

Title: **INS Romania's Experience with CAPI Data Collection for Household Statistical Surveys using Survey Solutions Platform**

Authors: **Ana-Maria Ciuhu, Romanian National Institute of Statistics,**
e-mail: anamaria.ciuhu@insse.ro

Silvia Pisciă, Romanian National Institute of Statistics, e-mail: silvia.pisica@insse.ro

Agenda item: **Future of Interview Modes and Interviewers**

Abstract:

In today's digital age, precise and efficient statistical data collection is essential for providing reliable and valuable information to governments, economies, and society as a whole. In this context, the National Institute of Statistics of Romania (INS Romania) has developed significant expertise in the use of CAPI technology for data collection through the Survey Solutions platform, developed by the World Bank.

This presentation aims to provide an in-depth perspective on how INS Romania has implemented and benefited from CAPI using the Survey Solutions platform. We will explore the advantages and challenges encountered in this transition to digital data collection in household statistical surveys, highlighting how this process has increased the efficiency, quality, and timeliness of the collected data.

Additionally, we will discuss the lessons learned by INS Romania in this process, including training and supervision strategies and data processing. INS Romania's experience can provide a solid foundation for other organizations interested in implementing CAPI technology in their statistical data collection efforts.

1. Introduction

In the evolving landscape of statistical data collection, the shift from traditional paper-based methods to digital tools has been important in enhancing data accuracy and processing speeds. INS Romania, recognizing these benefits, adopted the Survey Solutions platform, which has significantly transformed data collection strategy. This paper explores INS Romania's journey in integrating CAPI technology in household surveys and 2021 Population and Housing Census (PHC), focusing on the technical, operational, and methodological aspects of this transition.

2. Procedural Framework

2.1. Implementation Strategy

The implementation of Survey Solutions platform by INS Romania involved several stages, including adapting PAPI questionnaires to the CAPI method, implementing validation rules in the questionnaires, pilot testing, training personnel, and full-scale deployment. This section details the systematic approach taken by INS Romania, from the initial setup of the CAPI system to the ongoing adjustments and enhancements based on interviewers' feedback. The timeline of using CAPI method and Survey Solutions in Romanian NIS is described below.

CAPI method was first implemented for Information and Communication Technologies (ICT) survey in households, in 2018. The testing of the CAPI data collection method on the 2018 ICT survey in households was conducted from April 16 to April 30, 2018, on a sample of 2064 dwellings, in 160 PSUs out of the 792 constituting the sampling frame. Out of these PSUs involved in testing, 88 were located in urban areas and 72 in rural areas. During the test, 160 tablets were used and a total of 2065

households were surveyed. This stage of implementation involved training sessions with the staff from Statistical Territorial Units (DTS) i.e. the territorial units of INS Romania which carry on the data collection process. The staff received training on technical aspects of Survey Solutions and the use of tablets, filling out the questionnaire on the tablet, and the workflow of the Survey Solutions system.

The second survey switching to CAPI was the Labour Force Survey (LFS) i.e. the quarterly survey in 4th quarter of 2018. For testing on LFS, the territorial profile distribution of the test sample was the same described above for ICT. Upon concluding the testing phase, INS Romania setup an evaluation questionnaire, specifically designed to capture the interviewers' perspectives on the newly implemented data collection methodology. This initiative was not only integral to the process improvement but also served as a proof for the acceptance of the change among the operational staff. The responses were indicative of a strong endorsement, with nearly 70% of the operators expressing a positive reception towards the innovative approach. This substantial proportion underscores a successful transition and an encouraging alignment between the technology's capabilities and the operators' adaptability, marking a significant milestone in INS Romania's deployment to modernize its data collection processes. The CAPI data collection continued into the first quarter of 2019 across the 160 PSUs. Starting April 2019, the share of CAPI interviews increased to 60%, encompassing 496 PSUs. This increase this stage was natural, driven by the demand from the DTSSs. By the fourth quarter of 2020, the number of CAPI PSUs had expanded to 530 (out of 792), with the remaining ones continuing with the traditional PAPI (Paper-Assisted Personal Interviewing) approach. This progression reflects a clear trend towards the adoption of digital methods in data collection, illustrating a shift in the landscape of research methodologies employed by INS Romania.

