

Service communication et presse

450 route des Chappes – CS 50193

06904 Biot - Sophia Antipolis cedex

Tél. : +33 (0)4 93 00 81 21

Laurence.Grammare@eurecom.fr



Sophia Antipolis, FR, 13/03/2020

INFORMATION PRESSE

EURECOM met en place « CoronaCheck » : le premier système de vérification des faits pour les déclarations statistiques sur les coronavirus

Avec la propagation du virus COVID-19, la désinformation sur son origine, sa diffusion, ses effets, connaît actuellement un pic. Les affirmations erronées et trompeuses sont répandues sur les réseaux sociaux, mais aussi par les politiciens et les gouvernements, comme Trump qui affirmait que le taux de mortalité était plus faible que le taux effectif.

(https://en.wikipedia.org/wiki/Misinformation_related_to_the_2019%E2%80%9320_coronavirus_outbreak#United_States).

Étant donné le grand nombre de déclarations inexactes sur Internet, les vérificateurs de faits humains doivent les identifier, collecter les données pertinentes et effectuer les analyses nécessaires pour valider ou rejeter une déclaration. Malheureusement, étant donné la surcharge d'informations sur Internet, nous sommes maintenant confrontés à ce que l'O.M.S. appelle "Infodémie", qui rend impossible pour l'homme de vérifier manuellement chaque demande.

(<https://www.nytimes.com/2020/02/06/health/coronavirus-misinformation-social-media.html>).

L'équipe du professeur Papotti à EURECOM, France et du professeur Trummer à l'Université de Cornell USA, travaillent ensemble pour construire des systèmes qui vérifient automatiquement les allégations statistiques sur le coronavirus. Ce travail est rendu public aujourd'hui avec la publication de CoronaCheck (<https://coronacheck.eurecom.fr/>), un site web où, pour toute allégation statistique donnée, telle que "le 12 mars, les nouveaux cas n'ont augmenté que de 5% aux Etats-Unis", le système est capable de démontrer à partir des données officielles si elle est vraie ou fausse.

À partir d'une phrase, les systèmes utilisent des classificateurs de texte pour construire des fragments de requête qui sont ensuite exécutés sur les sources de données sous-jacentes. Le résultat des requêtes est utilisé pour étiqueter automatiquement la phrase comme correcte ou fausse. Le système prend déjà en charge une grande variété de demandes dans tous les pays touchés par le virus et

soutient les demandes concernant les cas avérés, les guérisons et les décès. En outre, le système apprend, grâce aux commentaires des utilisateurs, comment traiter les nouvelles demandes et exploiter les nouveaux ensembles de données. Les données sont collectées chaque jour à partir de sources officielles et les modèles d'IA qui soutiennent le système sont à la pointe du progrès en matière de compréhension du langage naturel et de génération de requêtes.

Les données sont collectées chaque jour auprès de sources officielles, notamment l'OMS, les CDC, l'ECDC, le NHC, le DXY, et les rapports des médias locaux.

"Nous invitons tout le monde à utiliser notre système pour augmenter le nombre de demandes que nous pouvons vérifier automatiquement. Notre objectif est de maintenir le système libre et ouvert afin d'aider les utilisateurs, tels que les journalistes, à vérifier les déclarations statistiques. Nous envisageons l'adoption de CoronaCheck par les réseaux sociaux pour identifier et vérifier automatiquement les déclarations diffusées en ligne et les soutenir dans la lutte contre la désinformation en ligne", a déclaré le professeur Paolo Papotti, chef du groupe "Qualité de l'information" au sein du département des sciences des données d'EURECOM.

L'initiative est soutenue par une Bourse de Recherche Google et par l'ANR. Pour plus d'informations, contactez : paolo.papotti@eurecom.fr ou communication@eurecom.fr

Paolo Papotti est Maître de Conférences à EURECOM, Sophia Antipolis, France, depuis 2017. Il a obtenu son doctorat à l'Université de Roma Tre (Italie) en 2007, et a occupé des postes de recherche à l'Institut de recherche informatique du Qatar (Qatar) et à l'Université d'État de l'Arizona (États-Unis). Ses recherches portent sur l'intégration des données et la qualité de l'information. Il est l'auteur de plus de 90 publications et son travail a été récompensé par deux citations "Best of the Conference" (SIGMOD 2009, VLDB 2016), un prix de la meilleure démo à SIGMOD 2015 et deux prix Google Faculty Research (2016, 2020). Il est rédacteur en chef adjoint du Journal of Data and Information Quality (JDIO) de l'ACM et des Actes du VLDB Endowment Journal (PVLDB).

EURECOM acteur clé de l'écosystème AI

EURECOM fait partie des grandes écoles d'ingénieurs françaises à vocation internationale c'est un Centre de Recherche renommé dans les sciences du numérique, à la pointe de la recherche et de l'innovation. Filiale du groupe Institut Mines Télécom, EURECOM est membre fondateur du Campus SophiaTech à Sophia Antipolis, le plus grand campus en Sciences et Technologies de l'Information des Alpes Maritimes et est membre du Pôle de Compétitivité Solutions Communicantes Sécurisées [SCS].

Les activités d'enseignement et de recherche sont organisées autour de 3 thématiques porteuses : sécurité numérique, systèmes de communications et Data Science.

Au travers de son département Data Science, EURECOM apporte notamment son expertise et sa réputation mondialement reconnues dans les domaines de l'apprentissage automatique, du traitement multimodal et de l'ingénierie des connaissances qui nourrissent les systèmes de l'intelligence artificielle, et se positionne comme un exemple à suivre pour l'interaction avec l'industrie.

Contact : communication@eurecom.fr, WWW.EURECOM.FR