UNECE

DESTATIS
Statistisches Bundesamt

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS
UNECE Expert Meeting on Statistical Data Confidentiality
26 to 28 September 2023, Wiesbaden, Germany

INF.1
13 February 2023

Information Notice No.1

I. PURPOSE AND TARGET AUDIENCE OF THE MEETING

1. The 2023 UNECE Expert meeting on Statistical Data Confidentiality will be hosted by the Federal Statistical Office of Germany (Destatis) in coordination with RheinMain University of Applied Sciences. It will take place in Wiesbaden, Germany on 26 to 28 September 2023, starting at 09:30 am on Tuesday 26 September, and ending in early afternoon on Thursday 28 September.

2. This meeting will be organized as part of the Conference of European Statisticians' work programme for 2023, within the context of the High-Level Group for the Modernisation of Official Statistics.

3. The objective of this expert meeting is to identify innovative approaches and best practices in statistical data confidentiality, and to provide a platform for practitioners to exchange experiences and foster collaboration in this area. In addition to the more traditional presentations, the agenda will include target-driven small group discussions and activities to identify best practices and new opportunities. Delegates will be asked to contribute to the development of internationally-coordinated work in the field of statistical data confidentiality. The meeting is primarily intended for experts from national and international statistical offices as well as invited academics dealing with statistical disclosure limitation.

II. AGENDA OF THE MEETING

4. The programme of this meeting will consist of the following substantive topics. (Further details about these topics can be found in the annex at the end of this Information Notice.):

- Innovative approaches in granting access to microdata for scientific purposes (new microdata access modes, services and tools);
- Producing useful microdata files; different approaches depending on different types of files;
- Risk assessment: Privacy, confidentiality, and disclosure vs utility;
- Output checking in research data centres;
- Challenges in publishing safe tables and maps;
- Software tools for statistical disclosure control;
- SDC communication, education, and training; and
- Other emerging issues.

III. PARTICIPATION AND ACCREDITATION

5. Representatives of all Member States of the United Nations and of interested intergovernmental organizations are welcome at this meeting. 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.**

6. 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 Germany in the country in which the participant resides, with a reference to the 2023 UNECE Expert meeting on Statistical Data Confidentiality.

A letter to facilitate obtaining a visa can be requested from the local contacts in Wiesbaden (contact details below).

7. Participants should register online by **28 August 2023** via following the link:

https://indico.un.org/e/SDC2023

8. Your timely registration will ensure that you receive any email broadcasts of information that we may send out to those who have registered.

9. Please note that if you do not already have a user account on the indico system, before you can register you may first need to create an account, and then activate it by clicking on the account activation link that is sent to you by email. Once you have a user account, and have logged in, you can submit your registration for this meeting. You will receive a notification confirming this, and then a further notification email when your registration is approved.

10. 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, although some support may be available directly from the UNESCO Chair in Data Privacy, as described below.

11. Although we anticipate that the COVID situation will not be particularly problematic at the time of this 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 Germany.

IV. FINANCIAL SUPPORT FOR PARTICIPATION

12. The UNESCO Chair in Data Privacy (http://unescoprivacychair.urv.cat) sponsors a limited number of travel grants for contributors and delegates from transition countries. For further information please directly contact unescoprivacychair@urv.cat (for attention of Moisès Pedrajas).

V. CALL FOR PAPERS, METHODS OF WORK AND OFFICIAL LANGUAGES

13. Participants are strongly encouraged to consider submitting an abstract summarising the content of their proposed contribution. These should cover one or more of the topics of the meeting programme. The official language of the meeting will be in English, and therefore all contributions should be submitted in English only. No translation or interpretation during the meeting will be provided.

14. Regarding abstracts received, the steering committee will send notification in due time about whether the submission is accepted or not and might request changes. Contributions should normally consist of a paper, plus an accompanying presentation. Other forms of contribution may be proposed. Information about the selection of contributions for the meeting, guidelines on formatting, and means of submission will be sent to authors by email. Please note that due to the nature of the meeting, it may not be possible to allocate time to all proposed contributions.

