



Series of conferences: New data and methods for generating official information

Sessions, November 3rd, 5th, 12th 2020,
February 4th and 9th, 2021

<https://www.inegi.org.mx/eventos/2020/newData/>
Program



Background:

The massive use of electronic devices and the internet generate a multiplicity of digital data, which makes it possible to strengthen the generation of official information both statistical and geographically. Specifically, using current data analytics, it is possible to statistically exploit large digital databases, generated by multiple activities, through various electronic devices and remote perception systems. Although, several challenges such as legal, privacy, financial, administrative, methodological, and technological ones, should be considered.

However, this possibility gives us the opportunity to explore and strengthen information on various issues, both in the economic and social areas. That is why, in this Series of Conferences we are reviewing some possibilities to improve the timely dissemination of economic information and to better detect and prevent health and security conditions.

Taking this into consideration, the National Institute of Statistics and Geography of Mexico and the Tec de Monterrey, had decided to join forces to organize this series of conferences through digital platforms, where selected topics on the use of novel sources and methodologies to improve the current information, will be presented.

General Objective:

To review the possibilities of accessing and exploiting new sources of data held by private information repositories, state and/or governmental entities, with clear operational application, to generate public interest information in economic and social matters.

Specific Objectives:

To review the current state of the art of Big Data use, generated by the banking and financial, commercial, and personal communications and telecommunications services sectors, among others.

To review the academic progress and the latest advances in statistical practices around the world.

Development:

The series of conferences will consist in thematic blocks, with speakers giving conferences on specific subjects. In the first block, non-traditional sources of information and innovative methodologies to improve the timely dissemination of short-term economic information will be reviewed. In the second block the focus will be on relevant sociodemographic issues, such as health and public safety.

All the conferences, as well as questions will be in English.

Audience:

These conferences are intended for official information generators, information users, academicians, and graduate students, as well as decision makers within the public and private sectors.

First Block: The timely dissemination of short-term economic information.

November 3rd

9:00- 10:00

9:00-9:05 Opening

Alejandro Poiré, Tec de Monterrey

9:05-9:30 Conference: “Types of Big Data for macroeconomic nowcasting “

Presentation by **Gian Luigi Mazzi**, Senior Consultant

Comments and moderation: **Francisco Corona**, INEGI

9:30-10:00 Q&A

November 5th

9:00- 10:00

9:00-9:30 Conference “The DEI: tracking economic activity daily during the lockdown”

Presentation by **António Rua**, Banco de Portugal/NOVA School of Business and Economics

Comments and moderation: **Gerardo Leyva**, INEGI

9:30-10:00 Q&A

10:00-11:00

10:00-10:30 Conference: “Consumer Mobility and Expenditure during the COVID-19 Containment: Evidence from French Transaction Data”

Presentation by **David Bounie**, **Youssef Camara**, Telecom Paris, Institut Polytechnique de Paris and **John W. Galbraith**, McGill University

Comments and moderation: **Carlos Lever**, Tec de Monterrey

10:30-11:11 Q&A

Second Block: Social topics

November 12th

9:00- 10:00

9:00-9:05 Opening

Julio A. Santaella, INEGI

9:05-9:30 Conference: "Quantifying beautiful places with deep learning"?

Are beautiful environments good for our health and happiness? In this talk, I will describe how millions of ratings from an online game called 'Scenic-or-Not' and a mobile app called 'Mappiness' have begun to offer new answers to this age-old question. I will explain how deep learning can help us understand whether beautiful places are simply natural places - or whether humans might be able to build beautiful places too.

Presentation by **Suzy Moat**, Data Science Lab, The University of Warwick/ The Alan Turing Institute

Comments and moderation: **Elio Villaseñor**, INEGI

9:30-10:00 Q&A

10:00-11:00

10:00-10:30 Conference: "Global COVID-19 Symptom Survey. A Facebook partnership"

This presentation discusses a partnership between Facebook and academic institutions to create a global COVID-19 symptom survey. The survey is available in 56 languages. A representative sample of Facebook users is invited on a daily basis to report on symptoms, social distancing behavior, mental health issues, and financial constraints. Facebook provides weights to reduce nonresponse and coverage bias. Privacy protection and disclosure avoidance mechanisms are implemented by both partners to meet global policy and industry requirements. Country and region-level statistics are published daily via dashboards, and microdata are available for researchers via data use agreements. Over 1 million responses are collected weekly. Operating on such a large scale is not without its challenges. The presentation will discuss those in detail.

