Research and Innovation Action

Social Sciences & Humanities Open Cloud

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

Deliverable D4.1 A sample management system for cross-national web survey

Dissemination Level	PU
Due Date of Deliverable	31/03/19 (M3)
Actual Submission Date	01/11/19
Work Package	WP 4 - Innovations in data production
Task	T4.1
Type	Software development specifications
Approval Status	
Version	V2.6
Number of Pages	p.1 – p.25

Abstract:

The document describes the technical specification of a sample management system for cross-national purposes, outlining the main functionalities of the system and describing the user roles related to access rights to panellists' contact information.

History

Version	Date	Reason	Revised by
1.0	2019/04/23	Initial version	GB - CJ
1.1	2019/05/06	Rewriting initial version	QA - GM
1.2	2019/05/09	Comments on previous version	CJ – GB
1.3	2019/05/16	Adding other comments	RF - GB - CJ
1.4	2019/05/20	Addressing the comments	QA - GM
2.0	2019/05/31	New version based on decisions from 05/22 City meeting	QA, GB, CJ, RF, GM, NS, ØS
2.1	2019/06/04	Integrating previous comments from CJ and GB in the new document	QA - GM
2.2	2019/06/06	Adding other comments	CJ - GM
2.3	2019/06/19	Addressing comments and amendments	QA - GM
2.4	2019/06/26	Adding other comments	CJ – GB
2.5	2019/07/08	Addressing comments	QA
2.6	2019/10/14	Adapting document to SSHOC template	GM - GB

Author List

Organisation	Name	Contact Information
ESS ERIC (City, University of London)	Rory Fitzgerald	r.fitzgerald@city.ac.uk
ESS ERIC (City, University of London)	Gianmaria Bottoni	gianmaria.bottoni@city.ac.uk
ESS ERIC (City, University of London)	Curtis Jessop	curtis.jessop@city.ac.uk
NSD	Øyvind Straume	Oyvind.Straume@nsd.no
Sciences Po	Quentin Agren	Quentin.Agren@siencespo.fr
Sciences Po	Geneviève Michaud	Genevieve.Michaud@sciencespo.fr
Sciences Po	Nicolas Sauger	Nicolas.Sauger@sciencespo.fr

Executive Summary

The European Social Survey ERIC (ESS) experimented with the world's first cross-national probability-based online panel, named CRONOS, recruited off the back of ESS Round 8. One of the biggest challenges of the project was the coordination and harmonised management of the sample in the three different European countries taking part in the project (Estonia, Slovenia and Great Britain. In particular, the need to restrict and enable access to panellists' information for subsets of the staff whilst also facilitating the correct level of access at the central level was identified as a challenge.

Currently, there is no "off-the-shelf" commercial software that is flexible enough to manage the complexity of a cross-national sample. For this reason, ESS ERIC in collaboration with Sciences Po decided to develop a sample management system (abbreviation SMS) capable of dealing with the complexity highlighted above as part of the SSHOC project.

This document describes the technical specification of a sample management system for cross-national purposes, outlining the main functionalities of the system and describing the user roles related to access rights to panellists' contact information.

Abbreviations and Acronyms

CRONOS	CROss-National Online Survey
DDI	Data Documentation Initiative
DPO	Data protection officer
ESS	European Social Survey
EVS	European Values Study
GGP	Generations & Gender Programme
QDDT	Questionnaire Design and Documentation Tool
QVDB	Question Variable Database
SHARE	Survey of Health, Ageing and Retirement in Europe
SMS	Sample management system
SSHOC	Social Sciences & Humanities Open Cloud
TMT	Translation management tool

