

IA HYBRIDE : *Quelles voies et quels potentiels après ChatGPT ?*

Jeudi 18 janvier 2024



Coordination scientifique :

David CORTES (AVIDENCE)

Laurent GOUZÈNES (KM2-CONSEIL)

Bernard MONNIER (MIM)

Guy-Alain NARBONI (IMPLEXE)

Jean ROHMER (Institut Frédrik Bull)



Renseignements, programme... :





**Compte-rendu du séminaire
du 18 janvier 2024**

**« IA HYBRIDE : Quelles voies et quels
potentiels après ChatGPT ? »**

Coordination scientifique :

**David Cortes, Laurent Gouzènes,
Bernard Monnier, Guy-Alain Narboni, et Jean Rohmer**

Table des matières

1/	<i>IA Hybride : quelques questions à se poser et quelques vœux</i>	3
	<i>État des lieux en IA : Symbolique / Connexionnisme / Hybridation / Interaction homme machine</i>	4
1/	<i>Formalismes logiques en IA : quelques exemples</i>	4
2/	<i>Leveraging Symbolic AI for XAI Purposes</i>	6
3/	<i>Comment faire le lien entre intelligence artificielle et créativité ?</i>	6
4/	<i>Quelques exemples d'hybridation connaissances et apprentissage pour l'interprétation d'images</i>	7
	<i>Hybridation symbolique / connexionnisme</i>	9
1/	<i>Améliorer les systèmes de classification neuronaux avec de la connaissance logique a priori</i>	9
2/	<i>XTRACTIS®, une IA Raisonnante Générale pour Découvrir des Connaissances Prédicatives Robustes</i>	10
3/	<i>IA Raisonnante Souveraine dans les Systèmes Critiques</i>	11
4/	<i>IA hybride dans la Défense et la Sécurité</i>	12
5/	<i>Expliquer les IA : de l'IA statistique à l'extraction de connaissance via une approche « régionale »</i>	13
6/	<i>Modèles NLP d'extraction d'information</i>	13
	<i>Les jeunes talents : Mon CV en deux minutes</i>	14
	<i>Hybridation homme/machine</i>	14
1/	<i>Borges and AI</i>	14
2/	<i>IA pour l'aviation de combat/hybridation homme/machine</i>	16
3/	<i>Des IAs et des Hommes – Pour une approche behavioriste du couple humain(s)-algorithme(s) dans les applications critiques</i>	16
	<i>Lien association X-IA</i>	18
	<i>Présentation de startups : ma startup</i>	18
1/	<i>Learning Robots : IA générative et éducation</i>	18
2/	<i>Muvraline : Green for AI and AI for green, with a hybrid AI. Telcos example</i>	18
3/	<i>DescartesMaus.ai: Dérisquer la décision dans un monde incertain grâce à l'IA</i>	19
4/	<i>Corolair : Former avec l'IA</i>	19
	<i>DEBAT sur l'IA hybride</i>	20

1/ IA Hybride : quelques questions à se poser et quelques vœux

[Vidéo ici.](#)

Un mot d'introduction et de bienvenu est réalisé par Bernard Monnier, Laurent Gouzènes et Jean Rohmer. Puis Bernard Monnier présente l'association Aristote et le sujet. Le terme IA a été introduit en 1956, et concrètement, il existe deux types d'IA : la connexionniste et la symbolique. Et aujourd'hui, le but est de lier les deux. « Au départ, l'IA qui existait n'était que symbolique, car on n'avait ni la donnée ni la puissance de calcul, puis la connexionniste a pu se développer, grâce à l'arrivée du Big Data, et du HPC », indique Bernard Monnier. Elle a eu des performances, en 2015, qui ont fait qu'elle est devenue opérationnelle, avec moins de 5% d'erreurs sur la reconnaissance de texte et la reconnaissance d'image. Après l'arrivée de ChatGPT 3 a fait qu'on s'est beaucoup amusé, notamment avec ses erreurs. « Mais ChatGPT 4, bien plus performant, ajouté au fait qu'OpenAI ne communique plus, et que le modèle est payant, nous fait penser que le modèle derrière, est un modèle hybride », ajoute-t-il. Rien n'est sûr, pour le moment... Puis il remercie toutes les personnes qui ont participé à l'élaboration du séminaire.

