

PROPOSITION DE SUJET DE THESE

Intitulé : Collaboration humain - système à base d'IA : comment ajuster son niveau de confiance pour optimiser les performances de l'équipe ?

Référence : **TIS-DTIS-2025-19**

(à rappeler dans toute correspondance)

Début de la thèse : Septembre/Octobre 2025

Date limite de candidature :

Mots clés

HAT, Confiance, XAI, SCAF

Profil et compétences recherchées

Grandes écoles, M2 Sciences Cognitives, IA ou SI
Connaissance de la démarche expérimentale
Programmation scientifique

Présentation du projet doctoral, contexte et objectif

Les systèmes à base d'IA deviennent de plus en plus présents dans notre quotidien et dans les activités professionnelles. Leur utilisation régulière alimente pour l'être humain un processus de construction d'une confiance dans le système qui peut modifier son comportement, que ce soit au niveau de sa supervision, de l'acceptabilité des propositions du système ou de la réalisation de sa propre tâche [1].

Dans le cadre du Système de Combat Aérien du Futur (SCAF) les pilotes du NGF (New Generation Fighter) vont devoir interagir avec différents systèmes à base d'IA possédant des niveaux d'autonomie potentiellement élevés (drones par exemple). L'opérateur restera décisionnaire des processus d'action et s'appuiera sur des informations provenant de ces différents systèmes à base d'IA dans un environnement complexe, dynamique et risqué. Une des problématiques sous-jacentes sera pour lui d'être à même d'évaluer la fiabilité des différentes informations qui vont guider ses choix. Or la montée de la menace cyber, de la guerre cognitive [2], et de la multiplication des sources d'information rend cette question de plus en plus ardue et critique. La question de la confiance dans ces sources d'information, de l'impact de cette confiance sur l'évaluation de la fiabilité des informations, et finalement de leurs contributions respectives à la décision finale de l'opérateur restent des points importants à comprendre pour mieux concevoir et mettre en œuvre ces systèmes complexes. C'est en effet la prise de décision du pilote du NGF au sein du collectif (agents artificiels et humains) qui restera un facteur clé pour la performance opérationnelle.

Des travaux de thèse récents menés au sein de l'équipe de recherche qui propose ce travail de thèse ([3], [4], [5]) ont exploré différentes composantes de cette problématique et s'appuient sur un état de l'art récent. Ces travaux expérimentaux, centrés sur le domaine militaire, ont permis de confirmer :

- que pour les opérateurs militaires la fiabilité perçue des agents avec qui ils coopèrent contribue fortement à la confiance qu'ils leur accordent [3], [5] ;
- que la confiance que l'opérateur accorde à un agent (humain ou artificiel) est un facteur qui modifie la manière dont il va ensuite le superviser et prendre en compte ses informations [3] ;
- qu'en cas d'aides multiples, la fiabilité perçue de l'agent et sa nature (humaine, artificielle plus ou moins évoluée) participent tous deux à la construction de la confiance dans ces agents [5] ;
- qu'en cas de données contradictoire entre les agents, ces confiances participent à l'arbitrage nécessaire pour la prise de décision de l'opérateur [5] ;
- que les interactions entre l'opérateur et le système d'aide vont modifier la confiance donnée au système [1] et permettent une co-construction de la décision [4].

Si les résultats expérimentaux récents de la littérature ont permis une meilleure compréhension des mécanismes liés à l'établissement de la confiance et à son impact sur le comportement de l'opérateur quand il travaille avec un agent artificiel à base d'IA, l'évolution de cette confiance et du comportement lors d'interactions répétées est moins modélisé et éprouvé expérimentalement. C'est pourtant un point important pour les opérateurs du SCAF qui vont interagir régulièrement avec certains systèmes militaires.

L'objectif de cette thèse est de contribuer à mieux comprendre et prendre en compte la manière dont la confiance dans les systèmes d'aides à base d'IA se construit, se maintient et vient modifier le comportement et les prises de décision au sein des équipes humains – agents à base d'IA. Elle visera à

identifier les leviers opérationnels utilisables, que ce soit au niveau de la conception des systèmes que de la doctrine ou de la formation, dans le cadre général du SCAF.

Les travaux seront réalisés au sein de l'équipe "Ingénierie Cognitive et Neurosciences Appliquées" de l'ONERA, sur la base aérienne de Salon de Provence, en collaboration étroite avec le CREA. Cette équipe pluridisciplinaire, dotée de moyens d'expérimentation avancés, utilise des concepts, modèles et outils issus des Sciences Cognitives pour la conception et l'évaluation des systèmes aéronautiques. Il est attendu de cette thèse un travail innovant qui comprendra une partie de modélisation et une partie expérimentale avec des opérateurs. Les grandes parties de l'étude seront donc :

Partie I : Etat de l'art

- a. Sur la confiance et en particulier sur son évolution au cours des interactions
- b. Sur l'évolution de la place des systèmes à base d'IA dans les opérations militaires
- c. Sur les notions acceptabilité et d'explicabilité (XIA) des systèmes à base d'IA mais aussi sur les moyens de reporter une fiabilité

Partie B : Modélisation et identification de principes

- a. Pour l'évolution de la confiance envers un système à base d'IA
- b. Pour la conception d'interfaces permettant une adaptation du niveau de confiance

Partie C : Expérimentation des principes pour des opérations de coopération humain – système à base d'IA en simulation avec des opérateurs.

Références

[1] Maille N, Amokrane-Ferka K, Leblanc B, Heulot N, (2024). Expérimentation de la confiance d'un utilisateur de système d'IA, Actes de la conférence CNIA 2024, La Rochelle, juillet 2024, pp 40-48.

[2] Luberrisse J, (2023). Cognitive Warfare in the Age of Unpeace – Strategies, defenses, and the new battlefield of the mind. Fortis Novum Mundum. ISBN 979-8-223-11789-6.

[3] Bollon Florent, (2020). Recognized Cyber Picture et confiance interpersonnelle : optimiser les activités collaboratives cyber militaires, (Doctoral dissertation, Université de Bordeaux)

[4] Metge Adrien (2022). Opérateurs et systèmes intelligents : se comprendre pour décider. Application à la supervision de drones (Doctoral dissertation, Université de Bordeaux).

[5] Muñoz Pauline (----) La confiance inter-agents comme outil de gestion de l'incertitude et levier de performance dans la décision en système sociotechnique complexe : le cas du SCAF (Doctorat en cours : soutenance prévue en décembre 2024)

Collaborations envisagées

CREA (co-encadrement avec Anne-Lise Marchand)

DGA Ev : département simulation sol

Laboratoire d'accueil à l'ONERA

Département : Traitement de l'information et Systèmes

Lieu (centre ONERA) : CSP

Contact : Nicolas Maille

Tél. : 04 90 17 01 31 Email : nicolas.maille@onera.fr

Directeur de thèse

Nom : Marchand Anne-Lise

Laboratoire : Crea (ecole de l'air)

Tél. :

Email : anne-lise.marchand@ecole-air.fr

Pour plus d'informations : <https://www.onera.fr/rejoindre-onera/la-formation-par-la-recherche>