15. The following **deadlines** and requirements apply:

(i) A short abstract of the proposed contribution should be submitted via the form available on the following webpage, as soon as possible and by 10 March 2023 at the latest:

https://indico.un.org/e/SDC2023

(ii) Any written papers must be supplied by **13 July 2023** at the latest. A link will be sent to the authors where documents can be uploaded.

(iii) Any presentation slides, videos or other electronic materials should be supplied by **4 September 2023** 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.

16. Papers will be made available online for this meeting before the meeting via the following website: https://unece.org/statistics/events/SDC2023. Presentations will also be added to that site after the meeting. Presentations will not be made available to delegates before the meeting, unless requested by the presenters.

17. Participants are encouraged to download the papers via the webpage and, where feasible, to use electronic devices to read papers in order to minimise paper use. These documents will not be distributed in the conference room.

VI. VENUE

18. The meeting will take place in Wiesbaden, Germany at:

Hochschule RheinMain
Building G
Kurt-Schumacher-Ring 18
65197 Wiesbaden
Germany

19. A second information notice with practical information will be shared with registered participants and on the meeting website in due time before the meeting.

VII. FURTHER INFORMATION

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

Secretariat of the United Nations Economic Commission for Europe:
Christopher Jones, email: *JonesC@un.org*

Local Contacts in Wiesbaden:
Sarah Gießing, email: *Sarah.Giessing@destatis.de*

DEADLINES

10 March 2023	**Abstract** or proposal for intended contribution
13 July 2023	**Paper**
28 August 2023	**Registration**
4 September 2023	**Presentation** or other material to be presented
26 to 28 September 2023	Meeting

VIII. ANNEX: EXPLANATORY NOTES TO THE AGENDA

1. Innovative approaches in granting access to microdata for scientific purposes (new microdata access modes, services and tools)

Description:

21. Several statistical agencies provide access to their microdata (e.g. for scientific purposes). Different modes of microdata access exist, such as release of (partially or fully) anonymised microdata files, onsite access (safe centres), remote access systems, remote program execution and remote analysis servers. An appropriate access mode should balance statistical confidentiality with the usefulness of information offered to the user community.

22. We invite papers discussing current practices, as well as innovative modes of access. Contributions to this topic could address both national solutions, as well as international trans-border access. Papers can be both from the perspective of the organization that provides access or from the perspective of the users of microdata.

23. Contributions to this topic could include:

- Case studies of remote access to microdata in virtual labs;
- Real-time confidentiality (solutions allowing users to query the data with confidentiality applied instantly, including querying systems and table builders);
- Data user perspectives;
- Efficient management of disclosure risk through the application of "five safes" approach (safe: projects, settings, data, outputs and people).

2. Producing useful microdata files; different approaches depending on different types of files

Description:

24. This topic will cover different approaches to generate anonymized microdata, including masking and synthetic data generation. Papers on anonymizing emerging types of microdata, such as mobility data, geo-referenced data or even unstructured individual data are also welcome. Given the current decentralization trend in computing, decentralized approaches to microdata protection are also of particular interest.

25. Items in this topic thus include:

- Microdata protection, including special types of data like mobility data; longitudinal data; geo-referenced/geospatial data; census data or linked data;
- Decentralized/local microdata protection; and/or
- Synthetic data.

3. Risk assessment: Privacy, confidentiality, and disclosure vs utility

Description:

26. Any data release, whether it consists of summaries, tables or microdata, inherently involves the risk that information about the respondents in the data set may be leaked. Risk assessment attempts to quantify the risk of leakage when releasing data/tables/summaries. Data release risk assessment has received considerable attention in the literature for over half a century. A number of measures have been proposed, including privacy measures, confidentiality measures, and disclosure measures (identity and value), for example. Yet, there

remains considerable confusion about what these measures mean and/or how they relate to one another. On the other hand, performing SDC implies the loss of information (due to suppression or perturbation of data). There are a number of methods for measuring it, e.g. assessing differences in distributions, variations or correlation between data before and after SDC. Some of them can have slightly universal form. Ultimately, a balance between minimization of the risk of disclosure and minimization of the information loss due to SDC should be established.

