

## Information Notice No.1

### I. DATE, LOCATION, AND OBJECTIVES

1. The 2022 UNECE Expert Meeting on Statistical Data Collection<sup>1</sup> will be hosted by Italian National Institute of Statistics in Rome, Italy from 26 to 28 October 2022, starting at 9:30am on Wednesday 26 October 2022, and ending by early afternoon on Friday 28 October 2022.

2. The meeting will be organized as part of the Conference of European Statisticians' work programme for 2022, within the context of the High-Level Group for the Modernisation of Official Statistics. The objective of this expert meeting is to identify innovative ways and best practices in statistical data collection, and to provide a platform for practitioners to exchange experiences and foster collaboration in this area. The target audience for the expert meeting includes senior and middle-level managers responsible for data collection activities and new data sources, across all statistical domains.

### II. AGENDA

3. In view of the rapidly changing socioeconomic environment, there is an imminent need to develop new techniques and solutions to enhance statistical data collection. On one hand, reliable and timely data are essential to support decision making in the new normal. On the other hand, it is important to improve respondents experience under different data collection modes. The expert meeting is an opportunity to jointly address these challenges, to share experiences, and to discuss how we can leverage cutting-edge ideas, approaches, and tools to enhance data collection practices.

4. The programme of the meeting will consist of the following substantive topics:

(i) **Automated Data Collection: New techniques and IT solutions to improve support and interaction with survey respondents**

The application of intelligent systems to support customer service activities acts as an amplifier of human capabilities to perform frequent and large-scale tasks. These systems are leveraging Artificial Intelligence (AI) combined with machine learning techniques to gain benefits in terms of reduced resources employed and training costs, improved requests handling, greater flexibility, and better control of burden on respondents. The adoption of possible automated data collection solutions has significant implications on the efficiency of the processes. These new techniques and innovative solutions can be applied at different stages of the investigation process. Particularly promising applications concern the issue of support and interaction with survey respondents. This topic will cover sub-topics such as:

- Automatic variable coding
- Outliers detection
- Automation of support and assistance to the survey units
- Standards for system to system data collection
- From questionnaire data collection to variable data collection

---

<sup>1</sup> Until 2020, it was called the Workshop on Statistical Data Collection.

- Legal aspects related to new IT solutions

**(ii) Responsive Design: Leveraging data collected in the early stages of a survey to inform data collection design in the later stages**

Responsive survey design is not a new concept having been proposed by Groves and Heeringa in 2006. The approach suggests that data collected during early phases of a survey can be used to inform decisions made about the design for later phases of data collection. This approach aims to reduce uncertainty in relation to response rates, reduce error, and control costs. Now that more National Statistical Organizations (NSOs) are introducing online first survey designs, and with enabling technology more advances, it is useful to reflect on the approach again. This topic will cover sub-topics such as:

- Adaptive/ responsive design principles
- Auxiliary data to create sample strata
- Design: static; fixed dynamic; variable dynamic
- Approach/ optimization strategies

**(iii) Respondent Care: Improving the experience of respondents from initial contact through to completion of the survey**

NSOs are experiencing a decline in response rates for business and household surveys alike. In order to stop the decline and reverse this trend, it is important to explore techniques that could improve the respondent experience with an end goal of increasing response. Identifying these techniques could come from focus groups with respondents to better understand what drives them to respond, ways of reducing burden, how we communicate with respondents, to name a few. This topic will cover sub-topics such as:

- Gamification
- Incentives or other motivation methods used
- Web data collection Portals and ease of access
- Communication with Respondents
- Follow-up strategies
- Interview Methods used
- Respondent feedback
- Managing and reducing statistical burden
- Respondent Ombudsman
- Response obligation and penalties management strategy
- Respondents' segmentation strategies: monitoring data collection, targeted communication, and recovery actions

**(iv) Managing Multimode Data Collection and Integrating Sources**

Multimode survey designs have been in use for a long time, to use collection procedures that produce the best possible data in the constraints of time and budgets. Multimode surveys combine different modes of data collection, including in-person, telephone, Internet, and mail. The use of multi-mode survey design is increasing, allowing statistical offices to exploit the potential offered by a combination of data collection modes, either to offset the weaknesses of a particular mode with the strengths of another or to try to reduce the overall costs of fieldwork, for example. At the same time, the last decade has seen a tremendous increase in the amount of additional data sources available to survey designers to complement and improve traditional collection methods. Nevertheless, these new sources pose a series of challenges concerning their integration. Many different types of data and different sources can be integrated, hence there is no one-size-fits-all approach for integrating data. In particular, appropriate quality frameworks for integrated data are important to maintain the value proposition of official statistics produced from integrating data sources. This topic will cover sub-topics such as:

- Data fusion: integrating multiple data sources to produce more consistent, accurate, and useful information

- New data sources - big data: such as scanner data; sensor data; social media data; smart phone data; web scraping data
- Administrative data sources: among which tax registers and non-probability samples
- Data donation: collecting data by asking people or companies to donate an export of specific data such as bank transactions or smart meter data
- Data crowdsourcing: citizen statistics
- Administrative data for multiple surveys
- Adaptive survey design: CAWI first, interviews if necessary
- Beyond CAPI: expanding multimode (CATI, CAWI, video mediated interviews etc.)

