



INVITATION PRESSE - PARIS – 4 JANVIER 2021

Invitation à la journée « Intelligence artificielle : l'ordinateur passe la barrière de la langue »

Mardi 12 janvier 2021
De 9h à 17h

Au siège du CNRS
3 rue Michel-Ange – 75016 Paris
(M° 9 ou 10, station Michel-Ange Auteuil)

Attention : places limitées, inscription obligatoire avant le 8 janvier auprès du bureau de presse du CNRS : presse@cnrs.fr

Les conférences seront également diffusées en ligne sur :
<https://webcast.in2p3.fr/live/gdr-tal>

Le traitement automatique des langues est une discipline de l'informatique mettant en œuvre des techniques et des outils de l'intelligence artificielle, conçus pour produire et traiter des données conséquentes autour des langues. Ce domaine scientifique a connu de nombreuses évolutions, principalement dues à la grande vague de l'apprentissage avec des réseaux de neurones profonds (ou *deep learning*). Fouille de données scientifiques sur le COVID-19, détection de sentiments dans des écrits et des "infox", diagnostic médical, traduction automatique, formation virtuelle, et passerelles avec la science-fiction, les scientifiques travaillant sur la langue écrite, orale ou signée du Groupement de recherche « Traitement automatique des langues », porté par le CNRS, vous proposent de suivre un cycle de conférences grand public, sur une journée, pour comprendre les défis scientifiques autour de ce domaine en mutation et en pleine expansion.



Inscriptions pour la presse :

RSVP avant le 8 janvier 2020 auprès du bureau de presse du CNRS : presse@cnrs.fr.

Attention :

En raison de la situation sanitaire, les places sont limitées. Le port du masque est obligatoire au siège du CNRS.

Par ailleurs, en raison du plan Vigipirate, une pièce d'identité vous sera demandée à l'entrée : CNI, passeport ou permis de conduire. **Les cartes de presse ne seront pas acceptées.**

Programme :

Introduction

Par Béatrice Daille, chercheuse de l'Université de Nantes au Laboratoire des sciences du numérique à Nantes (CNRS/École Centrale de Nantes/Université de Nantes/IMT Atlantique)
9h30

Fouille de publications scientifiques sur le COVID-19

Par Benoit Favre, chercheur d'Aix-Marseille Université au Laboratoire d'informatique et systèmes (CNRS/Aix-Marseille Université)
9h45 - 10h30

Au plus fort de l'épidémie de COVID-19, le *National Institutes of Health* (Etats-Unis) a recensé la publication ou pré-publication de 2 500 articles de recherche sur le sujet par semaine. Ces articles traitent de sujets divers comme les résultats de recherche fondamentaux sur les vaccins, les pratiques des services hospitaliers, ou encore l'impact du confinement sur notre société. Comment le corps médical peut-il « digérer » une telle quantité d'information afin d'en extraire des connaissances pertinentes pour la prise en compte au quotidien de l'épidémie ? Cette présentation montrera par quels moyens le traitement automatique du langage peut offrir une aide précieuse pour la veille scientifique, notamment grâce à des systèmes automatiques.

Ironie, sarcasme et sentiments : les algorithmes dans les subtilités de la langue

Par Farah Benamara, chercheuse de l'Université Toulouse III - Paul Sabatier à l'Institut de recherche en informatique de Toulouse (CNRS/Toulouse INP/Université Toulouse III Paul Sabatier)
10h30 – 11h15

Le langage évaluatif désigne toute expression du langage qui permet d'exprimer nos points de vue, opinions, émotions, souhaits à propos d'une personne, d'un objet ou d'un événement. Il est omniprésent dans les médias d'aujourd'hui (blogs, commentaires, forums, réseaux sociaux, etc.), ce qui rend crucial le développement d'outils pour extraire, synthétiser et comparer les



opinions exprimées par les internautes. Les applications sont multiples, telles que la prédiction des résultats de votes, l'analyse de la satisfaction clients ou encore l'aide à la modération de contenus. Dans ce contexte, l'exposé présentera les travaux sur la détection de l'ironie et du sarcasme ainsi que la détection de messages haineux.

