# Colloque RBP des 29-30 mars 2025. IA : avantages et défis

### Argumentaire général :

L’I.A. permet l’émergence de techniques assurément bénéfiques : faciliter le travail de beaucoup, améliorer la sécurité, aider les personnes handicapées, … etc. Cependant, ces intérêts vont de pair avec de fortes inquiétudes concernant, par exemple, les techniques de surveillance générale, la dépendance du travail humain aux algorithmes, la création de robots autonomes pouvant perpétrer des actes dangereux.

Pourra-t-on contrôler ces innovations techniques pour éviter l’asservissement de l’humanité ? Quels sont les critères qui permettraient de réguler l’utilisation des techniques issues de l’I.A. ? Comment éviter que les plateformes numériques pilotées par l’I.A. n’envahissent notre vie privée ? Comment éviter qu’elles ne deviennent des champs de bataille entre multinationales faisant beaucoup de victimes collatérales ? Que penser des prophéties faites par certains techno-progressistes concernant une I.A. qui surpasserait l’homme dans un bref avenir ? La théologie chrétienne est-elle une ressource pour penser l’humanité aux prises avec l’I.A. ? Ces questions seront l’enjeu du colloque proposé par le RBP les 29 et 30 mars 2025.

### Conférences

#### **L’impact de l’Intelligence artificielle sur la transformation des entreprises et de la société.** Yves CASEAU, Membre de l’académie des technologies

Cet exposé aborde la question de l’impact de l’Intelligence artificielle (IA), du point de vue de l’entreprise mais également du point de vue plus large de la société en s’intéressant aux questions éthiques soulevées par l’usage de l’IA. La première partie va définir ce qu’est l’intelligence artificielle telle que déployée dans les entreprises, les différentes formes qui sont mises en œuvre et les moyens de cette mise en œuvre. L’intelligence artificielle est multiple et les progrès récents de l’IA générative (que nous voyons avec Chat GPT) peuvent conduite à perdre de vue l’aspect “boîte à outils” et “combinaison hybride d’IA” qui sont essentielles dans les exemples les plus frappants de succès, tels que ceux de l’entreprise DeepMind (dont le fondateur Demis Hassabis vient de recevoir le prix Nobel de Chimie). La deuxième partie de l’exposé va s’intéresser aux questions éthiques de l’usage de l’IA selon les points de vue des différentes parties prenantes : l’entreprise, l’utilisateur et la société. Nous allons aborber les questions du contrôle, de la transparence et de l’auditabilité des systèmes logiciels fondés sur l’IA. Nous allons ensuite nous tourner vers des sujets propres à l’utilisateur : l’apprentissage et la coévolution (lorsque IA signifie “intelligence augmentée”), la crainte de perte d’emploi ou la perte de compétence face à un automate trop puissant. Nous terminerons cet exposé en abordant la question de l’impact de l’IA sur la planète, à la fois du point de vue du CO2 mais aussi de celui du renforcement des inégalités.

#### **« R****apports de l’IA avec la loi et le droit : h****ypothèse de la convergence générale dans l'ordinateur »** Arnaud BILLON, chercheur en éthique au Centre d’Etudes Avancées d’IBM-France et co-fondateur de « Technoréalisme »

L'histoire de la méthode des juristes montre comme un enfermement du dogme dans le code : ainsi le contrat bien écrit, ou la loi bien écrite (ceux qui répondent aux standards de la légistique) sont-ils des programmes logiciels ? On comprend que la modernisation de ce paradigme ne saurait avoir laissé indemnes les champs qui lui sont proches, comme la science ou l'économie politiques ; et sachant que ceci procède d'une reprise par la science juridique de la méthode scientifique, on se demande si l'on n'aurait pas affaire à une convergence plus générale, comme un paradigme implicite généralisé « si seulement nous vivions déjà dans un grand ordinateur ». La machine à gouverner historiquement se met en place, nous donnerons des raisons de le penser en retraçant l'histoire de la constitution de différentes machines ; et en interrogeant aussi les raisons du silence autour de cette progression : comment la pensée critique ne peut-elle pas être plus sensible à l'entrée dans la machine ? La machinerie spirituelle n'aurait-elle pas conditionner ce déport qui nous semble inéluctable ?

