



SOLEP

Société Luxembourgeoise de l'Evaluation
et de la Prospective

SOLEP asbl, the Société Luxembourgeoise de l'Evaluation et de la Prospective, marks its 10th anniversary in 2019. To celebrate, SOLEP invites you to a conference of interest to policy-makers, evaluation professionals, economists, statisticians, data analysts, researchers, and students.

The event will be held under the patronage of the Ministry of the Economy and Foreign Trade and the European Investment Bank, and is organised with the support of Chamber of Commerce, KPMG Luxembourg, Artemis Information Management and PwC Luxembourg.

BIG DATA ANALYTICS AND EVALUATION **Data Privacy, Use and Ethics**

Monday 4 March 2019
Luxembourg Chamber of Commerce
7, rue Alcide de Gasperi
09h00 - 17h30

The use of big data analytics in monitoring and evaluation is on the rise. What are the new analytical tools and how can they be applied for public decision-making? What are the implications to evaluation practitioners? What are the challenges ahead in terms of privacy protection and ethics? Two plenary sessions in the morning will be followed by three parallel technical workshops in the afternoon on the use of big data analytical tools for monitoring and evaluation.

8h30: Registration

9h00: Session 1. Big data analytics in evaluation: where do we stand and where are we headed?

- Keynote address by representatives of the Luxembourgish Government and Chamber of Commerce
- Presentation by SOLEP - "The use of big data analytics for policy insight: typology, benefits and challenges"
- Panel discussion and Q&A moderated by SOLEP:
 - Conchita D'Ambrosio and Andreas Zilian (University of Luxembourg's DRIVEN DTU – Doctoral Training Unit on Data-driven Computational Modelling and Applications)
 - Rick Davies (independent consultant)
 - Hansdeep Khaira (Evaluation Officer, Independent Office of Evaluation, IFAD – the United Nations International Fund for Agricultural Development)

This introductory session will set the scene for the rest of the day.

It will build on the **evaluation cases** and **associated benefits** discussed in [previous SOLEP workshops](#), supplemented with complementary examples. Key concepts will be also clarified following the United Nations (UN) Global Pulse initiative's typology. We will further look at in which regions of the world and in which sectors these analyses have been conducted, and we will discuss the **challenges** these new methodologies and applications have raised.

10h50: Coffee break

11h00: Session 2. Artificial Intelligence ethics and data privacy initiatives around the globe: progress and remaining challenges



- Presentation by Sophie Vulliet-Tavernier (CNIL – the French *Commission Nationale de l'Informatique et des Libertés*)
- Panel discussion and Q&A moderated by Danièle Lamarque (Member for Audit Quality Control at the European Court of Auditors):
 - Gerard Lommel (Commissioner of the Luxembourgish government for Data Protection in public administration)
 - Mark Cole (University of Luxembourg SnT – Interdisciplinary Centre for Security, Reliability and Trust)
 - Andreas Braun (consultant, PwC Luxembourg)

This second session will discuss cases of big data use and misuse in public decision-making from around the world. The cases discussed do not necessarily concern evaluation *per se* but pertain to the sphere of public policy and decision-making. Our guest speaker from CNIL will provide an outlook of efforts conducted around the globe to tackle these issues in terms of personal data protection and of multidisciplinary AI ethics initiatives.

A key objective of the following panel discussion is to understand:

- a. Data protection
 - How the examples seen across the two morning sessions, which pertain more (i) to development evaluation (and the use of data from developing countries) and (ii) to cases of data use/misuse that occurred mostly in non-European countries, could be replicated in Europe, most notably following the entry into force and application of the GDPR?
 - What are the remaining risks once GDPR provisions are applied?
- b. AI ethics
 - What balance to strike between innovation and regulation when it comes to AI?
 - What are regulators doing to support the adoption of, and demystify, AI?
 - How is AI embraced by regulators in practice?

