

# CERTIFICATION DE L'IA : OBSTACLES ET PERSPECTIVES

DAMOUR, Mathieu, [mathieu.damour@scalian.com](mailto:mathieu.damour@scalian.com), SCALIAN, Le Haillan

**Résumé :** *Depuis l'avènement de l'IA, les industriels des secteurs de l'automobile, de l'aéronautique et du ferroviaire n'ont pas caché leur intérêt pour l'intégration d'algorithmes de machine et deep learning dans des systèmes critiques. Pour ce faire, il conviendra de vérifier au préalable que de tels algorithmes garantiront la sécurité des biens et des personnes environnant le système. La mission de certification du projet DEEL a pour objectif d'explorer les pistes qui mèneront à la certification de l'intelligence artificielle dans des systèmes critiques. Un an après le lancement de la mission, nous revenons sur l'état actuel de nos réflexions.*

**Mots clés :** intelligence artificielle, machine learning, DEEL, certification.

## 1. Introduction

Le projet DEEL (*Dependable Explainable Learning*), porté pour son volet français par l'Institut Interdisciplinaire d'Intelligence Artificielle ANITI de Toulouse, cherche à valoriser une IA douée d'explicabilité et dans laquelle il sera possible de se fier. La mission de certification du projet DEEL rassemble des universitaires, des chercheurs, des partenaires industriels et militaires pour imaginer la méthode par laquelle il sera un jour possible de montrer qu'il est fiable et sécurisé d'embarquer un ou plusieurs algorithmes d'intelligence artificielle dans des systèmes critiques (i.e véhicules autonomes).

## 2. Méthodologie

Il existe aujourd'hui des centaines de protocoles de certification. Les composants logiciels dans l'aéronautique sont certifiés par la DO178c, tandis que l'ISO26262 couvre les défaillances électriques et informatiques dans le secteur automobile. Pour autant, aucune méthodologie ne permet de rassurer un utilisateur de voiture autonome quant à la capacité du modèle à reconnaître un piéton...à temps. Les membres du groupe de certification se réunissent mensuellement pour confronter leur expertise, leur approche de la certification, leur connaissance de l'IA afin d'explorer les limites des processus actuels, et les défis spécifiques à l'utilisation des méthodes de machine et deep learning.

Appuyés par une équipe de recherche dédiée, nous cherchons à identifier les nouveaux indicateurs et processus qui apporteront aux utilisateurs et aux autorités les gages de confiance suffisant pour intégrer de l'intelligence artificielle dans leurs habitudes courantes.

## 3. Originalité / perspective

La certification de l'IA est un enjeu mondial, à l'origine de nombreux groupes de discussion et dont la réussite est indispensable à l'industrialisation massive de systèmes critiques embarquant de l'intelligence artificielle. Notre groupe a d'ores et déjà cerné les limites des processus historiques, reposant sur la validation des spécification d'un cahier des charges. Nous avons également établi les différences de besoins et d'acceptation propres aux secteurs d'activités. Par exemple, l'aviation et l'automobile, par la spécificité de leur domaine d'usage respectif, ne posent pas les mêmes contraintes sur l'intégration de l'IA.

Les méthodes (e.g SOTIF, Assurances cases, overarching properties) proposées par différents acteurs industriels sont éprouvées, discutées et repensées. C'est en conclusion de ce travail préliminaire d'un an que nous nous efforçons de dresser une liste de propriétés (High Level Properties) qui, si elles sont portées par un algorithme, sont susceptibles d'apporter un degré de confiance éventuellement suffisant pour permettre son déploiement dans un système critique.