

**Aux représentantes
et représentants des médias**

COMMUNIQUÉ DE PRESSE

Les données de la COVID-19 sur une plateforme internationale

Neuchâtel, le 7 décembre 2020. **Présenter des données fiables et unifiées à l'échelle internationale sur l'évolution de la COVID-19 : tel est l'objectif de la plateforme COVID-19 Data Hub. Créée en avril dernier par Emanuele Guidotti, assistant doctorant à l'Institut d'analyse financière de l'UniNE, en collaboration avec David Ardia, professeur à HEC Montréal et ancien collaborateur de l'UniNE, elle fait appel à la coopération de la communauté scientifique des données partout dans le monde. A ce jour, les données ont été téléchargées 2,92 millions de fois.**

Accéder à des données internationales fiables sur la COVID-19, pas toujours facile. Emanuele Guidotti en a fait l'expérience en début d'année. « Travaillant pour un projet de recherche sur la COVID-19, j'ai réalisé la difficulté d'accéder aux données liées au virus. » En cause, l'hétérogénéité des sources : en fonction des pays, les informations sont divulguées dans des langages et formats différents. Pour les unifier, l'assistant doctorant a développé le tout premier prototype de la plateforme COVID-19 Data Hub au printemps. « Ce travail faisait partie initialement d'un document de recherche qui a été publié par la suite dans *Springer Nature* et présenté sur le site web du Joint Research Centre. » Grâce à la collaboration de David Ardia, la plateforme a reçu le soutien financier de l'Institut canadien pour la valorisation des données IVADO et de HEC Montréal.

Comme l'ont expliqué les deux chercheurs dans *Quartier Libre*, le journal indépendant des étudiant-e-s de l'Université de Montréal, « à chaque heure qui passe, les algorithmes de COVID-19 Data Hub téléchargent automatiquement les données les plus récentes des sources officielles de 193 pays, de 500 régions et d'environ 4000 villes ». « Une centaine de participant-e-s ont contribué à l'alimentation de la plateforme », précise Emanuele Guidotti. Les données sont vérifiées puis fusionnées dans une seule base de données, accessible gratuitement à tout le monde dans n'importe quel format. Depuis le lancement de la plateforme, les données ont été téléchargées 2,92 millions de fois, ce qui est « gigantesque », s'enthousiasment les deux chercheurs.

Le COVID-19 Data Hub va bientôt bénéficier du soutien financier du Consortium R, l'organisation mondiale qui promeut la croissance et le développement du logiciel R en tant que plateforme de premier plan pour la science des données et le calcul statistique. La maintenance du projet sera prise en charge par des professeur-e-s du Département de statistique et de biostatistique de la California State University, East Bay, avec qui Emanuele Guidotti et David Ardia ont commencé à travailler.

En savoir plus :

L'article: <https://bit.ly/3p9qGWx>

La plateforme: <https://covid19datahub.io>

Contacts :

*Emanuele Guidotti, assistant doctorant, Institut d'analyse financière,
Tél. +41 78 230 04 77 ; emanuele.guidotti@unine.ch*

*David Ardia, professeur à HEC Montréal,
Tél. +1 514 340-6103 ; david.ardia@hec.ca*