

# FOSSR

## FOSTERING OPEN SCIENCE IN SOCIAL SCIENCE RESEARCH

### DELIVERABLE 4.4

LOREDANA CERBARA<sup>1,\*</sup>,  
NICOLÒ MARCHESINI<sup>1</sup>, ANDREA GIOVANNI NUZZOLESE<sup>2</sup>,  
MARIO PAOLUCCI<sup>1</sup>, ROCCO PAOLILLO<sup>1</sup>,  
SARA ZUPPIROLI<sup>2</sup>

CNR - INSTITUTE FOR RESEARCH ON POPULATION AND SOCIAL POLICIES  
CNR - INSTITUTE OF COGNITIVE SCIENCES AND TECHNOLOGIES

26/02/2024

\* THE AUTHORS ARE LISTED ALPHABETICALLY AS THEY EQUALLY CONTRIBUTED TO THIS DELIVERABLE.



<i>Project Acronym</i>	<i>FOSSR</i>
Agreement number	IR0000008
Project Full Title	Fostering Open Science in Social Science Research
Funding Scheme	Public notice of the Ministry of University and Research for the presentation of project proposals for “Strengthening and creating Research Infrastructures” to be financed under the PNRR

## DELIVERABLE INFORMATION

<i>Deliverable Number</i>	<i>D4.4</i>
Deliverable Name	Development of the digital platform and integration with Open Cloud.
Dissemination level	Confidential
Funding Scheme	Public notice of the Ministry of University and Research for the presentation of project proposals for “Strengthening and creating Research Infrastructures” to be financed under the PNRR
Grant Agreement number	IR0000008
Contractual date of delivery	31/01/2024
Deliverable Leader	Andrea Giovanni Nuzzolese
WP/Task Responsible	Mario Paolucci (WP responsible) / Andrea Giovanni Nuzzolese (Task responsible)
Keywords	Online survey; Digital platform; Panel management; Panel Platform design
Abstract	<p>The preparation and implementation of multimodal interviews requires the availability of a complex application capable of interfacing with different types of users: system administrators, users with access authorisation for field monitoring, interviewers and interviewees. The latter are put in contact with the system through a dedicated application, an interface that allows direct communication with the interviewees but which also supports more extensive functions, such as the view of the activities carried out by each interviewee.</p> <p>On the main application side, the panel management and monitoring functions must be implemented, starting from the insertion of profiling data, but also the management of the interviews of each wave and the activity of the panellists.</p> <p>It is, therefore, a document necessary for the design and implementation phase of the main infrastructure management software in its part dedicated to the production of original data for the social sciences, which takes into account the needs, the</p>

	components already available, and the possible technical solutions that will support the final architecture of the product.
--	---

## DISCLAIMER

The content of this Deliverable reflects only the author's view. The European Commission and MUR are not responsible for any use that may be made of the information it contains.

## DOCUMENT HISTORY

<i>Date</i>	<i>Version</i>	<i>Contributor(s)</i>	<i>Description</i>	<i>Supervision</i>
19/01/2024	0.1	Sara Zuppiroli	Initial draft	All authors
22/01/2024	0.2	Sara Zuppiroli	Introduction	All authors
28/01/2024	0.3	Sara Zuppiroli	Selection of the web programming tool	All authors
29/01/2024	0.4	Rocco Paolillo	Customised features	All authors
31/01/2024	0.5	Rocco Paolillo	Realization Phase	All authors
02/02/2024	0.6	Sara Zuppiroli	Paint Diagrams	All authors
04/02/2024	0.7	Nicolò Marchesini	Tool Specification	All authors
05/02/2024	0.8	Nicolò Marchesini	Tool Specification	All authors
08/02/2024	0.9	Nicolò Marchesini	Tool Specification	All authors
09/02/2024	1.0	Sara Zuppiroli	Description Diagrams, Non functional requirements	All authors
21/02/2024	1.1	Nicolò Marchesini	Integration of "Tool Specification" section	All authors
21/02/2024	1.2	Nicolò Marchesini	English Revision	
21/02/2024	1.3	Loredana Cerbara	Abstract	All authors
21/02/2024	1.4	Loredana Cerbara	Revision of Tool specification and Customised features sections	All authors
26/02/2024	2.0	Loredana Cerbara, Nicolò Marchesini, Rocco Paolillo	General revision	All authors

## SUMMARY

<b>DELIVERABLE INFORMATION</b> .....	<b>1</b>
<b>DISCLAIMER</b> .....	<b>2</b>
<b>Document history</b> .....	<b>2</b>
<b>Summary</b> .....	<b>3</b>
<b>List of Figures</b> .....	<b>3</b>
<b>List of acronyms</b> .....	<b>3</b>
<b>1. Introduction</b> .....	<b>6</b>
1.1. FOSSR Objectives and Ambition .....	6
1.2. Purpose and scope of this document .....	7
<b>2. Tool specification</b> .....	<b>8</b>
2.1 Overview of all Process .....	11
2.1.1 Process to create Panel .....	12
2.1.2 Process Wave .....	14
2.3 Non-functional requirements: an overview of security .....	15
<b>3. Selection of the web programming tool</b> .....	<b>16</b>
<b>4. Customised features</b> .....	<b>18</b>
<b>5. REALISATION PHASE</b> .....	<b>19</b>
5.1 The ideal provider .....	19
5.2 System Security .....	19
<b>References</b> .....	<b>20</b>

## LIST OF FIGURES

Figure 1. Activity diagram of all process .....	12
Figure 2. Activity diagram Create Panel .....	13
Figure 3. Activity diagram wave .....	15

## LIST OF ACRONYMS

<i>ACRONYM</i>	<i>Description</i>
ANVUR	Agenzia nazionale di valutazione del sistema universitario e della ricerca
BOCCONI	Università degli studi di Milano Bocconi
CA	Consortium agreement