Each of the two testing phases on the ICT and LFS were followed by feedback surveys addressed to the interviewers, intended to reveal issues encountered in the field using CAPI data collection method. The results (INS Romania 2018; INS Romania 2019) highlight important issues regarding the tablets, Survey Solutions software and questionnaire. The benefits of engaging with end-users for feedback during the early stages of implementation are several:

- Tailored survey design: By involving interviewers early, the design of Survey Solutions can be tailored to meet the specific needs of those who will be using it most frequently. Interviewers can provide insights into the types of questions that are most effective, the layout that maximizes ease of use, and the functionalities that are essential in the field. This direct feedback ensures that the survey tool is practical and user-friendly for its primary users.
- Early detection of technical issues: Interviewers can identify technical issues that may not be visible in the design stage. Early feedback allows developers to address these issues before the tool is rolled out widely, saving time and resources that would otherwise be spent on post-launch troubleshooting and fixes.
- Increased interviewer engagement and satisfaction: When interviewers are involved in the development process, they are more likely to feel valued and invested in the outcome. This engagement can lead to increased satisfaction with the tool and greater enthusiasm for its deployment, which can improve data collection outcomes and reduce interviewer turnover.

Thus, engaging interviewers in the early stages of implementing Survey Solutions thus not only improves the tool itself but also enhances the efficiency of data collection efforts. This approach fosters a collaborative environment where the needs and experiences of those in the field are prioritized, leading to better outcomes in the surveys.

In 2021, along with the Regulation 1700/2019 (European Commission, 2019), the CAPI method became mandatory. Therefore, the LFS questionnaire was redesigned and the survey adopted full CAPI data collection. Consequently, other household surveys were implemented in Survey Solutions: yearly

EU statistics on income and living conditions survey (EU-SILC) starting 2021 round and 2022 Adult Education Survey (AES).

Another round of trainings was conducted in each county (NUTS3 level), summing up 42 trainings during the period of September - October 2022, covering the following topics:

- Implementation of the CAPI for the AES;
- CAPI data collection for household surveys using the Survey Solutions system;
- Ethics of statistical research in households;
- Case studies on LFS, EU-SILC, ICT questionnaires.

Other important statistical product – the PHC Round 2021 was fully implemented in CAPI and CAWI. Its data collection period in 2022 was splitted into: March 14 – May 27 online self-enumeration (CAWI) and May 31 – July 31 (CAPI). Several trainings were offered to the personnel of DTSS. They were designed as trainings for trainers, including information on configuring the tablets that will be used by enumerators for data collection, as well as the duties and activities of the census staff, both from an administrative and a Survey Solutions platform perspective, and involving not just aspects regarding technical issues, but also methodological aspects and their implementation in the questionnaire. The training materials was supplemented by a series of video tutorials for each role in Survey Solutions, designed to instruct both participants at the training and later, those who will serve as census staff in the field.