Table of Contents

Introduction

Large-scale cross-national probability-based on-line surveys require a modular infrastructure to effectively manage a cross-national sample with different characteristics and requirements. The European Social Survey (ESS) experimented with the world's first cross-national probability-based online panel, named CRONOS, recruited off the back of ESS Round 8. One of the biggest challenges of the project was the coordination and harmonised management of the sample in the three different European countries taking part in the project (Estonia, Slovenia and Great Britain). Major challenges included the harmonised management of incentives for respondents, the central management of different survey contacts (postal mails, emails and text messages) to panellists, the harmonisation of the surveys' translation from English into three different languages (Slovenian, Estonian and Russian) and issues related to the access to panellists information at country and central level. In particular, restricting access right and preventing subsets of the staff being able to access panellists' personal information proved challenging. There was no "off-the-shelf" commercial software that was flexible enough to manage the complexity of a cross-national sample. For this reason, ESS ERIC in collaboration with Sciences Po decided to develop a sample management system (abbreviation SMS) capable of dealing with the complexity highlighted above. This work is taking place in the context of the SSHOC H2020 award which aims to ensure that data produced will be made available according to the FAIR principles. This work is therefore driven by the need to facilitate a cross-national web panel and in turn produce FAIR data.

Background

A cross-national online panel is likely to be composed of four essential entities[1] (obviously plus the panellists) that communicate with each other. The entities include the central team, the participating national teams, the web-survey platform which administers the survey questions and a system to manage and store panellists' data. All the data flow between these subjects goes through several tools. Assuring an efficient and immediate communication between the entities is crucial to the success of a panel. In addition, such a tool should have the flexibility to react to contextual changes. For example, adding additional countries during the life cycle of the study should be possible without having to rebuild the tool.

Communication, automation and modularity are essential characteristics in order to deal with the complexity of cross-national probability-based surveys.

Another key element aimed for in this project is the input harmonisation of the survey features across all countries in order to maximise cross-country equivalence[2]. For a web panel this means having the possibility to rely on a central system to manage tasks related to panel administration, contact distribution (invitation, reminder and pre-notification) and fieldwork monitoring. At the same time, panel members need online accounts enabling them to update contact information when required and minimising manual updates being

---

[1] Obviously, these entities can be more or fewer, depending on how integrated they are

[2] Harmonization refers to the efforts to standardise inputs and outputs in multinational, multicultural, or multiregional surveys. The input-harmonization processes aims at achieving the comparability of the surveys and measures collected.

made following an email or phone call.

Currently commercial providers do not meet these requirements. Building an in-house infrastructure although initially costly will pay off in the longer term in terms of efficiency, efficacy and cost saving as there is no need for external providers. Whilst the tool is designed initially for the ESS it will be developed to maximise the chances that other large-scale cross-national survey infrastructures will want to utilise it in the longer term.

Scope

The biggest European social science infrastructures, ESS ERIC, EVS, SHARE ERIC and GGP are progressively experimenting with web interviewing. The web is seen as a mode to collect (additional) data in a more efficient way in terms of costs and time. The world's first cross-national probability-based input-harmonised web panel (CRONOS) was trialled by ESS ERIC in 2017 under the H2020 SUSTAIN grant (676166). On the basis of this first experiment it was clear that there was no commercial software available that met the needs of a cross-national online panel in terms of sample management. In fact, ad-hoc additional programming and manual processes were required to conduct CRONOS, adding additional costs and frustrations. The major issue was related to the need to restrict and enable access to panellists' information for subsets of the staff whilst also facilitating the correct level of access at the central level. For example, each country team should have access only to the contact information of their own panellists, and teams who do not need access should not be able to see contact information. ESS ERIC is therefore working with Sciences Po to adapt the existing sample management system (SMS) for their national web panel to meet the needs of cross-national projects under the SSHOC project. To develop the SMS, ESS and Sciences Po has consulted with other major infrastructures like EVS, SHARE and GGP in order to identify their specific needs and try and make the tool versatile enough to be used by all.

The project aims to design a web application that, paired with a survey platform, will enable the management of a cross-national sample for academic studies. In this context, an academic study (e.g. CRONOS) consists of a set of national samples to which surveys can be published centrally.

1. Technical Infrastructure

1.1 Survey Platform

The term "survey platform" refers to the software application used to design web questionnaires and publish them to a set of panellists that can use that to read the questions and then submit their answers online.

The development of a survey platform is out of the scope of the project, and the choice of a reliable third party service, namely Qualtrics, has been made as it works effectively with the SMS and allows multi language surveys to be administered (see Appendix 3 for more information related to the choice of Qualtrics).

