by M12 were uploaded and made available, a total of 12 deliverables, 1 report and 1 poster were uploaded in M1-12[38]. Key results are now easily shared with the wider EOSC for reuse. In addition, SSHOC can now effectively monitor the impact of these publications.

#### 2.2.1.3. TASK 2.3 SSHOC WEB PRESENCE

***MS3 SSHOC web platform launch*** was achieved in M2 of the project, one month before the deadline, to leverage on the buzz around the kick-off meeting in M3. This first iteration was aimed to raise awareness early on in the project on the activities planned and started to realise the social sciences and humanities' part of the European Open Science Cloud. This first iteration included information on the project, connection to EOSC, the marketplace and space for regular news and event updates. Registration for the SSHOC community and the regular newsletter were immediately set in place in this first version.

***MS4 SSHOC web platform second release*** was achieved in M6. This second release expanded the information on the SSHOC marketplace, SSHOC services, SSH landscape, SSHOC training, SSHOC in EOSC. The second release also included a downloadable Communication Kit and a section on publications, where deliverables are digested and published.

In M9 the training network was launched, with a registration mechanism from the SSHOC website, feeding into the GDPR compliant database of the project.

The preparation of ***MS5 SSHOC web platform continuous marketplace integration with SSHOC service offer 1/3: M9, M10, M12*** started in M9. With a dedicated task force with representatives from WP1, 2, 3, 6 and 7 to align with the activities, work and needs of the ongoing work in SSHOC. An inventory of services & tools was carried out across all tasks to define the catalogue of services for the website and the timeline of delivery of each of these services. In addition to the catalogue a page on "who benefits" will be added and all content will be updated according to the advances made in the project.

In Year 1 the website has received **9.600+ visits** from **4.900+ new users**.

### 2.2.2. Note on deviations from the plan and risk monitoring

The only deviation in WP2 was the delay of achieving the Milestone 2 and 5 (iterations of the SSHOC website. Milestone 2 was three months late due to technical issues and consortium decisions, and MS5 is now expected in February 2020.

Information on risk monitoring in WP2 for Year 1 is presented in the table below:

[38] SSHOC Zenodo Community: https://zenodo.org/communities/sshoc/

***Table 2**. Risk monitoring in WP2 - Year 1*

| Risks monitoring WP2 | | | | |
|---|---|---|---|---|
| No of risk (from the DoA) | Description | Have risk-mitigation measures been applied? | Did the risk materialise so far? | Comments |
| 2 | Lack of visibility through media channels (Low) | **Yes** | **Yes** | Targeted messages were sent directly via the multiplier channels of ESFRI landmarks and projects, LIBER and SSHOC liaisons with other EOSC projects. |
| 3 | Lack of engagement at outreach events (Medium) | **Yes** | **Yes** | Targeted messaging is used for each event to attract the envisioned stakeholders for engagement. Each event is promoted via the SSHOC channels and multiplied via partner networks. Interactive elements are included in events to collect input and feedback from stakeholders |
| 5 | Inefficient use of resources: survey creation, translation and storing (Medium) | **Yes** | No | Risk mitigation measure: Apart form the regular calls, additional task force calls were set up, messages via the internal communication platform and personal emails were sent to collect input. |

## 2.3. WP3 - Lifting Technologies and Services into the SSH Cloud

In the first year, most of the WP3 tasks started with their planned work by **executing inventories & surveys** concerning the subject matter of their activities, as such landscape overview of the state-of-the-art guarantees relevant implementation directions that take into account the research needs from the broad SSH field, thus ensuring the usability and added value of the chosen solutions. For example, Task 3.5 concerned with data interoperability made a survey with respect to the main (meta)data formats, and Task 3.4 concerned with citation of SSH resources made a thorough survey of the use of citation in the research data landscape. Moreover, contacts with other tasks in SSHOC and beyond were made to look for synergy and possible common software development.

### 2.3.1. WP3 progress

#### 2.3.1.1. TASK 3.1 MULTILINGUAL TERMINOLOGY

The main objective of Task 3.1 is to provide the **background resources and tools** to be used in the project to make data and services accessible and usable in SSH. In this respect, one of the main aims for SSHOC in the first year is to provide **multilingual metadata and taxonomies** to maximize the accessibility and to improve discovery by non-native speakers. To achieve this, a service will be developed to translate metadata in different languages a SSHOC solution for the vocabulary server and publication platform will be provided to publish the translated terminologies as a catalogue with download facilities and API service.

Based on a survey of already existing solutions and interviews in the community, a **list of criteria** and **features** that the publication platform should comply with concerning editing, linking and accessing the vocabularies, has been provided. The features, then, have been used to evaluate the different existing platforms which allow discovery of research data in SSH infrastructures and to provide recommendations for the SSHOC vocabulary publication platform. This process and results were published in the report on milestone 8 of the SSHOC project, ***MS8 The choice of a vocabulary and publication platform for SSHOC***[39].

In order to create multilingual metadata and taxonomies for discovery (starting from both the inventory and the interviews carried out in Task 3.5, as the main source of information), a first inventory of metadata schema with candidate vocabularies that are currently used and relevant for the research infrastructures involved in SSHOC and managed by the SSHOC main stakeholders has been provided in the report on ***MS10 Inventory of SSHOC metadata sets and taxonomies***[40]. Contacts with CUNI established who will

[39] At the time of this report all the Milestones reports are published within the SSHOC project document repository and available to all project partners.

[40] *MS10 Inventory of SSHOC metadata sets and taxonomies;* available in the SSHOC project document repository, available to project partners.

provide the HAMT tool suite which is being developed for T4.3 and will be used for the (pre)translation of metadata.

#### 2.3.1.2. Task 3.2 Selected SSH Ontologies and Vocabularies

Task 3.2 aims to foster the use of selected global ontologies in the social sciences and humanities, regarding occupational titles, educational categories, sectors of industry, geographical regions, food items, and religions. These ontologies allow to classify elements into standard global classifications, for example the ISCO classification of occupations and its derived social status or the NACE/ISIC classification of industries. These ontologies service the usage of vocabularies for classifying text corpora and predefined response categories for survey questions. The occupational ontologies will address the 20th and 21th century. The partners participating in the task aim to improve and optimize the above mentioned, partially existing multilingual ontologies and aim to develop use cases.

An intermediate result was produced for D3.4 Multilingual ontologies for Occupation, Industry, Regions and cities, Food items, and Religion, with use case. The focus in Year 1 has been on one of the two envisaged topics in Task 3.2 - the multilingual SSH ontology and vocabulary for the second half of the 20th century occupational titles.

#### 2.3.1.3. Task 3.3 Text & Data Mining

The main objective of Task 3.3. is to bring to practice and to adopt where needed the developed natural language processing solutions into the context of social sciences and humanities. In order to provide such demonstration of the algorithms and pipelines potential, in the first year of the project task participants focused on discussions and **analysis of the available datasets** for the MS11 report Selection and establishing access to data-sets for processing which has been shared in the project repository with project partners after achieving the milestone in December 2019. There is a mutual understanding within the task participants that this list can and will be extended as in the course of the project and through consultations with diverse communities more research questions can be brought to attention. Therefore, more opportunities to test and showcase the text and data mining performance can raise at the current work stage.

#### 2.3.1.4. Task 3.4 Making Data Findable by being Citable

The main objective of Task 3.4 is to foster Data Citation by providing a common mechanism to cite SSH data and build stronger links between data and publications. An expected side effect would be to enhance the reproducibility of SSH research, which is not very common nowadays. In this regard, an important point is to be able to give high visibility to the research data used in Social Science and Humanities following FAIR data principles. Deliverable ***D3.2 Inventory of SSH citation practices, and choice for SSHOC citation formats and implementation planning***[41] has been submitted. Task partners participated in the common

[41] SSHOC D3.2 Inventory of SSH citation practices, and choice for SSHOC citation formats and implementation planning was submitted to the EC on the 15 Nov 2019 and is awaiting approval at the time of this report - V1.0: 10.5281/zenodo.3595965.

SSHOC WP3/CLARIAH-DE workshop and liaised with WP6 to organize a round table of experts and a webinar about data citation. Contacts were established with Task 5.2 regarding the connection with Dataverse, with Task 8.2 about the topic of quality, and with WP7 about the integration into the MarketPlace.

#### 2.3.1.5. Task 3.5 Data and Metadata Interoperability Hub

Task 3.5 aims to address interoperability challenges within SSHOC by analysing the SSHOC work plan, following activities, and recommending solutions. In Year 1 the main focus was on the current status of interoperability within SSHOC and stakeholder infrastructures, and the initial collaboration with other tasks/work packages has started. On the basis of **interviews of 16 domain experts** on interoperability issues ***D3.1 Report on SSHOC (meta)data interoperability problems*** has been created. Work on MS14 (Inventory of existing (meta-)data format interoperability solutions) has started as well as work on possible conversion services. Collaboration with Task 3.6 on Switchboard (CNR/ISTI). Initial collaboration with Task 4.7[42] and WP7 was discussed.

#### 2.3.1.6. Task 3.6 Making Data Re-usable and Actionable

Task 3.6 is concerned with creating a SSHOC Switchboard and SSHOC virtual collection registry (VCR). It was intended to generalise and adapt the CLARIN Switchboard and VCR and adapt it to the needs of the other SSHOC communities. In Year 1, an inventorisation of the potential integration opportunities has been planned, as well as the collection of further functional requirements for the Switchboard and VCR. Also, potential collaborations with other SSHOC tasks and projects beyond SSHOC were to be investigated. Both planned milestones ***MS9 Inventory of services & tools, and domain of actionable data-types***, and ***MS12 Implementation plan SSHOC Switchboard and VCR services*** were achieved as planned. Although with a slightly changed scope for MS12 mainly limiting into the SSHOC Switchboard. The work on Task 3.6 is tightly connected with the Tasks 3.4. and 3.5 within the WP3, extending the collaboration to the Task 5.2, which was further discussed in the context of external collaboration with the German CLARIAH-DE project.

### 2.3.2. Note on deviations from the plan and risk monitoring

MS12 was achieved only in M12 (instead of M8), since it could benefit from inclusion of the outcomes of a common SSHOC workshop with the CLARIAH-DE project which is concerned with work on the same type of Switchboard service. The publication of D3.2 was also delayed due to small delays in setting up the writing team with the different partners. The work on MS11 was delayed profiting from a number of events for consultation with relevant communities, and also because of a prolonged hiring procedure at one of the task key partners. MS10, originally due M8, was completed at M10. It was postponed until after the CLARIN2019 Conference and the SSHOC Consortium meeting. These two events have been considered a good opportunity to have extensive discussions with SS Users at large and SSHOC colleagues from other SSHOC WPs.

[42] SSHOC Task 4.7 Modeling the SSHOC data life cycle

Information on risk monitoring in WP3 for Year 1 is presented in the table below:

***Table 3.*** *Risk monitoring in WP3 - Year 1*

| Risks monitoring WP3 | | | | |
|---|---|---|---|---|
| No of risk (from the DoA) | Description | Have risk-mitigation measures been applied? | Did the risk materialise so far? | Comments |
| 4 | Integration of Switchboard and VCR in specific catalogues | **Yes** | **Yes** | Risk-mitigation measure applied: the set of integration goals was selected in alignment with available developer teams. |
| 5 | Inefficient use of Resources | **Yes** | **Yes** | Risk mitigation measures applied:<br>- organisation of additional telcos, personal mails with invitation to submit contributions.<br>- Evaluation of translation quality per language pairs, e.g. the tools for the English-Czech pair is better than (most) human translators. |
| 6 | Too few available software developers | No | No | Risk-mitigation measure: aim at collaborating with other tasks making use of synergies for software development |
| New risk | Problems in hiring staff by some partners | **Yes** | **Yes** | Risk-mitigation measure applied: work plan for the Year 2 of the project includes more intense work involvement of the partners with delayed staff hiring. |
| New risk | Difficulty to decide on a specific infrastructure solution (infrastructure component, technology, methodology) | No | **Yes** | Risk-mitigation measure: Accept multiple solutions/recommendations |

## 2.4. WP4 - Innovations in Data Production

Work package 4 tasks are concerned with the data creation phase of data lifecycle and aims to develop, test and share innovative ways of producing data of interest to analysts and policy makers within the Social Sciences and Humanities with which to populate the Open Science Cloud.

In Year 1 of the SSHOC project Task 4.1 ("A sample management system for cross-national web surveys") team planned to publish a detailed specification of the sample management system (deliverable D4.1 due M3), while Task 4.2 ("Preparing tools for the use of Computer Assisted Translation") team planned to focus on collecting raw data i.e. questionnaires in the repositories of the survey infrastructures and the TMT software[43], in order to convert it to machine-readable data. The objective is to create a database that will have applications in the translation of survey questionnaire. This will help to undergo a transition from mainly human-based translation procedures to technology-intensive translation procedures.

Task 4.3 "Applying Computer Assisted Translation in Social Surveys" agreed on some key decisions e.g. languages, staff and methodological strategies to design experiments which will impact the implementation of the project. Planning the training of bilingual word embedding models (English-German) to be used in the verification of the wave 9 of the SHARE questionnaire, also took place.

For Task 4.4 ("Voice recorded interviews and audio analysis") planned activities were related to preparing guidelines to collect Audio capture data in social surveys (D4.12) while the focal point for Task 4.5 (on social policy APIs for social survey), was for survey methodologists and Social Scientists from the EVS and GGP to work together with data scientists and oral historians from CLARIN to develop a survey module specifically adapted to integrate audio recordings and their processing into traditional data collection processes.

In Task 4.6 "Semantic annotation of Heritage Science Data" mainly focused on the definition of the upcoming features of "Aïoli", a reality-based 3D annotation platform for the collaborative documentation of cultural heritage artefacts[44]

Finally, the activities of T4.7 ("Modeling the SSHOC data life cycle") for the first year of the project aimed at developing a first draft of a Reference Ontology for SSHOC.

### 2.4.1. WP4 progress

#### 2.4.1.1. Task 4.1. A sample management system for cross-national web surveys

The first key point of the task was to deliver a complete functional specification of the sample management system. This was achieved in a joint effort by Sciences Po development team and City University: an initial draft was proposed early 2019, amended through remote collaboration, and the major points finalised at

[43] Translation Management Tool

[44] www.aioli.cloud

a work meeting in London. The deliverable ***D4.1 A sample management system for cross-national web surveys*** was submitted to the EC on 1st November 2019, after another round of refinement.

Since the core functionalities were agreed upon early on, the development team could work on a **first software version** that implements a nominal survey lifecycle workflow. This release is now under internal testing at City University, and discussions about prioritised improvements are underway. A development roadmap will emerge, the next milestone being the test surveys scheduled for March 2020.

#### 2.4.1.2. Task 4.2. Preparing tools for the use of Computer Assisted Translation

Task 4.2 activities aim at improving the translation process of multilingual survey projects by transitioning to human based approaches to technology-based approaches. The first step was to **convert translations into text data (corpora)**. An interdisciplinary team (corpus and computational linguistics and survey methods) of **experts** was built with external collaborators in Norway. They provided expert feedback on building the multilingual corpus. **Key methodological aspects** have been decided in close collaboration with the external team. ESS UPF team has primarily worked on implementing such key decisions and processing questionnaire files into machine readable data:

**Data from the ESS and EVS**

Manual transformation of files and procedures to validate manual conversion have been implemented using Git Graphical User Interface via SourceTree software. Most recent rounds of both surveys will be extracted from the TMT in February 2020 (M16). ESS UPF team is in communication with partner CentERdata about the technical requirements of such data.

**Data from SHARE**

Close collaboration between teams in SHARE, CentERdata and ESS-UPF (leader of the Task) to produce an export of SHARE questions. SHARE is considered the most complex survey because the questionnaires are programmed, so joint efforts were needed to convert them to natural language.

**Data from GGP**

After studying the translation process in the two first editions of the GGP, ESS UPF team decided to only include the most recent wave and to get an extraction from the TMT in the spring of 2020 when most countries would finish the translation step. The reason is that only in this last round the process is similar to the other projects, including early-waves data.

#### 2.4.1.3. Task 4.3. Applying Computer Assisted Translation tools in Social Surveys

This Task has three key components: (1) test translation memories with data built in T4.2, (2) test machine translation in survey research and (3) developing a tool for translation verification.

The first part of the project started with a kick-off meeting in Barcelona (November2019) where project teams from ESS UPF and ESS GESIS met to discuss major **experimental design decisions**. The team decided on the **languages of the translation** memories/corpora. The final selection was three language pairs: English-German & English-Russian & English-French, to ensure (1) that relevant and frequent ESS/EVS languages are covered, (2) a coverage of divers language families, and (3) coverage of language by staff (existing and to be hired).

