 ******  

**Colloque interdisciplinaire** **du mercredi 8 au vendredi 10 novembre 2023**

**Université d’Artois, Maison de la recherche, Arras (62), France**

# Décision humaine, décision de l’IA

# Analyses interdisciplinaires sur le contrôle humain des systèmes d’IA

**Responsable scientifique** : Nathalie Nevejans, Professeure en droit privé, Titulaire de la Chaire IA Responsable (ANR-19-CHIA-0008), Université d’Artois.

**Comité scientifique** : Pierre Marquis, Yves Poullet, Zied Bouraoui, David Doat, Tiago de Lima, Véronique Steyer, Louis Vuarin, Srdjan Vesic, Yannick Meneceur, Renaud Bougeard, Grégory Bonnet, Marion Ho-dac, Mathieu Hainselin.

**Langue de travail** : français.

**Inscription pour le public :** pas de frais.

🡪 Présentiel : inscription **obligatoire** par un mail à [sarah.robaszkiewicz@univ-artois.fr](mailto:sarah.robaszkiewicz@univ-artois.fr) (indiquez pour objet : *inscription au colloque sur l’IA.* Précisez dans votre mail le(s) jour(s) ou demi-journée(s) de présence).

🡪 Distanciel (pas d’inscription) : assistez au colloque, discutez et posez vos questions en vous connectant directement sur la chaîne Artois.TV (<https://artoistv.univ-artois.fr/live/event/0006-decision-humaine-decision-de-lia/>).

Les développements de l’IA permettent aujourd’hui de créer des systèmes capables de prendre des décisions dans les domaines les plus variés (banque, santé, justice, services, infrastructures, ...). Celles-ci peuvent avoir un impact sur les personnes utilisant, soumises ou orientées vers le système d’IA. Consciente des risques potentiels pour la santé, la sécurité, l’environnement et les droits fondamentaux, l’Union européenne travaille sur une proposition de règlement sur l’IA (*AI Act*) depuis 2021)[[1]](#footnote-2), et souhaite renforcer la maîtrise par l’humain des systèmes d’IA. L’*AI Act* consacre ainsi l’exigence d’un « contrôle humain » à l’article 14[[2]](#footnote-3).

Ce contrôle humain doit être spécifiquement interrogé dans la phase de la prise de décision – autonome ou non – par un système d’IA, car c’est elle qui cristallise toutes les interrogations. D’une part, l’IA pose encore un certain nombre de questions délicates, comme les biais algorithmiques, la boîte noire, ou l’explicabilité. D’autre part, comme l’humain peut être vu comme un censeur de la solution/décision du système d’IA, un « codécideur » avec lui, ou un simple « validateur » du résultat du système, les aspects humains de ce contrôle doivent aussi être questionnés pour tenir compte des limites proprement humaines, comme les biais cognitifs, l’interférence des émotions, ou les réactions inadaptées de l’humain en situation critique. La question centrale sera donc : quelle place accorder à l’humain dans le processus décisionnel dans lequel un système d’IA intervient en tout ou partie ?

## Mercredi 8 novembre

##### Accueil : à partir de 8 h 30

##### Mots de bienvenue (9 h 00 à 9 h 15)

#### Pasquale Mammone, Président de l’Université d’Artois.

#### Hugues Hellio, Professeur en droit public, Directeur du Centre Droit Ethique et Procédures (CDEP -UR 2471).

### Ouverture du colloque

*Toutes les discussions du colloque seront animées par Frédéric Boisdron, Animateur YouTube en partenariat avec la Chaire IA Responsable, Consultant en robotique, IA et nouvelles technologiques.*

#### **Introduction – Décision humaine, décision de l’IA : enjeux et défis interdisciplinaires à l’aube de l’adoption de l’AI Act** (9 h 15 - 9 h 30), Nathalie Nevejans, Professeure en droit privé et IA, Titulaire de la chaire ANR de recherche et d'enseignement « IA Responsable », Centre Droit Ethique et Procédures (CDEP - UR 2471), Université d’Artois

# PARTIE 1 – PRISES DE DECISIONS HUMAINE ET ALGORITHMIQUE

## I – Le concept de décision questionné par l’IA (9 h 30 – 10 h 45)

##### Introduction

#### **La théorie de la décision en informatique**, Alexis Tsoukiàs, Directeur de Recherche du CNRS en Informatique au LAMSADE, PSL, Université Paris Dauphine

#### **La théorie de la décision en philosophie**, David Doat, Maître de conférences en philosophie, Université catholique de Lille (Technologie, vulnérabilités, éthique)