The Qualtrics survey platform also offers the functionality to send emails and text messages to panel members allowing the SMS to be focused on sample management.

Since survey lifecycle and sample management tasks are intertwined, an efficient sample management system (SMS) needs to interact closely with the survey platform. The necessary information exchange between the two services is carried out through the survey platform's web API. The SMS will be designed to cooperate with Qualtrics, and due to many specificities will not be able to operate with another survey platform. However not all the SMS functionalities are lost if another survey platform is selected. The SMS can still be used with other survey platforms, but some functionalities will not be available (see Appendix 1 for more information regarding the functionalities retained when the SMS is used with other survey platforms).

Different studies looking to use this SMS infrastructure may wish to use a different survey platform. To do so and to have it interact as closely with the SMS as outlined in this document would require a substantial re-engineering of the SMS infrastructure. Features such as the ability to import and export files should allow for data to be transferred to other survey software, but the integration will not be as close.

The operation of a cross-national web panel will require user interactions both on the survey platform and the sample management system (SMS), following a user role distribution outlined in this document.

1.2 Deployment environments

Two distinct environments (viz. running installations of the SMS) will be available: a pre-production environment to conduct tests, and a production environment in use for actual operations. The production environment will host the latest stable version of the application, and any added feature will first be staged in pre-production. Once development is complete, both environments will be functionally identical to allow for robust testing of new communications/surveys for that wave before live fieldwork. The testing environment in the SMS will correspond to a sandboxed area in Qualtrics.

Information systems administrators are responsible for managing the technical infrastructure on which the SMS runs, with reliability and data integrity as main objectives. The members of the software development team have the capacity to alter the behaviour of the application in both environments. They have both unlimited access to data in the pre-production and production environments but can only access production data when motivated by technical interventions, and access can be limited to only those who need it.

1.3 Technical components

This section briefly states the technologies chosen by the development team at Sciences Po to implement and distribute the SMS. Development will rely on Django and Vue.js as core frameworks, and the source code be versioned on a self-hosted Gitlab repository. The application will be packaged as a set of Docker containers to facilitate portability and re-use, in accordance with SSHOC general directives.

1.4 Source code distribution

The project is intended to be publicly released under an open-source license (yet to be determined), as soon as an adequate maturity level of code and documentation is reached. It will be made available from the SSHOC marketplace.

2. Security and GDPR

2.1 Infrastructure and communications security

All internet communications will be encrypted using the HTTPS protocol.

Application code will be versioned, and application data (panellists' personal contact data as well as the application's operational data) will be backed up at least on a daily basis, with a retention duration of 3 months.

Usual measures (access control, secure communications and storage) will be adopted to ensure the security of digital infrastructures. Strict access control will be enforced according to user roles (cf. 3. Users and roles).

2.2 GDPR compliance

All data collected by the SMS and Qualtrics will be kept within the EU, and privacy by design will guide development. As an example, only data strictly required for panel management operations and survey routing will be shared with the survey platform from the SMS.

Panellists' requests to access their data and/or to be forgotten will be addressed by their national coordinator, whose contact details will be available in the panellist portal (cf. 4.1.3 panellist) alongside procedure guidelines. Requests to **withdraw from the study** should be actioned by national coordinators setting a flag to ensure the panellist is no longer contacted. As for access and deletion requests, since they are likely to be rare and with potential side effects in the latter case, they will be treated on a case by case basis with help of the technical team.

The Qualtrics survey platform allows its customers to be GDPR compliant.

The ESS ERIC DPO has been providing advice regarding the development of this infrastructure. Each project using the system in the future must have a named DPO.

2.3 Documentation of data processing

The storage location and lifetime of all personal data processed under the operation of the sample management system should be documented.

All data stored in the scope of the infrastructure described in this document (including backups) should be securely deleted at the end of the project upon confirmation from the project lead.

3. Users and roles

For the operation of the online panel, user accounts are needed both on the survey platform and on the SMS. Overall, there are 6 different categories of users, each corresponding to a specific set of access rights to data and functionalities. Users will have individual credentials (identifier and password) required to access the application(s) on which they intervene.