ESS UPF team received **feedback from the ESS Translation Expert Panel** about ideal characteristics of a translation memory to be used in the translation of survey questionnaires. The Panel brings together experts in survey translation, experienced translators and survey researchers. Their feedback was important because it provided information about the features that a database storing past translations should have.

Regarding the second component, team members of ESS UPF and ESS GESIS were in regular contact to decide the languages of the experiment: English-German & English-Russian, to ensure testing among quite different language (families), given that machine translation (from English) may work better for some languages and less for others.

Current ongoing work, driven both by internal work at UPF and GESIS and by collaboration between the two (email, calls, document reviews) is about specification of research questions, item selection, searching and selecting indicators to measure features in the experiments, further specifying the tools & documentation environment, setting up a concrete roadmap for the experiments in German and Russian language.

The third component of the project started with the selection of the corpora (text data) for training automated models. A general-domain corpora publicly available and an in-domain corpora generated using translations from previous wave of the SHARE questionnaire were used. During the training of automated models, different state-of the art alternatives were created: Unsupervised Mapping, Semi-supervised Mapping and Identical Mapping. As survey texts have not been studied at length, it was important to have several alternatives to be evaluated. The SHARE team evaluated the **performance of the three models** and decided the latter mapping is slightly better than the other two.

#### 2.4.1.4. TASK 4.4. VOICE RECORDED INTERVIEWS AND AUDIO ANALYSIS

This task will collect audio data in the form of voice recorded interviews from the Generations and Gender Survey. This data will then be processed and analysed by colleagues at CLARIN to contribute with automatic speech recognition, speaker attribution, part-of-speech-tagging, named entity labelling and other tools. Survey methodologists and Social Scientists from the EVS and GGP work together with data scientists and oral historians from CLARIN to develop a survey module specifically adapted to integrate audio recordings and their processing into the traditional data collection process.

The researchers in the team discussed the various requirements of both survey infrastructures and linguistic infrastructures and the potential benefits of utilising services from each other's operations. Thus, a big focus of Year 1 of the task was the first scheduled deliverable, ***D4.12 Guidelines for the integration***

***of Audio Capture data in Survey Interviews*** which was submitted to the EC in December 2019. The guidelines will be taken forward and applied within an appropriate survey infrastructure in the remainder of the SSHOC project.

#### 2.4.1.5. TASK 4.5. SOCIAL POLICY APIS FOR SOCIAL SURVEYS

Task 4.5, led by KNAW, will integrate a policy imputation algorithm into social surveys. This will allow that when someone answers in a survey about their income and children, the algorithm would be able to infer whether it is likely that the person receives social benefits, and then ask, within the same interview, whether this is correct. During the first year, activities focused on the development of the **first social policy API tool**, currently hosted on the GGP server. The API allows users to query the OECD Family Policy Calculator in a machine-readable form. The data underlying the family policy calculator is large (more than 5GB) and is coupled within an algorithm which identifies policy benefits across a range of indicators in 35 countries from across the world. The data is compiled by Ministries within OECD member states and validated by economists at the OECD. The OECD is collaborative and supportive of publishing the API and is keen to see the work expand. The data and complexity of the algorithm mean that it is too large to distribute generally or readily load within traditional analytical software. The API allows users to extract responses from the calculator through an internet connection, making the data much more findable, accessible and interoperable than previously. There are a wide range of similar policy calculators in existence and the hope is to further include these through the course of the project.

#### 2.4.1.6. TASK 4.6. SEMANTIC ANNOTATION OF HERITAGE SCIENCE DATA

During the first year, activities on Task 4.6, led by CNRS, mainly focused on the definition of the upcoming features of "Aïoli", a reality-based 3D annotation platform for the collaborative documentation of cultural heritage artefacts[45]. During this year, CNRS managed to complete the **first beta-testing programme** which involved selected actors of cultural heritage scientific and professional community (archaeologists, architects, engineers, conservation scientists, curators and restorers, …). The aim of this initiative was not only to pinpoint technical difficulties, but also to collect feedback on potential features, to be included in the next version of the platform.

The above mentioned topics have been discussed during the WP4 meeting at the SSHOC Kick-off meeting (11-12 March 2019, Utrecht), the WP4 Workshop "Developing the SSHOC Reference Ontology" (21-22 May 2019, Heraklion, Crete), the 2nd SSHOC Consortium meeting (14-15 October 2019, Florence) as well as during several video calls with Task 4.6 members. All these meetings allowed the definition of some specifications on the upcoming features (to be implemented by CNRS) which concern the technical robustness of the platform, the collaboration framework, the managing of controlled vocabularies, the compatibility with CIDOC-CRM[46] (in collaboration with FORTH), as well as the interlinking with other tools for visualising hypothetical reconstruction of archaeological sites (in collaboration with CNR). The results

[45] aïoli. a reality-based 3D annotation platform for the collaborative documentation of cultural heritage artefacts (www.aioli.cloud)

[46] CIDOC-CRM: CIDOC Conceptual Reference Model (http://www.cidoc-crm.org/)

of these first works are reported in the first deliverable (***D4.16 Specification of the new features of the Aïoli platform***), submitted in November 2019. In December, CNRS worked on the definition of a first schema for the mapping of the Aïoli 2D/3D annotation process with CIDOC-CRM, as well as on the managing of controlled vocabularies.

#### 2.4.1.7. TASK 4.7. MODELING THE SSHOC DATA LIFE CYCLE

In the implementation phase of the task, FORTH[47] emphasized on having a workshop as early as possible. T4.7 team used the SSHOC Kick-off meeting in Utrecht for networking with scholars working in the Social Sciences and Humanities and tried to consolidate any contacts made in the following period, leading up to the workshop.

During the Kick-off meeting the team already circulated **a questionnaire** regarding data/metadata, methods and tools used by scholars working in the SSH and were given feedback regarding its relevance and information that needed to be added. The **updated version** of the questionnaire was shared with (i) work-package and task leaders of SSHOC with the request that it be shared with members involved in said work-package/task, and (ii) a few select SSHOC-consortium members, who had previously shown interest in participating in the workshop and/or helping out with the questionnaire. The idea was that we would be able to form an understanding regarding the metadata used across the SSH data lifecycle prior to the workshop, which could be used to inform the ontology. Despite the efforts, the response rate was much lower than anticipated. The **workshop** took place on 21-22 May 2019, hosted by FORTH. The team had advertised it through Basecamp, (internal SSHOC project communication platform), the SSHOC official website and by directly contacting work-package and task leaders.

Despite low attendance, the workshop proved quite fruitful: The ontologies CIDOC-CRM, CRMdig,[48] CRMsci[49] and the PARTHENOS Entities Model[50] were discussed, in order to see if they could be applied to model the research processes particular to the Social Sciences & Humanities. The decision was that even though they can serve as a starting point, the ontologies need to be further extended to form the **SSHOC Reference Ontology** - henceforth, SSHOCro.

The undertakings of the workshop were laid out in an internal Milestone Report ***M19 Consultation with SSH data producers***. The minutes (with links to the presentations) and any conclusions regarding the SSHOCro features, reached in the course of the workshop, can also be found in the SSHOC Document repository[51] and were available to all project partners.

To be able to **follow the workflows, tools and services used by scholars** working in SSH (i) T4.7 team directly contacted the Greek branch of CESSDA (SoDaNet) and arranged to participate in the CESSDA

[47] The Foundation of Research and Technology - Hellas is the lead partner involved in T4.7.
[48] CRMdig: CRM Digital; a model for provenance metadata (http://www.cidoc-crm.org/crmdig/)
[49] CRMsci: CRM Scientific Observation Model (http://www.cidoc-crm.org/crmsci/)
[50] For a definition of the PARTHENOS Entities Model: https://zenodo.org/record/2575465#.XjLx22gzaUk
[51] Internal SSHOC document repository contains the workshop results and conclusions: https://drive.google.com/drive/folders/1saYya5r1pJQRzLnS6KPa9mvb9u__kJaX

Training Workshop: Train the Trainer - Athens 2019[52], and (ii) attended the DARIAH Annual Event 2019[53]. The outcome of the workshop, a report on the progress of the SSHOCro (beta) and examples of the workflows were presented in the SSHOC 2nd Consortium meeting in Florence.

A **first draft of the model** with examples to illustrate the classes and properties linking them to one another was shared with SSHOC stakeholders for evaluation on 20 December 2019. A number of SSHOC consortium members were contacted directly via email, as they were involved in overlapping tasks - especially WP7 and T3.5.

### 2.4.2. Note on deviations from the plan and risk monitoring

The first targeted Milestone (MS16) was a complete extraction of survey questionnaires from the Translation Management Tool software (CentERdata). It has been moved to the 1st quarter of 2020, deviating from the original plan. An implication is that D4.3 Survey specific parallel corpora has also been moved to second quarter of 2020. Collecting the raw data has required more time than expected, but it does not affect the timeline of the project. The deliverable is only a prerequisite for using the data in other tasks planned after 2nd quarter of 2020.

In Task 4.5 Delays at data preparation stage in the OECD[54] meant that the deliverable is delayed by approximately one month, however this does not affect the timeline within the rest of the task, nor work anywhere else in SSHOC. The deliverable is only a prerequisite for integration in the GGP which is not due until M36 of the project.

In Task 4.7 D4.18 SSHOCro beta version was postponed two months. The extension gives the time necessary for reviewing as it clashed with the Christmas holidays. Furthermore, more time was needed for testing the meta-level schema with social science working research scenarios.

The feedback of consortium members working on the SSHOC Marketplace was also needed - a coordination with T7.1 and T7.4 being foreseen by the DoA. The delays are primarily due to a low response rate to the questionnaires distributed and a low workshop participation rate of planned and targeted scholars from the SSH and tool developers/service providers[55].

Information on risk monitoring in WP4 for Year 1 is presented in the table below:

[52] Athens 2019: http://sodanet.gr/en/training/workshops/cessda-training-workshop-athens-2019

[53] DARIAH Annual event 2019: https://www.dariah.eu/event/dariah-annual-event-2019/

[54] Organisation for Economic Co-operation and Development

[55] Despite the call for participation having been published quite early, participation was less than what anticipated – a number of conflicting events happened during or around the selected days, which were (almost) immediately preceded by the Easter Holiday. Unfortunately, rescheduling was not an option, as it would mean postponement of the workshop indefinitely. To engage more SSH researchers, the team might have to wait until after the summer holidays; in which case, the SSHOCro would fall behind causing problems to the partners working on the SSHOC Marketplace.

***Table 4.*** *Risk monitoring in WP4 - Year 1*

| Risks monitoring WP4 | | | | |
|---|---|---|---|---|
| No of risk (from the DoA) | Description | Have risk-mitigation measures been applied? | Did the risk materialise so far? | Comments |
| 4 | Insufficient cooperation from or access to SSH tool development teams | **Yes** | **Yes** | Although coordination of tasks started early, it took a time to establish collaboration between partners. Mitigation measures applied: Initiated contacts early and coordinate plans (T4.2), additional contacts established. |
| 6 | Too few available software developers | **Yes** | **Yes** | Scarcity of developers on the job market. Building technical team took longer than originally planned. Work was prioritized, rescheduled (T4.1 and 4.2) |
| 7 | Technical difficulties of complex software development | **Yes** | **Yes** | Mitigation measures applied: Software environment carefully tested involving users (partners in the Task T4.3). |
| 11 | Understaffing of research infrastructures | No | **Yes** | In T4.2 & T4.3 Staff has been a key challenge in the first year of the project, longer selection processes were needed |
| 1 | Difficulties in cooperating with other ERICs and Pan-European projects | **Yes** | No | Initiated contacts early and coordinated plans |
| New risk | Underestimated the time that the specification of the tool would require | **Yes** | **Yes** | Start development around core features without waiting for exhaustive specifications (T4.1 and T4.3) |
| New risk | Lack of engagement from the SSH community | **Yes** | **Yes** | In the Taak 4.7 there was a low response at Questionnaire circulation, despite the team's efforts and low participation at workshop planned due to conflicting events that could not be combined. Mitigation measures applied: Attendance of key events to meet partners and stakeholders directly |

## 2.5. WP5 - Innovations in Data Access

This work package facilitates innovations in data access and intends to provide tools and services for intelligently open data for the SSH domain to be incorporated into the EOSC cloud. It has a focus on data collected with and along social surveys and archaeological data.

For the period of M1 to M12, the aims of WP5 were one milestone on the laboratory processing of biomedical data (Task 5.1), two milestones for the specification of data repository services based on the Dataverse software (Task 5.2), one deliverable on ethical considerations in biomedical survey data (Task 5.1) and one deliverable on the specification plan for software applied to an archaeological case study. While not all tasks were involved in milestones and deliverables in this first reporting period, all tasks are scheduled to run over the whole project period from January 2019 to April 2022.

Communication was mainly done via Basecamp, as the project's main communication platform, and email. In 2019, two F2F meetings took place and were held at the consortium meetings in Utrecht, March 2019 and Florence, October 2019 to exchange across tasks and foster further collaboration among WP5 tasks and also beyond WP5 with other work packages.

### 2.5.1. WP5 progress

#### 2.5.1.1. Task 5.1 Legal, Ethical and Technological Issues of Access to Biomedical Data

Task 5.1 aims to develop strategies which provide access to innovative biomedical data collected in the context of social surveys and link them to traditional survey data. This will be exemplified with two cases: Dried Blood Spot Samples (DBSS) and accelerometer data, both collected in SHARE. DBSS data have been collected already in 2015, but require special ethics considerations and safeguards: This has been worked on in 2019, leading to the delivery of ***Milestone 21 Protocol of laboratory processing of DBSS data*** in July 2019, and the preparation of Deliverable 5.1: Guidelines for ethics considerations in making biomedical survey data, to be expected end of January 2020.

The collection of the accelerometer data started in October 2019 and is still continuing, in ten SHARE countries[56]. Task 5.1 team members have contacted the international Prospective Physical Activity, Sitting and Sleep Consortium (ProPASS)[57], to collaborate with other research disciplines in developing algorithms and processing strategies for the scientific use of this data. Researchers involved in this task met f2f on a regular basis in 2019 to ensure progress.

[56] The participating SHARE countries are Belgium, Czech Republic, Denmark, France, Germany, Italy, Poland, Slovenia, Spain and Sweden. Countries were selected based on geographical variation to ensure a geographical and cultural spread across Europe.

[57] ProPASS is an international research collaboration platform of cohorts using thigh-worn accelerometery to explore the effects of physical activity, posture, and sleep patterns on a wide range of health outcomes. https://www.propassconsortium.org/

#### 2.5.1.2. TASK 5.2 HOSTING AND SHARING DATA REPOSITORIES

Task 5.2 started with a virtual kick-off on 28 January 2019, followed up by a F2F meeting on 11-12 April at DANS in the Hague. During that meeting Milestone ***MS25 Selection of additional functionality for data repository service to be developed*** (List of functionalities that will be developed for repository service, based on common needs of CESSDA, DARIAH and CLARIN partners) was accomplished. Division of work and procedures for development were defined and agreed. The task members have an online task meeting every month to discuss developments and other issues.

Task 5.2 makes use of the Dataverse software[58] and has delivered the first version of the continuous integration/continuous pipeline (***MS24 Implementation of standard Dataverse installation on EOSC*** - Standard installation of repository service online). It runs on Jenkins, a free and open source automation server software, in the CESSDA cloud and allows to build all Docker[59] images and automatically deploys all dataverse components like database, search engine and the application itself.[60]

**Functionalities developed in 2019:**

- Integration of previewers. Spreadsheet viewer was developed by PSNC, tested by Harvard's Institute for Quantitative Social Science and integrated into the master branch, starting from version 4.18.[61]
- Other previewers are available in the SSHOC Dataverse toolbox, delivered as a separate Docker image (PDF, Spreadsheet, Images, text, html, video, audio, data explorer).[62]
- pyDataverse. AUSSDA is working on the backbone of DDI (Data Documentation Initiative metadata standard) import tool that will allow migration of XML files from NESSTAR to Dataverse.[63]
- DDI converter tool. Developed by DANS. It makes use of pyDataverse. It separates mappings from the conversion process.[64]
- Translation tool, weblate as a service, by DANS. Weblate based pipeline to get all translations of user interface and metadata and automatically synchronize them with file translations of Dataverse.[65]

[58] Dataverse is an open source web application to share, preserve, cite, explore, and analyse research data (https://dataverse.org/).
[59] Docker is a platform-based service that enables organizations to seamlessly build, share and run any application by delivering software in packages called containers (https://www.docker.com).
[60] Installation for preview is available at https://dataverse-sshoc-staging.cessda.eu/ (username / password: "cessda" / "CESSDA2018"). Source code is available at: https://bitbucket.org/account/user/cessda/projects/DVS (restricted CESSDA bitbucket repo) and https://github.com/Dans-labs/dataverse-kubernetes
[61] https://github.com/IQSS/dataverse/releases/tag/v4.18
[62] Source code available at: https://github.com/IQSS/dataverse-docker/tree/glam
[63] Source code: https://github.com/AUSSDA/pyDataverse
[64] Source code available at: https://github.com/IQSS/dataverse-ddi-converter-tool
[65] Source code: https://github.com/GlobalDataverseCommunityConsortium/weblate-docker; https://github.com/GlobalDataverseCommunityConsortium/dataverse-language-packs

- Integration of Apache Taverna in Dataverse workflow. ISTI is working on the integration of Taverna VRE (open source Workflow Management System) that allows Taverna to use Dataverse for storing and managing workflows.[66]
- Custom metadata fields support. UGOE and DANS are working on custom metadata support (SOLR and Kubernetes).[67]

#### 2.5.1.3. Task 5.3 Legal Issues of Innovative Data Access

Since the first Milestone of task 5.3 is set to be achieved in M14 (February 2020) there have been several Skype-meetings between the partners in the task: CNR, DNA and NSD. The team has also been in contact with representatives from DARIAH, as the work in Task 5.3 is somewhat aligned with the work of DARIAH and CLARIN in formulating a common code of conduct.