CAPI	Computer Assisted Personal Interviewing
CATI	Computer Assisted Telephone Interviewing
CAWI	Computer Assisted Web Interviewing
CESSDA	Consortium of European Social Science Data Archives
CID ETHIC	Centro Interdipartimentale per l'Etica e l'Integrità nella Ricerca
CQ	Core Questionnaire
CRUI	Conferenza dei Rettori delle Università italiane
DDI	Data Documentation Initiative
DX.Y	Deliverable X. Y (X refers to the WP and Y to the deliverable in the WP)
DSU	Dipartimento Scienze Umane e Sociali, Patrimonio Culturale
EAB	External advisory board
EC	Ethical committee
EOSC	European Open Science Cloud
EU	European union
ExCO	Executive committee
FM	Financial manager
GB	Governing board
GDPR	General Data Protection Regulation
ICAR	Istituto di Calcolo e Reti ad Alte Prestazioni
IM	Infrastructure manager
INAPP	Istituto nazionale per l'analisi delle politiche pubbliche
IOPP	Italian Online Probability Panel
IRPPS	Istituto di Ricerche sulla Popolazione e le Politiche Sociali
ISTC	Istituto di Scienze e Tecnologie della Cognizione,
ISTI	Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo"
JRU	Joint research units
KoM	Kick-off Meeting
KPI	Key performance indicators
LUISS	Libera università internazionale degli studi sociali Guido Carli
MEF	Ministero dell'Economia e delle Finanze
MISE	Ministero delle Imprese e del Made in Italy
MUR	Ministero dell'università e della ricerca
NRRP	The National Recovery and Resilience Plan (
PM	Project manager
PQ	Profiling Questionnaire
RI	Research Infrastructure
RISIS	Research infrastructure for research and innovation policy studies
SAB	Stakeholders advisory board
SB	Scientific advisory board
SC	Scientific Coordinator
SMC	Strategic management committee
UNIBO	Università degli Studi di Bologna
UNICATT	Università Cattolica del Sacro Cuore
UNIMI	Università degli Studi di Milano
UNISI	Università di Siena
UVR	Unità Valorizzazione della Ricerca



## 1. INTRODUCTION

The background mission of the FOSSR project relies on creating and strengthening awareness and knowledge of data and methodologies used in empirical social sciences among a wide audience through maintaining and reinforcing relevant Research Infrastructures (RIs) while fostering the development of a research and social environment conducive to an open, shared and simplified data access via innovative interfaces. The project will concretely contribute to the effective implementation of the ‘open science’ for social science researchers by providing innovative tools and services for research, a shared virtual environment for data access and a wide and varied package of training courses, sessions and programmes (FOSSR DoW, 2022). FOSSR adopts the common theme of developing Open Science in the Italian context to create a framework of tools and services for the social science scholar community involving the RIs in social sciences coordinated by CNR, namely CESSDA, SHARE and RISIS. The framework should be an integrated knowledge-sharing platform, a single point of access to all the tools and services made available by the Italian nodes of social science infrastructures.

FOSSR fosters the building of an Italian Open Science Cloud, along the lines of the European Open Science Cloud project, to integrate innovative services developed by the project for data collection, data curation and fairness and data analysis on economic and societal change.

FOSSR wants to promote to multiple audiences a widespread knowledge and awareness of the data and methodologies employed in empirical social science, fostering the growth of a broad societal environment favourable to further thriving of social science research in Italy, providing easy, open, streamlined access to social science data through innovative interfaces.

### 1.1. FOSSR Objectives and Ambition

FOSSR has the general aim of promoting, towards multiple audiences, a widespread knowledge and awareness of the data and methodologies employed in empirical social science, by providing (i) systematic and organised knowledge (also through summary harmonised data) about available social science data resources in Italian data archives, especially the CESSDA Archive, already object of the grand infrastructural proposal; (ii) resources supporting methodological advancement as to data collection and data analysis, especially important for RISIS to understand the design, the implementation, and the outcome of research and innovation policies, which can improve the robustness of empirical evidence produced for policy makers and to deal with new research questions, and (iii) tools and services to make publicly available advanced probability panels for longitudinal analyses to support important survey such as SHARE, complementing them with a network of online laboratories. The integration of this pool of resources shall concretely contribute to the realisation of open science for scholars in social sciences, going with an important program of scientific training for the production and analysis of social science based on FAIR empirical data.

One further key objective is setting a new generation of young researchers in social sciences by hiring several researchers and technologists with fixed time contracts, which will become highly skilled human resources in data science and data management in social science research, and by funding 20 PhD positions to train early career researchers in the field.

FOSSR is also aimed at fostering the growth of a broad societal environment favourable to the further thriving of social science research in Italy, providing easy, open, streamlined access to social science data through innovative interfaces (data exploration portals, time series, interactive visualisations), and online data analysis software aimed at students, along with

divulgarion resources about social science methodology. These aspects are particularly aimed at civil society organisations (NGOs, etc.), students, and ordinary citizens to foster a widespread societal awareness of social science data, results, and methods to promote easier and more user-friendly dissemination.

The main investment shall be in creating a suitable IT infrastructure and a network of data centres to provide researchers access online and onsite and to support the workflow of the proposed collaborative projects. One of the subsequent goals of FOSSR will be to develop innovative tools and services for data collection and analysis and to support participating users in acquiring and using advanced equipment and software for social science research needed for collaborative studies and projects. The achievement of the goals mentioned above can be reached through an online platform – the Open Cloud, also acting as a dissemination layer, intermediating between data producers, archives, and the broader shop floor of users (both scholars and stakeholders), assuring access intermediating between data producers, archives, and the broader shop floor of users (both scholars and stakeholders), assuring access to multiple software interfaces, geared at different audiences, methodological content, and training materials.

