UCLouvain

COMMUNIQUÉ DE PRESSE

8 janvier 2025

Recherche UCLouvain

Mesurer l'anonymat de nos données sur le web grâce à un logiciel UCLouvain révolutionnaire

EN BREF :

- L'émergence de l'**intelligence artificielle** a notamment pour conséquence la quasi-**impossibilité**, aujourd'hui, de **garantir l'anonymat** des personnes sur le web
- Des scientifiques de l'UCLouvain, Oxford University et Imperial College London ont développé un **modèle mathématique révolutionnaire** pour mieux comprendre les risques posés par l'IA et aider les régulateurs à **protéger la vie privée** des individus
- Exemple ? Cet outil devrait permettre de mieux réguler les **codes publicitaires** ou **trackers invisibles** qui permettent d'identifier les utilisateur·trices en ligne

INFOS : **HTTPS://WWW.OOA.WORLD/TAKE-THE-QUIZ**

CONTACT PRESSE : Julien Hendrickx, professeur de l'Ecole Polytechnique de l'UCLouvain : **+32 486 80 40 60**

L'**anonymat** est **essentiel pour protéger la liberté d'expression** et les droits numériques dans nos démocraties. Il repose sur l'absence d'**identification, de surveillance ou de traçabilité des individus**. Cependant, avec les avancées de l'intelligence artificielle, garantir cet anonymat devient de plus en plus difficile. **Julien Hendrickx**, professeur à l'Ecole polytechnique de l'UCLouvain, Yves-Alexandre de Montjoye, ingénieur UCLouvain et professeur associé à l'Imperial College London et Luc Rocher, ex-doctorant UCLouvain, professeur à la Oxford University ont mis au point un **nouveau modèle mathématique** pour **mieux comprendre les risques posés par l'IA** et **aider les régulateurs** à protéger la vie privée des individus. Les résultats de cette étude sont publiés dans la prestigieuse revue scientifique *Nature Communications*.

Dans une précédente recherche (2019), ces scientifiques étaient déjà parvenus à démontrer la **facilité à réidentifier les personnes** prétendument anonymisées sur le web, sur base de quelques informations partielles (âge, code postal, genre). Ce travail avait révélé l'ampleur des risques liés à la diffusion de données sensibles, même après anonymisation.

Dans cette nouvelle étude, les chercheurs proposent un **modèle innovant** (baptisé *modèle de correction Pitman-Yor (PYC))* qui évalue les **performances des techniques d'identification** à grande échelle, dans différents contextes d'application et de comportement. **Julien Hendrickx**, co-auteur et professeur à l'UCLouvain, explique : « *notre nouvel outil s'appuie sur les statistiques bayésiennes pour **apprendre à quel point les individus sont similaires**, et extrapoler la précision de l'identification à des populations plus importantes, avec une **performance jusqu'à 10 fois supérieure** aux règles précédentes. Ce travail fournit, pour la première fois, un cadre scientifique robuste permettant d'évaluer les techniques d'identification, pour les données à grande échelle*. »

L'objectif de cette recherche ? Aider à **mieux comprendre les risques posés par l'IA** et permettre aux **régulateurs de mieux protéger la vie privée** des personnes. Même si des réglementations telles que le RGPD encadrent strictement l'utilisation et le partage des données personnelles, les données anonymisées échappent à ces restrictions. Il était donc essentiel de déterminer si des données sont réellement anonymes ou peuvent être réidentifiées, afin de contribuer au respect à la vie privée.

Exemples ? Dans des **études médicales**, l'outil, développé à l'UCLouvain, peut aider à déterminer si des **informations sur des patient·es** pourraient être utilisées pour retrouver leur identité, et ainsi contribuer à empêcher cette identification. Dans la vie quotidienne, ce nouvel outil permet

également de **surveiller** (et donc contrer) la précision des **codes publicitaires et des trackers invisibles** qui identifient les utilisateur·trices en ligne à partir de petits détails, comme le fuseau horaire ou les paramètres de navigateur, une technique appelée "empreinte digitale de l'appareil".

Les scientifiques expliquent enfin comment cette méthode pourrait aider les organisations à **trouver un meilleur équilibre entre les avantages des technologies d'IA et la nécessité de protéger les données personnelles** des individus, rendant les interactions quotidiennes avec la technologie plus sûres et plus sécurisées. « *Nous pensons que ce travail constitue une étape cruciale vers le développement de méthodes rigoureuses pour évaluer les risques posés par les techniques d'IA de plus en plus avancées et la nature de l'identification des traces humaines en ligne. Nous espérons que ce travail sera d'une grande aide aux scientifiques, responsables de la protection des données, membres de comités d'éthique et autres praticiens qui cherchent à équilibrer le partage de données pour la recherche et la protection de la vie privée des citoyen·nes*. »

Service presse presse@uclouvain.be | LLN et institutionnel : **Isabelle Decoster** +32 486 42 62 20 et **Céline Gransard** +32 479 92 16 65 | Bruxelles : **Estelle Toscanucci** +32 495 48 35 49 | Hainaut : **Ugo Petropoulos** +32 498 42 15 13