The team has had a legal expert at NSD examine which **articles in the GDPR** are most relevant for research. The work will be built on the Nordic project led by NSD, that compared the GDPR and the national personal data acts in Norway, Denmark, Sweden and Finland[68]. In this project the provisions with highest relevance for research conditions in the new legal framework were selected and it was investigated if Norway, Sweden, Finland and Denmark had different interpretations and solutions – and if variations would influence conditions for research in general, data sharing and cross border research.

Going further, the team wishes to do the same with Italy and Germany as examples. Thereafter it will be examined how the GDPR is affecting the researchers in their daily lives, and researchers from different countries will be interviewed to get their point of view. An **interview guide** has been produced and the same guide will be used in different countries.

Furthermore, Task 5.3 has been in contact with different organizations in Europe working on the GDPR and research, such as the BBMRI-ERIC's work on a Health and Life Science Code of Conduct about potential collaborations. Task 5.3 has started planning the stakeholder workshop in cooperation with Task 6.2 *Task Fostering Communities: Engaging New & Existing Users* and intends to hold the workshop at the 4th SSHOC Consortium meeting.

#### 2.5.1.4. Task 5.4 Remote Access to Sensitive Data

Task 5.4 will enhance and extend the infrastructure for secure remote access to research data. The primary challenge is to address the policy and mutual agreement challenges of providing researcher access to sensitive, dispersed data, online across disciplinary and national borders. Work has advanced in several areas:

- There is a draft document defining **access conditions for sensitive data**. This was produced for GESIS (social science data) and is being extended and adapted for humanities data (CLARIN).

[66] Source code: https://github.com/vre4eic

[67] https://github.com/Dans-labs/dataverse-kubernetes/tree/85-custom-metadata/custom

[68] The full position paper of this initiative can be found here: https://nsd.no/personvernombud/dok/position-paper-new-legislation.pdf

- An inventory (spreadsheet) of technical and platform **solutions for sharing sensitive data** has been drafted. This includes refining the correct, relevant characteristics that need to be compared.
- An outline of the **minimal requirements for access to sensitive data** solutions has been prepared. This has been drafted for humanities data, and a parallel document for social sciences will be done before deciding if, or how, to integrate.
- Two **bilateral pilots** are being negotiated, GESIS-UKDA, and GESIS-IAB. Contracts for both have been sent to the relevant institutions, are being reviewed, and installation at UKDA should be possible in the first quarter of2020.

Significantly, sharpening the focus of what this task needs to deliver was succeeded, and this has resulted in a revised Deliverable D5.10[69] as the main output of our task: Requirements and Recommendations for Remote Secure Access in the Social Sciences and Humanities, with these subsections:

- Landscape and definitions of varieties of Remote Access
- Minimal requirements for RSA - (Data Confidentiality Levels for Sensitive Data)
- Platform Assessment and Recommendations - these two sections will be written based on Milestones 28 and 30, namely, Assessments of Existing Platforms, and Recommendations for a Platform for Further Expansion.

#### 2.5.1.5. Task 5.5 ESS as a Service: A Pilot making Cross-national Survey Data FAIR

In this task, by the example of ESS data, it will be showcased how FAIR open cloud storage data can be achieved for a large international longitudinal survey through well-defined procedures[70].

The first year of the project has encompassed the project initiation phase, to ensure a solid basis for the programming stage. A specification document has been drafted and will serve as the main reference document for Task 5.5. It will be updated continuously.

The work that has been carried out within the first year has demonstrated that there are no ready-made solutions that can be used for making ESS data available via the ESS repository. However, several systems were identified, which we will built upon and which will be implemented where suitable:

(1) The data and metadata will be structured in line with the international DDI standard
(2) DataCite/DOI will be used to ensure findability

[69] The revised description of deliverable is to be added to the 1st SSHOC Amendment

[70] NSD will develop a new repository for ESS data with DDI as conceptual standard. The repository will integrate data, metadata and para data. Use of unique identifiers for researchers and automated allocation of persistent identifiers will allow reuse of original and tailored datasets. With double entry user authentication, data will be available to users in a GDPR compliant and secure environment. The work will result in EOSC compatible ESS data offered as a service by spring 2022. By treating data at datum level automated linking between data and documentation at both input and output stages will be ensured and thus shortening the distance between survey and research.

(3) ORCID is intended to be used to identify researchers/authors/publications double level authentication in order to comply with the GDPR
(4) Building blocks from NSDs ongoing infrastructure development projects, funded by the Norwegian Research Council (repository from NORDi[71] project and basing on NSD's RAIRD[72] technology as used in the Microdata[73] and NORDi project) will be used.

Based on the specification document and the outcomes of a joint project meeting with the NORDi and Microdata projects, foreseen for the first week of February, developers will start building the repository during the first months of 2020.

#### 2.5.1.6. Task 5.6 Issues in Providing Open Data in Heritage Science and Archaeology

Beneficiary 8 UoY-ADS has worked with E-RIHS[74] and the broader Heritage Science and archaeological communities to **collect information** on the current information landscape, including the types of data and the challenges in making data FAIR across different European countries. This exercise forms the basis for D5.15 Report on opening access to research data in the Heritage Science and Archaeology domain (Open data in Heritage Science and Archaeology). In conjunction with UCL and SSHOC WP6, responsible for community engagement, the team has been **organising an international 3-day meeting** on data re-use to be held in York 20-22 April 2020. This is being held as a joint event between SSHOC and the SEADDA COST Action[75] and will provide an excellent means for SSHOC to engage with an **advanced domain-specific user community**.

The main effort towards the second part of this task is planned to begin soon after March 2020, to coincide with the appointment of a dedicated research fellow. In preparation for this, Beneficiary 20 NG has worked to ensure that the two data sources and the initial semantic mapping models, that will be the main focus of the task, are accessible and usable.

For the first data source, the Raphael Research Resource[76], this work has involved **updating the administration functions** of the original 2007 system, to accommodate developments in the PHP programming language and a few security concerns, in order for the graphical user interface to be

71 NORDi project - Norwegian Open Research Data Infrastructure, a major NSD (LTP 1.1.) project aiming at rebuilding the infrastructure for research data in Norway on the basis of the Open Access principle, thus making it easier to find, use and share research data.
72 RAIRD Project - Remote Access Infrastructure for Register Data by Statistics Norway (SSB) and Norwegian Social Science Data Services (NSD) which aim was to establish a national research infrastructure providing easy access to large amounts of rich high-quality statistical data for scientific research.
73 Microdata project: https://microdata.no/en/about/
74 E-RIHS is the European Research Infrastructure for Heritage Science - http://www.e-rihs.eu/. In SSHOC Project represented by Beneficiaries: CNR, DAI, CNRS, FORTH, and UCL.
75 Website of Cost Action on Saving European Archaeology from the Digital Dark Age:: https://www.cost.eu/actions/CA18128
76 Website of Raphael Research Resource: https://cima.ng-london.org.uk/documentation

accessible for evaluation. This work has also involved **updating the old SPARQL end-point** for the data held within the resource[77] to replace the triple store and migrate all of the code to GitHub, in order to make it easier for the new researcher to access and reformat all of the stored data[78]. The updating of the SPARQL[79] end-point was carried out in collaboration with an AHRC funded project examining Linked Conservation Data[80].

The second data source, a database of preparation layers[81], has not needed as much preparation, as it is actively being exploited as the output of the IPERION-CH[82] (H2020) project and still being developed. The initial semantic mapping models, again outputs of the IPERION-CH (H2020) project, were initially only available as a prototype description[83]. Work has been carried out to **convert and format** them as **re-usable and editable** tab separated text files. Additional effort has been carried out to start creating an open web resource to visualise these initial semantic models, with the intention of encouraging discussions around their further development. **Development of a new website** on **modelling examples** is in progress and will be specifically designed to dynamically generate interactive presentations of the initial semantic models, thus allowing for easy updates in the future[84]. Future work will involve populating a dedicated GitHub page[85] with the text version of the models to allow more open discussion and documentation (versioning) of their development and use within SSHOC.

#### 2.5.1.7. Task 5.7 Open Linked Data. Archaeological Case Study

Task 5.7 is about creating a use case within the field archaeology where research data is brought online and into the cloud. The goal in the first year was to **identify possible topics for the use case** and to formulate a plan for the creation of data and software based on this. Beneficiary 14, DAI, first suggested an archaeometric topic, namely the study of diseases in archaeology as conveyed by deformations of bone fragments. The DAI also wrote a **first draft of a plan** based on this topic, as a way to Deliverable D5.17 (Archaeology case study: Implementation plan for the software module). One point of the plan was the harmonization of vocabularies used in labs.

After the actual counterparts in the respective partner groups were identified, it turned out that the initially chosen topic would be a suboptimal match with the areas of expertise that Beneficiary CNR Rome and CNR Lecce could offer. The discussions among the project partners led to a **strong shift in focus** so that

77 Website of Linked Conservation Data Example: https://rdf.ng-london.org.uk/workshops/lcd
78 Source code available at: https://github.com/jpadfield/bg-lcd-example
79 SPARQL Protocol, RDF Query Language by World Wide Web Consortium (https://www.w3.org/TR/rdf-sparql-query/)
80 Website of Linked Conservation Data funded by AHRC: https://www.ligatus.org.uk/lcd/
81 Portal of collaborative research resource of Iperion: https://research.ng-london.org.uk/iperion/
82 http://www.iperionch.eu/
83 IPERION-CH Deliverable D8.5: Completed example of prototype designs for integration of various types of documentation and analytical data generated for a single object: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c8d1c525&appId=PPGMS
84 Website (under development): https://rdf.ng-london.org.uk/modelling/
85 https://github.com/jpadfield/cidoc-crm.examples

the project partners' expertise in virtual reconstruction technologies now plays a central role. Correspondingly, on the DAI side it was decided that the best fit would be to combine this with further development of their ChronOntology gazetteer[86] for temporal terms, moving the focus from research data to norm data. A main goal is now to **connect existing infrastructures** and to create easy-to-use workflows. As a consequence, the earlier draft of the project plan needed to be rewritten completely, which led to a delay in D5.17 Implementation plan for the software module (Archaeology case study). The new version of the plan is being actively worked on. Since the use case is supposed to be of high value for concepts such as Linked Open Data, another focus was put on data interoperability. More specifically, work is underway to align D5.17 with the findings of the Interoperability Deliverable from the E-RIHS PP project.

### 2.5.2. Note on deviations from the plan and risk monitoring

D5.1 could not be finished in M10 because of a delay in the processing of the SHARE biomarker data, caused by the necessity of additional validation tests as well as by practical obstacles in the application of the GDPR to these types of data. D5.2: Data access protocol for DBSS data, linked to survey data, conforming FAIR principles (Access to biomedical data) could possibly be affected by this delay, as that deliverable builds upon D5.1. This could cause a delay of about three months. No other deliverables are expected to be affected and we do not perceive any other negative consequences of this delay for Work Package 5.

D5.17 had been delayed due to a shift in focus for the archaeological use case, where the project partners' expertise in virtual reconstruction technologies now play a bigger role in the design of the user interface. In addition, since another focus of the archaeological use case will be on data interoperability, it was decided to allow more time for aligning D5.17 with the findings of the Interoperability Deliverable from the E-RIHS PP project, which was due at the end of 2019 but has been delayed to early 2020. No negative consequences of this delay were perceived for the rest of the Work Package or for the second Deliverable (D5.18) of Task 5.7, which will build on D5.17 but is due only in M38.

Information on risk monitoring in WP5 for Year 1 is presented in the table below:

***Table 5.*** *Risk monitoring in WP5 - Year 1*

| **Risks monitoring WP5** | | | | |
|---|---|---|---|---|
| No of risk (from the DoA) | Description | Have risk-mitigation measures been applied? | Did the risk materialise so far? | Comments |
| 7 | Technical difficulties of complex software development | **Yes** | No | Mitigation measure applied: hiring scientific developers |
| 8 | Unforeseen legal (GDPR and national) barriers to data sharing | No | No | |

[86] ChronOntology gazetteer web service, see: https://chronontology.dainst.org/info/about

## 2.6. WP6 - Fostering Communities, Empowering Users, & Building Expertise

Main goal of WP6 in the first year of the project was to set the basis and initiate the procedures of raising awareness, familiarising with the project, broadening the SSHOC network of user communities and empowering them through building expertise.

As planned, WP6 through T6.1 and in close collaboration with all other tasks delivered a sound and data-driven engagement strategy, as well as a building expertise strategy. Both documents have been designed and delivered in a way that gives emphasis on strategic coordination and synergetic approaches across stakeholder groups and communities, taking into account the current stakeholder landscape that is relevant for SSHOC, existing networks and tools, as well as gaps that need to be addressed and possible risks. At this period an inventory of existing learning materials has been delivered as planned and explained below under T6.3.

The milestone of this period on launching the community engagement strategy has been successfully achieved, although with a minor delay, in order to take into account, the kick-off of relevant project tasks, EOSC developments and events.

Running horizontally through the project, WP6 aimed and managed to establish close collaboration with all other work packaged and especially WP2, in hosting shared meetings, co-organising activities and co-establishing procedures across the project and co-create supporting materials and tools to facilitate planning and tracking of dissemination, engagement and training activities throughout the project.

### 2.6.1. WP6 progress

#### 2.6.1.1. Task 6.1 Mapping the landscape and developing strategies to foster communities & build expertise

Task 6.1 formally started in M1. A key focus of Task 6.1 was to develop coherent and effective strategies to empower user communities with the necessary knowledge, skills, and expertise required to appropriately, effectively, and efficiently use and contribute to the SSHOC resources - both from disciplinary and cross-disciplinary perspectives. Task 6.1 worked on two deliverables in the past 12 months and it formally ended in December after the completion of the two deliverables. T6.1 will, however, continue to regularly check the progress on the KPIs identified in the strategies through the monitoring of activities of the other tasks under WP6.

***D6.1 Community Engagement Strategy*** was submitted on M7, with one-month extension, to take into account other project tasks. D6.1 outline how the project engages with the SSHOC community coherently and efficiently, thereby fostering community engagement, developing a data-sharing culture which follows the FAIR principles, and giving user communities knowledge, skills and expertise so that they can use and contribute to SSHOC resources.

The strategy explains the motivation for engaging with the identified stakeholder groups as in D2.1 SSHOC Overall Communication and Outreach Plan, outlines key messages and value proposition for each group and describes the most effective channels and approaches to be used.

***D6.2 Building Expertise Strategy*** was submitted on M12 and builds upon the D6.1 Community Engagement Strategy, describes an initial plan for building expertise among user communities and provides a framework for SSHOC training activities by identifying and contextualizing training initiatives, gaps and target audiences and by outlining training methodologies and planned training-related activities. All project partners involved in T6.1 contributed to the content creation of these strategies according to their assigned project activities and area of expertise.