## 1.2. Purpose and scope of this document

Social science is one of the branches of science devoted to studying societies and the relationships among individuals within those societies. The branches of social science include anthropology, economics, political science, psychology, and sociology. Population growth and the complexity of modern society have made the social sciences of utmost importance. Thanks to them, it is possible to thoroughly understand social dynamics to act by preventing or solving related problems. This approach requires much information to build models to simulate and understand individual and societal behaviours properly. Therefore, the interaction of the scientific community and the ability of scholars to access this information quickly and efficiently is of great importance.

FOSSR aims to create an Italian Open Science Cloud for the Social Sciences, providing innovative tools and services to investigate issues related to contemporary societies' economic and societal change. To achieve such a result, referring to RIs in the social sciences that make FAIR-type data available is of primary importance. This type of data will be obtained by the RIs involved in social sciences coordinated by CNR, namely:

- CESSDA (Consortium of European Social Science Data Archives) provides the scientific community with facilities, tools, datasets, and certified services to conduct research activities of excellence in the social sciences domain.
- SHARE (Survey of Health, Aging, and Retirement in Europe) is an interdisciplinary and longitudinal survey on the economic, social, health, and well-being conditions of the 50+ population in twenty-seven European countries (plus Israel).
- RISIS (Research Infrastructure for Research and Innovation (R&I) Policy Studies), which provides data and services to support the development of a new generation of analyses and indicators for the study of science, technology, and innovation processes based on three main perspectives: actors involved to understand the role they play, topics addressed to understand the directionality of the R&I efforts and geography of science and innovation.

FOSSR shall incorporate tools (hardware and software) and methods functional to research practices traceable to the paradigms of e-science, behavioural economics, and computational social sciences. In operational terms, the functionalities for which important innovations are expected are data collection, data integration, data curation, data sharing, creating a survey



facilitator, constructing a social listening structure, and activating an artificial population facility.

In operational terms, this framework should be an integrated knowledge-sharing platform, a single point of access to all the tools and services made available by the Italian nodes of social science infrastructures.

FOSSR has the general aim of promoting, towards multiple audiences, widespread knowledge and awareness of the data and methodologies employed in empirical social science by providing (i) systematic and organised knowledge about available social science data resources in Italian data archives; (ii) resources supporting methodological advancement as to data collection and data analysis, (iii) tools and services to make publicly available advanced probability panels for longitudinal analyses to support important survey. The integration of this pool of resources shall concretely contribute to the realisation of open science for scholars in social sciences, going with an essential program of scientific training for the production and analysis of social science based on FAIR empirical data.

This document aims to describe activity 4, which involves the analysis of requirements for the design and implementation of multi-modality interview software and the management of panellists. The document addresses the **milestone “Implementation of multi-mode interviewing”**, due at the month 15 of the project. The remainder of the deliverable is organised as follows: Chapter 2 presents a brief description of our high-level requirements (HLRs); Chapter 3 presents a list of possible software to cover these requirements, among which we identify the software chosen as a tool for managing the surveys and the panel; Chapter 4 shows the customisations identified for the FOSSR project; finally, Chapter 5 addresses the choices to ensure the security of the system and which parameters to use for choosing the company that will support the implementation of the customisations.

This deliverable will serve as an implementation guideline during the development step.

## 2. TOOL SPECIFICATION

In the creation of the infrastructure for social sciences within FOSSR, i.e. the Italian Online Probability Panel (IOPP) representing the Italian population, several High-Level Requirements (HLRs) have been identified (detailed in Deliverable 4.1A “Definition of the strategy for building the digital platform using Open Cloud”<sup>1</sup>) to managing and monitoring the panellists’ data as well as managing the surveys.

Specifically, they are grouped into five main categories:

- HLR-01: Panel Design.

The tool to create and monitor the panel, i.e. each individual part of the representative sample followed over time, should have a system to set, save, and edit individual attributes as well as to map potential changes and save them if needed and approved by the central hub. Changing shouldn’t be an overwriting but a new attribute added to the individual. For instance, place of residence, marital status, number of children, or current profession are individual attributes evolving. Through the system, the life histories of individuals should be able to be tracked.

- HLR-02: Panel Management

The system must collect data regarding the panellists recruitment phase (CAPI) and the proper survey waves (CAWI).

---

<sup>1</sup> Compared to deliverable 4.1A, the names of the HLRs have been updated to fit more closely with their description.

Concerning the recruitment phase of panellists, the system must collect data about the CAPI interview. Such pieces of information regard i) the set of information about how the interview has been conducted, such as the number of contacts needed to interview the participants or timing of the interview; ii) individual consensus to partake in the online longitudinal panel according to the GDPR and project data protection policy; iii) set of information coming from the individual profile survey regarding the main socio-demographic characteristics, e.g., sex, age, marital status, no. of children, or profession, and such characteristics will become the attributes of each panellist involved in the longitudinal panel.

As soon as the individuals agree to join the panel, they will fill the waves in through an online questionnaire. The system should: i) alert each panellist of the new wave to fill in by a certain deadline; ii) send alerts (via email and social profiles) to complete the new survey; iii) create summary statistics about the panellist participation in the longitudinal panel; iv) notify the central hub as soon as participants are lacking one or more inclusion criteria for being part of the online panel (such as relocation abroad or exceeding the age limit); v) allowing the central hub to delete panellists.

- **HLR-03: Sub-sample management**

Since the representative sample of the longitudinal panel will allow targeting specific survey modules based on the research needs, the system should i) permit the survey managers to filter panellists by attributes; ii) send the invitation to the specific wave to the sub-sample; iii) registering the responses to the sub-modules without altering the general flow and architecture of records for the whole panel; iv) producing summary statistics for the sub-modules.