3.1 SMS roles

The following 3 roles concern user accounts on the SMS. Each user account will hold exactly one of these roles. Which categories of user will have access to which variables will be defined in the specification for the initial sample file load.

3.1.1. Headquarters Study Administrators (HQ)

Study administrators are part of the central study team (often referred to as "headquarters"), and are responsible for a range of survey and sample management tasks as outlined in 4.1.1 Headquarters tasks.

Study administrators may only access and update all panel members' sample information except their contact details, across all national samples.

3.1.2 National Coordinators (NC)

National coordinators are assigned to a national sample, in the scope of which they are responsible for sample management and content translation tasks as outlined in 4.1.2 National Coordinator tasks 4.1.2 National Coordinator tasks.

Most importantly, national coordinators can access and update all of panel members' sample data (including contact details) within their sample.

3.1.3 Panellists

A panellist belongs to a national sample, part of an international survey. A language attribute will be known for each panellist (initially specified in the uploaded sample data, (cf. 4.2.1 Batch uploading sample data), that will determine the language of surveys and contacts being issued. Each panellist will use email or phone number and password as login credentials to access a dedicated portal (cf. 4.1.3 panellist). The email address/phone number and password used to log in may be changed. The two main tasks carried out through this portal are:

- to access available surveys
- to view/edit personal information

Parts of the panellist portal conveying contextual information and guidelines may be accessible without authentication.

3.2 Qualtrics roles

The following roles concern user accounts on the survey platform. A user account on Qualtrics may perform several roles. The list of Qualtrics user accounts with associated roles will be specified by the study management team and created upon request by the development team.

A member of the study management team may need to have both a SMS account and a Qualtrics account. None of the roles will be able to see the sample information pushed across from the SMS to Qualtrics in the Qualtrics software by default. However, access can be manually given to individual Qualtrics users on ad-hoc basis.

3.2.1 Survey Developers

Survey developers have access rights to Qualtrics limited to survey development activities: creating surveys, adding questions to surveys (worded in the source language), defining the survey logic (filters, branches). The Qualtrics survey developer role could be cumulated with either NC or HQ SMS roles.

3.2.2 Data Managers

Data Managers have access limited to survey answer data collected through Qualtrics. For example, a data manager account on Qualtrics could be granted to someone working at the study's data archive.

3.2.3 Content editors

Survey related written content (e.g. contact messages, survey question translations) is drafted on Qualtrics, thereby leveraging the existing edition and translation interface. This category of tasks is performed by content editors. As an example, a NC account in the SMS may be cumulated with a Content editor account on Qualtrics used for survey translation.

4. Main functionalities

The functionalities mentioned below are to be understood as performed by the pairing of the sample management system and survey platform. When the survey platform may be leveraged, the choice to delegate a given functionality to either of those services should be examined closely and may be subject to change. However, the general rationale guiding this repartition may be stated as: the source of truth for survey-related content is the survey platform; the source of truth for sample management data is the SMS; all operations are triggered from the SMS (Example: panellist data is selectively pushed by the SMS to Qualtrics; contact messages are drafted on Qualtrics, and displayed in the SMS with translation progress hints; contact sending is triggered from the SMS and executed by Qualtrics).

4.1 User tasks

Once authenticated, SMS users access a set of functionalities through a web interface tailored to their role. Qualtrics users (assuming they don't cumulate a sample management system user account) may only access the survey platform, on which they can perform tasks allowed by their privileges.

HQ, NC and panellist user interfaces are worded in English. The English text strings used in the panellist interface can be made available by the development team. If translations of this content are provided, the panellist user interface can be internationalized.

4.1.1 Headquarters

As part of the initial setup of the application, a set of HQ user accounts will be created by the development team, as requested by the study's central management. From then on, headquarters study administrator is autonomous to take control of the application by performing the following tasks:

◆ creating national coordinators accounts

◆ creating national samples

    ◊ named entities representing a group of respondents (empty of data at this stage)

◆ assigning national coordinators to samples

◆ viewing the list of all surveys available on Qualtrics

    ◊ Including hints on translation progress

◆ publishing surveys to samples

    ◊ publication is a prerogative of HQ users, even for national waves

    ◊ samples are the smallest units to which a survey can be published

        ▪ it is not possible to target a subset of respondents within a sample via the SMS. This would need to be done via Qualtrics.