#### **La théorie de la décision en sciences de gestion**, Véronique Steyer, Maîtresse de conférences en sciences de gestion, i3-CRG, Ecole polytechnique, et Louis Vuarin, Chercheur en post-doctorat en sciences de gestion à Telecom Paris (SES-I3) et Chargé d’enseignement à Polytechnique (CRG-I3)

##### 10 h 15 – 10 h 45 : Discussion

##### 10 h 45 – 11 h 15 : Pause-café

## II – Les processus et outils impliqués dans la prise de décision (11 h 15 – 12 h 45)

##### Introduction

#### **Neuroanatomie fonctionnelle de la prise de la décision**, Renaud Bougeard, Docteur en médecine, spécialiste en neurochirurgie, Expert judiciaire près la cour d’appel de Lyon, Doctorant en droit

#### **La prise de décision basée sur l’argumentation : principes, avantages et défis**, Srdjan Vesic, Docteur en informatique, Chargé de recherche CNRS en informatique, Centre de recherche en informatique de Lens (CRIL), Chaire IA Responsable, Université d’Artois

#### **La prise de décision basée sur le Machine Learning : incertitude, arbitrage et causalité,** Zied Bouraroui, Maître de conférences HDR en informatique, Centre de recherche en informatique de Lens (CRIL), Université d’Artois

##### 12 h 00 – 12 h 30 : Discussion

##### 12 h 30 : Déjeuner

## III – Les fondements sous-tendant les décisions (14 h 30 – 15 h 30)

##### Introduction

#### **La liberté de décider, entre droit et technique**, Liane Huttner, Maîtresse de conférences en droit du numérique, Paris Saclay

#### **Agents moraux artificiels - Décider sur et avec des valeurs morales**, Grégory Bonnet, Maître de conférences HDR en informatique, Université de Caen Normandie, Groupe de Recherche en Informatique, Image et Instrumentation de Caen (GREYC)

##### 15 h 00 – 15 h 30 : Discussion

## Synthèse de la journée (15 h 30 – 16 h 30)

##### 16 h 30 : Café et petit buffet

## Jeudi 9 novembre

##### Accueil : à partir de 8 h 30

# PARTIE 2 – CONTROLE HUMAIN DES DECISIONS DU SYSTEME D’IA

##### 9 h – 9 h 15 : Ouverture

## I – Les limites spécifiques de l’IA impactant la prise de décision (9 h 15 – 11 h 30)

##### Introduction

#### **Les limitations techniques du Machine Learning**, Mélanie Gornet, Ingénieure de l’ISAE-SUPAERO en sciences des données, Diplômée de SciencesPo Paris en relations internationales, Doctorante en droit à Télécom Paris, Institut Polytechnique de Paris

#### **Biais, discrimination et IA**, Eduard Fosch Villaronga, Maître de conférences en droit et robotique, Leiden University, Pays-Bas

#### **Petit état de l’art des problématiques liées au raisonnement causal en IA**, Tiago de Lima, Maître de conférences en informatique, Centre de recherche en informatique de Lens (CRIL), Chaire IA Responsable, Université d’Artois

##### 10 h 00 – 10 h 30 : Pause-café

#### **Enjeux et approches de l'IA explicable**, Pierre Marquis, Professeur en informatique, Directeur du Centre de recherche en informatique de Lens (CRIL - UMR CNRS 8188), Titulaire de la chaire ANR de recherche et d'enseignement « IA EXPEKCTATION », Université d’Artois

#### **Transparence et explicabilité**, Michael Lognoul, Assistant-doctorant et chercheur senior en droit à l’UNamur (CRIDS, NADI), Enseignant à la faculté de droit de l’UCLy et à l’Université d’Artois (DU Responsable de l’éthique de l’IA)

##### 11 h 00 – 11 h 30 : Discussion

##### 11 h 30 : Déjeuner

## II – Le contrôle humain de l’IA et sa mise en œuvre (13 h 30 – 16 h 00)

##### Introduction

#### **Sur la contrôlabilité et la dangerosité des IA**, Gérard Le Lann, Directeur de recherche émérite INRIA, Ingénieur ENSEEIHT, Docteur d’Etat en Informatique

1. **L’exigence légale de contrôle humain sur l’intelligence artificielle : dans quels desseins ?**, Noémi Bontridder, Chercheuse au Centre de recherche en information, droit et société (CRIDS) de l’Université de Namur, Membre du groupe de travail sur l’éthique de l’information du Programme Information Pour Tous de l’UNESCO, Enseignante vacataire à l’Université d’Artois (DU Responsable de l’éthique de l’IA) et à ISEN Lille – Junia.