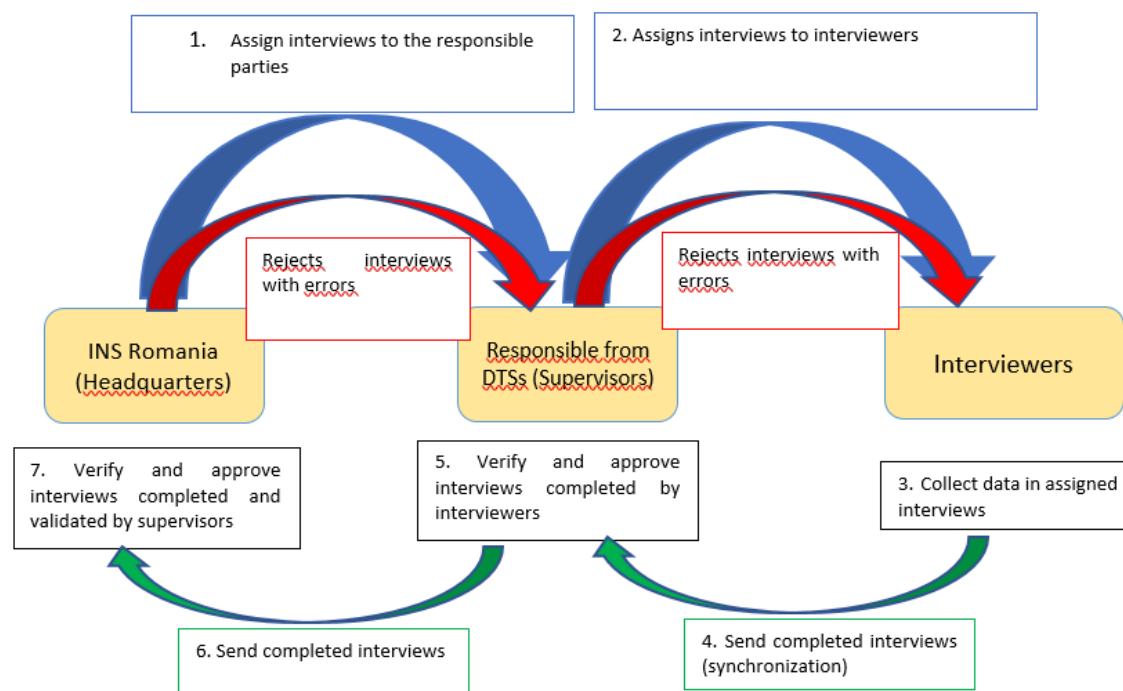
Currently, four out of the six statistical surveys in the social statistics are already implemented and are conducted exclusively via CAPI in all 792 PSUs, including the Adult Education Survey (AES) and the European Health Interview Survey (EHIS); the other two surveys, the Household Budget Survey (HBS) and the Time Use Survey (TUS), are in the first phase of implementation.

2.2. Operational workflow

The workflow of the data collection process used by the INS Romania according to Survey Solutions capabilities and technical characteristics is described below:

1. Assign interviews to the responsible parties: The process starts at INS headquarters, where interviews are allocated to the appropriate county supervisors.
2. Allocate interviews to interviewers: The supervisors then assign these interviews to the individual interviewers who will conduct them.
3. Collect data in assigned interviews: Interviewers go out into the field to collect data by conducting the assigned interviews.
4. Send completed interviews (synchronization): Once the interviews are completed, interviewers send the data back to their supervisors for review. This step is typically done through a synchronization process in a CAPI system.
5. Verify and approve interviews completed by interviewers: County supervisors review the interviews submitted by interviewers. They check for completeness and accuracy.
6. Send completed interviews: After approval, the completed and verified interviews are sent back to the INS headquarters.
7. Verify and approve interviews completed and validated by supervisors: At the INS headquarters, there is a final verification and approval process for the interviews that have been validated by the supervisors.

Figure 1. Workflow of data collection process used by the INS Romania according to Survey Solutions



Source: Authors' conception based on Survey Solutions documentation (World Bank, 2024)

The diagram also indicates a feedback loop where incorrect or incomplete interviews are rejected at various stages:

- Interviews that don't meet the standards are rejected by the county supervisors and sent back to the interviewers for correction or re-conducting.
- Similarly, if the INS headquarters finds issues with the interviews validated by the supervisors, they can also reject and return them for further review or correction.

This process ensures that the data collected is accurate and reliable, as it undergoes multiple layers of checks and approvals before being finalized.

Discussions and lessons learned

INS Romania's experience offers valuable insights into the do's and don'ts of adopting digital data collection methods. The lessons learned cover various aspects such as the importance of training, and the benefits of engaging with end-users for feedback during the early stages of implementation.

In the transition to modern data collection methods, the electronic questionnaire has emerged as a significantly versatile tool, offering a wider array of functionalities over the traditional paper-based questionnaires. Such functionalities include automated jumps between questions and personalized error messages, which allow for a more dynamic and responsive surveying experience. Consequently, a CAPI questionnaire can be entirely reconceptualized in ways that a PAPI questionnaire cannot. During the transition period, the electronic questionnaire served as a bridge solution, balancing two key requirements. Firstly, it was designed to closely reflect the paper questionnaire, in order to facilitate a smoother transition for the operators to the new digital collection mode. This was essential because entirely different questionnaires may measure the same phenomenon differently. Secondly, there was an imperative to harness the full potential of the electronic system's capabilities. For the future, INS Romania proposes a transition to multimodal data collection by gradually introducing CAWI (Computer-Assisted Web Interviewing) as an alternative method also for household surveys.