Puis vient Jean Rohmer ([ressources ici](#))

Il introduit plusieurs citations connues sur l'IA :

- « Comme tout bon informaticien, j'ai horreur des nombres », Alain Colmerauer
- « Everytime I fire a linguist, performance of the system improves », Frederic Jelinek
- « En matière de traduction automatique, tous les problèmes sont résolus », Bernard Vauquois (1967)
- « Deep Learning is going to be able to do everything », Geoff Hinton
- « Plus c'est explicable, moins c'est efficace », Yann Le Cun
- « Good linguists would save tons of CO2 », JR

Selon lui, l'IA connexionniste ridiculise l'IA symbolique depuis l'arrivée de ChatGPT, car elle a réussi là où l'IA symbolique a toujours « ramé », de manière surprenante (en passant par les nombres, en transformant les textes en matrices, en jetant les données après utilisation...) et même si des critiques apparaissent nettement (l'indiscipline de l'IA connexionniste, la question écologique...). Mais son ascension a permis de parler d'IA, alors qu'au début des années 2000, personne n'osait plus vraiment indiquer « spécialiste en IA » sur son CV.

Il revient ensuite sur la notion d'hybridation, qui peut être réalisée de différentes manières, et autant amener au meilleur comme au pire des deux mondes. Au bout du compte, trouvera-t-on dans le domaine une foultitude d'hybridation possible, ou un concept qui prendra le dessus sur tous les autres ?

Il classe alors dans un tableau les différentes oppositions entre les deux types d'IA, qui se différencient de deux manières différentes : le vécu et la pensée, les chiffres et les lettres.

Voici le tableau qui résume ces oppositions :

LE VRAI CHAMP DE BATAILLE IAS: IA SYMBOLIQUE / IAC: IA CONNEXIONNISTE

Source des données Type de traitement	DU VECU ! DONNEES APPRENTISSAGE	DE LA PENSEE ! STRUCTURES, SCIENCE LOIS, MODELES
DES CHIFFRES ! RESEAU DE NEURONES CALCUL ARITHMETIQUE	IAC CHATGPT TRADUCTION AUTOMATIQUE MODERNE	SIMULATION NUMERIQUE CAO CALCUL SCIENTIFIQUE
DES LETTRES ! LOGIQUE RAISONNEMENT	MOTEUR DE RECHERCHE RAISONNEMENT ANALOGIQUE RAISONNEMENT A BASE DE CAS	IAS SYSTEME EXPERT PROLOG TRADUCTION AUTOMATIQUE ANCIENNE

Journée ARISTOTE IA Hybride 18 Janvier 2024

12

Enfin, il liste les différents obstacles à l'hybridation, notamment le fait que les deux systèmes sont « orthogonaux » et intrinsèquement très différents, en insistant particulièrement sur le fait que l'IA connexionniste gère en grande partie des adresses de données dans un réseau de neurones, quand l'IA symbolique gère des valeurs de données.

Il évoque ensuite les différences d'infrastructures nécessaires à chaque type. Et il affirme qu'il n'est pas évident que les GPU soient les systèmes les plus adaptés pour l'IA hybride, car les deux types d'IA ne reposent pas du tout sur les mêmes principes de base au niveau informatique.

Il conclut notamment sur ses vœux pour l'IA hybride en 2024 et l'avenir du domaine. Ce qu'il souhaite avant tout, c'est que le secteur « converge » et ne passe pas « de deux camps à 50 chapelles » afin de ne pas multiplier les points de vue et complexifier encore tout cela.

État des lieux en IA : Symbolique / Connexionnisme / Hybridation / Interaction homme machine

1/ Formalismes logiques en IA : quelques exemples

Par Stéphane Demri, LMF – CNRS

[Vidéo ici](#)

[Ressources ici](#)

Stéphane Demri revient sur les différentes logiques en œuvre dans l'IA symbolique. Car c'est un champ large de la recherche sous représenté dans les médias, selon lui. Il décrit l'IA Symbolique comme le fait de représenter du savoir par des symboles et d'établir des règles pour manipuler ces symboles. Il décrit les actes de séminaire de la conférence KR23, la vingtième International Conference on Principles of Knowledge Representation and Reasoning, qui s'est tenue à Rhodes, en Grèce, dans lesquels on

retrouve des thèmes comme la description logique, les agents multi-systèmes, ou les stratégies de raisonnement.