#### ***Contrôle humain et « due process »****, Winston Maxwell, Directeur des Law & Technology Studies, JD (Cornell), Docteur en sciences économiques, HDR en droit, Ancien avocat New York et Paris, Télécom Paris*

##### 14 h 15 – 14 h 45 : Pause-café

#### **Le contrôle humain lors du déploiement l’IA à l'hôpital**, David Gruson, Professeur à la Chaire santé de Sciences Po Paris, Ancien élève de l’Ecole nationale d’administration et de l’Ecole des hautes études en santé publique, Directeur du Programme Santé du Groupe Luminess, Fondateur d'Ethik-IA

#### **La nécessité d’une évaluation et d’un contrôle collectif de l’IA**, Yves Poullet, Professeur émérite et Recteur honoraire de Faculté de droit de l'UNamur, Professeur associé à l’UC Lille, Ancien directeur du CRID, Membre de la chambre juridictionnelle de l’autorité belge de la protection des données, membre de la CCF d’Interpol, Expert auprès de l'UNESCO, Membre de l'Académie Royale des Sciences, des lettres et des beaux-arts de Belgique (Classe Technologies et Société)

#### **Contrôle humain de l’IA et normalisation**, Marion Ho-Dac, Professeure de droit privé, Centre Droit Ethique et Procédures (CDEP - UR 2471), Université d'Artois, Membre de l’AI Transparency Institute, et Baptiste Martinez, Docteur en droit privé, Ingénieur de recherche, Université d’Artois

##### 15 h 30 – 16 h 00 : Discussion

## Synthèse de la journée (16 h 00 – 17 h 00)

##### 17 h 00 : Café et petit buffet

## Vendredi 10 novembre

##### Accueil : à partir de 8 h 30

# PARTIE 2 – CONTROLE HUMAIN DES DECISIONS DU SYSTEME D’IA (suite et fin)

## III – Les facteurs humains de complexité (9 h 00 – 10 h 30)

##### Introduction

#### **Les biais humains et leurs conséquences au regard du contrôle humain**, Mathieu Hainselin, Maître de conférences HDR en psychologie expérimentale, Université de Picardie Jules Verne

#### **Décider en situation d’incertitude**, Albert Moukheiber, Docteur en neurosciences, Psychologue clinicien et Chargé de cours

#### **Évaluation de la capacité humaine à détecter une solution proposée par l’IA : une étude exploratoire**, José Samaniego Cho, Docteur en Psychologie Cognitive, Chercheur en post-doctorat au sein de la Chaire IA Responsable, Université d’Artois

#### **Analyse juridique de l’influence exercée sur l’humain par certains systèmes d’IA,** Alison Blondeau, Doctorante en droit de l’IA, Chargée d'enseignement à la Faculté de droit de Douai, Membre associé de la Chaire IA Responsable (équipe de recherche), Centre Droit Ethique et Procédures (CDEP - UR 2471 ), Université d’Artois

##### 10 h 00 – 10 h 30 : Discussion

##### 10 h 30 – 11 h 00 : Pause-café

## IV – Décision du SIA, recours juridiques et responsabilités (11 h 00 – 12 h 00)

##### Introduction

#### **La contestation humaine ex post des décisions totalement ou partiellement algorithmiques**, Pascal Alix, Avocat associé, Doctorant en droit de l’IA

#### **Quelle responsabilité civile pour l’IA : une responsabilité multi-régime ou une responsabilité intersectorielle ?**, Christian Byk, Secrétaire général de l’Association internationale droit, éthique et science, Chercheur associé à l’Institut des sciences juridique et philosophique de la Sorbonne (Paris I-CNRS), Représentant de la France au Comité intergouvernemental de bioéthique de l’UNESCO

#### **Perspectives européennes et internationales de la régulation de l’IA**, Yannick Meneceur, Conseil de l'Europe, Chef de la division centrale, Direction générale des droits humains et de l’État de droit, Maître de conférences associé en droit (sous réserve)

##### 11 h 45 – 12 h 00 : Discussion

## Synthèse de la matinée (12 h 00 – 12 h 30)

##### 12 h 30 : Fin du colloque pour le public – Cocktail

*Ce colloque a pour partenaires*

 ******

1. Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l’intelligence artificielle (Législation sur l’IA), 21 avril 2021, COM(2021) 206 final, 2021/0106 (COD) : https://eur-lex.europa.eu/legal-content/FR/ALL/?uri=CELEX:52021PC0206 [↑](#footnote-ref-2)
2. Dernières évolutions en date : Amendements du Parlement européen à l’*AI Act*, 14 juin 2023 : https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236\_FR.html [↑](#footnote-ref-3)