- **HLR-04 Survey design**

The system will allow the design of the online survey step-by-step, including different types of questions and answers, usage of images or videos, and sizing the survey to different devices potentially used to fill in. Additionally, the system should integrate such management with panel management.

- **HLR-05 Panel Maintenance**

Panellists will receive a minimum of five waves per year. The architecture system will record each single wave subsequently and not overwrite previous data. The system will constantly monitor responses, producing statistical reports and breaking down information by individual, sampling stratum, and other structural attributes. Such analysis will be the basis for planning further strategies to retain panellists and limit potential drop-outs, such as defining specific recalls (for instance, via social media) or re-shaping the incentive formula.

Additionally, to foster the panel life, the system will produce a basic mortality analysis on the panellists to plan refreshment sample strategies, and it will be able to import such refreshment samples into the system accordingly.

Regarding panel maintenance, the system under development will also include the purchase or realisation of an app that the individual panellist will use to respond to waves from the mobile application. This app, in addition to providing a secure tool for data collection, will be implemented to issue notifications for each available wave and as a tool for self-updating profiling data if it changes from the profiling questionnaire filled in by the panellist. The central hub will approve self-editing profiling data before being registered in the database.

To support the implementation of the HLRs mentioned above as a multi-mode platform, the software being identified must meet the following requirements (detailed in Deliverable 3.3 “Preparation for and implementation of multi-mode interviewing”):

- **HLR-01 Creation and management of user groups**

Since different types of users will have access to the platform - e.g., FOSSR central hub, partners’ employees, and external users - the platform must permit the creation and management of different types of users grouped by superuser.

Each user in a group is defined by characteristics and privileges which allow that group's users to perform or not perform specific actions.

Inserting a user into a group can occur through the self-registration of the individual user.

The FOSSR team could constitute a particular user group composed of administrators and high-level users, while external users placed in a different group should have limited options.

Among the privileges that can be assigned to a user in a group, the following can be identified: the possibility of creating, modifying, using, publishing, and testing surveys; the ability to create new users within a specific group; the possibility of creating interaction methods such as calls and forums.

- HLR-02 Survey management

The software will allow the creation and editing of questions and sub-questions in the various possible types:

- Yes/no questions
- Multichoice questions (w/ comments)
- Scales/Array (e.g. Likert) and Array of text
- Array by column and two/three scales
- Array with dropdown
- Dropdown questions;
- Ranking questions
- Date/time questions;
- Numerical and MULTInumerical inputs
- Bootstrap questions
- List/Radio questions (w/ comments)
- Text questions (short and long answers)
- Multiple texts
- Management of arrangements for the questionnaire's accessibility, multilingual question management, and conditional logical submission of questions.

Furthermore, creating a new survey will be possible by importing an existing one from another survey system/software.

- HLR-03 Interview management

Concerning the single interview, or wave, the software will allow the administrator and the superuser to add, delete, or edit the participants' list to set up the type of performed interview (e.g., CAPI online or CAWI from desktop or mobile device), recording collected data according to the general scheme.

- HLR-04 Receive Data survey

The software will record the participants' answers to the different surveys they are exposed to, counting the number of waves, rounds, or sub-rounds if the survey targets a specific sub-sample. Each participant should have the option to save the survey, log out, and log in at different times within the given compilation time frame. The data will be recorded once the survey is completed or the time frame expires.

## 2.1 Overview of all Process

The activity diagram in Figure 1 represents the sequence of high-level activities in the annual waves. The first year of activity will involve creating the panel and designing the Profiling Questionnaire (PQ) and the Core Questionnaire (CQ). The PQ will be administrated during the panellists' recruitment phase for the initial and refreshment samples. The CQ, defined by the FOSSR central team, will comprise at least five sections (one per wave) and will cover crucially social topics such as working conditions, family issues, migration, or climate change. The CQ will be administrated annually to the panellists in the same order throughout the year. In the

first year after the panel establishment, the CQ will be administered through the five waves. From the following year, each wave will be associated with a specific survey module resulting from the open call addressed to the FOSSR stakeholders community. The open call and related applications will be managed and scheduled annually by the FOSSR central hub. Once the scheduling for a wave is completed and the questionnaire start date approaches, the questionnaire will be sent to the panellists. Each panellist, upon responding, will be presented with their profile page and the project's privacy statement. Panellists who do not respond are monitored and prompted to respond through various selected and personalised communication channels. Upon completion, the FOSSR group will review, correct, and make the data open and then distribute it to the various commissioning research groups. Concurrently, a group will oversee and manage the panel. At the end of the annual waves, a new check and update will be conducted on the panellists.

