



I

Intelligence Artificielle (IA) et transparence : prédictibilité, sentiment de contrôle et confiance

Soutenance de thèse – PAGLIARI Marine

12 Décembre 2023 à 14h00

Salle Roland Garros, École de l'air et de l'espace, BA 701, Salon de Provence

Devant le jury composé de :

Pr. Bert TIMMERMANS, University of Aberdeen King's College, Ecosse, Rapporteur
Dr. Emilie CASPAR, Ghent University Universiteitstraat, Belgique, Rapporteuse
Pr. Elisabeth PACHERIE, Institut Jean Nicod, Paris, Examinatrice
Dr. Sophie BOUTON, Institut Pasteur, Paris, Examinatrice
Pr. Lionel BRUNEL, Université Paul-Valéry Montpellier, Montpellier, Examinateur
Pr. Valerian CHAMBON, Laboratoire Parole et Langage, Aix-en-Provence, Directeur de thèse
Dr. Bruno BERBERIAN, ICNA/ONERA, Salon de Provence, Co-encadrant de thèse

Résumé

L'émergence de systèmes autonomes a considérablement transformé notre manière d'interagir avec le monde qui nous entoure et les prévisions pour l'avenir indiquent que cette tendance va continuer à se développer, en particulier avec l'avènement de l'Intelligence Artificielle (IA). Les systèmes autonomes avec lesquels nous collaborons exigent une adaptation de nos compétences cognitives et comportementales afin de nous synchroniser efficacement avec eux. Par conséquent, il devient essentiel de comprendre comment notre cognition s'est adaptée, et continue de s'adapter, aux caractéristiques de cette automatisation accrue. L'une des questions centrales de cette thèse s'articule autour des fonctions cognitives mobilisées par les nouvelles formes d'interactions entre opérateurs humains et systèmes autonomes. Dans ce travail de thèse, nous nous sommes intéressés à l'influence de ces nouvelles formes d'interaction sur l'expérience subjective de contrôle des opérateurs humains, i.e. le sentiment d'agentivité. Le sentiment d'agentivité fait référence au sentiment de contrôler les actions que nous effectuons et les conséquences de ces actions dans le monde (Haggard & Chambon, 2012). L'objectif de cette thèse était de (i) caractériser l'impact de l'IA sur la cognition humaine, en particulier sur le sentiment d'agentivité des opérateurs humains, et de (ii) définir et implémenter de nouveaux formats d'« explication », afin de compenser l'influence de l'opacité des systèmes sur l'émergence du sentiment d'agentivité des opérateurs.

Pour ce faire, nous avons mené quatre études expérimentales. Nos résultats indiquent que plus l'autonomie de l'IA est élevée, plus le sentiment de contrôle et de responsabilité des opérateurs diminuent. La communication d'explications permet toutefois l'adaptation des opérateurs humains à la tâche en cours, ainsi que l'augmentation concomitante de leur sentiment d'agentivité (Études 1 et 2). En outre, nous montrons que lorsque l'IA fournit des explications, les opérateurs humains leur attribuent une responsabilité accrue (Étude 3), et que ces explications permettent d'accéder à des indices supplémentaires qui favorisent l'émergence d'un sentiment de contrôle plus fiable (Étude 4). Enfin, nos résultats soulignent l'importance de maintenir l'engagement de l'opérateur humain dans la tâche à réaliser, afin de préserver son sentiment de responsabilité dans l'interaction avec le système artificiel.

Cette thèse souligne l'importance d'intégrer la cognition humaine dans le développement des partenaires artificiels à venir. Les modèles de cognition humaine peuvent en effet contribuer à une caractérisation plus fine des informations sur lesquelles repose le développement d'interactions qualitatives, de manière à proposer des solutions qui garantissent une meilleure collaboration entre opérateurs humains et systèmes artificiels autonomes.

Mots clés

Intelligence Artificielle Explicable (XIA), Niveau d'autonomie, Sentiment d'agentivité, Responsabilité, Intentions

Vous êtes invité à rejoindre la web-conférence JITSY via le lien ci-dessous :

<https://rdv.onera.fr/SoutenanceMarinePAGLIARI>