#### 2.6.1.2. Task 6.2 Fostering Communities: Engaging New & Existing Users

Building on the needs defined in the D6.1 Community engagement strategy, task 6.2 launched the strategy by initiating procedures on getting SSHOC stakeholders engaged in the development of SSHOC services, as well as raising awareness on tools and services being developed. These procedures were described in ***Milestone 38 Launch Community engagement strategy.*** In order for the task to reach a maximum output of its work, it was needed to collect detailed information from all relevant project tasks on their concrete products, their timeline of development and brainstorm on appropriate ways to engage stakeholders in the development of the SSHOC services, also based on information as planned in the Grant Agreement. For this purpose, an internal survey was conducted in September 2019. Results of the survey were presented at the 2nd SSHOC consortium meeting in Florence, with more detailed information on T6.2's aims and requesting further collaboration. The results of the survey and the goals set in the GA were used as a basis for more detailed interviews with the project task leaders and the establishment of permanent links between them and T6.2 participants, in order to regularly monitor progress and needs.

Task 6.2 aims to organize a minimum of 6 geographically distributed workshops, targeting SSHOC stakeholders. Interviews with task leaders helped to clarify their needs and set up actual planning and developments, but also identify primary target audiences and disciplines at this stage of the project. Having in mind targeted events identified by WP2 and WP6, a list of **events/conferences was suggested**, where SSHOC workshops for 2020 can potentially take place. Holding SSHOC workshops at larger international events in Europe, where target audiences are already expected to participate, will ensure and maximize impact. Geographical distribution is also taken into account. In this sense, **abstracts of proposals of workshops** for raising awareness have already been submitted to the team. Through the process of identification of needs, it was agreed to keep a level of flexibility in terms of number of activities and formats, in order to better cover the needs of the project and the targeted audiences, but in accordance with the set KPIs. T6.2 will also support the organization of awareness raising webinars (minimum 6). The planning of first webinars (at least 3 in 2020) has started.

Among events to be organized by T6.2, in collaboration with WP2, are also the midterm Stakeholder Forum (mid 2020) and the Final Conference (end 2021). A **purpose and scope document** for the Stakeholder Forum has been prepared in cooperation with WP2. In order to reach a larger audience, we proposed to

collocate with a larger event. The review of possible events and planning started in M11. On the basis of the successful SSHOC event that took place before the RDA 14th Plenary, **close collaboration with other EOSC clusters** is being explored for these two events.

#### 2.6.1.3. Task 6.3 Empowering Users: Training Materials and Online Learning Paths

The main focus for the first months was on collecting data for preparing and submitting the deliverable ***D6.7 Inventory of existing learning materials***, due on M12. As the deliverable D6.2 Building Expertise Strategy informs the work for T6.3, there was also involvement by this task.

For the purpose of collecting information about training materials for the inventory, a bespoke simple **web application** was developed, to allow user-friendly data entry, as well as good navigation through and analysis of the collected data, serving as empirical basis for the D6.7 deliverable. Starting with M6, training material sources were continually manually collected. The resulting set of more than 30 sources of training material served as starting point for the inventory. To fill the database a curation sprint was organised, starting in M10. All project partners involved in Task 6.3 joined for this and picked selected materials from training material sources. In preparing the curation sprint the other tasks of WP6 were involved in sharing experiences.

For D6.7 an analysis of this collection was made, providing a qualitative and quantitative **overview of the landscape of training materials in the SSH cluster**. This distant view shows potential for interconnecting information. On detail level we identified corresponding training material, that may serve as a starting point for the creation of learning paths. This also supports the identification of SSHOC-scope training materials. D6.7 was successfully finished and submitted on 19/12/2019.

The inventory of training sources and materials as well as the analysis will support the organising of the training events in WP6. It also serves as an important input for the work in WP7 on the marketplace. Much of the identified training material is being considered to be integrated into the SSHOC marketplace platform. This is something that will be discussed more in detail with WP7 in the future.

The next focus is the development of a strategy for the provision of training material for the SSHOC infrastructure. This builds up on the outcomes of D6.2. **The analysis of the training material** shows a good potential to establish learning paths. How this can be solved on a technical as well as organisational layer will be the main task for the next few months. Connections to other WPs are already in discussion, main line of interaction being the training events, co-organized by WP6, but with content coming from the individual WPs and the accompanying training material being picked up by T6.3. Based on D6.7 there is also a **gap analysis** in progress, to identify which new training materials need to be developed.

#### 2.6.1.4. Task 6.4 Building Expertise: the SSHOC Training Network

The first activity of T6.4 is to **identify organizations** to become the initial nodes of the SSHOC Train-the-trainer Network and assess their training capabilities/potential. The SSHOC Train-the-trainer network was launched at the Open Science Fair in Porto in September 2019. The **start of the training community** was

announced on the SSHOC website, where interested parties can also find the sign-up form to join the network. The launch of the training community was also presented in a **YouTube video**. Since the launch of the network, 22 members have already joined the SSHOC training network.

Prior to the launch in M9, WP6 partners were also present at the LIBER2019 conference in June 2019 and the Digital Humanities 2019 conference in July 2019, where the SSHOC training activities and the Train-the-trainer network were presented, in collaboration with the LIBER working groups on Digital skills for library staff and researchers and on Digital Humanities and digital cultural heritage. In addition, at several occasions where SSHOC engagement and training activities were presented during the first year of the project, community input was gathered through mentimeter on potential existing training nodes that may be of interest for our network. This feedback will be used in the upcoming year to further develop and tailor the network to build and strengthen our community.

The second activity of task 6.4 is to develop a SSHOC train-the-trainer toolkit. The toolkit will consist of various elements, such as training resources, exercises, guidance for setting up training and standard evaluation forms for SSH trainers. In preparation of the toolkit, WP6 has worked on **an inventory of existing training materials** (T6.3, D6.7). An overview of existing training hubs, materials and topics they are currently covering was created, as well as an **overview of SSHOC-relevant training initiatives** (T6.1, D6.2). The hubs include training initiatives of projects such as EOSC-Hub, CLARIN, Parthenos, CESSDA, OpenAIRE and Elixir. The results of these efforts will be used to further guide the development of the train-the-trainer toolkit.

#### 2.6.1.5. Task 6.5 Coordinating targeted training in the Social Sciences and Humanities

The main activity of Task 6.5 includes the coordination of training events with the main goal of maximizing the impact of SSHOC on the SSH communities. These events on six predefined topics from the GA are being delivered in person at workshops, and online through webinars that are related to the topic of the workshop.

The first months of Task 6.5 were dedicated to the **identification of best themes**, **audiences, lecturers** that would fall under each of the six broadly determined topics. This was followed by **identification of best venues** for a specific event. In order to maximize the outreach, the events are collocated at other events that are central to the targeted community wherever possible and built on the cooperation with other training initiatives and research networks. At this stage, the T6.5 partners have leveraged the outcomes of the Building expertise strategy (D6.2) and engaged in close collaboration among all the T6.5 partners who cover different fields in the SSH domain and are thus able to highlight the main needs of specific communities. The result was an outline of training events in the project's timeframe. The plan envisages six workshops to be delivered by the end of 2020 and six webinars by the first half of 2021. Two of these events have been delivered to date with their follow-up webinars to be conducted in spring 2020. With the help of WP2, Task 6.5 ensures timely promotion of the events as well as publication of post-event blogs that bring an overview of the event highlights and list useful links to SSHOC and external resources.

### 2.6.2. Note on deviations from the plan and risk monitoring

Small deviations from the WP6 plan took place, as stated above. D6.1 was delayed from M6 to M7, in order to take into account relevant activities of other SSHOC work packages, as well as general EOSC developments. MS38 was also delayed from M7 to M9 for the same reasons.

Information on risk monitoring in WP3 for Year 1 is presented in the table below:

***Table 6.*** *Risk monitoring in WP6 - Year 1*

| **Risks monitoring WP6** | | | | |
|---|---|---|---|---|
| No of risk (from the DoA) | Description | Have risk-mitigation measures been applied? | Did the risk materialise so far? | Comments |
| 3 | Lack of engagement at outreach events | **Yes** | **Yes** | WP6 introduced interactive sessions to maximize engagement at events. |
| 5 | Inefficient use of resources: survey creation, translation and storing | **Yes** | No | Risk not applying for the moment. However, to prevent this in WP6 relevant databases were set up, an internal survey circulated and regular contact established with all WPs to identify engagement existing networks, tools and needs for all project partners/WPs. |
| 9 | Delay of crucial EOSC cloud infrastructure outside the project | **Yes** | No | D6.1 and MS38 had minor delays to take into account EOSC development and there is the provision of periodical monitoring of KPIs to prevent this situation |
| 14 | Delay in development of SSHOC services might affect training schedule | **Yes** | **Yes** | Certain tasks did not provide any output suitable for training activities during the development of targeted training program meaning that their resources are currently not supported by targeted training, but continuous communication is held with all WPs in order to promote as many SSHOC resources as possible. |
| 15 | Geographical distribution of partners not sufficient to ensure European coverage for national training nodes | **Yes** | No | The call to join the SSHOC training community is open to the public and the geographical distribution of awareness raising and training activities aim to prevent this issue. |
| 16 | Specialized training to be delivered by partners outside WP6 due to their expertise | **Yes** | **Yes** | Experts needed to be scouted across all WPs in order to obtain well-distributed expertise, engage various SSH communities and ensure broad outreach. Additional efforts were invested to bring all the involved parties on the same page (both in terms of technical/methodological expertise as well as in terms of training expertise). |

## 2.7. WP7 - Creating the SSH Open Marketplace

The first year of the WP7 work was mainly focused on Task 7.1 "User requirements, Conceptual Model and System Architecture of the SSH Open Marketplace". The team collected user requirements and use cases that allowed the design of the conceptual model and the system architecture of the SSH Open Marketplace. All other WP7 tasks contributed as well.

As the SSH Open Marketplace will be populated with existing sources and catalogues, the collection and prioritisation of sources to harvest was a central aspect of the work in Task 7.3. Based on the first definitions of the requirements and architecture, the development work also started in Task 7.2. Finally, the first discussions regarding the curation workflow and governance organisation of the portal began as well in Task 7.4. All these activities were explained and summarised in the deliverable 7.1 "System specification for the SSH Open Marketplace", delivered as planned in September 2019.

### 2.7.1. WP7 progress

#### 2.7.1.1. Task 7.1 User requirements, Conceptual Model and System Architecture of the SSH Open Marketplace

In order to specify accurate user requirements for the SSH Open Marketplace, Task 7.1 partners decided to **collect use cases** from previous work and conduct **interviews with SSH researchers** to refine user needs. Based on those materials, **user stories were extracted and prioritised** before being translated into technical requirements to inform the design of the conceptual model and system architecture. A mapping between diverse sources and their metadata formats as well as the **first SSH Open Marketplace data model** were also created under this task.

The SSH Open Marketplace data model was designed to be generic and flexible enough to be refined through configuration during runtime. Furthermore, based on a WP3 deliverable (D3.1 - Metadata interoperability), an analysis of the different controlled vocabularies to rely on for the description of the SSH Open Marketplace contents has been made. Finally, based on design decisions and user requirements, **technical components and their implementation** were described, allowing Task 7.2 to start the development work.

#### 2.7.1.2. Task 7.2 Development of the Marketplace Application

Task 7.2 is dedicated to the development of the SSH Open Marketplace based on the system specification (***D7.1 System specification*** submitted in October 2019) produced in T7.1. The team of T7.2 worked closely with T7.1 on formulating the specification. Pursuing a **user-centric design**, crucial input for the specification were requirements from potential users collected in the first months of the project. In this phase also development guidelines were collected, and alternatives of the overall system architecture were discussed.

Once the specification was finished, actual **development of the system** began by defining the REST-API[87] which serves as a contract between the server application and the client. Agreeing on a traditional development path with well-established technologies, the server component is a Java-Spring application, the client application is being implemented in JavaScript using the React framework[88]. The **code base** is maintained collaboratively in git repositories[89] hosted by one of the partners, UGOE. Parallel to the basic development work on the server and client component (skeleton and basic wiring), an initial UI-design[90] has been proposed by an UI-expert from one of the partners, PSNC.

In parallel to the development of the application, in collaboration with T7.3, **potential data sources** were identified and prioritised. Data about tools from TAPoR[91] were selected as the first source to be processed. A **mapping** has been defined from the data model of TAPoR to the data model of the SSH Open Marketplace, based on which transformation pipeline was implemented using the ETL-framework Unified Views[92], part of the semantic suite PoolParty[93] provided by one of the partners, SWC. The pipeline reads the TAPoR API, extracts the data, generates the MP-compatible data structure and pushes the converted data against the API of the Marketplace. The pipeline can be run regularly, allowing continuous updates from the source.

After the first three months of implementation T7.2 was able to deliver an **MVP – a minimal viable product** – i.e. a running application which demonstrates the connectivity between the main components and exhibits minimal functionality. It implements the proposed UI-design and allows for a simple full-text search in the data and detail view for individual items. All this functionality of the client is based on the fully implemented API of the server-side component. The API also allowed to ingest the first real dataset (tools from TAPoR), validating the suitability of the API definition.

87 Web service API (application programming interface) that adheres to the REST (representational state transfer) architectural constraints.
88 React is a JavaScript library for building user interfaces.
89 https://gitlab.gwdg.de/sshoc
90 User Interface design
91 http://tapor.ca/
92 Unified Views (https://unifiedviews.eu/) is an open source Extract-Transform-Load (ETL) framework that allows users to define, execute, monitor, debug, schedule, and share Resource Description Framework (RDF) data processing tasks.
93 https://www.poolparty.biz/

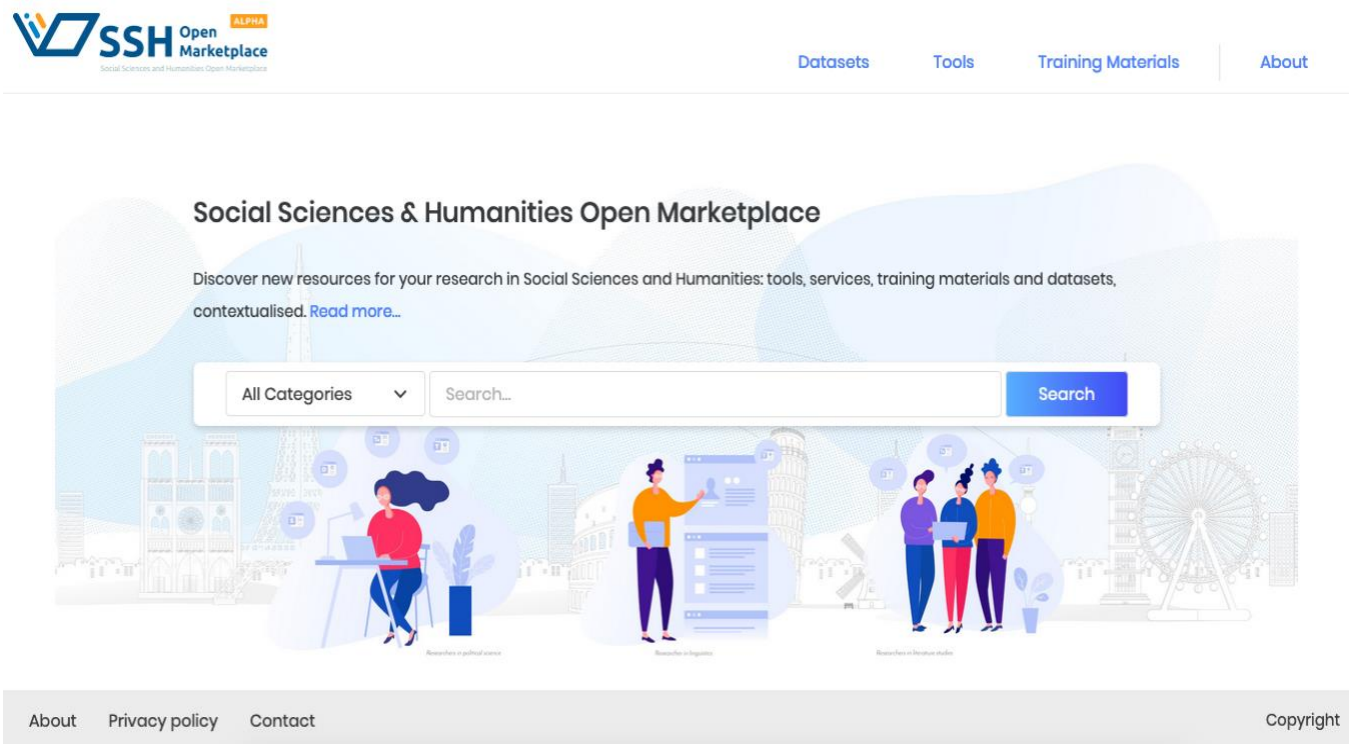


***Figure 3**. SSH Open Marketplace. Minimal Viable Product. December 2019*

#### 2.7.1.3. TASK 7.3 MARKETPLACE INTEROPERABILITY

As the SSH Open Marketplace is not the only location where information about SSH-related resources is stored, it is important to take into account the necessity to avoid double work in keeping information synchronised between the various sources of information. Likewise, the possibility to export (parts of) the data stored in the SSH Open Marketplace into other databases, provides a good reason for making an overview of all the data sources and targets. This is what Task 7.3 is about.