◊ a survey can be published in one action to a selection of samples

◆ viewing the list of contact messages drafted on Qualtrics

◊ including hints of translation progress

◆ triggering sending of email/text messages to a selection of samples

◊ limit recipients based on survey completion status (e.g. reminders)

▪ it is not possible to target a subset of respondents within a sample when sending is triggered from SMS. This would need to be done via Qualtrics.

◆ importing additional sample variables in pre-defined slots

◊ 50 slots ("n1" through "n50") that can accept numeric values are provisioned

◊ 50 slots ("t1" through "t50") that can accept text values are provisioned

◊ HQ users fill those slots as needed by importing CSV files indexed by panellist ID

◊ Slots may be overwritten through time

◊ The provenance of slot contents at any given time is documented outside the SMS

◆ viewing and exporting a subset of panellist data fields

◊ data fields viewable/exportable by HQ are part of the CSV format specification

◆ following survey completion statistics during fieldwork

◊ ex: for survey X published 3 days ago to 1000 panellists, 50% have started and 25% have finished

4.1.2 National Coordinators

National coordinator (NC) accounts are created by HQ users. NC users may only perform actions and view data within the samples they have been assigned by HQ users. Given those restrictions, they may perform the following tasks:

◆ importing panellist data for their national sample as a CSV file

　　◊ CSV format specification must include:

　　　　▪ list of column names

　　　　▪ for each column specify

　　　　　　• data type

　　　　　　• data validation to be enforced

　　　　　　• if viewable by (in addition to NC)

　　　　　　　　· panellist

　　　　　　　　· HQ

　　　　　　• if editable by panellist

　　　　　　• if shared with Qualtrics (fed-forward data)

　　　　▪ provide CSV sample with fake data

　　　　　　• one valid

　　　　　　• one with typical errors that should be caught by validation

◆ viewing/updating panellist data records

◆ updating panellist data records in bulk

　　◊ same CSV format as specified for import

　　◊ columns and rows may be omitted

　　◊ all cells are set to the provided value

　　　　▪ ex: a cell left blank sets the value to a blank

◆ exporting national sample data as .csv files (for postal mail merging or pre- and post-fieldwork snapshots)

　　◊ same output format as specified for import

　　◊ As a first step, all columns are included (use external tool for narrowing down)

　　◊ At a later stage, allow to select the subset of columns to export

4.1.3 Panellists

Panellist accounts are created when data is first imported to a sample by a NC user. A panellist belongs to exactly one sample. Initial login credentials are sent to panellist by email, text message, or letter, and the first required action is to change the temporarily attributed password. The SMS eventually aims to accommodate password recovery for panel members for whom either email or phone numbers are available (via email or text message). The development process however will first target a SMS that requires email as a contact mode, to subsequently add support for mobile phone alternative. Panel members without email addresses or phone numbers must request this information by contacting their NCs.

Once logged into the portal, the panellist may perform the following tasks:

◆ view and update personal data

◊ which data fields are displayed to and editable by panellists is part of the CSV specification

◆ update preferences, such as preferred mode of contact

◊ the self-editable data fields is part of the CSV specification

◆ access their NC contact details, for requests regarding:

◊ non-editable data fields (date of birth, gender)

◊ opting-out or GDPR-related requests

◆ access direct links to answer available surveys

◆ access contextual information related to the study

◊ static content, that must be provided by the study team at the beginning of the project (ex: FAQ and help sections, legal mentions regarding GDPR...)

4.2 Sample data operations

Sample data management operations allow National Coordinators to keep panellists personal data safe and up to date, Headquarters Study Managers to create, update and monitor sample variables as required to execute fieldwork designs.

4.2.1 Batch upload

National sample data files (containing panel members' data from their recruitment survey) will be initially uploaded by national coordinators. Uploads for different samples can occur independently. Within a sample, the upload can be done incrementally, with several (possibly overlapping, in the case of updates) files.

All uploaded files follow the same structure: the chosen format is CSV, and the column names and types are part of the specification.

