



SOLEP

Société Luxembourgeoise de l'Évaluation
et de la Prospective

En 2019, la Société luxembourgeoise de l'évaluation et de la prospective (SOLEP) asbl fête ses 10 ans. À cette occasion, la SOLEP vous invite à sa nouvelle conférence d'intérêt pour les décideurs politiques, professionnels de l'évaluation, économistes, statisticiens, analystes de données, chercheurs et étudiants.

L'évènement se déroulera sous le patronage du Ministère de l'Économie et du Commerce extérieur et la Banque européenne d'investissement, en partenariat avec la Chambre de commerce, KPMG Luxembourg, PwC Luxembourg et Artemis Information Management.

ANALYSE DE BIG DATA ET EVALUATION

Confidentialité, utilisation et éthique des données

Lundi 4 mars 2019

Chambre de commerce du Luxembourg

7, rue Alcide de Gasperi

09h00 - 17h30

L'utilisation de l'analyse de big data dans le suivi et l'évaluation est à la hausse. Quels sont ces nouveaux outils d'analyse et comment sont-ils appliqués à la prise des décisions publiques? Qu'est-ce que cela signifie pour les praticiens de l'évaluation? Quels sont les défis à venir en termes de protection de la vie privée et d'éthique? Les séances plénières du matin seront suivies par des ateliers techniques dans l'après-midi sur les outils d'analyse de big data pour le suivi et l'évaluation.

8h30 : Inscription des participants

9h00: Séance 1 - Analyse de big data dans l'évaluation: où en sommes-nous et où allons-nous ?

- o Introduction par des représentants du Gouvernement luxembourgeois et de la Chambre de commerce
- o Présentation par la SOLEP - L'utilisation de l'analyse de big data pour l'éclairage des politiques publiques : typologie, bénéfices et défis
- o Panel et questions-réponses, modérés par la SOLEP :
 - Conchita D'Ambrosio et Andreas Zilian (DRIVEN DTU de l'Université du Luxembourg - Unité de formation doctorale sur le pilotage des données, modélisation et applications informatiques)
 - Rick Davies (Consultant indépendant)
 - Hansdeep Khaira (Spécialiste de l'évaluation, Bureau indépendant de l'évaluation (IOE) du IFAD – Fonds international de développement agricole des Nations Unies)

Cette session d'introduction préparera le contexte pour le reste de la journée.

Elle sera basée sur des **exemples d'évaluation** et les **bénéfices y associés** qui ont été discutés lors des ateliers précédents de la SOLEP. Ils seront complétés par d'autres exemples si nécessaire. Les concepts clés seront également clarifiés conformément à la typologie de l'initiative des Nations Unies (Global Pulse). Nous examinerons plus en détails dans quelles régions du Monde et dans quels secteurs ces analyses ont été effectuées. Nous discuterons enfin des préoccupations soulevées par ces nouvelles méthodologies et leurs applications.

10h50 : Pause-café

11h00: Séance 2 - Initiatives en matière d'éthique et de protection des données de l'intelligence artificielle (IA) à travers le monde: progrès et défis à venir

- o Présentation par la CNIL – la Commission Nationale de l'Informatique et des Libertés française



SOLEP

Société Luxembourgeoise de l'Évaluation
et de la Prospective

- o Panel et questions-réponses animés par Mme Danièle Lamarque, membre pour le contrôle qualité de l'audit de la Cour des comptes européenne :
 - Gérard Lommel (Commissaire du gouvernement luxembourgeois pour la protection des données dans l'administration publique)
 - Mark Cole (Université du Luxembourg, SnT - Centre interdisciplinaire pour la sécurité, la fiabilité et la confiance)
 - Maria Axente (Programme d'intelligence artificielle de PwC)

L'objectif de la deuxième séance est de discuter des cas d'utilisation et d'utilisation abusive du Big Data dans la prise de décision publique et ce, dans le monde entier. Ces derniers cas concernent pas nécessairement l'évaluation en tant que telle, mais relèvent du domaine de la politique publique et de la prise de décision. Ce faisant, notre conférencier invité de la CNIL donnera un aperçu de ce qui est fait (ou non) dans le monde pour s'attaquer à ces problèmes en termes de protection des données à caractère personnel et d'initiatives multidisciplinaires sur l'éthique de l'IA.

Un des objectifs clés de la table ronde suivante est de comprendre:

a. Concernant la protection des données :

- Comment les exemples observés au cours des deux séances du matin, qui concernent davantage (i) l'évaluation du développement (et l'utilisation de données provenant de pays en développement) et (ii) les cas d'utilisation / utilisation abusive de données survenus principalement dans des pays non européens, pourraient être reproduits en Europe, notamment après l'entrée en vigueur et l'application du RGPD?
- Quels sont les risques restants une fois que les dispositions du RGPD sont appliquées?

b. Concernant l'éthique de l'IA

- Quel équilibre trouver entre innovation et réglementation en matière d'IA?
- Que font les régulateurs pour soutenir l'adoption et démystifier l'IA?
- Comment les organismes de réglementation adoptent-ils l'AI en pratique?