Mon ordinateur est-il un bon psy ? Le traitement automatique du langage au service du diagnostic médical

Par Maxime Amblard, chercheur de l'Université de Lorraine au Laboratoire lorrain de recherche en informatique et ses applications (CNRS/Université de Lorraine/Inria)
11h30 – 12h15

Dans le secteur de la santé mentale, l'analyse automatique de la parole du patient peut jouer un rôle important. À l'interface de la psychiatrie-psychopathologie, de la linguistique, de la sémantique formelle et des sciences du numérique, les scientifiques étudient les troubles du langage et de la pensée, pour définir un modèle cognitif et sémantique des troubles du discours. L'intelligence artificielle et le traitement automatique des langues sont utilisés pour produire des indices proposés aux professionnels de santé pour la détection précoce, le diagnostic et le suivi des patients avec l'objectif, à terme, de proposer des solutions logicielles permettant d'aider les médecins dans leur travail quotidien.

La traduction automatique à l'heure des réseaux de neurones

Par François Yvon, chercheur du CNRS au Laboratoire interdisciplinaire des sciences du numérique (CNRS/Université Paris-Saclay)
12h15 – 13h00

La question de la traduction automatique n'est pas entièrement résolue. Cette intervention vise à faire un état des lieux concernant les derniers développements dans ce domaine, en particulier les avancées récentes obtenues grâce à l'adoption de nouvelles architectures d'apprentissage profond (ou *deep learning*). Les nouvelles thématiques qui émergent sous l'impulsion des grands acteurs de l'Internet, par exemple la traduction multilingue (depuis et vers de multiples langues), seront également discutées.

Le dialogue virtuel au service de la formation

Par Magalie Ochs, chercheuse d'Aix-Marseille Université au Laboratoire d'informatique et systèmes (CNRS/Aix-Marseille Université)
14h15 – 15h00

Aujourd'hui, les personnages virtuels sont de plus en plus utilisés dans le domaine de la formation. Utilisés pour la formation des compétences sociales des individus, ils permettent de les immerger dans un environnement de réalité virtuelle et de simuler une gestion d'équipe, un harcèlement scolaire ou des phobies sociales par exemple. Cet exposé présentera une plateforme de réalité virtuelle permettant aux médecins de simuler une situation d'annonce d'événements indésirables graves face à un patient virtuel autonome. A travers un casque de



réalité virtuelle, le médecin peut dialoguer avec une patiente virtuelle humanoïde réaliste (expressions d'émotion, gestes de douleur). D'autres recherches sont en cours pour la formation à la prise de parole en public face à une audience virtuelle et la sensibilisation aux discriminations sociales.

Quand la science-fiction inspire les technologies de la langue

Par Frédéric Landragin, chercheur du CNRS au laboratoire Langues, textes, traitements informatiques et cognition (CNRS/ENS Paris/Université Sorbonne Nouvelle)
15h00 – 15h45

L'ordinateur parlant HAL de *2001, l'Odyssée de l'espace* (Kubrick, 1968) a fasciné nombre de spectateurs, dont certains sont devenus par la suite chercheurs en intelligence artificielle. Une manière de boucler la boucle, car HAL, imaginé en collaboration avec des chercheurs à l'époque, préfigurait les chatbots et robots de compagnie que l'on voit se développer ces dernières années. Dans cette présentation, quelques inventions stimulantes de la science-fiction seront passées en revue, certaines bien connues comme le droïde de protocole C3PO dans *Star Wars*, d'autres moins connues mais tout aussi amusantes, comme le comprimé informatico-translateur à avaler¹, traducteur automatique universel d'une efficacité redoutable qui, au-delà de sa formule magique, pose des questions essentielles sur les enjeux de la traduction automatique et du traitement automatique des langues en général.

Les algorithmes à la chasse aux infos

Par Vincent Claveau, chercheur du CNRS à l'Institut des recherches en informatique et systèmes aléatoires (CNRS/ENS Rennes/Inria/INSA Rennes/IMT Atlantique/Université de Bretagne-Sud/Université de Rennes 1)
15h45 – 16h30

Les fausses informations sont très présentes sur le Web et les réseaux sociaux, avec des conséquences potentiellement graves pour notre société. D'un point de vue informatique, ces *fake news* peuvent prendre différentes formes, et leur détection nécessite alors de disposer d'outils adaptés à chacune de ces formes. Le traitement automatique de langues est utilisé pour la détection de certaines *fake news*, parfois à l'aide d'analyse des images et vidéos associées. Le dévoiement d'algorithmes d'apprentissage profond (*deep learning*) fait également émerger de nouvelles menaces, avec des contremesures à mettre en place.





Note

¹ Voir Les Voyages électriques d'Ijon Tichy de Stanislas Lem.

Contact

Presse CNRS | Alexiane Agullo | presse@cnrs.fr