Il présente ensuite une première classe de logique importante à ses yeux : le raisonnement par ontologies, qui introduit des concepts, des classes, des objets, des relations entre les objets et les classes etc. Afin de spécifier un domaine particulier et pouvoir faire de la classification. De relier les objets, les concepts, ou voir quels sont les axiomes vérifiés dans cette ontologie. Ce type de raisonnement est très utilisé dans le domaine médical où on souhaite classifier des individus, des maladies, etc. On peut alors s'interroger sur ces ontologies : comment les définir, les spécifier, raisonner avec, mais aussi des questions plus larges, comme comment faire si ces ontologies sont inconsistantes, ou qui doivent être réparées, comme ajouter des nouveaux concepts sans modifier toutes les anciennes inférences ? Et plus largement, quelle nouvelle connaissance peut-on tirer de cette ontologie qui n'a pas été préalablement spécifié ?

Il revient ensuite sur les logiques de description, qui sont des langages formels pour définir des ontologies. Définir des notions de classes, des relations entre les objets... Chercher à améliorer l'expressivité des langages, tout en restant décidable etc. C'est ainsi qu'il existe beaucoup de travaux sur les logiques de description. Il donne ensuite un exemple formel avant de passer à la description logique avec des domaines concrets, qui permet d'ajouter des données avec des propriétés concrètes aux ontologies. Cela complique énormément les modèles, car cela revient à travailler sur des domaines qui sont potentiellement infinis. Stéphane Demri décrit alors plusieurs approches pour réussir à traiter ces domaines concrets. Lui, travaille sur la dernière approche, avec des automates.

Methods for Handling Concrete Domains

- ▶ Tableaux-based **calculi** for " ω -admissible" domains. [Lutz & Miličić, JAR 2007]
- ▶ Translation into decidable **logics** [Carapelle & Turhan, ECAI'16]
 - Decidability of concept satisfiability problem w.r.t. general TBoxes for $\mathcal{ALC}(\mathbb{N}, <, =, (=)_n)_{n \in \mathbb{N}}$.
- ▶ Translation into **automata**-based problems.
 - Concept satisfiability problem w.r.t. general TBoxes for $\mathcal{ALC}(\mathbb{N}, <, =, (=)_n)_{n \in \mathbb{N}}$ in EXPTIME. [Labai & Ortiz & Šimkus, KR'20]
 - Concept satisfiability problem w.r.t. general TBoxes for $\mathcal{ALC}(\{0, 1\}^*, <_{pre})$ in EXPTIME. [Demri & Quaas, JELIA'23]

(constraint automata for data words or data tree)

Il passe ensuite à la deuxième partie de son exposé, des logiques qui permettent de décrire les capacités stratégiques de coalition d'agents dans un système multi-agents, et notamment une, la méthode ATL pour Alternating-time temporal logic, qui peut être appliqué à des comportements de robot, par exemple. C'est une logique qui va exprimer dans le langage lui-même, l'existence de stratégies pour une coalition, vérifiant un objectif fixe.

Il conclut en posant trois questions :

- Peut-on utiliser des formalismes appliqués à la représentation de connaissances pour faire autre chose, par exemple, de l'apprentissage ?
- De nouvelles théories arithmétiques doivent-elles apparaître pour la vérification de réseaux de neurones ?
- Enfin, alors que de nouveaux problèmes logiques émergent, liés à l'apprentissage, en trouvera-t-on d'autres ?

2/ Leveraging Symbolic AI for XAI Purposes

Par Pierre Marquis, CRIL-CNRS

[Vidéo ici](#)

[Ressources ici](#)

Pierre Marquis est spécialisé en IA symbolique, et il travaille dans ce domaine pour améliorer l'explication de l'IA hybride. Il revient tout d'abord sur le besoin d'hybridation à l'heure actuelle. Si le Deep Machine Learning (ML) est à la base de la révolution de l'IA depuis 10 ans, il insiste sur le fait que le Deep ML n'est pas pareil que le ML, et qu'il existe également du ML symbolique, depuis des décennies, et que c'est toujours un domaine de recherche très actif.