12h20 : Allocution de clôture par le SOLEP

12h30 : Pause déjeuner

14h00 : Ateliers techniques en parallèle

- o **Atelier 1 - La modélisation de prévision en tant qu'outil d'évaluation**, présenté par le Dr Rick Davies (Consultant indépendant)

Les outils d'analyse mégadonnées (Big data) peuvent s'avérer utiles en travaillant avec de petites données (Small data), un scénario plus courant pour de nombreux évaluateurs. Un ensemble d'outils, appelé modélisation prédictive, utilise des algorithmes pour identifier les attributs de cas associés à des résultats d'intérêt. Des exemples seront donnés sur l'utilité de la modélisation prédictive à toutes les étapes des programmes d'aide au développement: sélection, mise en œuvre, évaluation et synthèse de projets. Ces exemples porteront principalement sur la ré-analyse des jeux de données existants. Cela contribue à extraire la valeur des investissements existants et à relever le défi de la reproductibilité auquel sont confrontées de nombreuses découvertes en sciences sociales. Ce faisant, les participants découvriront deux outils logiciels: Rapid Miner Studio et EvalC3. Un certain nombre de distinctions conceptuelles importantes seront encore clarifiées, y compris celles entre les modèles prédictifs et explicatifs et les enquêtes fondées sur la théorie par rapport aux données.

- o **Atelier 2 – L'utilisation d'images et de signaux satellitaires pour faciliter la prise de décision**



- **Réaliser des évaluations d'impact dans des contextes difficiles en utilisant l'analyse géospatiale**, présenté par Hansdeep Khaira (Spécialiste de l'évaluation, Bureau indépendant de l'évaluation (IOE) de l'IFAD)

Cette présentation expose les résultats d'une évaluation d'impact d'un projet de réhabilitation de l'irrigation visant à améliorer les revenus de l'agriculture plusieurs années après un conflit en République de Géorgie. Il montrera comment une combinaison de techniques quasi expérimentales et de collecte de données de télédétection utilisant des images satellitaires a été utilisée pour estimer le développement de la végétation dans les zones traitées et non traitées. Il montrera en outre comment la triangulation avec d'autres données obtenues lors de visites sur le terrain et d'enquêtes auprès des ménages menées dans les mêmes zones non conflictuelles a permis de confirmer les connaissances acquises grâce aux données collectées géographiquement et de déterminer les effets du projet sur les moyens de subsistance des agriculteurs grâce à une augmentation de la production agricole. Un certain temps sera également consacré à comprendre comment la méthodologie a été complètement automatisée en développant un algorithme dans le logiciel statistique open source R pour une réutilisation future plus systématique.

- **Estimation des précipitations en temps réel à l'aide de signaux de communication par satellite**, présenté par M. R. Bhavani Shankar (Chercheur, Université du Luxembourg SnT)

Cette présentation décrit l'utilisation opportuniste des signaux de communication par satellite pour l'estimation des précipitations et présente les résultats du projet RAFAEL (Estimation des précipitations utilisant des données de signalisation d'un réseau de communication par satellite) financé par le FNR et mené à SnT. L'idée principale de RAFAEL est d'extraire des informations sur les précipitations des données de signalisation entre satellites et terminaux utilisateurs en utilisant des techniques avancées d'apprentissage automatique. Il existe plus de 300 000 terminaux de satellites en Europe et 2 millions dans le monde. Ceux-ci peuvent être transformés en capteurs de mesure de précipitations fiables et en temps réel. Les données de SES sont utilisées dans ce projet et des algorithmes avancés de traitement du signal et d'apprentissage sont utilisés pour transformer les signaux satellites en estimations de précipitations. L'estimation ponctuelle et précise des précipitations fournit au gouvernement des informations cruciales dans des domaines tels que le calcul du risque d'inondation éclair, l'agriculture, l'érosion des sols, etc. Une carte des crues éclair basée sur des estimations de précipitations sera également présentée pour démontrer l'utilité du projet.

- **Atelier 3 - L'intelligence des médias sociaux dans les évaluations et la recherche**, présenté par Raquel de Luis Iglesias (Consultante, Ipsos MORI) et avec Dr Jun Pang (Chercheur, DRIVEN DTU et SnT, Université du Luxembourg)

L'objectif de cette session est de présenter des cas dans lesquels la conférencière et ses collègues d'Ipsos ont utilisé le Big Data dans des évaluations ou des études de recherche. Ces cas illustreront le potentiel de l'intelligence sociale dans l'évaluation et la manière de surmonter les obstacles techniques à l'obtention de données. Un exemple sur la manière dont l'intelligence sociale a été utilisée pour mesurer les résultats d'un rapport de suivi sur l'éducation dans le monde sera présenté. L'analyse des médias sociaux via la plateforme Synthesio a aidé l'équipe d'évaluation à explorer le volume de la conversation et son emplacement, à la fois en termes de canaux et d'emplacement géographique, qui étaient les influenceurs et le sentiment de ceux qui commentaient le rapport. Dans cette évaluation, la bibliométrie et l'analyse de texte ont également été utilisées pour approfondir les connaissances. L'atelier montrera comment les données de l'intelligence sociale peuvent être triangulées avec les données recueillies à l'aide de méthodes traditionnelles pour fournir des résultats approfondis et fondés sur des preuves.



SOLEP

Société Luxembourgeoise de l'Évaluation
et de la Prospective

Cet atelier sera précédé d'une conférence par le Dr Jun Pang du DRIVEN DTU et du SnT de l'Université du Luxembourg. Le Dr Pang présentera un nouveau projet de doctorat supervisé par DRIVEN DTU dans lequel des données d'Instagram seront examinées afin d'étudier l'engagement social et la structure spatiale-urbaine.

16h00 : Pause-café

16h15 : Conclusions des ateliers en session plénière

17h15 : Allocution de clôture



SOLEP

Société Luxembourgeoise de l'Évaluation
et de la Prospective