12h20: Closing remarks by SOLEP

12h30: Lunch break

14h00: Parallel technical workshops

- **Workshop 1. Prediction modelling as an evaluation tool**, presented by Dr Rick Davies (independent consultant)

Big data analytic tools can be useful when working with small data, a more common scenario for many evaluators. One set of tools, known as predictive modelling, uses algorithms to find attributes of cases associated with outcomes of interest. Examples will be given of how prediction modelling can be useful in all stages of development aid programmes: project selection, implementation, evaluation and synthesis. These examples will focus largely on the re-analysis of existing datasets. Doing so helps extract value from existing investments and addresses the replicability challenge facing many social science findings. In the process, participants will be introduced to two software tools: Rapid Miner Studio and EvalC3. A number of important conceptual distinctions will further be clarified, including those between predictive and explanatory models and theory versus data led inquiry.

- **Workshop 2. The use of satellite imagery and signals to support decision-making**



- **Conducting impact evaluations in difficult contexts using geospatial analysis**, presented by Hansdeep Khaira (Evaluation Officer, Independent Office of Evaluation of IFAD)

This presentation shares the insights of an impact evaluation of an irrigation rehabilitation project aimed at improving farming income several years following a conflict in the Republic of Georgia. It will show how a combination of quasi-experimental techniques with remote sensing data collection using satellite imagery were used to estimate vegetation development in treated and non-treated areas. It will further show how triangulation with other data obtained from field visit and household survey administered in the same non-conflict areas helped confirm the insights gained from the geo-collected data and ascertain the effects of the project on farmers livelihoods through an increase in agricultural production. Time will also be spent understanding how the methodology has been completely automatized by developing an algorithm in open source statistical software R for future, more systematic reuse.

- **Estimating real-time rainfall using satellite communication signals**, presented by M. R. Bhavani Shankar (Research Scientist, University of Luxembourg SnT)

This presentation describes the opportunistic use of satellite communication signals for rainfall estimation and presents the findings of FNR funded RAFAEL (Rainfall estimation using signalling data of satellite communication network) project pursued at SnT. The main idea of RAFAEL is to extract rainfall information from the signalling data between satellites and user terminals by employing advanced machine learning techniques. There are more than 300,000 satellite terminals across Europe and 2 million worldwide; these can be transformed into reliable and real time rainfall measurement sensors. The data from SES is used in this project and advanced signal processing and learning algorithms are used to transform the satellite signals into rainfall estimates. The timely and accurate rainfall estimate provides crucial information for government in sectors including flash flood risk calculation, agriculture, soil erosion, etc. A flash flood map based on rainfall estimates will also be shown to demonstrate the utility of the project.

- **Workshop 3. Social media intelligence in evaluation and research**, presented by Raquel de Luis Iglesias (consultant, Ipsos MORI) with Dr Jun Pang (Research Scientist, DRIVEN DTU and SnT, University of Luxembourg)

The objective of this session is to showcase examples in which the speaker and her colleagues at Ipsos have used big data in evaluations or research studies. These cases will illustrate the potential of social intelligence in evaluation and how technical barriers to obtain data can be overcome. An example on how social intelligence was used to measure the outcomes of a monitoring report on education across the world will be shown. Analysis of social media through the platform Synthesio helped the evaluation team explore the volume of conversation and its location, both in terms of channels and geographic location, who the influencers were and the sentiment of those commenting on the report. In this evaluation, bibliometrics and text analytics were also used for further insight. The workshop will show how data from social intelligence can be triangulated with data collected through traditional methods to provide in-depth and evidence-based findings.

This workshop will be preceded by a talk by **Dr Jun Pang**, who will present a new DRIVEN DTU-supervised PhD project, where Instagram data will be explored to study social engagement and urban spatial structure.



SOLEP

Société Luxembourgeoise de l'Evaluation
et de la Prospective

16h00: Coffee break

16h15: Workshops conclusions in plenary session

17h15: Closing remarks by SOLEP