Selon Pierre Marquis, l'hybridation est nécessaire car aucun des deux types d'IA n'est suffisant. L'IA symbolique, par exemple, est limitée. Elle s'avère efficace dans le raisonnement, mais nettement moins dans les tâches de perception. En revanche, l'apprentissage profond (Deep ML) est efficace pour percevoir, reconnaître, classifier des données... Mais il l'est nettement moins dans les tâches de raisonnement ou pour générer du savoir transférable. Il n'y a aucune assurance que les prédictions soient correctes à 100%, et il reste très sensible aux données de départ. Le DML reste un ensemble de boîtes noires, opaques, et cela manque de sens commun. En outre, il peut être « trompé », très facilement. L'IA hybride consistera donc à prendre le meilleur des deux. Car une IA « trustworthy », de confiance, est obligatoire pour les applications à haut risque où à niveau de sécurité critique comme dans le domaine militaire. Il revient ensuite sur les notions liées à la compréhension des IA. Travailler avec une IA dans le but qu'elle devienne « trustworthy » fait appel à de nombreux concepts : l'interprétabilité, l'explicabilité, la transparence, la confidentialité, la fairness, la reliability, la safety... Mais cela devient primordial au-delà du domaine scientifique car « expliquer les décisions prises par un modèle est devenu une question juridique. » C'est ainsi que se développent les Xplanable IA, où l'on se concentre sur l'interprétabilité et l'explicabilité. Il détaille ensuite les propriétés à expliquer, puis revient sur particulièrement l'enjeu des données dans les modèles de machine learning (ML), en différenciant les données subsymboliques (comme un pixel) et les données symboliques (une table de données, une valeur, un attribut...)

3/ Comment faire le lien entre intelligence artificielle et créativité ?

Par Wendy Mackay, DataIA, Inria

[Vidéo ici](#)

Spécialiste des interactions humain-machine, Wendy Mackay parlera essentiellement de la créativité, et du rôle de l'IA dans ce domaine. Son métier consiste à aider les utilisateurs de différentes technologies et qui les poussent dans leurs limites. Elle travaille avec des professionnels de la créativité, afin de les aider à faire ce qu'ils désirent dans leur domaine. Son expertise est « l'interaction » perçue comme un phénomène à part entière, puis elle exécute des va-et-vient entre la théorie et l'observation de ce dit phénomène. Selon elle, et en suivant le principe des « sciences de l'artificiel », on ne peut pas étudier les choses que l'on a nous-mêmes créées. Car nous sommes confrontés à nos propres biais dans l'étude. Il y a entre la théorie et l'expérience, un « Artefact », qui est le fruit d'un design. Et il est très difficile de le contourner. On ne peut donc pas prétendre à une compréhension absolue du phénomène en lui-même, et nous sommes contraints à faire des aller et retours entre tous ses éléments.

Elle détaille alors la théorie générative de l'interaction, un concept qui a été publié récemment. Elle souligne la différence de la génération avec le terme « générative » utilisé en IA, car elle, parle de la génération de choses qui n'ont pas existé préalablement. Mais cette théorie ne suffit pas. Il va falloir aboutir à la création de concepts, pertinents pour la conception de systèmes interactifs. Ils sont tirés de la théorie vérifiée dans le monde réel. A partir de ces concepts, on va déduire des principes génératifs (actionnable), implémentable. Avec eux, on va pouvoir analyser les « artefacts » existant au prisme de ces concepts. Cette observation peut se réaliser via un prisme analytique, et un prisme critique : est ce que c'est une bonne chose. Que l'on va ensuite étendre aux nouveaux artefacts à étudier et vérifier ce qui est faisable ou non. Par exemple, cela permet de s'interroger sur les effets court-termes ou long-termes de certaines interactions. « Comme le like Facebook, ou sur un réseau social, qui à court-terme peut être bénéfique, mais à long-terme, créé une addiction pour l'utilisateur », illustre-t-elle.

Elle détaille ensuite les liens entre l'IA et l'approche d'analyse d'interaction Humain-Machine (IHM) et revient notamment sur la boucle Humain-Machine, toujours présente dans tous les systèmes. Ce qui posera in fine un problème si on mesure l'efficacité du modèle sans analyser l'impact du dit système sur l'être humain. « Par exemple, les humains peuvent n'être vus que comme des sources de données ou des sources d'erreurs pour l'algorithme, et ce la pose un problème. Car il faut mesurer également et ne pas négliger l'effet de l'algorithme sur l'humain. C'est ainsi que le système sera plus avantageux si on perçoit l'algorithme comme une source d'information pour les utilisateurs. » détaille-t-elle. En cela, la question à se poser au départ n'est pas la même : l'algorithme est-il plus performant que l'humain ou ce dernier est-il mieux servi par l'algorithme ? La question posée définit l'approche que l'on a sur le système.