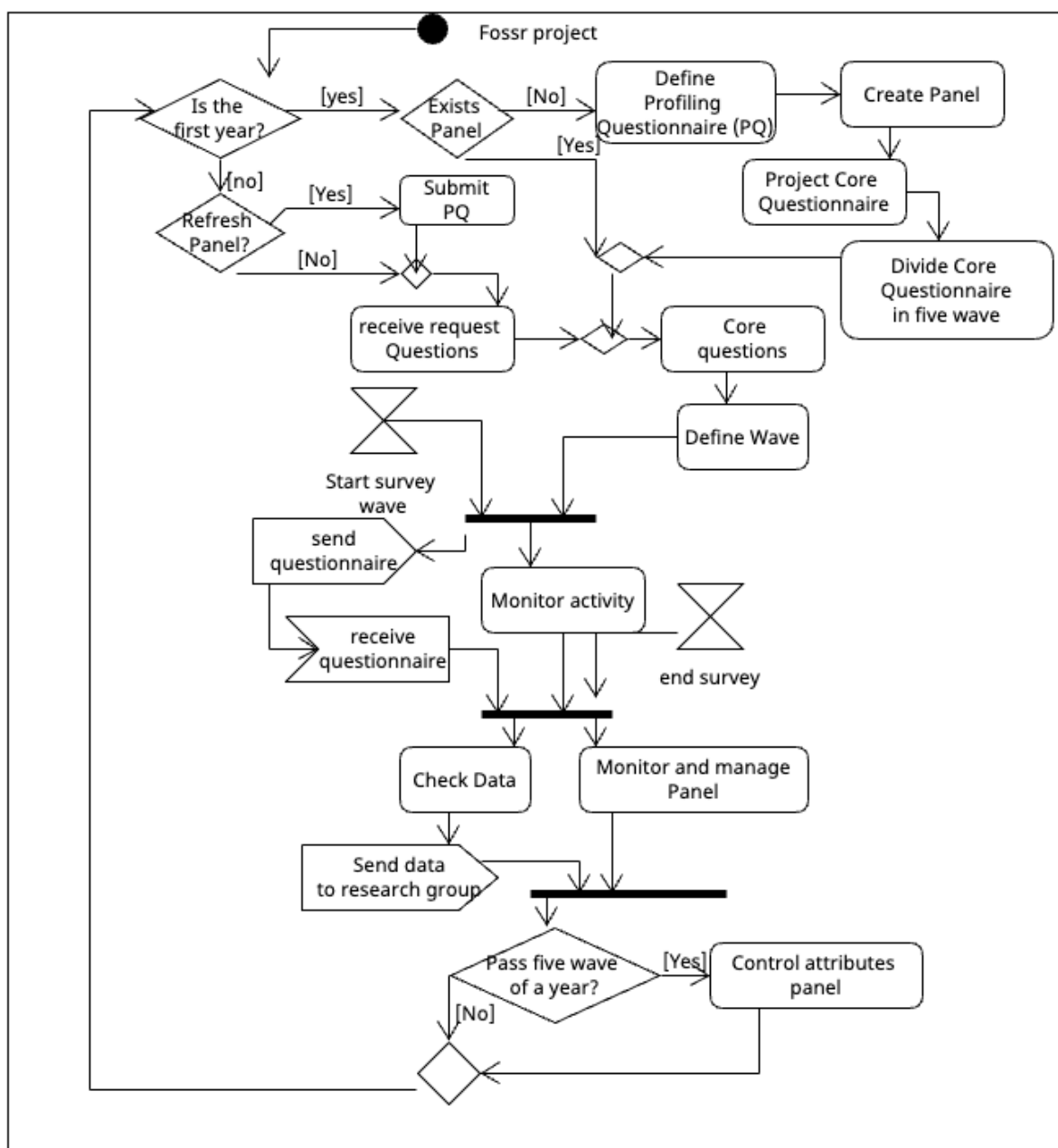


Figure 1. Activity diagram of the whole process

### *2.1.1 PROCESS TO CREATE PANEL*

The activity diagram in Figure 2 represents the sequence of activities to create the panel. Three main actors are involved in this activity: FOSSR, Istat, and a survey agency. The FOSSR group is initially tasked with deciding the characteristics related to the sample of 30,000 individuals to be requested from Istat. Istat will handle the selection of these individuals by providing FOSSR with a list of people complete with their name, surname, and residential address. At the same time, FOSSR defines the Profiling Questionnaire (PQ) and the Data Processing Policy. Upon receiving the data from Istat, FOSSR will import them into LimeSurvey and create personal tokens for the first access to the survey system. This will be followed by sending a letter containing the personal token and instructions for accessing the initial PQ. Upon first access, each member must sign the data processing consent. It is important at this stage to have all consents personally signed. Any member who refuses to consent to data processing will be removed from the system. The survey agency will contact individuals who do not have first access to the system for a personal interview at the possible panellist's residence. The panellists' contact, control, and profile insertion operations are entrusted to the survey agency. The data collected from the PQ will then be imported into LimeSurvey in the panellists' registry. This will enable the selection of subgroups of panellists to whom various investigations planned by different research groups can be submitted.