At the WP7 meeting in June, a first **candidate list of sources to import from** was presented and discussed. This has been collaboratively extended afterwards. Certain candidates, like the EOSC portal marketplace, have also been analysed further in direct communication with the people responsible in the EOSC-hub project. The result is a **shortlist of sources**, with prioritised entries, and where available pointers to APIs or documentation on accessing the data stores. This list, in combination with the findings on what constitutes a reasonable granularity and quality level:

- has been used to populate the **first alpha version of the Marketplace** (task 7.2) and as common ground to discuss curation (task 7.4)
- will be used as a starting point to draft a **report on Marketplace Interoperability** (target: M24), which in turn will serve as a base for the final deliverable of this task, D7.3 (M36).

#### 2.7.1.4. Task 7.4 Governance: Population, Curation & Sustainability of the SSH Open Marketplace

Task 7.4 works on three deliverables. First, D7.6 Resources for Marketplace content description, in other words a survey on existing social sciences and humanities vocabularies in Europe has to be carried out for M24 (31/12/2020). The **draft of the survey** is currently completed. A collaboration with WP2 has been made in order to propose the best dissemination strategy possible.

Secondly, the governance of the SSH Open Marketplace will be defined in deliverable D7.5 Marketplace - Governance, for M24 (31/12/2020). This work evolves in close cooperation with the technical construction and evolution of the SSH Open Marketplace. Discussions are undertaken to define the best **governance and business models** possible.

Last, but not least, the **curation of the SSH Open Marketplace** will be settled through D7.4 Data population & curation for M36 (31/12/2021). As for D7.5, the curation also evolves as the SSH Open Marketplace takes shape. A close collaboration with WP7's other task forces is therefore necessary.

### 2.7.2. Note on deviations from the plan and risk monitoring

Some large European portals were mentioned in the Grant Agreement, as potential examples WP7 could use to compare and learn from large scale aggregators (European Data Portal, the German Mobility Marketplace, the Industrial Data Space, and the Data Market Austria, Big Data Europe Aggregator). D7.1 chose to focus more on domain-relevant portals (like data models used by TAPoR, SSK, CIDOC, Scholarly Ontology, or the DARIAH in-kind contribution tool) and on EOSC initiatives (EOSC-Hub, eInfraCentral and OpenAire catalogues) to analyse reference models and metadata schemas used.

Within WP7 the recruitment process on behalf of UGOE was substantially delayed through 2019. Reasons for this lay outside the partner's accountability (tight employment market). Although an early coordination of this issue with the WP lead allowed for a pragmatic handling (redistribution of tasks and responsibilities), this nevertheless was an influence, particularly regarding Task 7.1.

Information on risk monitoring in WP7 for Year 1 is presented in the table below:

***Table 7.*** *Risk monitoring in WP7 - Year 1*

| **Risks monitoring WP7** | | | | |
|---|---|---|---|---|
| No of risk (from the DoA) | Description | Have risk-mitigation measures been applied? | Did the risk materialise so far? | Comments |
| 6 | Too few available software developers | **Yes** | No | Mitigation measres applied: Work prioritied and reschedules, build on existing tools and past experience, involve users in testing |
| 7 | Technical difficulties of complex software development | **Yes** | No | Hired scientific developers highly experienced in SSH |

## 2.8. WP8 - Governance / Sustainability / Quality Assurance

The focus of WP8 is on the elaboration of an efficient framework to cover a number of issues related to the governance of the SSH cloud and its successful integration into the EOSC, including sustainability, quality and ethical requirements for research data, cross-disciplinary coordination and cooperation. With the participation of 5 ESFRI Landmarks and ESFRI 1 Project in the ESFRI Social and Cultural Innovation (S&CI) landscape, WP8 works to implement a robust model for shared governance (T8.1); elaborate a quality-assurance scheme for SSH and Heritage Science research data, engage with and support data providers and data sources during the certification process (T8.2); elaborate a shared legal and ethical framework for public research data sharing and reuse (T8.3); foster cross-disciplinary collaborations and development of shared resources at the European level by liaising with other disciplinary clusters (T8.4).

In the first year most of the WP8 activities where focussed on **mapping the landscape** to explore possible synergies with all the relevant actors within and outside the ESFRI boundaries (T8.1; T8.4) and effectively implement a number of actions to consolidate the most promising ones (T8.4). Another relevant target for WP8 during the first year of the project was to establish the first version of the SSHOC Certification framework for repositories (T8.2).

### 2.8.1. WP8 progress

#### 2.8.1.1. Task 8.1 Governance & Sustainability

Task 8.1 started its activities on M01. In the first year the attention was focussed on the **mapping of related connections** encompassing EOSC and national initiatives (EOSC-Nordic, Italian ICDI, etc.). During the first year, part of the activities were devoted to mapping and exploiting cross-task and cross-WP commonalities and priorities, and to develop a proactive attitude in addressing crucial issues around the cooperation and interaction between relevant actors, at global, ESRI, national and regional levels: results of this activities (in collaboration with T8.4) will constitute the ground for the establishment of an internal information management platform, initially deployed by CNR and further developed, extended and integrated (also with other similar tools used in SSHOC ) during the project.

#### 2.8.1.2. Task 8.2 Trust & Quality Assurance

Task 8.2 started its activities on M03 and focussed its activities during the first year on the elaboration of a Report on "*Certification plan for SSHOC repositories*" [Deliverable D8.2 due in M12]. This document is relevant to the SSH ERICs and to repositories within SSHOC communities. In line with the aims of Task 8.2, the report specifies modes of support in building trust and helping repositories reach Trusted Digital Repos certification. Moreover, this work lays the ground for the SSHOC work on *trust* that is needed in order to facilitate the adoption of standards and the FAIR principles in SSH data repositories.

As a general approach, T8.2 charted the current *trust landscape* within the SSHOC communities (i.e.: CESSDA ERIC, CLARIN ERIC, DARIAH ERIC and E-RIHS) and considered the legacy of the PARTHENOS

thematic cluster. Due to the diversity of repositories within the SSHOC communities, a flexible yet sustainable approach to *trust* is needed that is adaptable to a wide variety of data infrastructures.

Among the main activities for T8.2 the mapping and selection processes (for both standards, institutions and repositories) were probably the most relevant: before **selecting the CoreTrustSeal**[94] as the standard trusted digital repository certification reference, other existing Trust and Certification approaches were taken into consideration. T8.2 identified organisations for which the CoreTrustSeal requirements were applicable, with the aim to further develop the CoreTrustSeal framework to better support a variety of repositories and selected repositories that will be the main focus of the support activities provided at later stages in the project. Even though there are no direct dependencies with other SSHOC tasks, task 8.2 decided to align its activities to both SSHOC and existing EOSC-related efforts promoting trust and the FAIR principles (i.e.: FAIRsFAIR[95] etc.)

#### 2.8.1.3. Task 8.3 Legal and Ethical Issues

Task 8.3 started its activities on M10 as planned. The focus of the Task is to work towards the elaboration of a legal and ethical framework for SSH research to facilitate empirical trans-national research. During the first period of the project activities has - and will be - devoted to **analyze the GDPR landscape** at European level (in coordination with T5.3) and set the conditions to establish concrete test beds in WP9 around data deposit and reuse. The task will also take in consideration the work done in other ERICS (BBMRI etc.) also leveraging on the information gathered and managed by T8.1 and T8.4.

#### 2.8.1.4. Task 8.4 Overarching clusters

Task 8.4 started its activities on M01 and focussed on paving the path towards the **elaboration of a coordination platform** to manage the the collaboration with other clusters (starting but not limited to: Biodiversity, Computation, Biology) and ease the dialogue along several directions (including political, organisational, etc.) also beyond the EU borders, towards future global collaborations (i.e.: with CELAC - Caribbean and Latin America countries).

In this context, Task 8.4 started several activities with different actors (ESFRI, the RDA, CELAC, EOSC) to be further developed in concrete actions (i.e.: EU-CELAC ResInfra, starting on February 2020) during the project. As a next step, an **internal platform** to manage the information gathered from T8.4 and other tasks (for the moment, mostly T8.1) has been developed by CNR-OVI and will be further implemented during the project and integrated with tools used by other WPs.

### 2.8.2. Note on deviations from the plan and contingencies

Task 8.2 asked for an extension to the elaboration of the Report on Certification plan for SSHOC repositories (D8.2) in order to allow a better level of alignment with FAIRsFAIR project's certification support work, CESSDA Trust work, and the E-RIHS quality work (to be available in February 2020). However,

94 Core Trust seal website: https://www.coretrustseal.org/
95 FAIRsFAIR project website: https://www.fairsfair.eu/

this delay will not produce any negative consequences for the task, the WP and the SSHOC project, since no other activities, milestones or deliverables will be affected.

Information on risk monitoring in WP8 for Year 1 is presented in the table below:

***Table 8.*** *Risk monitoring in WP8 - Year 1*

| **Risks monitoring WP8** | | | | |
|---|---|---|---|---|
| No of risk (from the DoA) | Description | Have risk-mitigation measures been applied? | Did the risk materialise so far? | Comments |
| 5 | Inefficient use of resources: survey creation, translation and storing | **Yes** | **Yes** | Realised a mapping of available tools developed and used projectwide, planned the integration of the information management platform with other SSHOC tools |
| 8 | Unforeseen legal (GDPR and national) barriers to data sharing | **Yes** | **Yes** | This is especially related to T8.3 which started on M10 |
| 11 | Understaffing of research infrastructures | **Yes** | **Yes** | Rescheduled the work; changed the rules and procedures after the 1st project's year |
| 12 | Imbalance between research communities in the governance: disciplines poorly/over represented | **Yes** | **Yes** | WP8 activities experienced a substancial balance in governance and representation |
| 13 | Overlapping of work with other clusters | **Yes** | No | Possible common actions involving other clusters have been investigated in various contexts. The establishment of communication platform will streamline the process |

## 2.9. WP9 - Data Communities

The focus of WP9, led by UNOTT, is Data Communities, and in particular the communities of scholars in Ethnic and Migration Studies (Task 9.2), Electoral Studies (Task 9.3), and Heritage Science (Task 9.4). In addition, it focuses on identifying opportunities and obstacles for a wide variety of scholarly communities in the social sciences and humanities – particularly those that are not represented via ERICs in SSHOC – for engaging with and making use of the infrastructure, tools, products and services under development by EOSC/SSHOC (Task 9.1). Year 1 intends to inventory said general opportunities and obstacles (Task 9.1), and to progress with the work plan of Task 9.2. Most of the components of the work plans forming Tasks 9.3 and 9.4 are foreseen for Year 2 and the following years.

A novelty regarding WP9 also impacting the whole SSHOC project are the links identified between SSHOC *WP9 Data Communities,* and the EURhisFIRM, an EC funded project aiming to "design a world-class research infrastructure to connect, collect, collate, align, and share detailed, reliable, and standardized long-term company-level data for Europe"[96], adding up to the existing European Value Studies and Migration Studies pilots already in the project. After communicating with EURhisFIRM partners, SSHOC Consortium decided that connecting the two projects would be of great benefit. Followed by lengthy discussions, the agreed model of joining was to add 3 core institutions forming the EURHISFIRM Executive Board: PSE, House of Finance of Goethe University Frankfurt, and the University of Antwerp to the SSHOC Consortium. The change was decided to be added to the 1st SSHOC Amendment planned for early 2020.

For the EURHISFIRM project, its membership in the SSHOC will enhance EURhisFIRM's integration into an EU-wide data community, which will, in turn, enable EURHISFIRM to adapt its model to this cutting-edge experience. Specifically, EURHISFIRM will benefit by sharing and learning from ongoing innovations such as the continuous developments with regards to cloud ready SSH tools, technical work for matching heterogeneous data sets, as well as working with the interoperability principles of FAIR (findable, accessible, interoperable, reusable) data.

### 2.9.1. WP9 progress

#### 2.9.1.1. Task 9.1 Identifying shared and unique challenges for SSH data communities; Evaluation and Usability Report

The main component of this task in the first year of the project was to produce a report identifying shared and unique challenges to user communities attempting to contribute to SSHOC (deliverable ***D9.1 Report on challenges user communities face when attempting to contribute to SSHOC*** *submitted in November 2019*). Follow-up reports on the same topic have to be delivered approximately at mid-term and towards the end of the entire project, but then with the additional focus of evaluating the usability of existing and developed procedures, tools and services from the perspectives of the user communities involved in this

[96] EURhisFIRM project website: https://eurhisfirm.eu/index.php/the-project-2/

WP. In this respect WP9 is to collaborate closely with WP6. The specifics of this will have to be worked out in the first trimester of 2020 in order to realise the midterm report.

#### 2.9.1.2. TASK 9.2 ETHNIC AND MIGRATION STUDIES

This task consists of two components: Component a) Making data Findable: surveys to ethnic and migrant minorities (EMMs) across Europe. ETHMIGSURVEYDATA designs the compilation of a database of quantitative surveys undertaken with EMM (sub)samples across Europe. Sciences Po leads this task on behalf of and in coordination with the relevant community and its COST Action network[97]. The work in this Task is to consolidate this database in a way that complies with the FAIR principles. This will involve compliance to current international standards (DDI, Dublin Core, RDA recommendations, etc.); embedding the database in an infrastructure ensuring its sustainability and updating; enabling producers of existing and future data to add metadata autonomously; ensuring accessibility and linkability to OpenAire[98], the CESSDA Data Catalogue[99], and SSHOC; and Integration in H2020 CROSS-MIGRATION[100]. Progress on this task (due date M22) has been satisfactory. In coordination with the COST Action Ethmigsurveydata, the team has **prototyped** and is now close to a **Beta version of the online metadata registry**[101]. Although this pre-Beta version still only contains surveys for Croatia and Norway there is now complete information for 26 European countries. Currently, the quality control for these datasets is being undertaken and those will be uploaded gradually until December 2020 for the most part. For the quality control tasks a very detailed version control documentation was prepared that, alongside the coding instructions and the coding templates will be made available on the online registry. Thorough preparations started so that the registry will include an API that is DDI-compliant so that the metadata can be harvested.

Component b) of this Task involves (in coordination with WP3) a pilot of inclusion of ethnic and migrant minority survey data in the European Question Bank. This task develops **a pilot** assessing the feasibility of a full stream of the European Question Bank (EQB) dedicated to the ethnic and migrant minority (EMM) survey data uncovered in component (a) of his Task. This component is due in M36 of the project. This pilot assesses the volume of this material not yet archived. Jointly with Task 3.1 in WP3 technical solutions will be evaluated for integrating in the EQB this large stream of surveys and survey questions. Thus, this task joins existing European infrastructures and the data research community it serves. Specifically, this pilot involves (in addition to the requirements listed under component (a): finding solutions for the automated retrieval from PDF or other document formats of questions in multiple languages; connecting the EQB with the ETHMIGSURVEYDATA metadata (see component (a); exploring how existing and future data producers can update the EQB autonomously. As far as progress of this task is concerned, the **planning stage** for this task has started in the second half of Year 1. This includes a **selection of topics**

97 See https://ethmigsurveydatahub.eu/ for more information about this COST action, and https://www.cost.eu/ for more general information about the COST program

98 OpenAire network website: https://www.openaire.eu/

99 CESSDA Data Catalogue: https://datacatalogue.cessda.eu/

100 CROSS-MIGRATION is a H2020 project (2018-2020) which resulted in Migration Research Hub (https://crossmigration.eu)

101 see here: https://ethmig.bhdevelopment.ro/surveys

on which to undertake **initial pilots** for the extraction of questions, finding a work plan with GESIS on how to proceed to link the EMM questionnaire data to EQB and the use of Colectica[102] at Sciences Po to ensure that all the process is done following DDI codebook.

#### 2.9.1.3. TASK 9.3 DATA COMMUNITY PROJECT: ELECTORAL STUDIES

This Task aims to generate an Open Research Knowledge Graph in the field of Electoral Studies to be included in the SSHOC. It starts with the subfield of electoral behaviour and motivations of citizens (other subfields, e.g., behaviour of political parties and elites, or aggregate outcomes, can be added later). Taxonomies and ontologies underlying the Knowledge Graph will be developed with active involvement of the user community. The project will harvest, integrate, interlink, analyse national, European and cross-national data on citizens' electoral behaviour (mainly survey data) and relevant additional data (e.g. contextual data) from existing (open) sources and repositories. This task builds upon existing foundational work that has been conducted by e.g., the European Voter[103] and the COST-TEV projects[104], and the European Election Studies. This project intends (i) to create a new repository of election-relevant data and information (including continuous update mechanisms) interlinked using Linked (Open) Data principles and technologies, (ii) to allow analysis and visualisation (making use of W3C standards RDF and SPARQL[105]). The overall results will be integrated in an election studies analytics dashboard to access analytics and visualisations in an easy to use manner, to export results, and to query the data (for more advanced users) via a data API for further analysis using common statpacks such as STATA, SPSS or R[106].