Presentation by **Frauke Kreuter**, Joint Program in Survey Methodology, University of Maryland

Comments and moderation: **Hugo Fuentes**, Tec de Monterrey

10:30-11:00 Q&A

February 4th, 2021

9:00- 10:00

9:00- 9:05 Opening

INEGI

9:05-9:30 Conference: "Small area estimation from survey data: applications in criminology"

The geographical study of crime and related phenomena at the neighbourhood or smaller location or area level has traditionally relied on the use of crime reported data to the police. Criminologists traditionally have relied on surveys at the city, regional and national level to compensate for the biases and limitations of police data (not all crime is reported to the police and important attitudinal and behavioural elements related by theory to crime and criminal justice are not captured by this administrative data), but present limitations if we want to drill down at larger scale of geographical resolution. New forms of data, whether coming from public health data (Fulchand, 2020), other untapped large administrative datasets (O'Brien et al. 2015), or crowdsource data (Buil-Gil & Solymosi, 2020) have been advocated as a potential solution for some of these issues. Aside from these, the use of statistical methods to derive unplanned area estimates from surveys can also be a fruitful line of work. Most surveys are designed to record samples which only allow producing reliable direct estimates (weighted means or totals) at the level of very large areas, such as groups of states in the US or Police Force Areas in the UK, while direct estimates tend to be unreliable at the level of small geographies. Small area estimation techniques borrow strength from related areas to produce reliable small area estimates. In this presentation, we provide a general introduction to these techniques and use a couple of examples to illustrate how they can be used to both compensate for and illustrate the limitations of traditional police data. We finish the presentation highlighting our future work in this area.

Presentation by **Juanjo Medina Ariza**, University of Seville
David Buil-Gil, The University of Manchester

Comments and moderation: **Grisel Ayllón Aragón**, Tec de Monterrey

9:30-10:00 Q&A

10:00-11:00

10:00-10:30 Conference: "New forms of data for public safety"

Crowdsourcing and civic technologies offer new forms of data, facilitating insight into people's experiences with public safety. These data sources allow us to follow people along their routine activities at unprecedented scales. Researchers can transform such insight into learning about experiences of place and crime. In this talk I will discuss my work on making use of such crowdsourced data to better understand topics of: fear of crime, disorder, sexual harassment, and police use of social media. I will introduce methods for data collection, and issues with data analysis. The aim is to illustrate the potential of using such approaches to better understand the situational factors associated with crime and fear of crime as experienced by people in their everyday lives, and encourage discussion of the strengths, limitations, and ethical and policy implementations of these emerging methodological approaches.

Presentation by **Reka Solymosi**, The University of Manchester

Comments and moderation: **José Guillermo Castillo**, INEGI

10:30-11:00 Q&A

February 9th, 2021

10:00- 11-00

10:00- 10:05 Opening

INEGI

10:05-10:40 Conference: “Ethical issues of AI and Big Data”

Presentation by **David Leslie**, The Alan Turing Institute

Comments and moderation: **Miguel Ángel Toro**, Tec de Monterrey

10:40-11:00 Q&A

Speakers:



Gian Luigi Mazzi. After more than 30 years as Eurostat permanent staff and 2 years as Technical Director at GOPA Luxembourg (part of the GOPA group), Gian Luigi Mazzi is now Senior Consultant in Statistics and Econometrics. His main areas of interest are: time-series analysis and modelling, business cycle analysis, nowcasting/forecasting, big data, seasonal adjustment and uncertainty in official statistics. Gian Luigi Mazzi is author of more than 150 working papers, papers and chapters in books as well as editor of some books, handbooks and guidelines.



António Rua Rua is a senior economist in the Economics and Research Department of the Portuguese Central Bank, Banco de Portugal.