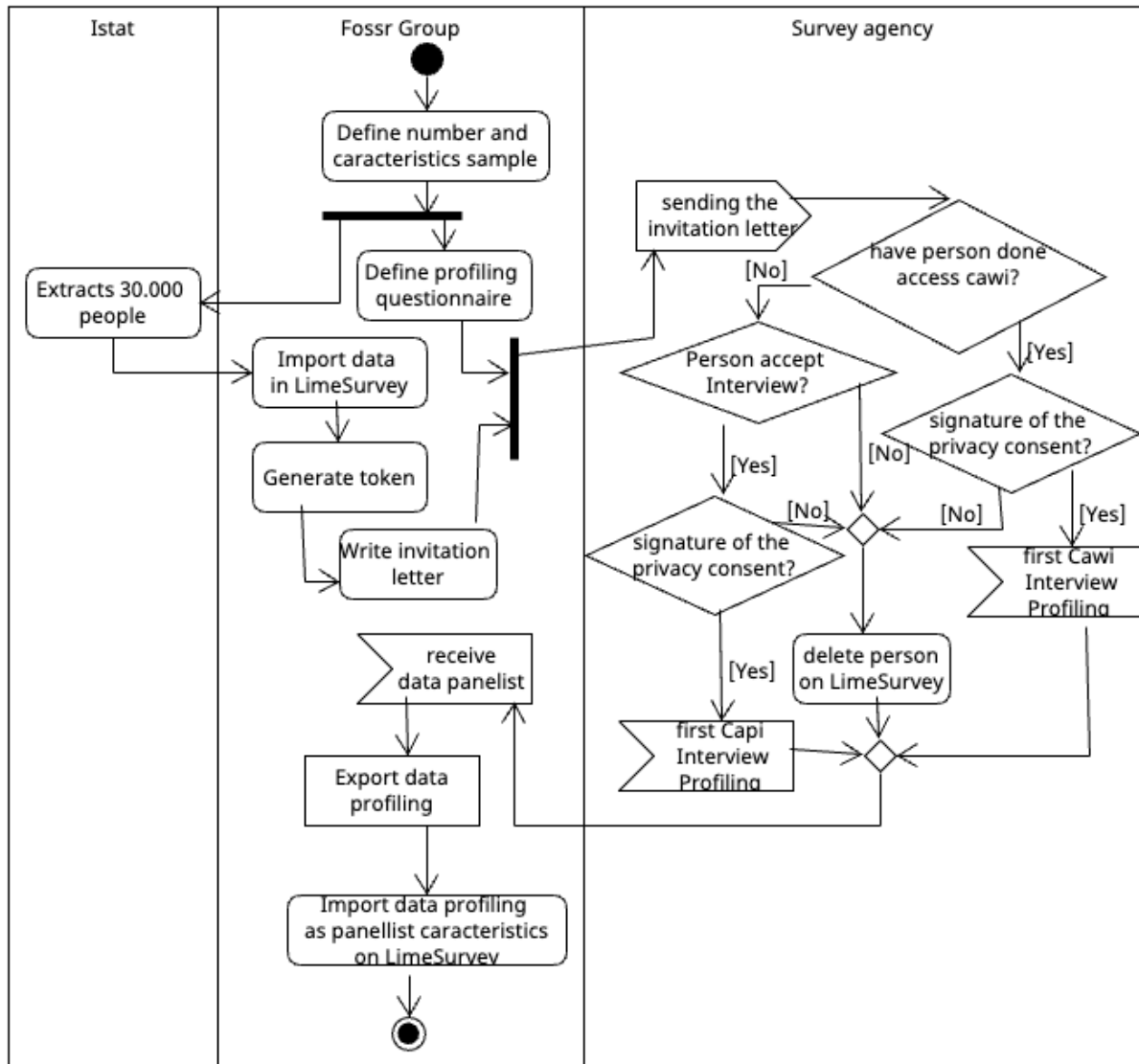


Figure 2. Activity diagram Create Panel

### 2.1.2 PROCESS WAVE

The activity diagram in Figure 3 represents the tasks that must be carried out for each survey wave. The different research groups will submit their investigation requests to the dedicated FOSSR group, which will decide whether to accept and then schedule the investigation in a specific wave or reject the request. Then, all privacy information will be prepared for each research group, and their respective questions will be imported. Finally, categories of selected panellists from different specific modules in the wave will be chosen. This way, each wave will present different questionnaires with their respective investigation questions. Once the survey questions have been entered into the system, panellist groups have been created, and the investigation start date has been reached, the system will send the questionnaires to all panellists. Before starting the investigation, the user's profile will be presented to identify if any selection characteristics have changed and, therefore be able to submit the correct questionnaire still. (For example, if a specific set of questions is only sent to single people, but in the group

of people identified in the panel, some have become engaged/married, and their data has changed). If a panellist does not respond to the questionnaire, the investigation end-date has not yet been reached (time response wave < M), or the maximum alert number to be sent has not been reached (reminder sent < N), the panellist will receive a reminder to complete the survey for that wave.