The task starts with Deliverable D9.6, a **Demarcation Report**, that was due in M06 and was delayed as a consequence of the delay in D9.1. This report demarcates the user community Election Studies and actors therein and segments thereof to be contacted, recruited and trained for active collaboration in co-production of knowledge graph. It also inventories existing data, publications and services, and reports the near non-existence of taxonomies, ontologies and controlled vocabularies. D9.6 is expected to be delivered in M13, and is being used in the development of D9.7 (Design of the Knowledge Graph), the delivery of which has been postponed to M15 because parts of it overlap with the subsequent D9.9 (Delivery of user-validated knowledge graph, due in M26).

[102] Colectica tool for design, document, and publishing statistical data and survey research using open data standards. Website: https://www.colectica.com/

[103] For the European Voter project, see Thomassen, J.J.A. The European Voter; A comparative study of modern democracies. Oxford: Oxford University Press, 2005, and the associated European Voter Database, archived in GESIS under number ZA3911 (see https://www.gesis.org/en/services/data-analysis/more-data-to-analyze/more-international-data/the-european-voter-project).

[104] For the COST-TEV project see: https://www.mzes.uni-mannheim.de/d7/en/projects/the-true-european-voter-a-strategy-for-analysing-the-prospects-of-european-electoral-democracy-that-includes-the-west.

[105] SPARQL and RDT - graph data formats for representing information in the Web (see: https://www.w3.org/TR/rdf-sparql-query/)

[106] STATA, SPSS and R - standard integrated software packages that provide data manipulation, visualization, statistics, and reproducible reporting

### 2.9.2. Note on deviations from the plan and risk monitoring

D9.1 was delayed from April (M4) to November (M11) of the project, owing to circumstances affecting the author, but the delay has not generated delays in other work packages. It did cause a delay in D9.6 which was produced by the same author.

Both *D9.6 Demarcation Report* and *D9.7 Design of Knowledge Graph* were delayed and still pending for submission, both being in review process. This delay has had no consequences for work in other work packages, or for subsequent work in WP9.

Information on risk monitoring in WP9 for Year 1 is presented in the table below:

***Table 9.*** *Risk monitoring in WP9 - Year 1*

| **Risks monitoring WP9** | | | | |
|---|---|---|---|---|
| No of risk (from the DoA) | Description | Have risk-mitigation measures been applied? | Did the risk materialise so far? | Comments |
| 6 | Too few available software developers | No | No | In WP9 this risk may at earliest materialise in the course of the work for D9.9 (Production of Knowledge Graph and Election Studies Analytics Dashboard); this risk, and mitigating measures, will be discussed in the course of the work for D9.7 (Design of the Knowledge Graph), and will be commented upon in the report of that deliverable. |
| 7 | Technical difficulties of complex software development | No | No | *The same as for Risk 6* |
| 8 | Unforeseen legal (GDPR and national) barriers to data sharing | No | No | Likelihood of this risks is low, particularly where use is being made of data that are already in the public domain and where these problems have therefore already been addressed (if they existed at all). |
| 10 | Low acceptance or understanding of the new data access tools, guidelines and services by users and public | No | No | D9.8 is one of the places to assess which practical measures will help to mitigate these risks. How successful it will be to overcome (or even prevent) these risks can only be assessed after the completion of D9.4 and D9.9, and will be reported systematically in D9.10. |
| New Risk | Vulnerability to obstructing personal circumstances in small teams where work to be done cannot be transferred to another team member | **Yes** | **Yes** | It happened so far to Ben.13 UNOTT, and to Sciences Po (associate temporarily away on maternity leave). This risk is inherently unpreventable as it is uncertain whether and when obstructing personal circumstances will occur; mitigation by intensive contact with CO and, where possible and necessary, rescheduling. In WP9, delivery of D9.1 and D9.6 were rescheduled, and delay communicated with the PO. This risk is most severe for units (WPs, Tasks, etc) in which only a small number of people are involved, and where, consequently, the possibilities for handing over a particular task to someone else are limited. |

# 3. Summary of progress and delivery

As explained in the previous chapters, SSHOC project managed to achieve **13 Milestones**, and submit **13 Deliverables** in Year 1. This has shown to be a demanding task, and the Consortium finds it to be a success, as most of the delays were due to overlapping in Deliverable planning, ending up with several being scheduled for the same date, or overlapping with the holiday season. There are still 9 pending Deliverables postponed mostly to early 2020, which are either in preparation or review process, and none of the delays caused any significant or negative consequence for the rest of the work. All of the delays were communicated to the PO. The following tables show a summary of the Milestones achieved and Deliverables submitted in Year 1:

***Table 10.*** *Milestones planned and achieved in Year 1 of SSHOC Project*

| Year 1 2019 | | | | | | |
|---|---|---|---|---|---|---|
| **MILESTONES** | | | | | | |
| **WP No.** | **Task** | **No.** | **Title** | **Lead partner** | **Due Date (month)** | **Actual delivery** |
| WP 1 | 1.1 | MS 1 | Integrated project management and reporting tools and templates | CESSDA ERIC | 4 | 3 |
| WP 2 | 2.3 | MS 2 | Community database creation & management | TRUST-IT | 1 | 4 |
| WP 2 | 2.3 | MS 3 | SSHOC Web platform Launch | TRUST-IT | 3 | 2 |
| WP 2 | 2.3 | MS 4 | SSHOC web platform second release | TRUST-IT | 6 | 7 |
| WP 2 | 2.3 | MS 5 | SSHOC web platform continuous marketplace integration with SSHOC service offer 1/3: M9, M10, M12 | TRUST-IT | 12 | *pending* |
| WP 3 | 3.5 | MS 8 | Choice of vocabulary publication platform for SSHOC | CLARIN | 6 | 7 |
| WP 3 | 3,6 | MS 9 | Inventory of services & tools, and domain of actionable data-types | CLARIN | 8 | 8 |
| WP 3 | 3.1 | MS 10 | Inventory of SSHOC metadata sets and taxonomies | CNR | 8 | 10 |
| WP 3 | 3.3 | MS 11 | Selection and establishing access to data-sets for processing | CLARIN | 8 | *pending* |
| WP 3 | 3.6 | MS 12 | Implementation plan SSHOC Switchboard and VCR services | CLARIN | 10 | 10 |
| WP 4 | 4.2 | MS 16 | Survey questions extracted from TMT and delivered in machine readable format, to be used as input for T4.3 and T3.1 | CentERdata | 2 | *pending* |
| WP 4 | 4.7 | MS 19 | Consultation with SSH data producers | FORTH | 6 | 5 |
| WP 5 | 5.1 | MS 21 | Protocol of laboratory processing of DBSS data | SHARE ERIC | 6 | 7 |
| WP 5 | 5.2 | MS 24 | Implementation of standard Dataverse installation on EOSC | KNAW | 4 | 7 |
| WP 5 | 5.2 | MS 25 | Selection of additional functionality for data repository service to be developed | KNAW | 6 | 4 |
| WP 6 | 6.2 | MS 38 | Launch Community engagement strategy | LIBER | 7 | 10 |

***Table 11.*** *Deliverables planned and delivered in Year 1 of SSHOC Project*

| Year 1 2019 | | | | | | |
|---|---|---|---|---|---|---|
| **DELIVERABLES** | | | | | | |
| **WP No.** | **Relative number in WP** | **Del. No.** | **Title** | **Lead partner** | **Expected submission date** | **Actual delivery date** |
| WP1 | D1.1 | D1 | Project Management Plan | CESSDA ERIC | 30.04.2019. | 26.06.2019. |
| WP1 | D1.2 | D2 | Quality Assurance & Risk Assessment Plan | CESSDA ERIC | 31.08.2019. | 19.12.2019. |
| WP1 | D1.6 | D6 | Data Management Plan - DMP | CESSDA ERIC | 30.06.2019. | *postponed* |
| WP2 | D2.1 | D7 | SSHOC overall communication and outreach plan | TRUST-IT | 31.03.2019. | 05.04.2019 |
| WP3 | D3.1 | D10 | Report on SSHOC (meta-) data interoperability problems | CESSDA ERIC | 30.06.2019. | 03.07.2019. |
| WP3 | D3.2 | D11 | Inventory of SSH citation practices, and choice for SSHOC citation formats and implementation planning | CNRS | 30.06.2019. | 15.11.2019. |
| WP4 | D4.1 | D20 | Detailed specification of the sample management system | ESS ERIC | 31.03.2019. | 01.11.2019. |
| WP4 | D4.3 | D22 | Survey specific parallel corpora | ESS ERIC | 31.12.2019. | *postponed* |
| WP4 | D4.12 | D31 | Audio Survey Modules - Guidelines for the integration of Audio Capture data in Survey Interviews | KNAW | 31.12.2019 | 30.12.2019. |
| WP4 | D4.14 | D33 | Policy API tool | KNAW | 31.12.2019. | *postponed* |
| WP4 | D4.16 | D35 | Specification of the new features of the Aïoli platform | CNRS | 30.06.2019. | 18.11.2019. |
| WP4 | D4.18 | D37 | SSHOCro beta version | FORTH | 31.12.2019. | *postponed* |
| WP5 | D5.1 | D40 | Guidelines for ethics considerations in making biomedical survey data FAIR (Access to biomedical data) | SHARE ERIC | 31.08.2019. | *postponed* |
| WP5 | D5.17 | D56 | Implementation plan for the software module (Archaeology case study) | DAI | 31.10.2019. | *postponed* |
| WP6 | D6.1 | D60 | Community engagement strategy | LIBER | 30.06.2019. | 31.07.2019. |
| WP6 | D6.2 | D61 | Building expertise strategy | LIBER | 31.12.2019. | 17.12.2019. |
| WP6 | D6.7 | D66 | Inventory of existing learning materials | DARIAH ERIC | 31.12.2019. | 19.12.2019. |
| WP7 | D7.1 | D75 | System specification | DARIAH ERIC | 30.09.2019. | 02.10.2019. |
| WP8 | D8.2 | D82 | Certification plan for SSHOC repositories | CESSDA ERIC | 31.12.2019. | *postponed* |
| WP9 | D9.1 | D87 | Report on challenges user communities face when attempting to contribute to SSHOC | UNOTT | 30.04.2019. | 18.11.2019. |
| WP9 | D9.6 | D92 | Demarcation Report | UNOTT | 30.06.2019. | *postponed* |
| WP9 | D9.7 | D93 | Design of Knowledge Graph | SWC | 31.12.2019. | *postponed* |

# 4. Use of resources and budget expenditure

The first year of the project is 30% of the project duration (12/40 months). However, out of 42 Tasks in the project, 4 of them have started after M2 (February), and 8 of them have started later in the year. Looking at the work and achievements planned presented by Milestones and Deliverables, it can be concluded that 15 out of 49 Milestones were scheduled (30%) and 22 out of 101 deliverables were planned (21%) for Y1 and not all have been delivered by December 2019. That is in line with the Use of resources reported by the partners in Year 1, where in total about 240 person-months were used (14%) of the total person-months allocated to the project (1754 PM). Translated to budget expenditure, about 1,9 million €) was spent (13%) out of the total budget of 14,5 million €. The expectation is that Year 2 and 3 of the project (2020-2021) will be more demanding, as all the tasks are active, and most of the delivery is planned in that period. The tables below represent the summary of the Use of Resources in Year 1 per partner[107]:

***Table 12.*** *Use of resources: Person-months reported per partner per work package M1-M12*

| Year 1 2019 | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TOTAL SPENT PM | | | | | | | | | | | | |
| **No.** | **Partner** | **WP1** | **WP2** | **WP3** | **WP4** | **WP5** | **WP6** | **WP7** | **WP8** | **WP9** | **Total** | **% spent** |
| 1 | CESSDA ERIC | 8.33 | - | - | - | - | - | 0.04 | 0.06 | - | 8.43 | *20%* |
| 1.1 | NSD | - | - | - | - | 0.68 | - | - | - | - | 0.68 | *3%* |
| 1.2 | UKDS | 0.04 | - | 2.51 | - | 0.78 | 0.36 | - | 0.56 | - | 4.25 | *17%* |
| 1.3 | UL-ADP | - | 1 | - | - | - | 1.44 | - | - | - | 2.44 | *13%* |
| 1.4 | GESIS | - | - | 3 | - | 2.04 | - | - | - | - | 5.04 | *17%* |
| 1.5 | SND | 0.16 | - | 1.53 | - | - | - | - | 0.62 | - | 2.31 | *16%* |
| 1.6 | UTA-FSD | 0.14 | - | 5.3 | - | - | - | - | 0.95 | - | 6.39 | *26%* |
| 1.7 | FORS | 0.5 | - | - | - | 0.28 | - | - | - | - | 0.78 | *9%* |
| 1.8 | DNA | 0.38 | - | - | - | 0.06 | - | - | - | - | 0.44 | *7%* |
| 1.9 | AUSSDA | 0.22 | 2.48 | - | - | 2.47 | - | - | - | - | 5.17 | *28%* |
| 2 | ESS ERIC | - | - | - | - | - | - | - | - | - | - | - |
| 2.1 | City | 0.77 | - | - | 5.37 | - | - | - | - | - | 6.14 | *22%* |
| 2.2 | GESIS | 0.15 | - | - | 3.28 | - | - | - | - | - | 3.43 | *30%* |
| 2.3 | UPF | 0.1 | - | - | 10.25 | - | - | - | - | - | 10.35 | *22%* |
| 2.4 | NSD | - | - | - | - | 1.59 | - | - | - | - | 1.59 | *3%* |
| 3 | SHARE ERIC | - | - | - | - | - | - | - | - | - | - | - |
| 3.1 | MPISOC | 0.1 | - | - | - | 2.7 | - | - | 0.19 | - | 2.99 | *3%* |
| 3.2 | UniVe | - | - | 6.69 | - | - | - | - | - | - | 6.69 | *19%* |
| 4 | CLARIN ERIC | 0.99 | 0.75 | 15.21 | - | 0.48 | 0.02 | 1.98 | 1.06 | - | 20.49 | *17%* |
| 4.1 | Athena | - | - | 2.93 | - | - | - | - | - | - | 2.93 | *14%* |
| 4.2 | EKUT | - | - | 1.35 | - | - | - | - | - | - | 1.35 | *5%* |
| 4.3 | CUNI | 0.63 | - | 4.63 | 3.14 | - | - | - | - | - | 8.4 | *29%* |
| 4.4 | UL-FF | 0.22 | - | - | - | - | 1.18 | - | - | - | 1.4 | *5%* |
| 4.5 | RUN | - | - | - | 0.5 | 1.9 | - | - | - | - | 2.4 | *13%* |
| 5 | DARIAH ERIC | 0.14 | 0.11 | - | - | - | 0.22 | 6.28 | - | - | 6.75 | *26%* |
| 5.1 | PSNC | 0.11 | - | - | - | 2.24 | - | 3.08 | 1 | - | 6.43 | *10%* |
| 5.2 | OEAW | - | - | 0.09 | - | - | 1.49 | 7.3 | - | - | 8.88 | *12%* |
| 5.3 | UGOE | - | - | - | - | - | - | 1 | - | - | 1 | *1%* |
| 6 | LIBER | 0.82 | 2.21 | 0.58 | - | - | 7.17 | - | - | - | 10.79 | *26%* |
| 7 | KNAW - GGP | 0.26 | - | 2.93 | - | 5.37 | 1.59 | - | 1.69 | - | 11.84 | *17%* |
| 8 | UoY - ADS | - | - | - | - | 2.64 | - | - | - | - | 2.64 | *26%* |
| 9 | TiU | 0.45 | - | 0.06 | 0.28 | - | - | - | - | - | 0.79 | *3%* |
| 10 | TRUST-IT | 0.38 | 5.65 | - | - | - | 0.35 | 0.34 | 0.15 | - | 6.87 | *27%* |
| 10.1 | Trust-SRL | 0.17 | 4.89 | - | - | - | 1.27 | 0.12 | 0.07 | - | 6.52 | *26%* |
| 10.2 | COMMpla | 0.12 | 3.56 | - | - | - | 1.2 | 0.32 | 0.16 | - | 5.36 | *43%* |
| 11 | SWC | 0.14 | - | - | - | - | - | 4.66 | - | 3.34 | 8.14 | *27%* |
| 12 | SCIENCES PO | 0.13 | - | - | 9.99 | - | - | - | - | 5.52 | 15.64 | *15%* |
| 13 | UNOTT | - | - | - | - | - | - | - | - | - | - | *0%* |
| 14 | DAI | 0.1 | - | 1.25 | 0.75 | 3.5 | - | - | - | - | 5.6 | *29%* |
| 15 | CNRS | 0.04 | 0.53 | 2.68 | 1.47 | - | 0.88 | 8.55 | - | - | 14.15 | *17%* |
| 16 | CNR | - | - | - | - | - | - | - | - | - | - | *0%* |
| 17 | UCL | 0.16 | - | - | - | - | 3.85 | - | - | - | 4.01 | *42%* |
| 18 | FORTH | - | - | - | 8.83 | - | - | - | - | - | 8.83 | *38%* |
| 19 | CentERdata | 0.42 | - | 1.25 | 5.13 | 4.42 | - | - | - | - | 11.22 | *10%* |
| 20 | NG | 0.47 | - | - | - | 0.15 | - | - | - | - | 0.62 | *5%* |
| **Total** | | **16.65** | **21.18** | **51.99** | **48.99** | **31.3** | **21.02** | **33.67** | **6.51** | **8.86** | **240.17** | ***14%*** |

[107] The Use of resources summary presents an estimation reported by partners in Y1. Two of the beneficiaries have not been able to report their use due to administrative issues, which are being addressed.