#### **Machines intelligentes ? Rôle et place dans le salut chrétien ?** Jean-Marc MOSCHETTA, enseignant-chercheur à SUPAERO et chercheur à l'Institut Catholique de Toulouse

Avec l'intelligence artificielle générative, nous faisons désormais face à l'émergence d'une intelligence non biologique qu'il ne suffit plus de considérer comme une simple simulation ou un perroquet du langage. Le fait que les rouages algorithmiques utilisés dans les systèmes d'intelligence artificielle ne pensent pas en eux-mêmes ne les disqualifient pas pour autant, car ce qui importe, c'est la relation que nous entretenons avec eux et la force de proposition qu'ils représentent pour résoudre nos problèmes. La vexation que leur performance représente aujourd'hui pour l'intelligence humaine tient en partie à ce que ces systèmes insinuent que l'intelligence est une construction, qu'elle est "artificielle par nature", n'ayant aucun besoin pour se construire du concept d'homme. Sur le plan théologique, cela repose la question du statut spécifique de l'humanité dans le plan de salut de Dieu, du regard proprement chrétien que nous posons sur l'intelligence biologique ou machinique et enfin du rôle que nous devons faire jouer à ces systèmes pour participer à "l'eucharistisation du cosmos". Nous essaierons, à distance de la fascination technolâtre, de mettre en évidence qu'à travers les réels problèmes sociétaux liés à l'usage immodéré de l'intelligence artificielle, ce sont nos relations entre humains, entre l'homme et la terre, entre l'homme et Dieu qui doivent attirer notre attention et faire l'objet du travail de l'Esprit Saint, travail qui inclut désormais nos machines à penser. Tout dépendra du fait qu'avec l'intelligence artificielle, nous nous laissions transformer par Dieu pour lui devenir semblables.

### Ateliers-carrefours

#### **Les corrélats neurobiologiques de la conscience : NBIC, neurotechnologies et neuroéthique** Laure TABOUY neuroscientifique et éthicienne à Paris-Saclay

Comment à travers la question de la conscience, la convergence NBIC (neuroscience, de l’informatique, de l’ingénierie et des technologies) et les neurotechnologies interrogent notre humanité et le sens d'être un être humain ? Cet atelier a pour but d’aborder les questions et les dilemmes éthiques qui émergent de la convergence NBIC et des neurotechnologies, à travers la question de la conscience. L’idée est de faire émerger, par le dialogue et la controverse différents points de vue en s’intéressant aux questions éthiques soulevées par l’usage des neurotechnologies connectées à l’IA ayant pour objectif de comprendre la conscience.

Les recherches, les progrès et les innovations en neurosciences et en IA et autour de la convergence neurosciences / informatique / ingénierie / NBIC remettent en question l’approche et la conception matérialiste, techno-centrée, et déterministe des neurosciences, de la pensée et de la conscience. La première partie va définir ce que sont la convergence NBIC et les neurotechnologies, telle qu’elles sont déployées dans les entreprises et les laboratoires de recherches, les différentes neurotechnologies qui sont mises en œuvre et les moyens de cette mise en œuvre. Les neurotechnologies sont des dispositifs utilisés pour étudier la structure et le fonctionnement du cerveau. La convergence des neurosciences, de l'informatique et de l'ingénierie, ainsi que leur sophistication croissante sur le marché mondial, les rendent plus miniaturisés, plus efficaces et plus puissants. Cela accentue la porosité des frontières entre les usages médicaux et non médicaux, civils et militaires, ainsi que la diversité des objectifs, des utilisations et des investissements. Elles sont porteuses d'espoir, mais leur capacité d'influence ou de manipulation est inquiétante.

Considérant l'éthique comme une critique constructive visant à protéger notre avenir de nos actions présentes, et sur la base de la méthode Design Fiction, la deuxième partie de l’atelier consiste à repenser l’éthique à la lumière de la convergence des neurosciences, de l'informatique et de l'ingénierie, en développant une « *éthique par conception* » basée sur le terrain, calibrée et appliquée à ces neurotechnologies. Nous terminerons cet atelier en abordant la question de l’impact des neurotechnologies sur nos relations mais aussi la planète, à la fois du point de vue énergétique mais aussi de celui de la justice sociale.