He is also Associate Professor at NOVA School of Business and Economics, Portugal. His major fields of expertise are time series econometrics and quantitative macroeconomics, with particular emphasis on forecasting and empirical macro and finance. He holds a Ph.D. from the Technical University of Lisbon a M.Sc. and a Bachelor's degree from the *Universidade Nova de Lisboa, Portugal*.



John W. Galbraith has a Ph.D. in econometrics from Oxford University and has been a professor in the Department of Economics at McGill University since 1987. His fields of interest include time series econometrics, and empirical macroeconomics and finance.

His recent work concerns maximum horizons for forecasts with statistical models in the fields of macroeconomics and financial volatility and estimating long-memory models and GARCH models.

David Bounie is full Professor of Economics and the Head of the department of Economics and Social Sciences at Telecom Paris, Institut Polytechnique de Paris. His research is interested in digital economics and more precisely in how digital technologies are transforming the finance industry (digital payments, cryptocurrencies, artificial intelligence) in developed and developing countries (financial inclusion).

He collaborates with colleagues in Economics, Statistics and Computer sciences from universities, public institutions and private research companies. Beyond academic journals, some of his works have been published in central banks' working paper series such as the European Central Bank, the Federal Reserve Bank of Boston, the Bank of Canada, De Nederlandsche Bank, the Deutsche Bundesbank, the Austrian Central Bank, and the Colombian Central Bank.

He is the co-founder of the Chair Digital Finance (with Marianne Verdier, Univ. Paris 2 Panthéon-Assas) in partnership with Institut Louis Bachelier, Cartes Bancaires CB and La Banque Postale, and the interdisciplinary ANR Chair in Artificial Intelligence (AI) "Explainable AI for Anti-Money Laundering" with Winston Maxwell (Law) and Stéphan Cléménçon (Applied Mathematics) in partnership with PWC, Dataiku, and ACPR/Banque de France.

He serves as an Associate Editor for Electronic Commerce Research and Application.

Youssef Camara is currently a PhD candidate in Economics and Data Science. The aim of his study is to use the big data of bank card transactions to investigate various economic research questions, particularly to make macroeconomic forecasts. To do so, he uses both traditional methods and new techniques from machine learning.

He works on the topics related to statistical and economic modeling, prospective studies (forecasting/machine learning), program/impact evaluation, and he does appropriate statistical tests before making recommendations.



Suzy Moat is Professor of Behavioural Science at Warwick Business School, where she co-directs the Data Science Lab. She is also a Fellow of The Alan Turing Institute. Her research investigates whether data on our usage of the Internet, from sources such as Google, Wikipedia and Flickr, can help us measure and even predict human behaviour in the real world.

Moat's work touches on problems as diverse as linking online behaviour to stock market moves (with Preis, Curme, Stanley, et al.), estimating crowd sizes (with Botta and Preis) and evaluating whether the beauty of the environment we live in might affect our health (with Seresinhe and Preis). The results of her research have been featured by television, radio and press worldwide, by outlets such as CNN, BBC, The Guardian, Wall Street Journal, New Scientist and Wired.

Moat studied Computer Science at UCL and Psychology at the University of Edinburgh and won a series of prizes during her studies. With her collaborator and Data Science Lab co-director Tobias Preis, she led an online course on using big data to measure and predict human behaviour which has attracted over 25,000 learners to date. Moat has also acted as an advisor to government and public bodies on the predictive capabilities of big data.



Frauke Kreuter is Director of the Joint Program in Survey Methodology at the University of Maryland, USA; Professor of Statistics and Methodology at the University of Mannheim; and head of the Statistical Methods Research Department at the Institute for Employment Research (IAB) in Nürnberg, Germany. She received her Master in Sociology from the University of Mannheim, Germany, and her PhD in Survey Methodology from the University of Konstanz. Before joining the University of Maryland she held a postdoc at the UCLA Statistics Department. Her research focuses on sampling and measurement errors in complex surveys. In her work at JPSM she maintains strong ties to the Federal Statistical System and served in advisor roles for the National Center for Educational Statistics and the Bureau of Labor Statistics.