4.2.2 Bulk update panellist data

National coordinators and headquarters study administrators will be able to update panellist data records in bulk (as described in and tasks).

The upload functionality should work such that for all cases & variables included in the file, the data in the SMS are over-written by the content in the uploaded file (including being 'blanked'). Cases & variables not included in the uploaded file should not be updated.

An additional upload functionality for more ad-hoc updates, where column 1 indicates the case to be updated, column 2 the variable, and column 3 the new variable will be considered at a later stage should the requirement become clear.

4.2.3 Individually update panellist data

Besides performing bulk updates by importing CSV files, national coordinators will also be able to view and edit panellists' contact details and preferences from the SMS user interface (upon request from a panellist, or at their own initiative to correct errors for example).

4.2.4 Additional variables

A number of slots to accommodate extra numeric and string variables not known before fieldwork will be provided (cf. 4.1.1 Headquarters tasks).

The variables may then be assigned purposes at the discretion of the headquarters study administrators, and allocated values through the bulk updating of panellist of data (4.2.3). The purpose of these variables may (or may not) change between waves, and should be clearly documented, with exports of the data stored for future documentation.

4.2.5 Ongoing computation of values

The information in the SMS will change over time as people's circumstances change. Some sample variables may be derived from those changing variables and SMS should automatically update them. For example,

someone adding a mobile phone number may make them eligible for text message reminders, so any sample flags relating to this would need to change. Or, an age variable may be computed from a panel member's data of birth – as the panel member's birthday is passed the age values should automatically update.

4.2.6 Sample checking

A set of rules for variables included in the SMS will be defined as part of the specification for the initial .csv file load (as outlined in 4.1.2). For pre-defined variables, these can be more complex (e.g. an email address must contain '@', a username, a domain name, a TLD..., or 'date_of_birth' must contain dates as YYYYMMDD, with 1900 <= YYYY <= 2011.

Checks for generic slots can only cover type and range: e.g. 'n1' only contains integers between 1 and 10 and 't20' only contains text strings no longer than 20 characters.

Additional ad-hoc checking will be conducted by exporting and manually reviewing data.

4.3 Contact Management

4.3.1 Batch SMS and emails

It should be possible to compose, translate and send email and text messages to a selection of the sample. Messages will be drafted and translated on Qualtrics by content editors. From the SMS, headquarters will view contact messages available on Qualtrics, follow the translation progress, and trigger the sending of the e-mails.

The SMS keeps records of all contacts sent. Delivery outcomes will be extracted from Qualtrics.

Triggering email contacts via the SMS will be the first implementation target, after which text message support will be added.

As a first step, batch sending will target samples as a whole. Restricting recipients to sample subsets based on survey completion status may be implemented at a later stage.

Mail-merge will be performed by Qualtrics, and may hence include all data fields pushed to the survey platform.

4.3.2 Letters

Sending letters to panellists may only be done externally, by exporting sample postal addresses, and other required mail-merge fields (provided these variables are among the CSV upload or slots) as a .csv file (cf. 4.5 Data and Paradata exports) and using usual mail-merge features available in any word processing software (LibreOffice or Microsoft Word).

Tracking recipients of postal mails must be done manually through processes external to the SMS.

4.4 Survey Management

The link between the sample management application and the Survey Platform Qualtrics is done via the

Qualtrics Application Programming Interface (API). The ability for information to be passed between the two systems is therefore limited to features available in that said API.

4.4.1 Making surveys available to panel members

Once a study has been set up and tested on the survey platform, a survey can be published through the SMS interface to a selection of national samples. Note that survey publication only targets active panel members.

4.4.2 Pushing sample data towards the Survey Platform

The panellist data specification (described in 4.1.2 National Coordinators) defines the subset of personal data to be shared with Qualtrics for survey routing and mail-merge purposes. This selection will not change during the operation of the application and should therefore be broad enough to meet the potential needs of all surveys.

For example, a "gender" variable could be sent from the sample management application to the survey platform to allow for gender-based filters in a questionnaire. Shared data is synced every time an update is made in the SMS.

4.4.3 Pulling data from the Survey Platform to the Sample Management System

Updating panellists' profile by pulling data from survey results must be excluded. This limitation concerns the content of survey answers, not paradata such as survey completion status, which will be available in the SMS.