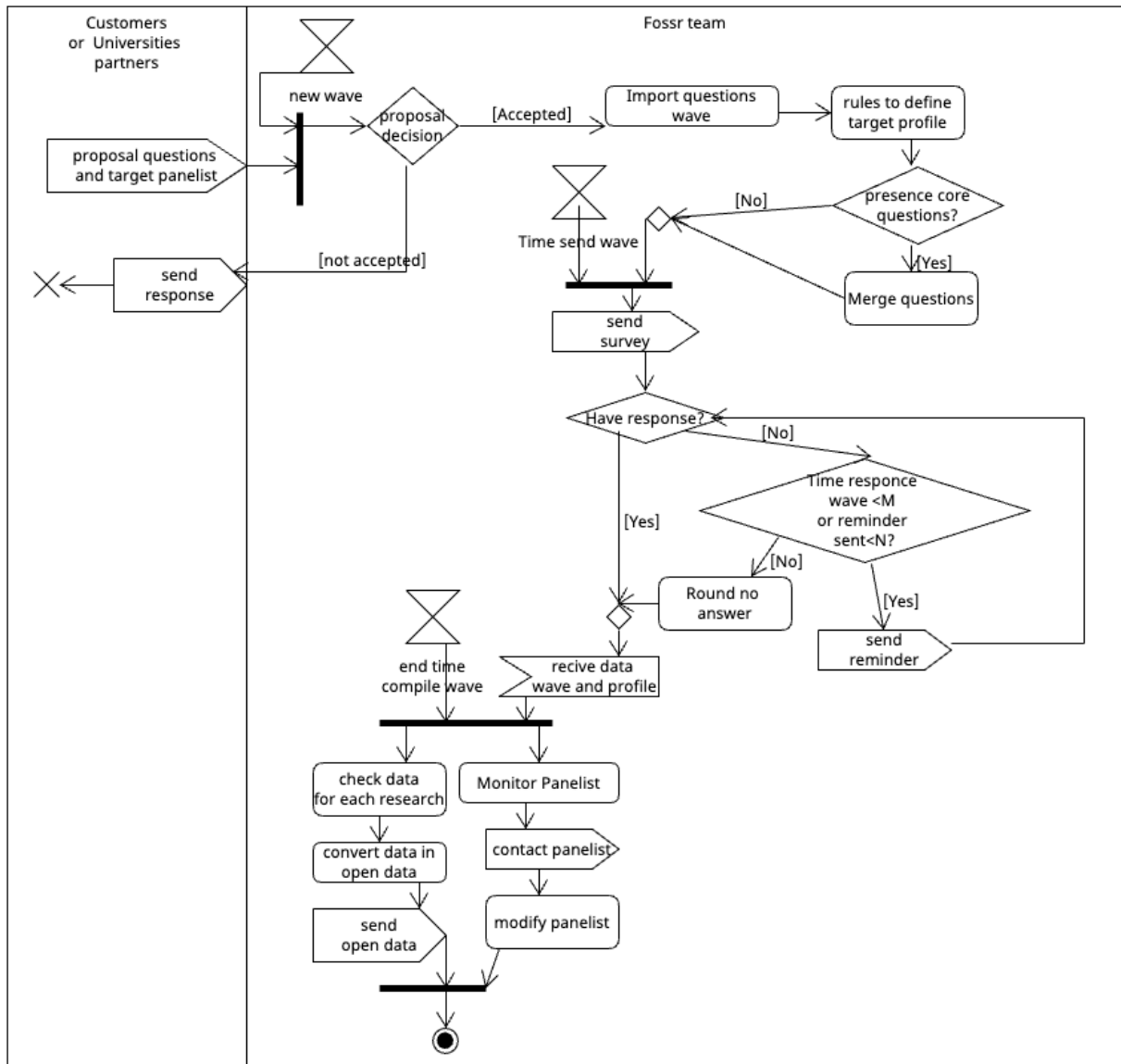


Figure 3. Activity diagram wave

### 2.3 Non-functional requirements: an overview of security

Among the most relevant non-functional requirements, the following requests related to system security have been highlighted:



- Maintenance of data in an anonymous form
- Ensuring the identity of the panellist connecting to the system for both the privacy consent acceptance operation and questionnaire response.
- The use of secure protocols for message exchange over the network

Request for certificates to authenticate the servers where the service will reside.

The need to keep the service active and retrieve data, considering the possibility of physical machine failures or an overload of requests.

The resolution proposal is discussed in Chapter 5.

### 3. SELECTION OF THE WEB PROGRAMMING TOOL

In our exploration of tools for managing online panels and surveys (see deliverables 3.2 and 4.1), we meticulously assessed several options, namely LimeSurvey, Blaise, JD eSurvey, Quick Survey, Tell Form, EU Survey, and DataCTRL by Centerdata. In our evaluation of survey software options, a comprehensive set of features was considered to ensure that the chosen tool aligns seamlessly with our requirements. The following key characteristics were scrutinised to gauge the software's suitability for managing online panels and executing surveys:

Questions:

- **Type and Characteristics:** The capability to select various question types (open, closed, multiple-choice, table format, etc...), designate questions as mandatory, and include instructions.
- **Importing and Exporting Questions:** The ability to import questions from external sources (repositories, other questionnaires) and export questions to create a searchable repository.
- **Code List Management:** The capacity to import code lists from different sources (CSV, repository, VC) and reuse them across multiple questions.

Structure:

- **Multiple Questionnaires:** Effective management of multiple questionnaires associated with the same study.
- **Sections, Subsections, and Loops:** The ability to organise questions into sections and subsections for efficient grouping and incorporate loops to repeat sets of questions (e.g., events).

Routing and Validation:

- **Filter Definition:** The possibility of defining filters for questions or groups of questions.
- **Exporting Flow:** Capability to export the questionnaire flow in structured formats (e.g., JSON) and searchable formats (e.g., PDF).
- **Validation Systems:** Presence of filter validation systems, flow tests, and validation of questions and response methods.



#### Operationalisation:

- **Concept Relationship:** Ability to define relationships between concepts, operational definitions, and variables.
- **Visualization:** Displaying possible operational definitions for a concept and visualising relationships graphically.

#### Administration:

- **Method Differentiation:** Ability to differentiate questionnaire elements based on administration methods (CAPI, CATI, CAWI).
- **Multilingual Support:** Capability to administer the same questionnaire in multiple languages, including translation.
- **Codecard Generation:** Generating code cards for certain administration methods (CAPI) if required.

#### Import and Export:

- **Questionnaire Import:** Possibility of importing questionnaires from other software, standards (e.g., DDI), and non-standardized formats (DOCX, PDF).
- **Questionnaire Export:** Ability to export the questionnaire to other software and standards (DDI), with flexibility in export formats based on the administration method.

#### Documentation:

- **Metadata Addition:** Inclusion of descriptive metadata for study and questionnaire documentation.

We opted for LimeSurvey as it comprehensively covers most of the critical features outlined. Its open-source license allows for customisation to meet our specific requirements and the flexibility to extend functionalities beyond its existing capabilities, ensuring a tailored solution for our online panel management and survey execution needs.

In addition, LimeSurvey is grounded in its capacity to offer full control through server-based installation, ensuring enhanced data management and security measures. This aligns seamlessly with our commitment to flexibility and customisation, allowing us to tailor the tool precisely to our needs.

LimeSurvey emerged as the ideal solution, encapsulating the key attributes required for managing online panels and executing surveys while providing the flexibility and control necessary for our unique application environment.

The decision to use LimeSurvey as the application supporting panel and survey management will enable us, thanks to its open license, to enhance the sharing of the extended tool during the dissemination and sharing phase of the expected results in FOSSR.