Opportunities and Challenges of Using CAPI Method and Survey Solutions

Below we are presenting some opportunities and challenges faced by INS Romania in the process of adopting and using CAPI method implemented with Survey Solutions.

Opportunities:

1. Integration of technology in data collection: The use of tablets and the Survey Solutions software represents a step forward in integrating technology into data collection processes, aiming to improve efficiency and accuracy.
2. GPS functionality for precise location tracking: The capability to use GPS for locating addresses during surveys can enhance the quality of spatial data collected, allowing for more accurate and geographically precise data analysis.
3. Digital data submission and storage: The method enables data to be directly submitted and stored digitally, eliminating the need for paper-based data collection and manual data entry, thus potentially decreasing errors and increasing data processing speed.
4. Preloading data: The system allows for the preloading of data that is retained from one wave to another in the case of rotational surveys, and the preloading of data from administrative sources in the case of the PHC.
5. Translation into Romanian: The software's translation into Romanian for the platforms for Headquarters, Supervisors and Interviewers provides opportunities for overcoming language barriers, enhancing usability for non-English speaking operators. The PHC 2021 questionnaire was translated into 17 languages (INS Romania, 2022) for the minorities in Romania and because it addresses to all residents of Romania, overcoming the situation of those who have been in the country for less than 12 months and do not yet have a strong command of the Romanian language.

Challenges:

1. Tablet-related issues:

- Battery discharging, notably due to GPS use, and slow charging times, impacting fieldwork efficiency.
- Instances of tablets freezing during GPS location tracking and repeated resets during interviews, leading to data loss.
- Problems with SIM card connectivity and devices not turning on, requiring replacements and logistical adjustments.

2. Survey Solutions Software Limitations:

- Inability of DTSs to automatically generate reports needed for documentation and payments.
- Supervisors not being able to view submitted forms, affecting the data validation and review process.
- Delayed visibility of rejected and corrected questionnaires, complicating the data correction flow.

3. Questionnaire Design and Management Issues:

- Complexities in date of birth entry, with the format requiring significant time and effort to input.
- Refusal by some households to disclose income information, affecting data completeness.
- Lengthy and numerous questions, alongside technical terms requiring explanations, which may complicate the interview process.

4. Technical and Operational Difficulties:

- Challenges with GPS location tracking, including time-consuming attempts and visibility issues in direct sunlight.
- Weather conditions and lack of internet connectivity mostly in rural areas affecting the use of tablets outdoors and the completion of surveys.

The integration of technology through CAPI and Survey Solutions in data collection processes offers significant opportunities for improving efficiency and data quality. However, the transition also presents challenges related to hardware reliability, software usability, questionnaire design, and operational execution. Addressing these challenges is important for harnessing the full potential of digital data collection methods.

Conclusions

INS Romania's adoption of CAPI technology using the Survey Solutions platform exemplifies the significant advantages that digital data collection tools can offer. While challenges remain, the overall positive impact on the efficiency, accuracy, and timeliness of data collection is undeniable. This case study serves as a robust model for other statistical organizations considering a similar transition.

References

European Commission (2019) Regulation (EU) 2019/1700 of the European Parliament and of the Council of 10 October 2019 establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples, <https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32019R1700>

INS Romania (2018) Testing the CAPI data collection method for the ICT survey in households – 2018, https://insse.ro/cms/files/metoda_CAPI/Raport_evaluare_CAPI_si_anexe.rar

INS Romania (2019) Testing the CAPI method of data collection for the LFS survey - fourth quarter of 2018, https://insse.ro/cms/files/metoda_CAPI/Raport_evaluare_testare_metoda_CAPI_de_colectare_a_datelor_pentru_ancheta_AMIGO_trIV_2018.rar

INS Romania (2022) Questionnaires of PHC round 2021, <https://www.recensamantromania.ro/rpl-2021/instrumentar-2021/>

World Bank (2024) Survey Solutions version 23.09, <https://mysurvey.solutions/en/>