

PRIVACY IN STATISTICAL DATABASES 2022

Paris, France - September 21-23, 2022

PROGRAM

Wednesday, 21 September

08h30 Registration

09h15 Opening Session: William E. Winkler in memoriam

Session: Privacy Models

09h40 "Privacy Analysis with a Distributed Transition System and a data-wise metric", *Siva Anantharaman, Sabine Frittella and Benjamin Nguyen.*

10h05 "Multivariate Mean Comparison under Differential Privacy", *Martin Dunsche, Tim Kutta and Holger Dette*

10h30 "a,m-warden: A syntactic privacy notion for continuous data publishing against probabilistic attacks", *Adrian Tobar Nicolau, Javier Parra Arnau and Jordi Forné **

10h45 Break

11h05 "Asking The Proper Question: Adjusting Queries To Statistical Procedures Under Differential Privacy", *Tomer Shoham and Yosef Rinott*

11h30 "Towards integrally private clustering: overlapping clusters for high privacy guarantees", *Vicenc Torra*

Session: Machine learning and privacy

11h55 "Membership Inference Attack Against Principal Component Analysis", *Qualid Zari, Javier Parra-Arnau, Ayşe Ünsal, Thorsten Strufe and Melek Önen*

12h20 "When Machine Learning Models Leak: An Exploration of Synthetic Training Data", *Manel Slokom, Peter-Paul de Wolf and Martha Larson*

12h45 Lunch break

Session: Case studies I

14h45 "A Note on the Misinterpretation of the US Census Re-identification Attack", *Paul Francis*

15h10 "A Re-examination of the Census Bureau Reconstruction and Reidentification Attack", *Krish Muralidhar*

15h35 "Quality Assessment of the 2014 to 2019 National Survey on Drug Use and Health (NSDUH) Public Use Files", *Neeraja Sathe, Feng Yu, Lanting Dai, Devon Cribb, Kathryn Spagnola, Ana Saravia, Rong Cai and Jennifer Hoening*

16h00 "Privacy in Practice: Latest Achievements of the Eustat SDC Group", *Marina Ayestaran, Marta Mas and Ana Miranda*

16h25 Break

Session: Microdata protection

16h45 "An optimization-based decomposition heuristic for the microaggregation problem", *Jordi Castro, Claudio Gentile and Enric Spagnolo-Arrizabalaga*

17h10 "Optimising Twin Uniform Distribution for Multiplicative Noise Data Masking", *Pauline O'Shaughnessy, Harrison Rowles, John Brackenbury and Yan-Xia Lin **

17h25 "Optimization of Outlier Protection for Different Subpopulations in Multivariate Datasets", *Anna Oganian, Mehtab Iqbal and Goran Lesaja **

17h40 "Joint secondary anonymization of categorical and numerical variables in sensitive time-series microdata", *Eugenia Koblents Lapteva and Alberto Lorenzo Megía **

PRIVACY IN STATISTICAL DATABASES 2022

Paris, France - September 21-23, 2022

Thursday, 22 September

Session: Disclosure risk assessment and record linkage

- 09h00** "The risk of disclosure when reporting commonly used univariate statistics", *Ben Derrick, Elizabeth Green, Felix Ritchie and Paul White*
- 09h25** "Hausdorff Distance: A Powerful Tool for Matching Households and Individuals in Historical Censuses", *Thais Pacheco Menezes, Michael Fop and Thomas Brendan Murphy **
- 09h40** "Measuring reidentification risk in (pseudo)anonymized datasets: the QaR method and software", *Michel Bera, Kimon Spiliopoulos, Antonis Spinakis and Photis Stavropoulos **

Session: Unstructured and mobility data

- 09h55** "Automatic evaluation of disclosure risks of text anonymization methods", *Benet Manzanares-Salor, David Sanchez and Pierre Lison*
- 10h20** "Generation of Synthetic Trajectory Microdata from Language Models", *Alberto Blanco-Justicia, Najeeb Moharram Jebreel, Jesús A. Manjón and Josep Domingo-Ferrer*

10h45 Break

Session: Synthetic data

- 11h05** "Synthetic Individual Income Tax Data: Methodology, Utility, and Privacy Implications", *Claire Bowen, Victoria Bryant, Leonard Burman, John Czajka, Surachai Khitatrakun, Graham MacDonald, Robert McClelland, Livia Mucciolo, Madeline Pickens, Kyle Ueyama, Aaron Williams, Doug Wissoker and Noah Zwiefel*
- 11h30** "On integrating the number of synthetic data sets m into the a priori synthesis approach", *James Jackson, Robin Mitra, Brian Francis and Iain Dove*
- 11h55** "Challenges in Measuring Utility for Fully Synthetic Data", *Jörg Drechsler*
- 12h20** "Utility Assessment of Synthetic Data Generation Methods", *Md Sakib Nizam Khan, Niklas Reje and Sonja Buchegger **

12h35 Lunch break

Session: Synthetic data (cont.)

- 14h35** "Comparing the Utility and Disclosure Risk of Synthetic Data with Samples of Microdata", *Claire Little, Mark Elliot and Richard Allmendinger*
- 15h00** "Utility and Disclosure Risk for Differentially Private Synthetic Categorical Data", *Gillian Raab*
- 15h25** "Generative Modeling of Complex Data", *Nicolas Grislain, Luca Canale and Johan Leduc **

16h10 Guide visit to the "Musée des arts et métiers"

TBD Gala Dinner

PRIVACY IN STATISTICAL DATABASES 2022

Paris, France - September 21-23, 2022

Friday, 23 September

Session: Tabular data

- 09h00** "Perspectives for Tabular Data Protection – How About Synthetic Data?", *Felix Geyer, Reinhard Tent, Michel Reiffert and Sarah Giessing*
- 09h25** "On Privacy of Multidimensional Data Against Aggregate Knowledge Attacks", *Ala Eddine Laouir and Abdessamad Imine*
- 09h50** "Synthetic Decimal Numbers as a Flexible Tool for Suppression of Post-published Tabular Data", *Øyvind Langsrud and Hege Marie Bøvelstad*

Session: Case studies II

- 10h15** "How Adversarial Assumptions Influence Re-identification Risk Measures: A COVID-19 Case Study", *Xinmeng Zhang, Zhiyu Wan, Chao Yan, Thomas Brown, Weiyi Xia, Aris Gkoulalas-Divanis, Murat Kantarcioglu and Bradley Malin*
- 10h40** "Privacy Protection for Youth Risk Behavior Using Bayesian Data Synthesis: A Case Study to the YRBS", *Yixiao Cao and Jingchen Hu **

10h55 Break

Session: Privacy-Preserving Protocols

- 11h15** "Tit-for-Tat Disclosure of a Binding Sequence of User Analyses in Safe Data Access Centers", *Josep Domingo-Ferrer*
- 11h40** "Secure and non-interactive k-NN classifier using symmetric fully homomorphic encryption", *Yulliwas Ameur, Rezak Aziz, Vincent Audigier and Samia Bouzefrane*
- 12h05** "Anonymisation and Fingerprinting of Microdata: A Genetic Algorithm for Finding the Optimal Set for Data Distribution", *Tanja Šarčević and Rudolf Mayer **
- 12h20** "Anonymized data collection on platform workers", *Joris Duguépéroux, Gaspard Defréville, Lucas Verney and Nicolas Deffieux **

12h35 Closing remarks

All the presentations include 5 minutes for questions

(*) Short presentation. Paper in USB proceedings