The importing of any other data (e.g. survey answers) would need to be done by exporting them from Qualtrics and then uploading them into the 'additional variables' slots in the SMS.

4.5 Data and Paradata exports

4.5.1 Sample data and paradata exports

The SMS provides privacy-compliant export formats suitable for matching respondent data and survey answers, based on a shared anonymous identifier.

As a typical use case, a sample data export (i.e. a snapshot) could be made at the exact moment a fieldwork begins, in order to keep track of accurate sample data corresponding to a specified fieldwork.

Appendix 1: The functionalities of the SMS with survey platforms different from Qualtrics

The functionalities outlined above need to be understood as performed by the pairing of the sample management system and the survey platform (i.e. Qualtrics). However, it is obviously possible to use the SMS with other survey platform providers. Not all of the main functionalities will be available using a different web survey platform from Qualtrics. However, some of the main basic functionalities of the SMS can be retained. The table 1 shows a list based on the functionalities described in the previous sections divided by those that are retained and those that are not available when using the SMS with another web survey platform.

Table 1 – Functionalities available with other survey platforms

	Functionalities	
	Retained	*Not available*
USER TASKS		
Headquarters	Creating national coordinators accounts	Viewing the list of all surveys available on Qualtrics • Including hints on translation progress
	Creating national samples	Publishing surveys to samples
	Assigning national coordinators to samples	Viewing the list of contact messages drafted on Qualtrics
	Importing additional sample variables in pre-defined slots	Triggering sending of email/text messages to a selection of samples
	Viewing and exporting a subset of panellist data fields	Following survey completion statistics during fieldwork
National Coordinators	Importing panellist data for their national sample as a CSV file	
	Viewing/updating panellist data records	
	Updating panellist data records in bulk	
	Exporting national sample data as .csv files (for postal mail merging or pre- and post-fieldwork snapshots)	
Panellists	View and update personal data	Access direct links to answer available surveys
	Update preferences, such as preferred mode of contact	
	Access their NC contact details, for requests regarding	

	Access contextual information related to the study	
SAMPLE DATA OPERATIONS	Batch upload	
	Bulk update panellist data	
	Individually update panellist data	
	Additional variables	
	Ongoing computation of values	
	Sample checking	
CONTACT MANAGEMENT	Letters	Batch SMS and emails
SURVEY MANAGEMENT		Making surveys available to panel members
		Pushing sample data towards the Survey Platform

Appendix 2: Workflow examples

This section outlines how some common tasks would be carried-out from start to finish, using the notions introduced above.

A21 Survey routing based on external data

Suppose that a survey needs to be routed depending on the level of education of the respondent, information that is not available in the initial panellist data. First, headquarters use the available data slot 's1' to import a numerical variable encoding level of education (say, integers from 1 to 5). This has the effect of pushing this data to Qualtrics (all generic slots are shared with the survey platform), associating metadata 's1' to each anonymous respondent record. The survey developer (Qualtrics role) is then free to use 's1' to specify survey routing. This assumes that the survey developer is aware of the externally documented convention that 's1' currently holds an encoding of education level.

A22 Drafting and sending a contact email

Headquarters begin by drafting an English version of the contact message, entitled 'Welcome survey!', through the Qualtrics interface accessed with a [Content Editor account](#).

The message contains the placeholder "`${l://SurveyURL}`" that will be replaced by an individual link to answer the survey. Likewise, any personal data shared with Qualtrics could be embedded (such as first name for informal personalized greeting).

In the SMS, the French National Coordinator sees that a new message has been created, that does not yet have a French translation. The NC connects to Qualtrics and translates the message in French.

As national coordinators contribute their translations of the content, HQ can follow the overall progress in the SMS.

Once all languages are available, HQ picks a selection of samples, and triggers sending. This has the effect of making Qualtrics deliver a personalized message to every respondent of the selected samples.

Panellists may access the survey either from the link inserted in their invitation email, or from the (same) link presented in the "Available surveys" section of the panellist portal.