***Table 13.*** *Use of budget reported per partner*

| Year 1 2019 | | | | | | | |
|---|---|---|---|---|---|---|---|
| **TOTAL SPENT BUDGET** | | | | | | | |
| **No.** | **Partner** | **Personnel €** | **Sub-contracting €** | **Other costs €** | **Indirect costs €** | **Total Spent €** | **% spent** |
| 1 | CESSDA ERIC | 68,887.26 | - | 40,894.68 | 27,445.49 | 137,227.43 | *32.72%* |
| 1.1 | NSD | 5,653.74 | - | 2,545.85 | 2,049.90 | 10,249.49 | *3.59%* |
| 1.2 | UKDS | 25,872.69 | - | 1,169.80 | 6,760.62 | 33,803.11 | *12.24%* |
| 1.3 | UL-ADP | 8,954.32 | - | 1,103.28 | 2,514.40 | 12,572.00 | *14.78%* |
| 1.4 | GESIS | 32,714.55 | - | 1,016.58 | 8,432.78 | 42,163.91 | *14.64%* |
| 1.5 | SND | 13,622.32 | - | 2,311.64 | 3,983.49 | 19,917.45 | *12.85%* |
| 1.6 | UTA-FSD | 43,010.88 | - | 1,743.67 | 11,188.64 | 55,943.19 | *24.47%* |
| 1.7 | FORS | 9,817.35 | - | 2,543.24 | 3,090.15 | 15,450.74 | *13.58%* |
| 1.8 | DNA | 3,605.73 | - | 1,332.65 | 1,234.60 | 6,172.98 | *8.33%* |
| 1.9 | AUSSDA | 24,774.25 | - | 1,427.74 | 6,550.50 | 32,752.49 | *21.83%* |
| 2 | ESS ERIC | - | - | - | - | - | *0.00%* |
| 2.1 | City | 31,792.83 | - | 4,001.41 | 8,948.56 | 44,742.80 | *17.34%* |
| 2.2 | GESIS | 18,143.30 | - | 2,999.13 | 5,285.61 | 26,428.04 | *18.85%* |
| 2.3 | UPF | 35,220.80 | - | 8,441.14 | 10,915.49 | 54,577.43 | *16.48%* |
| 2.4 | NSD | 15,255.15 | - | 2,551.35 | 4,451.63 | 22,258.13 | *3.86%* |
| 3 | SHARE ERIC | - | - | - | - | - | - |
| 3.1 | MPISOC | 30,783.99 | - | 1,483.49 | 8,066.87 | 40,334.35 | *5.25%* |
| 3.2 | UniVe | 30,467.61 | - | - | 7,616.90 | 38,084.51 | *17.89%* |
| 4 | CLARIN ERIC | 155,816.20 | - | 22,684.06 | 43,406.82 | 221,907.08 | *19.48%* |
| 4.1 | Athena | 12,826.17 | - | 3,641.36 | 4,116.88 | 20,584.41 | *14.77%* |
| 4.2 | EKUT | 8,304.01 | - | 1,822.33 | 2,531.59 | 12,657.93 | *5.07%* |
| 4.3 | CUNI | 19,273.46 | - | 2,237.60 | 5,377.77 | 26,888.83 | *14.64%* |
| 4.4 | UL-FF | 3,255.88 | - | 2,170.52 | 1,356.60 | 6,783.00 | *4.73%* |
| 4.5 | RUN | 21,695.00 | - | 96.62 | 5,447.91 | 27,239.53 | *15.73%* |
| 5 | DARIAH ERIC | 31,021.65 | - | 5,327.07 | 9,087.18 | 45,435.90 | *24.85%* |
| 5.1 | PSNC | 41,647.09 | - | 4,486.64 | 11,533.43 | 57,667.16 | *14.02%* |
| 5.2 | OEAW | 38,534.47 | - | 4,884.18 | 10,854.66 | 54,273.31 | *11.37%* |
| 5.3 | UGOE | 5,063.03 | - | 951.45 | 1,503.62 | 7,518.10 | *1.25%* |
| 6 | LIBER | 63,324.24 | - | 7,684.47 | 17,752.18 | 88,760.89 | *19.68%* |
| 7 | KNAW - GGP | 93,729.73 | - | 2,250.41 | 23,995.04 | 119,975.18 | *15.31%* |
| 8 | UoY - ADS | 14,851.17 | - | 1,229.15 | 4,020.08 | 20,100.40 | *21.39%* |
| 9 | TiU | 6,098.89 | - | 993.07 | 1,772.99 | 8,864.95 | *3.17%* |
| 10 | TRUST-IT | 34,239.66 | - | 5,765.03 | 10,001.17 | 50,005.86 | *19.30%* |
| 10.1 | Trust-SRL | 17,736.62 | - | 728.44 | 4,616.27 | 23,081.33 | *14.98%* |
| 10.2 | COMMpla | 16,394.85 | - | - | 4,098.71 | 20,493.56 | *27.61%* |
| 11 | SWC | 48,865.04 | - | 3,771.48 | 13,159.13 | 65,795.65 | *27.21%* |
| 12 | SCIENCES PO | 78,423.61 | - | 8,449.97 | 21,718.40 | 108,591.98 | *14.38%* |
| 13 | UNOTT | - | - | - | - | - | *0.00%* |
| 14 | DAI | 41,641.11 | - | 455.28 | 10,524.10 | 52,620.49 | *35.85%* |
| 15 | CNRS | 67,196.01 | - | 5,809.93 | 18,251.49 | 91,257.43 | *15.54%* |
| 16 | CNR | - | - | - | - | - | *0.00%* |
| 17 | UCL | 20,663.81 | - | 759.11 | 5,355.74 | 26,778.66 | *26.88%* |
| 18 | FORTH | 16,189.84 | - | 3,070.45 | 4,815.07 | 24,075.36 | *16.46%* |
| 19 | CentERdata | 81,755.95 | - | 2,308.87 | 21,016.21 | 105,081.03 | *10.01%* |
| 20 | NG | 3,023.07 | - | 874.53 | 974.40 | 4,872.00 | *4.38%* |
| **Total** | | **1,340,147.33** | **-** | **168,017.67** | **375,823.01** | **1,883,988.01** | ***13.03%*** |

# 5. Conclusions

In the first 12 months, all the Tasks in the SSHOC project initiated their work, gathered their teams, revised the description of action to set up a feasible timeline, and produced their first deliverables. Most work plans for M1-M12 were on schedule. Some minor adjustments were needed in Deliverables and Milestones schedules due to technical difficulties, overlap of delivery dates, or overlap with the holiday season. Minor administrative issues have been discussed with some Beneficiaries in order to be able to meet the requirements for regular internal reporting. However, no significant delay has happened so far that would interrupt work in the Work Packages, nor was any insurmountable obstacle encountered. All the risks have been identified at an early stage and did not affect the ongoing work, as mitigation measures have been applied where needed.

A number of important outputs were already produced, especially plans and strategies for continuation of work and collaboration of partners in general. The team working on the major project deliverable - Social Sciences and Humanities Open Marketplace, has started with the conceptual model and the system architecture of the SSH Marketplace and has delivered by now a minimal viable product (MVP), a running application which demonstrates the connectivity between the main components and exhibits minimal functionality.

The dissemination of the project was focused on reaching the target audiences, and, in the following period, the results directly stemming from deliverables will be digested in communication activities and tailored to target the identified stakeholders, while other activities will focus on connecting to the end-user communities. In addition, connections with other cluster projects and EOSC relevant bodies have been established as planned.

Thus, the Consortium considers the project to be on track and has successfully started the second year of its implementation.

# 6. Appendix – *Tables of events*

| Year 1 | | | | | | |
|---|---|---|---|---|---|---|
| **SSHOC organised & co-organised events** | | | | | | |
| **Conference** | **Event type** | **Title** | **Date** | **No. Attendees** | **Target Stakeholder** | **Link** |
| SSHOC KOM, Utrecht, The Netherlands | Panel discussion | Where Research Infrastructures meet | 10/03/2019 | 100 | SSHOC & EOSC cluster projects | https://www.sshopencloud.eu/news/where-social-sciences-and-humanities-meet-european-open-science-cloud |
| SSHOC workshop, Crete, Greece | workshop | Developing the SSHOC Reference Ontology | 21-22/05/2019 | | Scientific Community | https://www.sshopencloud.eu/sshoc-wp4-workshop-developing-sshoc-reference-ontology |
| LIBER 2019, Dublin, Ireland | workshop | Social Sciences & Humanities Open Cloud: What's In It For Research Libraries? | 26/06/2019 | 60 | Scientific Community | https://www.sshopencloud.eu/liber-annual-event-workshop-social-sciences-humanities-open-cloud-what's-it-research-libraries |
| DH 2019 Utrecht, Netherlands | workshop | The Case Of Interview Data – A multidisciplinary approach to the use of technology in research using interview methods | 09/07/2019 | | Scientific Community | https://www.sshopencloud.eu/digital-humanities-2019-workshop-case-interview-data |
| CLARIN annual conference in Leipzig , Germany | workshop | The Case Of Interview Data – A multidisciplinary approach to the use of technology in research using interview methods | 30/09/2019 | | Scientific Community | https://www.sshopencloud.eu/sshoc-workshop-using-corpora-implementing-validation-workflows-combine-quantity-and-quality |
| RDA Plenary 14, Helsinki, Finland | workshop | EOSC services, collaborations, and RDA | 21/10/2019 | 40 | Scientific Community | https://www.sshopencloud.eu/eosc-services-collaborations-and-rda |
| SSHOC 2nd consortium meeting, Florence, Italy | EOSC ecosystem project presentations slot | Bigger picture and links to - EOSC secretariat, FAIRsFAIR, RDA Europe, EOSChub | 14/10/2019 | 70 | SSHOC & EOSC ecosystem | https://www.sshopencloud.eu/2nd-sshoc-consortium-meeting |
| DARIAH-FI hosted workshop, Finland | workshop | DARIAH-FI workshop "Reuse & sustainability: Open Science and Social Sciences and Humanities research in-frastructures" | 23/10/2019 | | Scientific Community | https://www.sshopencloud.eu/dariah-fi-work%C2%ADshop-"re%C2%ADuse-sustainability-open-science-and-so%C2%ADcial-sci%C2%ADen%C2%ADces-and-hu%C2%ADman%C2%ADit%C2%ADies |
| TEI training tutorial | Tutorial | TEI training tutorial | 28/11/2019 | 40 | Scientific Community | https://ufal.mff.cuni.cz/sites/default/files/hp/tutorial-3.pdf |

| Year 1 | | | | | | |
|---|---|---|---|---|---|---|
| **SSHOC attended events** | | | | | | |
| **Conference** | **Activity** | **Title** | **Date** | **Strake-holders targeted** | **Target Stake-Holder** | **Link** |
| Society for Historical Archaeology, USA | presentation | Roundtable presentation in workshop on Big Data synthesis | 10/01/ 2019 | 20 | Scientific Community | |
| Esfri Ris and EOSC Workshop, London | Workshop presentation | | 30/01/ 2019 | | RIs | https://www.esfri.eu/esfri-ris-eosc-workshop-speakers-ron-dekker |
| Eurise Workshop | Session participation | | 13/03/ 2019 | | RI's | https://www.sshopencloud.eu/eurise-workshop-software-sustainability-within-research-infrastructures |
| 4th International ESS Conference | Conference participation | | 15-17/04/ 2019 | | Scientific Community | https://www.sshopencloud.eu/4th-international-ess-conference-turbulent-times-europe-instability-insecurity-and-inequality |
| CSDI (Comparative studies and implementation workshop), Warsaw | Participation at a Workshop | Designing a Sample Management System for use in a cross-national on-line web panel: initial thinking and ideas | 18-20/03/ 2019 | 75 | Scientific Community | https://csdiworkshop.org/wp-content/uploads/2019/05/Designing-a-Sample-Management-System-for-use-in-a-cross-national-on-line-web-panel.pdf |
| RDA Plenary 13, Philadelphia, US | Presentation | Social Sciences & Humanities Open Cloud | 04/04/ 2019 | 75 | Scientific Community | https://zenodo.org/record/3610082#.XiGFjP5KiCg |
| EOSC hub week 2019 | Session participation | EOSC for Social Sciences and Humanities | 10-12/04/ 2019 | 100 | EOSC ecosystem | https://sshopencloud.eu/eosc-hub-week |
| | Poster presentation | Realising the Social Sciences and Humanities part of the EOSC | 10-12/04/ 2019 | 100 | EOSC ecosystem | https://www.eosc-hub.eu/events/eosc-hub-week-2019/posters# and https://doi.org/10.5281/zenodo.3626960 |
| 4th International ESS Conference | Conference participation | | 15-17/04/ 2019 | | Scientific Community | https://www.sshopencloud.eu/4th-international-ess-conference-turbulent-times-europe-instability-insecurity-and-inequality |
| Computer Applications in Archaeology Annual Conference, Krakow, Poland | Presentation | Digital Infrastructures for Archaeology: Past, Present and Future directions | 24/04/ 2019 | 50 | Scientific Community | https://archaeologydataservice.ac.uk/resources/images/presentations/2019/PDFa/jdr_introduction_CAA_ARIADNEplus_2019.pdf |
| ArchAIDE final conference, Pisa, Italy | Presentation | Open Data and ARCHAIDE | 13-15/05/ 2019 | 70 | Scientific Community | https://archaeologydataservice.ac.uk/resources/images/presentations/2019/PDFa/jdr_Open_Data_and_ArchAIDE.pdf |

| | | | | | | |
|---|---|---|---|---|---|---|
| DARIAH 2019, Warsaw, Poland | Presentation | | 13-17/05/ 2019 | | Scientific Community | https://dariah-ae-2019.sciencesconf.org/261285 |
| C25XVIII International Scientific and Practical conference "BUILDING OF INFORMATION SOCIETY: RESOURCES AND TECHNOLOGIES, Kyiv, Ukraine | Presentation | Building an electronic repository and archives in the European Open Science Cloud Tykhonov, Vyacheslav | 21/05/ 2019 | | Scientific Community | https://zenodo.org/record/3598601#.XiBWLi1x85g |
| Dataverse Community meeting at Harvard University, USA. | Presentation | pyDataverse | 20-21/06/ 2019 | | Scientific Community | https://zenodo.org/record/3265128 |
| | Presentation | Dataverse in the European Open Science Cloud Tykhonov, V; Conzett, P.; Wittenberg, M. SSHOC project presented during Dataverse Community meeting at Harvard University, USA. | 20-21/06/ 2019 | | Scientific Community | https://zenodo.org/record/3598611#.XiA9GS1x-L4 |
| IIIF Conference 2019, Goettingen, Germany | Participation at a conference | | 24-28/06/ 2019 | 300 | Scientific Community | https://iiif.io/event/2019/goettingen/ |
| LIBER 2019, Dublin, Ireland | poster presentation | Realising the Social Sciences and Humanities part of the European Open Science Cloud - Liber | 26-28/06/ 2019 | 450 | Libraries & Archives | https://doi.org/10.5281/zenodo.3633016 |
| DH2019 Conference | SSHOC booth | | 9-12/07/ 2019 | | Scientific Community | https://dh2019.adho.org |
| Synergies for Europe's Research Infrastructures in the Social Sciences: Showcasing the SERISS project, Zagreb, Croatia | Presentation + video | Realising the Social Sciences and Humanities part of the EOSC | 15-19/07/ 2019 | 100 | RIs | https://www.sshopencloud.eu/4th-international-ess-conference-turbulent-times-europe-instability-insecurity-and-inequality |
| "Culture & Technology" European Summer University in Digital Humanities | Lecture | Language as social and cultural data | 02/08/ 2019 | | Scientific Community | https://www.sshopencloud.eu/culture-technology-european-summer-university-digital-humanities |
| European Archaeology Conference, Bern, Switserland | presentation | Making Archaeological Data FAIR | 05/09/ 2019 | 30 | Scientific Community | https://archaeologydataservice.ac.uk/resources/images/presentations/2019/PDFa/jdr_EAA_2019.pdf |
| COST Action CA16111 (ETHMIGSURVEYDATA) meeting, Rome, Italy | Participation in activities organized jointly with other EU project(s) | International Ethnic and Immigrant Minorities Survey Data Network (ETHMIGSURVEYDATA) | 8-9/09/ 2019 | 50 | Scientific Community | https://ethmigsurveydatahub.eu/executive-wg-meeting-and-mc-meeting/ |
| DG CNECT & RTD meeting | Workshop presentation | Building EOSC through H2020 projects | 9-10/09/ 2019 | | EC and EOSC ecosystem | https://www.sshopencloud.eu/building-eosc-through-h2020-projects |