## 4. CUSTOMISED FEATURES

Though LimeSurvey was the best solution, we identified some customised features to be added to the software. Such additional features are needed to integrate multimodal interviews within the application environment. The additional customised features were the content of the procurement call:

1. *Management of the online panel*: automatic acquisition of attributes of the participants collected via a profiling questionnaire; attributes need to be editable (created, modified, deleted).
2. *Maintenance of the online panel*: information on response rates of participants; remainder alert for participants via social messaging (WhatsApp, Telegram or Messenger) or App developed as part of the procurement procedure; diverse formats of data for uploading or downloading must be guaranteed, with the option to integrate previously stored data; it should be secured the update of participants' information via administered questionnaires (or self-administered questionnaires), not affecting the previously stored data, so to generate a historical record for each participant.
3. *Subgroups management*: a tool is required to identify the participants' membership to subgroups, to facilitate group-targeted interviews.
4. *Attributes and surveys alignment*: the customisation of questionnaire items at the procurement call should be unaffected by future interventions.
5. *Authorization*: the development of new extensions should be integrated with the current LimeSurvey authorisation system and not interfere with it.
6. *Panel monitoring*: the panel must be integrated with information regarding the activities of each participant (i.e. number of interviews, remainders, earned incentives) without affecting the system's functioning.
7. *Development of an interface for management and communication with participants*: a new interface needs to be developed - both in the form of a separate application and as additional features of LimeSurvey-that will enable the participants to trace records of their activities in the survey, acquired incentives, and a messaging service between survey administrators and participants, which can be used for remainders alert.
8. *Interviews monitoring*: possibility to produce a report with the response rate of participants sorted by attributes.
9. *Upload and download*: response to surveys collected should be downloaded or uploaded from and to repositories administered by systems of survey management or data management (e.g. Dataverse<sup>2</sup>) external to LimeSurvey and available in popular formats such as CSV, TSV, JSON, XML.
10. *Documentation*: a System Development Kit (SDK) should be made available for extensions and other customisations; the software should be accessible via REST API via HTTP(s) requests described via the OpenAPI<sup>3</sup> specifications. Metadata documentation must comply with the Italian DCAT Application Profile v 1.1<sup>4</sup> and the

---

<sup>2</sup> <https://dataverse.org/>

<sup>3</sup> <https://www.openapis.org/>

<sup>4</sup> <https://www.dati.gov.it/content/dcat-ap-it-v10-profilo-italiano-dcat-ap-0>

second or higher version of Data Documentation Initiative<sup>5</sup> (DDI); metadata documentation must also comply with CESSDA ontologies and vocabularies<sup>6</sup>

11. *General requirements*: the software must be developed with component-based architecture, satisfying the requirements of availability, scalability, security, reliability, and usability. The software must be responsive, accessible, adaptable to different types of devices and blind-friendly. The software should be multi-lingual, including English, French, German and Slovenian, with the default language being Italian. A customisable yet modifiable template must be developed.

## 5. REALISATION PHASE

### 5.1 The ideal provider

The ideal provider should have a proven knowledge of the LimeSurvey software. Being a Featured Partner of LimeSurvey is an additional asset. In their offer, the provider should suggest how they will address each customised feature requested in the previous chapter. Additional criteria of selection are suggestions to facilitate a smart implementation of the software from features already available in the current version of LimeSurvey. We value strategies that increase the interoperability between devices and the CNR servers, foster the engagement of participants, and facilitate incentive distribution within the fulfilment of the customised features required. Particular attention must be devoted to the system's security and data protection from hacking, e.g., with proven penetration tests. In the realisation phase of the project, high collaboration with the provider is expected to accommodate possible requests, e.g. implementing a testing environment and a production environment. It is desirable to guarantee high personalisation of the service, for instance, integrating the image of CNR within messages or the App interface. The provider should enable the client to have independent access to the functioning of the software to run modifications autonomously when necessary. A clear technical assistance plan should be proposed throughout the collaboration. The ideal provider would stand for proven similar collaborations with other companies, additional ancillary services to the task (e.g. tailored information acquisition tools), and a variety of professionalities in their team for the whole project.

### 5.2 System Security

The LimeSurvey system will be implemented within the array of services of FOSSR in collaboration with FOSSR WP6 and WP7. It will be hosted in containers for virtualisation environments. Several security policies and tools will be implemented for the protection of the data of participants. Services and security measures suggested by the provider should consider them:

- *Control of user access via identification and access management*. An option is to implement a *Single Sign On* system connecting to the Identity Providers of universities, large research centers, or national systems (e.g. Public System for Digital Identity, SPID).

---

<sup>5</sup> <https://ddialliance.org/>

<sup>6</sup> <https://doi.org/10.5281/zenodo.7528240>

- *Encryption for securing data in transit and data stored.* Secure communication protocols such as HTTP(s) will be implemented.
- *Monitoring, Intrusion Detection, and Audit to secure against intrusion activities.* Activities will include regular auditing activities and system logs, intrusion detection systems (IDS) and intrusion prevention systems (IPS) to detect suspicious activities and plan interventions. Firewalls will be implemented to control traffic in and out of the cloud system. A multi-level approach will be implemented using perimeter firewalls.
- *Emergency planning and recovery.* Procedures will be implemented to deal with emergencies such as data loss, security breaches, or service interruptions. Activities will include regular backups, data and service redundancies with the two first-level nodes hosting replicas of data and services of the second-level nodes they clone. This guarantees the availability of services in case of failures in the system. Further activities to this goal include regular updates and patches and the prevention of malware, phishing, and other cyber threats e.g. via antivirus and anti-malware.
- *Monitoring of the physical servers.* Physical servers with access control, video surveillance, and intrusion detection systems.

## REFERENCES

Zuppoli, S., et al (2023). *Deliverable 4.1A, Definition of the strategy for building the digital platform using Open Cloud*, Fostering Open Science in Social Science Research - FOSSR project.

Cerbara, L., et al (2023). *Deliverable 3.3, Preparation for and implementation of multi-mode interviewing*, Fostering Open Science in Social Science Research - FOSSR project.