Appendix 3: Advocating the choice of Qualtrics as a survey platform

The purpose of this document is to justify the choice of Qualtrics as the survey platform to be used alongside the sample management system, in the context of the SSHOC project.

The ambition is not to perform a comprehensive benchmark of all existing solutions, but rather to perform due diligence by advancing the key arguments that make Qualtrics a satisfactory candidate. The justification is expressed from different viewpoints: software capabilities, questionnaire design, and FAIR post-production.

A31 Background

Qualtrics is a web-based survey platform that has been in activity since 2002, and ranked 7 in Forbe's "Cloud 100" list of 2018 . It is "well-known" to be the world leading survey platform solution and is adopted by major organisations (some of which, like the MIT, are academic).

Although it primarily advertises corporate use cases, evidence of scientific publications in quantitative statistical analysis citing Qualtrics shows that is being successfully used in an academic context as well.

A32 Software integration

A321 Cloud service

Qualtrics is a cloud-based service, meaning that apart from the license, *no additional human and financial resources are needed to ensure hosting the software* (and handling issues of scaling and data persistence). This contrasts with Blaise for example that needs to be self-hosted on a dedicated Windows server. In the experience of SciencesPo, the latter setup has proven brittle and with high maintenance costs, requiring very specific technical skills.

A322 Comprehensive and well-documented web API

To be efficient, the sample management system must be able to communicate with the survey platform to programmatically perform actions on behalf of users. For instance, a national coordinator must be able, through the SMS, to publish a questionnaire available on the survey platform to a panel of respondents. These types of interactions between the two web applications are only possible if a sufficiently fine-grained web API is exposed by the survey platform.

Qualtrics has a well-documented and extensive web API, that allows access by programmatic means to a large subset of the functionalities a Qualtrics user would access through the regular web interface. We have assessed that the functionality thus exposed to third party software by Qualtrics is sufficient to accommodate the functional needs of the sample management system.

A323 Data privacy

Importantly, the interactions between SMS and Qualtrics do not require sharing sensitive personal data with the platform. This allows survey answers and respondent data to live in two distinct information systems with a neat boundary, providing a high level of privacy protection. Qualtrics exposes its functionalities to the SMS while enabling such a privacy-preserving data boundary.

Besides, Qualtrics enables its customers to beGDPR compliant [GDPR compliant](), and ensures that data centres (as well as backup data centres) are located in the EU which is very beneficial in terms of GDPR.

A324 Responsive customer support

Furthermore, this API is actively maintained, and the customer support has proved to be very responsive in our interactions so far (response within 24 hours), both on giving usage advice and fixing reported bugs. During the CRONOS-1 project the user support from Questback in Norway was not sufficient.

A33 Questionnaire Design

Unlike Blaise which requires mastering a dedicated programming language, Qualtrics provides a user-friendly web interface to design questionnaires.

This does not come at the expense of many features: a wide range of question types are supported and can be flexibly combined (100+ question types, embedded data, branching, display logic, randomization, etc.)

Question-level functionalities open the possibility for more advanced paradata collection and respondent tracking on the SMS side, even though it is beyond its current scope.

A34 Fair post-production

Neither survey questionnaires or answers are locked into the platform: both can be exported in a variety of structured formats for further processing and diffusion.

One could be concerned about the inability to export questionnaires and variable documentation in DDI format. However, after consulting with partners developing the SERISS tool suite (TMT, QVDB, QDDT), it appears that nothing prevents exploiting the structured output exported by Qualtrics to produce DDI documentation. However, this is beyond the scope of the sample management system.

A35 Cost/Benefit analysis

With regard to the aforementioned benefits, both on the level of software integration, questionnaire design and overall reliability, Qualtrics stands out as a cost-effective solution. Expenses are limited to a yearly license of 5k€, which is cheaper than the corresponding Blaise license. Most importantly, no extra resources for software hosting and specific survey programming skills are needed.

A36 Conclusion

Qualtrics is a reliable survey platform, offering the opportunity of efficient software coupling with a sample management system without sacrificing privacy requirements. It is a cost-effective solution, not requiring technical skills to operate. However, it is accepted that it may not be suitable for the most complex surveys and in particular for complex panel studies like SHARE.