ESSnet Smart Surveys
2020-2021

Intermediate meeting 6-7 October 2020

Work Package 3: Development of a conceptual framework, reference architecture and technical specifications for a European platform for trusted smart surveys

Issues to be discussed

Claudia De Vitiis – ISTAT - Italy

Open issues and challenges for WP3

General issues

- What we have presented on the work done so far meets Eurostat expectations?

- We have doubts on the boundaries of the conceptual framework :

  - For mixed mode survey designs (smart + traditional data collection): Should the framework include the processing of data deriving only from the smart part of the sample and data collection or also from the combined data set (from the complete survey process)?

- We understand that the general goal is harmonizing the smart part of the surveys, being aware that countries may not be harmonized on the traditional part of the survey

Smart survey methodology

- Are trusted smart surveys meant to stream data to NSIs for real time processing and analysis or batch processing with complete survey data?
- Does the automation of trusted smart surveys also envision the automated publication of results? Or in other words where should the automated production of official statistics start and end?

Technical Infrastructure

- Should we document the legacy technical components that still exist in smart surveys, e.g., collection of data through the web interface?
  - We understand that smart surveys should collect most of the data without user interaction – should we write detailed description only on this issues?
- Do you expect to include in technical documentation a detailed description of technical requirements?
  - E.g., platform – iOS, Android; security issues – recommended cryptography, algorithms etc.?
- Are there any preferred data stores we should recommend in our work?
  - For example, the use of Elastic Search and Hadoop components instead of traditional SQL (used now in selected Smart Surveys).

Integration in existing architectural framework

- [**Issue: expected outcome**] Focus on a subset of business functions and model the related process steps, or provide an overview of the whole statistical process?

Preservation of privacy and transparency

- How to manage risk with different interpretation of GDPR in member countries?
  - Example: privacy concerns with mobile apps that use mobile sensors data have increased during the COVID-19 crisis and we can see different interpretation of national privacy commissioners on legality of Corona tracking apps – see Norway case*. How much variability do we need to incorporate in Privacy guidelines?
- Impact of ECJ decision on EU-US shield agreement on the smart data collection. How to mitigate the risk that Smart Surveys might not be allowed to use certain features in Android and iOS?
  - Apple and Google require users to give them certain data (such as phone number, OS version, timestamp, location) to use smart features (gps, camera, compass, accelerator, microphone) in mobile applications which can result in user's data stored out of EU. How much focus does ESSNet project need to invest in mitigation activities?

* https://gdpr.report/news/2020/06/17/norway-drop-contact-tracing-app-after-amnesty-privacy-investigation/

Metadata and process auditability

- Priorities: The field of metadata is vast, we have to focus on specific aspects. Suggested choices are:
  - - data structures
    - sensor data
    - process steps
    - privacy issues
- Practical examples

  - domain models or at least terminologies
  - process descriptions
  - privacy requirements
- Use case : Toy TUS use case: survey on smart phone combining timestamped data collected through a form (type of activity) and mobile positioning data (GPS/Galileo).
  Is it relevant?

Thank you for your attention !