27. For this topic, we welcome papers that address:

- Relationships between the different measures of risk, information loss and/or an exposition of these measures;
- Development of new measures that go beyond existing measures;
- Trade-off between risk and utility (opposite of the information loss) in data releases;
- Record linkage for disclosure risk assessment;
- Privacy models in official statistics: k-anonymity and extensions, differential privacy, etc.;
- Risk assessment procedures for the 2020 census; and/or
- Any other topic related to risk and utility measures.

4. **Output checking in research data centres**

Description:

28. Recently, more and more statistical institutes offer researchers the option to analyse more detailed microdata (than the protected data leaving the institute) in a safe setting called a research data centre. In this topic papers are welcomed that describe challenges to check the output in research data centres. The more users one has in such centres the more output is being produced. Thus simple rules and procedures will be necessary to make sure that the output released is safe and will not intervene with the regular output of the statistical institute that facilitates the research data centre. Aspects that can be included in papers in this topic are e.g. educating the researchers before and during their statistical activities, practical checks to be implemented and how to manage keeping track of all output checks over projects and time.

5. **Challenges in publishing safe tables and maps**

Description:

29. This topic aims at discussing the various challenges of developing, adapting or parametrizing methods to handle disclosure risks of data presented in aggregated formats like statistical tables, or, in case of georeferenced data, as statistical maps, so that they can be published safely in these formats.

30. We invite papers dealing with challenges and innovation of traditional practices like cell suppression techniques, as well as papers discussing alternative approaches, hybrid strategies, or with a particular focus on publishing safe maps.

31. Topics of interest include:

- Advantages and challenges of combining methods (e.g., Targeted Record Swapping and Cell Key Method), studying risk and utility for combined methods
- Utility and risk measurement; approaches to automate balancing risk and utility
- The relative advantages of suppression versus perturbation techniques.
- Improving state of the art algorithms like, e.g. for cell suppression;
- Methods and algorithms to handle group attribute disclosure risks using cell suppression.

6. **Software tools for statistical disclosure control**

Description:

32. The final stage in the process concerning statistical data confidentiality is the implementation of SDC methods. Either because of the complexity of the methods, or the volume of the data to be protected, it is preferred to have dedicated software to apply such methods. Contributions can be either as a paper or possibly in a "walk around" fashion (where a live demonstration can be given), allowing participants to discuss the features they would most desire to be available in SDC tools software.

33. Possible areas of contribution include:

- General purpose SDC software;
- New implementations (or even prototypes); and/or
- Automated output checking.

7. **SDC communication, education, and training**

Description:

34. This topic covers communication, education and training on statistical disclosure control and SDC methodology for different user groups. Data users should be informed and aware of applied SDC methods especially if perturbative methods are used. Researchers or other microdata users need training on methods applied to microdata and how to make safe outputs. SDC practitioners such as data analysts or statisticians need training on the impact of different SDC methods on data in order to make informed decisions on how to protect their statistics or other aggregated outputs.

35. We invite papers discussing current practices or new ideas on effective SDC related communication, education or training for different target groups and different type of data users.

8. **Other emerging issues**

Description:

36. While statistical institutes have traditionally dealt with survey sample data, they now also have to deal with other types of data, including administrative data, unstructured text, event-based data, network data, and detailed geo-referenced data. In addition to the heterogeneity of data types, some of these data can be privately held, and there exists the need for designing access techniques able to make the process for treating the data for statistical purposes to be privacy-preserving. Alongside "traditional" methods and techniques for data protection there is also increasing interest in new(er) techniques, such as perturbative methods and privacy preserving techniques, for example.

37. The aim of this topic is to bring together new techniques and methodologies, applied to (old and) new types of data, and to discuss issues concerning the design, the implementation and the communication of these new techniques.

38. Contributions to this topic could include:

- Perturbative methods (use cases, opportunities and threats, communication);
- Privacy preserving techniques;
    (i) Secure data sharing;
    (ii) Secure multi party computation;
    (iii) Privacy preserving record linkage;
- Privacy and A.I./machine learning;

- The relationship between input and output privacy;
- Anonymity in network data;
- SDC and visualisations;
    (i) Visualising confidentiality measures;
    (ii) Confidentialising visualisations; and/or
- Privacy and geo-referenced data.