

*in cooperation with*



POZNAŃ UNIVERSITY  
OF ECONOMICS  
AND BUSINESS

## JOINT UNECE/EUROSTAT EXPERT MEETING ON STATISTICAL DATA CONFIDENTIALITY

1-3 December 2021

Auditorium C

3rd Floor, Centrum Edukacyjne Usług Elektronicznych (CEUE) building  
ul. Towarowa 55, 61-896 Poznań, Poland

29 November 2021

### Tentative timetable

#### DAY 1 – Wednesday 1 December

Local time	Item
08:45	Registration of those arriving in Poznań & Connection for remote participants
09:30	<b>Opening of the meeting by hosts:</b> - Maciej Żukowski, Rector of Poznań University of Economics and Business - Dominik Rozkrut, President of Statistics Poland <b>Introduction to HLG-MOS and Statistical Data Confidentiality (Taeke Gjaltema)</b> - Taeke Gjaltema (UNECE)
10:00	<b>Introducing the procedure for questions and answers</b>
10:05	<b>Topic: Access to microdata</b> Session Organizers: Aleksandra Bujnowska (Eurostat) and Eric Schulte Nordholt (Statistics Netherlands)
10:15	<b>Access to different kinds of Statistics Netherlands' microdata</b> Eric Schulte Nordholt (Statistics Netherlands)
10:30	<b>Creating ready-made research datasets from national administrative registers</b> Päivi Kankaanranta (Statistics Finland)
10:45	<b>Fingerprinting relational data</b> Tanja Šarčević (SBA Research)
11:00	<b>Questions to the authors</b>
11:15	<b>Break (20 minutes)</b>
11:35	<b>Shedding light on the legal approach to aggregate data under the GDPR &amp; the FFDR</b> Emanuela Podda (Università di Bologna)
11:50	<b>Transnational access to confidential microdata: Progress and impact for research</b> Maria Alkhoury (Centre d'accès sécurisé aux données / Secure Data Hub)
12:05	<b>Microdata access where we are and where we need to go</b> Elizabeth Green (University of the West of England)
12:20	<b>Questions to the authors</b>
12:35	<b>Lunch break</b>
13:50	<b>Topic: Access to microdata (contd)</b>
13:55	<b>Microdata access services coping with COVID-19 lockdown</b> Natalia Volkow (National Institute of Statistics and Geography)
14:10	<b>Questions to the author and general discussion</b>

## JOINT UNECE/EUROSTAT EXPERT MEETING ON STATISTICAL DATA CONFIDENTIALITY

### DAY 1 – Continued

14:30	<b>Topic: Risk assessment: Privacy, confidentiality, and disclosure</b> Session Organizers: Josep Domingo-Ferrer (Universitat Rovira i Virgili), Krish Muralidhar (University of Oklahoma)
14:40	<b>Statistical disclosure control for machine learning models</b> Felix Ritchie (University of the West of England)
14:55	<b>The trade-off between the risk of disclosure and data utility in SDC – a case of data from a survey of accidents at work</b> Andrzej Młodak (Statistical Office in Poznań)
15:10	<b>Using machine learning to assist output checking</b> Josep Domingo-Ferrer (Universitat Rovira i Virgili)
15:25	<b>Questions to the authors</b>
15:40	<b>Break (20 minutes)</b>
16:00	<b>Disclosure metrics born from statistical evaluations of data utility</b> Devyani Biswal (University of Ottawa)
16:15	<b>Risk assessment procedures for the 2020 U.S. census</b> David Van Riper (University of Minnesota)
16:30	<b>Questions to the authors</b>
16:45	<b>Proposal for a risk assessment scale for privacy risks in the disclosure of statistical information</b> Jesús González López (National Institute of Statistics and Geography)
17:00	<b>Database reconstruction is very difficult in practice</b> Krishnamurty Muralidhar (University of Oklahoma)
17:15	<b>Questions to the authors and general discussion</b>
17:40	<b>Topic: Other Emerging issues</b> Session Organizers: Peter-Paul de Wolf (Statistics Netherlands) and Janika Tarkoma (Statistics Finland)
17:45	<b>Some of entity resolution and the impacts on personal privacy</b> Rebecca Steorts (Duke University)
18:00	<b>Questions to the author</b>
18:05	<b>End of Day 1</b>

### DAY 2 – Thursday 2 December

08:45	<i>Connection for remote participants</i>
09:00	<b>Reflections from day 1</b>
09:15	<b>Topic: Software tools for statistical data confidentiality</b> Session Organizers: Peter-Paul de Wolf (Statistics Netherlands) and Andrzej Młodak (Poznań Statistical Office)
09:25	<b>Suppression of directly-disclosive cells in frequency tables</b> Daniel Lupp (Statistics Norway)
09:40	<b>Introducing a graphical user interface for creating the metadata governing the secondary cell suppression process</b> Michel Reiffert (Destatis)
09:55	<b>Assessing, visualizing and improving the utility of synthetic data</b> Gillian Raab (Scottish Centre for Administrative Data Research)