| | | | | | | |
|---|---|---|---|---|---|---|
| CLARIN 2019, leipzig, Germany | poster presentation | Realising the Social Sciences and Humanities for the European Open Science Cloud | 30/09/-02/10/ 2019 | | Scientific Community | https://doi.org/10.5281/zenodo.3633039 |
| Time Machine conference'19, Dresden, Germany | presentation | Running Dataverse repository in the European Open Science Cloud (EOSC), Running Dataverse repository in the European Open Science Cloud (EOSC) T. Vyacheslav | 12/10/ 2019 | | Scientific Community | |
| DDI workshop at GESIS | Participation at a Workshop | | 17/10/ 2019 | | Scientific Community | |
| OS FAIR 2019, Porto, Portugal | poster presentation | Social Sciences and Humanities Community Building: Training - Open Science FAIR | 16-18/10/ 2019 | | Scientific Community | https://doi.org/10.5281/zenodo.3633059 |
| | presentation | Open Science Training for Library Staff & Researchers | 16-18/10/ 2019 | 60 | Scientific Community | https://zenodo.org/record/3427305#.XiAnCf5KiCg |
| RDA P14 - EOSC at RDA, Helsinki, Finland | Participation to a Workshop | The International Research Community Contributing to the EOSC | 21-25/10/ 2019 | 250 | EOSC ecosystem | https://www.sshopencloud.eu/international-research-data-community-contributing-eosc |
| | Presentation | Building Dataverse Communities that follow RDA Best Practices for Data Sharing and Management. Part of Presentation at RDA 2019 in Helsinki, FI by V.Tykhonov | 21-25/10/ 2019 | | Scientific Community | https://zenodo.org/record/3598658#.XiA71y1x-L4 |
| | Booth | EOSC KIOSK | 21-25/10/ 2019 | | Scientific Community | |
| CESSDA's Mentorship Programme Promotes Knowledge Exchange between European Data Archives | Participation at a Workshop | | 5-6/11/ 2019 | | Libraries & Archives | https://www.cessda.eu/News-Events/News/CESSDA/CESSDA-s-Mentorship-Programme-Promotes-Knowledge-Exchange-between-European-Data-Archives |
| DESIR Final Event, Zagreb, Croatia | Participation at conference | | 5-6/11/ 2019 | 60 | Scientific Community | https://www.dariah.eu/2019/10/30/desir-final-event-and-dariah-general-assembly-in-november/ |
| Open Access Ambassadors Conference | Presentation at Conference | | 10-11/12/ 2019 | | Scientific Community | https://oambassadors.mpdl.mpg.de/ |

| | | | | | | |
|---|---|---|---|---|---|---|
| CLARIAH-DE 1st General Assembly, Tubingen, Germany | Participation AT conference | | 14-15/11/ 2019 | 50 | Scientific Community | https://hdl.handle.net/10037/15901 |
| SHAREcare, Amsterdam, The Netherlands http://sharecare.nu/presentations-amsterdam/ | Presentation | | 22/11/ 2019 | 50 | Scientific Community | https://www.slideshare.net/ShareCareX/07-reusable-padfield |
| Europeana 2019 | Workshop presentation | New perspectives for Digital Cultural Heritage | 27/11/ 2019 | | Scientific Community | https://www.sshopencloud.eu/europeana-2019-eosc-digital-cultural-heritage |
| EOSC symposium 2019 | Panel discussion | Social & Cultural Data – taking the Users' Perspective | 27-29/11/ 2019 | 200 | EOSC ecosystem | https://www.sshopencloud.eu/eosc-symposium-2019 |
| | Poster presentation | A marketplace for the Social Sciences and Humanities | 27-29/11/ 2019 | 200 | EOSC ecosystem | https://doi.org/10.5281/zenodo.3550957 |
| | Booth | EOSC KIOSK | 27-29/11/ 2019 | 200 | EOSC ecosystem | |
| EDDI conference | presentation | Enrichment of DDI support in the Dataverse data repository. Tykhonov, S.; Wittenberg, M. | 12/01/ 2019 | | Scientific Community | https://zenodo.org/record/3600093#.XiA-bS1x-L4 |
| TRIPLE kick-off meeting | Presentation at Conference | | 04/12/ 2019 | | EOSC ecosystem | https://operas.hypotheses.org/3007 |

# 7. References

**WP1**

- Basecamp: Project Management & Team Communication Software: https://basecamp.com/
- The internal SSHOC Project repository stored at CESSDA Shared Drive and accessible to all project partners: https://drive.google.com/drive/shared-drives
- Details on all the tools provided for the project are explained in the Milestone 1 report shared with partners in the SSHOC document repository
- Amir Spahic, Ivana Ilijasic Versic, Franciska De Jong, & Cees van der Eijk. (2019). SSHOC D1.1 Project Management Plan (Version v1.3). DOI: 10.5281/zenodo.3595441
- Project website of EURhisFIRM EC funded project: https://eurhisfirm.eu/
- EOSC Life project website: https://www.eosc-life.eu
- ESCAPE project website: https://projectescape.eu
- ENVRI-FAIR website: https://envri.eu/about-envri-fair/
- PANOSC project website: https://www.panosc.eu
- Martina Drascic, Ivana Ilijasic Versic, Carsten Thiel, Daan Broeder, Gianmaria Bottoni, Marieke Willems, … Vasso Kalaitzi. (2019). SSHOC D1.2 Quality Assurance & Risk Assessment Plan (Version V1.2). DOI: 10.5281/zenodo.3595453
- Internal Wiki page available to SSHOC partners: https://sites.google.com/cessda.eu/sshocwiki/

**WP2**

- SSHOC Website publication of Deliverables produced: https://www.sshopencloud.eu/publications/deliverables
- SSHOC newsletter: https://www.sshopencloud.eu/newsletter/sshoc-newsletter
- Publication in EOSC-Hub magazine: https://www.eosc-hub.eu/sites/default/files/Magazine%233_Web_version%20Final-compressed.pdf
- Ilijašić Veršić, I. und Ausserhofer, J. (2019) „Die Sozial- und Geisteswissenschaften und ihre Interoperabilität mit der European Open Science Cloud: Was ist SSHOC?“, Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare, 72(2), S. 383-391. doi: 10.31263/voebm.v72i2.3216.
- SSHOC Twitter: https://twitter.com/SSHOpenCloud
- SSHOC LinkedIn: https://www.linkedin.com/company/18997546/
- SSHOC Flicker: https://www.flickr.com/photos/170844433@N02/
- SSHOC Youtube channel: https://www.youtube.com/channel/UCw-mY8v84yeHW2z4KG3ZLtA
- SSHOC Slideshare: https://www.slideshare.net/SSHOCInitiative
- SSHOC logo pack: https://www.sshopencloud.eu/logo-pack
- SSHOC brand guidelines: https://www.sshopencloud.eu/sites/default/files/SSHOC%20Guideline_booklet.pdf
- SSHOC communications kit: https://www.sshopencloud.eu/communications-kit
- General SSHOC video: https://www.youtube.com/watch?v=UP7dBiV3-yQ
- Interview with Ron Dekker, CESSDA Director: https://www.youtube.com/watch?v=gUWEragkniw

- Interview with Franciska de Jong, CLARIN-ERIC executive director: https://www.youtube.com/watch?v=cgJbOgfXvzY
- Interview with Astrid Verheusen, LIBER executive director: https://www.youtube.com/watch?v=eqoCi1Cu2yY
- Interview with Frank Fischer, DARIAH co-director: https://www.youtube.com/watch?v=xfWuho7tVfA
- Launch of the SSHOC Training Community - Interview with Ellen Leenarts - DANS-KNAW: https://www.youtube.com/watch?v=DWuHgQGKhdA
- SSHOC flyer: https://www.sshopencloud.eu/sites/default/files/SSHOC_flyer_A5_October2019.pdf
- SSHOC banner: https://www.sshopencloud.eu/sites/default/files/sshoc_banner_0.pdf
- https://www.sshopencloud.eu/sites/default/files/SSHOC%20Post-IT_76x76_August2019_V0.2_0.pdf
- SSHOC Launch : https://www.sshopencloud.eu/news/"sshoc"-social-sciences-humanities-open-cloud-acronym-rememberMarketplace System specification: https://www.sshopencloud.eu/sites/default/files/PR_SSHOC%20Releases%20System%20Specification%20for%20SSH%20Open%20Marketplace_201912.pdf
- News on SSHOC website: https://sshopencloud.eu/news/my-cloud-your-cloud-our-cloud
- News on SSHOC website: https://www.sshopencloud.eu/news/eosc-esfri-cluster-projects-rda%C2%A0connecting-commonalities-and-collaborative-solutions-community
- News on SSHOC website: https://sshopencloud.eu/rda-14th-plenary-data-makes-difference
- News on SSHOC website: https://sshopencloud.eu/eosc-symposium-2019
- SSHOC Dashboard: https://datastudio.google.com/reporting/1JTtWjrE7ee81DbpkWZc6596J22VNQncR
- SSHOC ZENODO Community: https://zenodo.org/communities/sshoc/

**WP3**

- SSHOC Milestone 10 report: MS10 Inventory of SSHOC metadata sets and taxonomies; available in the SSHOC project document repository, available to project partners
- Nicolas Larrousse, Daan Broeder, Jan Brase, Cesare Concordia, & Vasso Kalaitzi. (2019). SSHOC D3.2 Inventory of SSH citation practices, and choice for SSHOC citation formats and implementation planning (Version v1.0). DOI: 10.5281/zenodo.3595965.

**WP4**

- Aïoli", a reality-based 3D annotation platform for the collaborative documentation of cultural heritage artefacts: www.aioli.cloud
- CIDOC-CRM: CIDOC Conceptual Reference Model: http://www.cidoc-crm.org/
- CRMdig: CRM Digital; a model for provenance metadata: http://www.cidoc-crm.org/crmdig/
- CRMsci: CRM Scientific Observation Model: http://www.cidoc-crm.org/crmsci/
- Athens 2019: http://sodanet.gr/en/training/workshops/cessda-training-workshop-athens-2019
- DARIAH Annual event 2019: https://www.dariah.eu/event/dariah-annual-event-2019/

**WP5**

- Raphael Research Resource: https://cima.ng-london.org.uk/documentation

- GitHub LCD example: https://github.com/jpadfield/bg-lcd-example
- AHRC: https://ahrc.ukri.org/
- Linked Conservation Data: https://www.ligatus.org.uk/lcd/
- IPERION-CH and second data source: https://research.ng-london.org.uk/iperion/
- Presentation on IPERION-CH D.8.5: Completed example of prototype designs for integration of various types of documentation and analytical data generated for a single object: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c8d1c525&appId=PPGMS
- New website (development in process): https://rdf.ng-london.org.uk/modelling/
- Github CIDOC-CRM: https://github.com/jpadfield/cidoc-crm.examples
- Dataverse (https://dataverse.org/)
- Docker (https://www.docker.com)
- SSHOC dataverse on GITHUB (release): https://github.com/IQSS/dataverse/releases/tag/v4.18
- Vigdis Kvalheim, Marianne Høgetveit Myhren. New legislation – a unique opportunity for harmonising the legal framework for research in the Nordic countries. NSD. Vår ref.: 201600326/10 MHM/INGU Arkiv: 153; Website: https://nsd.no/personvernombud/dok/position-paper-new-legislation.pdf
- Microdata project: https://microdata.no/en/about/
- E-RIHS - European Research Infrastructure for Heritage Science - http://www.e-rihs.eu/.
- Website of Cost Action on Saving European Archaeology from the Digital Dark Age: https://www.cost.eu/actions/CA18128
- Website of Raphael Research Resource: https://cima.ng-london.org.uk/documentation
- Website of Linked Conservation Data Example: https://rdf.ng-london.org.uk/workshops/lcd
- Website of Linked Conservation Data funded by AHRC: https://www.ligatus.org.uk/lcd/
- Portal of collaborative research resource of Iperion: https://research.ng-london.org.uk/iperion/ and http://www.iperionch.eu/
- PERION-CH Deliverable D8.5: Completed example of prototype designs for integration of various types of documentation and analytical data generated for a single object: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c8d1c525&appId=PPGMS
- ChronOntology gazetteer web service, see: https://chronontology.dainst.org/info/about

**WP6**

- V. Kalaitzi; M. Torma; F. Grant; I. Vipavc Brvar; D. Fišer; K. Pahor de Maiti; C. Petitfils; M. Durco; M. Willems; 2019; SSHOC D6.1 Community Engagement Strategy; https://zenodo.org/record/3592244#.XiWeIPlKipp
- M. Willems; L. Marino; S. Muschella; A. Schwabe; J. Ausserhofer; E. Smith; V. Kalaitzi; I. Vipavc Brvar; 2019; D2.1 SSHOC Overall Communication and Outreach Plan; https://sshopencloud.eu/sites/default/files/D2.1_SSHOC_Overall_Communication_Outreach_Plan_v1.0.pdf

- M. Torma; V. Kalaitzi; E. Dijk; M. Wittenberg; M. Willems; D. Fišer; K. Pahor de Maiti; M. Durco; I. Vipavc Brvar; 2019; D6.2 SSHOC Building Expertise Strategy; https://zenodo.org/record/3595996#.XiWf-flKipp
- Event; EOSC services, Collaborations, and RDA; 2019; https://www.sshopencloud.eu/eosc-services-collaborations-and-rda
- M. Durco; K. Illmayer; L. Barbot; 2019; D6.7 SSHOC Inventory of existing learning materials; https://zenodo.org/record/3596003#.XiWgM_lKipo
- SSHOC SSH Training Community registration form; https://www.shopencloud.eu/join-ssh-training-community
- LIBER working group on digital skills for library staff and researchers; https://libereurope.eu/strategy/digital-skills-services/digitalskills/
- LIBER working group on digital humanities and digital cultural heritage; https://libereurope.eu/strategy/digital-skills-services/digitalhumanities/
- YouTube video; Launch of the SSHOC Training Community - Interview with Ellen Leenarts at OS Fair 2019; https://www.youtube.com/watch?time_continue=1&v=DWuHgQGKhdA&feature=emb_logo
- Event; Working with interview data: SSHOC workshop on a multidisciplinary approach to the use of technology in research; 2019; https://sshopencloud.eu/news/working-interview-data-sshoc-workshop-multidisciplinary-approach-use-technology-research
- Event; Using corpora for implementing validation: SSHOC masterclass on workflows that combine quantity and quality; 2019; https://sshopencloud.eu/news/using-corpora-implementing-validation-sshoc-masterclass

**WP7**

- The SSHOC code base**:** https://gitlab.gwdg.de/sshoc
- TAPoR website: http://tapor.ca/
- Unified Views website: https://unifiedviews.eu/
- PoolParty website: https://www.poolparty.biz/

**WP8**

- Core Trust seal website: https://www.coretrustseal.org/
- FAIRsFAIR project website: https://www.fairsfair.eu/

**WP9**

- EURhisFIRM project website: https://eurhisfirm.eu/index.php/the-project-2/
- ETHMIGSURVEYDATA website: https://ethmigsurveydatahub.eu/
- OpenAire network website: https://www.openaire.eu/

# List of Figures



# List of Tables