(v) **Rapid Data Collection Strategies and Real Time Indicators: Filling the information gaps of the rapidly changing environment**

The recent health crisis has significantly increased the need to fill the information gaps relating to the strategies adopted by people and businesses to face the crisis and manage the recovery. The need to quickly fill the information gaps requires the use of innovative information sources and the carrying out of "rapid" statistical surveys. This statistical approach and the strategies adopted concern both survey planning and its actual management. Technical and methodological characteristics of rapid surveys involve actions taken to increase the propensity by units to participate in the survey but also the quality of the final validated data. This topic will cover sub-topics such as:

- Experimenting and evaluating rapid data collection strategies
- Real time indicators
- New statistical sources
- Technologies for rapid data collection

5. In addition to the sessions based on the submitted contributions, the agenda will include target-driven small group discussions on lessons learned and to identify topics for future work. Delegates will be asked to contribute to the development of internationally-coordinated work in the field of statistical data collection.

6. The Organizing Committee of the meeting will make the final decision with regard to the acceptance of proposals for contributions. The Organizing Committee consists of the following members:

- Andrea Ascheri, Eurostat
- Pasquale Papa, Istat, Italy
- Ian O'Sullivan, ONS, United Kingdom
- Susan Oudshoorn and Leonne Hollanders, Statistics Netherlands
- Lise Rivais, Statistics Canada
- Paulo Saraiva dos Santos, Statistics Portugal
- Ahmet Sari, Turkstat, Turkey

### **III. PARTICIPATION AND ACCREDITATION**

7. Representatives of Member States of the United Nations and of interested intergovernmental organizations are welcome at this meeting, as are researchers at universities or research institutes working on statistical data editing applicable in the context of official statistics. Participants representing non-governmental organizations in a consultative status with the United Nations Economic and Social Council may also attend. **All participants must be accredited by the competent authorities of their country or international organization.**

8. All participants attending the meeting are requested to have a valid passport and, if required, a visa. Applications for visas should be made as soon as possible to the Embassy of Italy in the country in which the participant resides, with a reference to the 2022 UNECE Expert Meeting on Statistical Data Collection. A letter

to facilitate obtaining a visa can be requested from the local contacts in Italy (see contact details under section VII).

9. Participants and/or their offices are requested to make their own travel arrangements and hotel reservations. The UNECE Secretariat regrets not being able to offer any financial assistance regarding travel and accommodation arrangements.

10. Participants should register online by **31 July 2022** via the following form:

<https://forms.office.com/r/TuJBd0jcbV>

11. Although we anticipate an improved COVID situation by the time of the meeting, participants must ensure they have insurance to cover all associated eventualities, including cancellation of flight and hotels if the public health situation deteriorates. Participants attend at their own risk, and should stay up-to-date with any requirements that may be needed for those who travel to Italy, including for example any documentation that might be required to demonstrate vaccination against COVID.

#### **IV. CALL FOR CONTRIBUTIONS**

12. Participants are strongly encouraged to consider submitting an abstract that summarises the content of their proposed contribution. These should cover one or more of the topics in the meeting programme. Those who contribute are also encouraged to share links to any code repositories containing code, if applicable (e.g. in GitHub). Please note that the official language of the meeting will be English, and therefore all contributions should be submitted in English only. No translation or interpretation during the meeting will be provided.

13. A short abstract of the proposed contribution should be submitted as soon as possible and by **30 June 2022** at the latest, via the following form: <https://forms.office.com/r/iMXKprtKYG>

14. The Organizing Committee will notify in due time whether the submission is accepted or not and might request changes. Information about the selection of contributions for the meeting, guidelines on formatting, and means of submission will be sent to authors by email. Contributions should normally consist of a paper, plus an accompanying presentation. Other forms of contributions may be proposed. Please note that as we finalize the agenda, it may not be possible to allocate time to all proposed contributions.

15. Any written papers must be provided by **31 August 2022** at the latest. A link will be sent to the authors where documents can be uploaded.

16. Any presentation slides, videos or other electronic materials should be provided by **31 August 2022** at the latest. Any equipment required for practical demonstrations must be provided by the participants. A link will be sent to the authors where presentations can be uploaded.

#### **V. DOCUMENTATION, METHODS OF WORK AND OFFICIAL LANGUAGE**

17. The expert meeting will consist of presentations, small group discussions to exchange experiences and to identify lessons learned, and interactive discussions to identify future work in the area of data collection. Participants are expected to actively contribute to the discussions.

18. Prior to the meeting, abstracts and papers will be made available in pdf format on the following website: <https://statswiki.unece.org/display/Collection/2022+Data+Collection+Expert+Meeting>. After the meeting, papers and presentations will be made available via the meeting webpage: <https://unece.org/statistics/events/DC2022>. Presentations will not be made available to delegates before the meeting.

19. Participants are encouraged to download the papers from the website and, where feasible, to use electronic devices to read papers in order to minimise paper use. Documents posted on the website before the meeting will **not** be distributed in the conference room.

## VI. VENUE

20. The expert meeting is planned to take place in Rome, Italy at:

Istat – Istituto nazionale di statistica  
Via Cesare Balbo, 16 – 00184 Roma  
Tel. +39 06 46731

21. A second information notice with practical information will be shared with registered participants, and loaded onto the meeting website in due time before the meeting.

## VII. FURTHER INFORMATION

22. For further information you may contact the following organisers:

Statistical Division, United Nations Economic Commission for Europe

Mr. Wai Kit Si Tou, email: [sitou@un.org](mailto:sitou@un.org)

Mr. Taeke Gjaltema, email: [taeke.gjaltema@un.org](mailto:taeke.gjaltema@un.org)

Local contact in Italy

Mr. Pasquale Papa, email: [papa@istat.it](mailto:papa@istat.it)

### DEADLINES

30 June 2022	<b>Abstract</b> or proposal for intended contribution
31 July 2022	<b>Registration</b>
31 August 2022	<b>Paper</b> or detailed abstract if no paper is provided
31 August 2022	<b>Presentation</b> (draft)
17 October 2022	Final versions of presentations
26 to 28 October 2022	Expert meeting