#### **Être humain au temps de l’IA : quelle place pour l’expérience vécue ?** Mathieu GUILLERMIN, philosophe Université Catholique de Lyon, Unité de recherche Sciences et Humanités,

Les progrès des technologies d’intelligence artificielle ont été considérables depuis une quinzaine d’années. Les machines sont de plus en plus capables d’imiter l’humain. La question de la spécificité de l’humain et des limites à tracer quant aux tâches qu’il est acceptable de déléguer aux machines se pose avec une insistance croissante. De ce point de vue, il semble particulièrement important de ne pas s’intéresser qu’aux résultats et produits finis (des humains et des machines), mais aussi à ce qui se joue en amont, avec en particulier la question de l’expérience vécue, de la conscience phénoménale. Dans certains cas, ce qui est vécu semble essentiel (expression de compassion, expérience d’apprentissage, création artistique …). Mais quel est le statut de nos discours à propos de l’expérience et de la conscience phénoménale ? Comment en parle-t-on « sérieusement » ou « scientifiquement » ? Une question épistémologique lourde de conséquences éthiques.

#### **L'IA peut-elle faire des mathématiques ?** Clément DENIS, docteur en mathématiques à l'Université Aix-Marseille, «

Pour étudier les conséquences sociales, philosophiques et théologiques de l'intelligence artificielle, il est nécessaire de commencer par bien comprendre ce qu'elle peut – et pourra – faire. C'est l'objet de cet atelier, qui aborde la question des capacités de l'IA en mathématiques, et de façon plus générale en sciences.

Depuis la publication en 1976 de la première preuve véritablement assistée par ordinateur, plusieurs techniques ont été développées pour aider les mathématiciens : simulations numériques et logiciels de calcul formel, puis plus récemment assistants de preuve tels que le pionnier Coq ou Lean. Ces logiciels fournissent une aide précieuse, voire indispensable, mais ne peuvent proposer de solutions créatives. Depuis quelques années, les IA génératives offrent des perspectives nouvelles. Depuis le lancement de GPT-3 en 2020, les LLM (*Large Language Models*) sont en effet passés des multiplications à deux chiffres aux problèmes d'Olympiades de Mathématiques Internationales, et progressent régulièrement sur les bancs d'essai tels que *FrontierMaths*.

La question de la capacité de ces nouvelles IAs à produire des mathématiques en situation réelle peut se poser à différents niveaux : rédaction de solutions d'exercices de lycée, résolution de problèmes de licence, ou contribution à des articles de recherche. Nous présenterons plusieurs exemples d'applications à des niveaux variés, du lycée à la recherche – avec des résultats déjà impressionnants – et nous inviterons les participants à partager leurs expériences concernant l'utilisation de ces modèles.

Nous aborderons également la question de leurs limites : le problème des hallucinations et de la fiabilité des réponses apportées d'une part, et la capacité à résoudre des problèmes absents du corpus d'entraînement d'autre part. Enfin, nous proposerons quelques perspectives sur l'évolution des IAs dans les prochaines années.

#### **IA, autonomisation des machines et théologie** Philippe GAGNON, chercheur à l'Université Catholique de Lille, Chaire Sciences et Technosciences à l'heure de l'écologie intégrale

Les discours entendus face à la montée de l’intelligence artificielle et de ses usages, ainsi que les préoccupations qu’ils révèlent, nous orientent spontanément vers une question relative à des performances cognitives implémentées dans des machines qui les ferait aussi intelligentes que nous, si ce n’est nous dépassant. Nous voyons donc une sorte de compétition à l’œuvre, qui se ressource à des présupposés à l’égard de la performance et d’une certaine possession de la puissance. Après avoir exploré les limites de ce scénario, nous suggérerons qu’une herméneutique de la nature peut se déceler dans le projet même de la science. Mais surtout, nous voudrions montrer que, théologiquement, un déplacement est sans doute nécessaire qui nous fasse apprécier davantage un vieux problème, très peu discuté sur le sujet, qui voudrait que l’homme, à l’égal de Dieu, soit capable de reproduire son image. Ce sont les conséquences de ce mythe ou obéit ou enfreint qui seront déterminantes au terme.