In addition to her academic work Dr. Kreuter is the Founder of the International Program for Survey and Data Science, developed in response to the increasing demand from researchers and practitioners for the appropriate methods and right tools to face a changing data environment; Co-Founder of the Coleridge Initiative, whose goal is to accelerate data-driven research and policy around human beings and their interactions for program management, policy development, and scholarly purposes by enabling efficient, effective, and secure access to sensitive data about society and the economy. coleridgeinitiative.org; and Co-Founder of the German language podcast Dig Deep.



Prof. Medina is a Talentia Senior Distinguished Researcher at the Department of Criminal Law and Crime Science at the University of Seville. Previously he was a Professor of Quantitative Criminology and Head of the Criminology Department at the University of Manchester. Prof. Medina recent research has focused on the use of machine learning applications to the prediction of domestic abuse by police agencies and the geography of crime in the metropolitan area of Barcelona.



David Buil-Gil is a Research Fellow at the Department of Criminology of the University of Manchester, UK. His areas of interest cover geographic criminology, small area estimation applications in criminology, measurement error in crime data, emotions about crime, perceptions about the police, new methods for data collection, and cybercrime. He holds a PhD in Criminology from the University of Manchester, a MA in Crime Analysis and Prevention from Miguel Hernandez University and a Bachelor's in Criminology from Autonomous University of Barcelona.



Reka Solymosi is a Lecturer in Quantitative Methods at the Department of Criminology at University of Manchester. She is a member of the Software Sustainability Institute and an Honorary Lecturer at University College London. Her research focuses on making use of new forms of data to gain insight into people's behaviour and subjective experiences, particularly focusing on crime, victimisation, transport, and spatial research. She is also dedicated to promoting data literacy in social science study and practice.



David Leslie is the Ethics Theme Lead at the Alan Turing Institute. Before joining the Turing, he taught at Princeton's University Center for Human Values, where he also participated in the UCHV's 2017-2018 research collaboration with Princeton's Center for Information Technology Policy on "Technology Ethics, Political Philosophy and Human Values: Ethical Dilemmas in AI Governance." Prior to teaching at Princeton, David held academic appointments at Yale's programme in Ethics, Politics and Economics and at Harvard's Committee on Degrees in Social Studies, where he received over a dozen teaching awards including the 2014 Stanley Hoffman Prize for Teaching Excellence. He was also a 2017-2018 Mellon-Sawyer Fellow in Technology and the Humanities at Boston University and a 2018-2019 Fellow at MIT's Dalai Lama Center for Ethics and Transformative Values.

David now serves as an elected member of the Bureau of the Council of Europe's Ad Hoc Committee on Artificial Intelligence (CAHAI). He is on the editorial board of the *Harvard Data Science Review (HDSR)* and is a founding editor of the Springer journal, *AI and Ethics*. He is the author of the UK Government's official guidance on the responsible design and implementation of AI systems in the public sector, *Understanding artificial intelligence ethics and safety* (2019) and a principal co-author of *Explaining decisions made with AI* (2020), a co-badged guidance on AI explainability published by the Information Commissioner's Office and The Alan Turing Institute. He is also Principal Investigator of a UKRI-funded project called *PATH-AI: Mapping an Intercultural Path to Privacy, Agency and Trust in Human-AI Ecosystems*, which is a research collaboration with RIKEN, one of Japan's National Research and Development Institutes founded in 1917.

David was a Principal Investigator and lead co-author of the NESTA-funded *Ethics review of machine learning in children's social care* (2020). His other recent publications include the HDSR article "Tackling COVID-19 through responsible AI innovation: Five steps in the right direction" (2020) and *Understanding bias in facial recognition technologies: An explainer* (2020). David is also a co-author of *Mind the gap: how to fill the equality and AI accountability gap in an automated world* (2020), the Final Report of the Institute for the Future of Work's Equality Task Force. In his shorter writings, David has explored subjects such as the life and work of Alan Turing, the Ofqual fiasco, the history of facial recognition systems and the conceptual foundations of AI for popular outlets from the BBC to Nature.