**JOINT UNECE/EUROSTAT EXPERT MEETING ON STATISTICAL DATA  
CONFIDENTIALITY**

**DAY 2 – Continued**

<b>10:10</b>	<b>Questions to the authors</b>
<b>10:25</b>	<b>Private linear regression: Can we scale up with Big Data?</b> Giuseppe Bruno (Bank of Italy)
<b>10:40</b>	<b>Automatic checking of research outputs</b> Marco Stocchi (Eurostat)
<b>10:55</b>	<b>Questions to the authors and general discussion</b>
<b>11:15</b>	<b>Break (20 minutes)</b>
<b>11:35</b>	<b>Topic: Microdata protection</b> Session Organizers: Josep Domingo-Ferrer (Universitat Rovira i Virgili), Krish Muralidhar (University of Oklahoma)
<b>11:45</b>	<b>Accounting for longitudinal data structures when disseminating synthetic data to the public</b> Joerg Drechsler (Institute for Employment Research)
<b>12:00</b>	<b>AI-based privacy preserving census(like) data publication</b> Johannes Gussenbauer (Statistics Austria)
<b>12:15</b>	<b>Generating tabular data using generative adversarial networks with differential privacy</b> Giacomo Astolfi (European Central Bank)
<b>12:30</b>	<b>Questions to the authors</b>
<b>12:45</b>	<b>Lunch break</b>
<b>14:00</b>	<b>Topic: Microdata protection (contd)</b>
<b>14:05</b>	<b>Generative adversarial networks for synthetic data generation: A comparative study</b> Claire Little (University of Manchester)
<b>14:20</b>	<b>Data access modernization in National Statistical Offices through synthetic data, the HLG-MOS guide</b> Kenza Sallier (Statistics Canada)
<b>14:35</b>	<b>Extreme value protection adjustment for different subpopulations in complex data sets</b> Anna Oganian (National Center for Health Statistics)
<b>14:50</b>	<b>Questions to the authors and general discussion</b>
<b>15:25</b>	<b>Break (20 minutes)</b>
<b>15:35</b>	<b>Topic: Tabular data</b> Session Organizers: Steven Thomas (Statistics Canada) and Sarah Giessing (Destatis)
<b>15:45</b>	<b>Differential privacy and noisy confidentiality concepts for European population statistics</b> Fabian Bach (Eurostat)
<b>16:00</b>	<b>Fair risk-utility comparison of tabular perturbation methods by post-processing to expected frequencies</b> Øyvind Langsrud (Statistics Norway)
<b>16:15</b>	<b>Suppression or perturbation?</b> Wim Kloek (Eurostat)
<b>16:30</b>	<b>Questions to the authors</b>
<b>16:45</b>	<b>Considerations to deal with the frozen cell problem in Tau-Argus Modular</b> Sarah Giessing (Destatis)
<b>17:00</b>	<b>Increasing utility of economic statistical information</b> Steven Thomas (Statistics Canada)
<b>17:15</b>	<b>Questions to the authors and general discussion</b>
<b>17:35</b>	<b>End of Day 2</b>

**JOINT UNECE/EUROSTAT EXPERT MEETING ON STATISTICAL DATA  
CONFIDENTIALITY**

**DAY 3 – Friday 3 December**

<b>08:45</b>	<i>Connection for remote participants</i>
<b>09:00</b>	<b>Reflections from day 2</b>
<b>09:15</b>	<b>Topic: Other Emerging issues (Contd)</b> Session Organizers: Peter-Paul de Wolf (Statistics Netherlands) and Janika Tarkoma (Statistics Finland)
<b>09:25</b>	<b>A proof-of-concept solution for secure processing of mobile network operator data for official statistics</b> Fabio Ricciato, (Eurostat)
<b>09:40</b>	<b>Modelling data environments within PROV to assist anonymisation decision-making</b> Mark Elliot (University of Manchester)
<b>09:55</b>	<b>Structural uniqueness in network data</b> Marieke de Vries (Statistics Netherlands)
<b>10:10</b>	<b>Questions to the authors and general discussion</b>
<b>10:40</b>	<b>Updates on the HLG project on input privacy preservation</b> Dennis Ramondt (Statistics Netherlands)
<b>10:55</b>	<b>Questions to the presenter</b>
<b>11:00</b>	<b>Results of the discussions on future work</b>
<b>11:30</b>	<b>Conclusion of the meeting</b>
<b>11:45</b>	<b>End of the meeting</b>