

ENJEUX ÉTHIQUES DE L'INTELLIGENCE ARTIFICIELLE

FFCNCERCERXBD16

PRIX : 2 090 €

DURÉE : 2 JOURS

Pauses et déjeuners offerts



PANORAMA



ATELIER



FAISABLE À DISTANCE



RÉALISABLE EN ANGLAIS

PRÉSENTATION

IA générative, destruction d'emplois, robots tueurs, véhicules autonomes, justice prédictive : l'intelligence artificielle génère des angoisses et appelle à un encadrement réglementaire pour garantir son caractère éthique.

Les enjeux éthiques de l'IA concernent une multitude de concepts différents. Tantôt il est question de la discrimination raciale d'un logiciel de recrutement; tantôt de l'avenir de l'humanité face à des robots plus intelligents que les humains. Aussi passionnantes soient-elles, les discussions sur la singularité technologique et l'avenir de l'humanité restent éloignées de la réalité de l'IA d'aujourd'hui, même si l'avènement de modèles de très grande taille (du type ChatGPT) soulève de nouvelles inquiétudes.

OBJECTIFS

- Examiner les trois principaux problèmes éthiques et juridiques qui entourent les algorithmes : les biais algorithmiques, l'explicabilité et la responsabilité Identifier les enjeux sociétaux tels que l'impact de l'intelligence artificielle sur l'emploi, le copyright, et la démocratie
- Examiner les questions et les solutions techniques, de gouvernance et de régulation qui encadreront l'IA dans différents cas d'usage (santé, transport, finance, justice, défense, administration publique)
- Distinguer les risques et les mesures de protection pour participer à l'élaboration d'une étude d'impact IA éthique d'un système de gestion des risques requis par le futur règlement européen AI Act
- Identifier les outils indispensables à l'implémentation de modèles et d'architectures adaptés aux problèmes considérés

PROGRAMME

Introduction

- En quoi l'IA est-elle différente d'autres transformations technologiques ?
- Qu'entend-on par « éthique », et quelle est la différence entre l'éthique et le droit (RGPD, droits fondamentaux, devoir de loyauté, etc.)
- Rappel des principaux types d'IA : symbolique, apprentissage machine (machine learning), approches hybrides, leurs forces et faiblesses en matière éthique
- Les principaux accords et recommandations internationaux sur l'IA éthique
- Ligne directrice du groupe d'experts (HLEG) de la Commission Européenne
- Recommandation OCD
- Principes Asilomar

Intelligence artificielle et avenir du travail

- IA et « petites mains » qui étiquettent et nettoient les données
- Trappe Malthusienne, innovation et effet de l'IA sur l'emploi

Grands sujets pour l'IA éthique

- Biais algorithmiques
- Explicabilité des algorithmes
- Responsabilité
- Cas particulier des modèles de fondation (par ex. ChatGPT)

Cas d'usage

- Voiture autonome
- IA en tant qu'outil de recrutement
- Reconnaissance faciale dans les aéroports
- Détection de fraude ou de criminalité IA en tant qu'outil de recrutement

Comment concevoir une approche éthique « by design » ?

- Solutions technologiques pour intégrer l'équité dans l'algorithme
- Solutions de gouvernance
- Étude d'impact
- Analyse de risques
- Prise de responsabilité interne (accountability)
- Supervision d'autorités de régulation

Synthèse et conclusion

PUBLIC/PRÉREQUIS

Cette formation s'adresse autant aux ingénieurs et data scientists qu'aux responsables data, product owners, responsables innovation et transformation digitale et investisseurs qui seront impliqués dans les projets IA susceptibles d'avoir un impact sur les personnes ou sur la société.

Une formation de base sur la RGPD est souhaitable pour tirer pleinement profit de la formation.

RESPONSABLE(S)

Tiphaine VIARD

Maîtresse de conférences à Télécom Paris. Ses travaux de recherche portent sur les enjeux de l'IA, à travers son coût environnemental, son coût humain, et son impact sur la régulation du numérique.

MODALITÉS PÉDAGOGIQUES

Les enseignants privilégient une approche interactive (dialogue socratique) et les études de cas, avec un niveau de participation élevé des stagiaires. Les enseignements seront en français, mais l'anglais sera très présent.