C'est ainsi qu'elle arrive à la notion de « partenariat » avec l'exemple du déplacement. L'utilisateur peut mettre en place différentes stratégies : la délégation de contrôles (il prend un taxi). Le contrôle total (il prend sa moto) ou le partenariat (il prend un cheval). Dans ce dernier cas, les deux partagent le contrôle, avec des intelligences et des avantages différents qui vont s'associer. L'objectif étant ainsi de créer ce type d'interaction entre les humains et les systèmes.

Elle donne alors des exemples d'outils à l'interaction très limitée, comme un smartphone, ou un ordinateur. Si les systèmes sont puissants, les interactions sont très faibles : pointer, déplacer, écarter... Mais pour cuisiner, modeler, jouer du violon... les interactions sont bien plus nombreuses, et dans ces outils l'idée d'apprentissage de l'interaction est plus forte : on peut progresser. Elle détaille alors le principe de commandes gestuelles via un téléphone mobile, qui permet de créer un cycle d'apprentissage, où l'on peut ensuite développer ses propres commandes.

Enfin elle pointe le fait que tous les débats sur l'explicabilité des IA sont des discussions qui visent à expliquer l'algorithme, mais n'intéressent in fine que très peu l'utilisateur. Car au fond, un utilisateur cherche d'un système à obtenir la meilleure version de lui-même et non pas celui qui va le remplacer.

4/ Quelques exemples d'hybridation connaissances et apprentissage pour l'interprétation d'images

Par Céline Hudelot, Laboratoire MICS Centrale Supélec.

[Vidéo ici](#)

[Ressource ici](#)

Céline Hudelot est professeur d'informatique à Paris Saclay, au sein du laboratoire MICS. L'idée de sa présentation est de donner une vue globale de l'IA neuro symbolique. L'idée est de lier les approches neuronales de l'IA avec les approches symboliques. Elle s'appuie sur le papier célèbre de Dominique Cardon : La revanche des neurones, qui se base sur deux assertions : le raisonnement humain et le savoir sont complexes à représenter, mais du savoir peut-être implicitement contenu dans un ensemble de données. Cette approche est connexionniste.

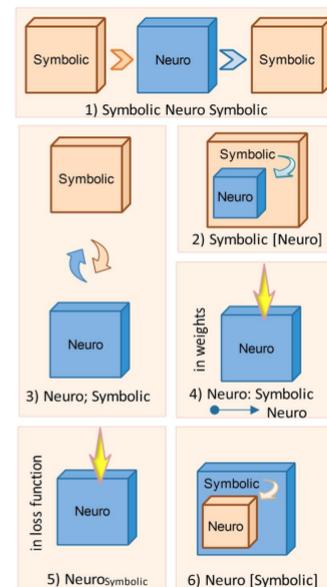
Avec l'IA symbolique, plutôt que d'utiliser des statistiques, on va utiliser des symboles pour représenter la connaissance (même incomplète) et le monde. Le but de l'hybridation est d'utiliser le meilleur des deux mondes. Pourquoi ? Pour combler le manque de raisonnement de haut niveau dans la première approche, et que les Deep Neuronals Models (DNM) sont des boîtes noires qui peuvent être trompées, ou que leur approche doit être orientée par une éthique ou une politique précise.

Cette approche hybride sera la troisième génération de systèmes d'IA. On parle alors d'IA neuro symbolique, d'apprentissage informé, ou de méthode formelle de représentation de la connaissance avec l'apprentissage.

Elle détaille ensuite les différentes manières de réaliser cette hybridation. Et revient sur les propositions de taxonomies proposées par Kautz qu'elle détaille, une par une. (ci-dessous)

Kautz's Neurosymbolic taxonomy

- **Symbolic Neuro Symbolic**
- **Symbolic[Neuro]**
- **Neuro;Symbolic**
- **Neuro :Symbolic → Neuro**
- **Neuro_{Symbolic}**
- **Neuro[Symbolic]**



Source : Hassan et al, 22

Henri Kautz, The Third AI Summer : AAAI Robert S. Englemore Memorial Lecture¹

Puis décrit ensuite d'autres formes de taxonomies : celle de Rueden, qui permettent de revenir sur les détails de l'intégration de symbolique au cœur du réseau de neurones, dans les tuyaux du machine learning : explicité les sources, les représentations, les fins attendues, et les bénéfices de l'intégration.

Puis Céline Hudelot présente ensuite brièvement quelques travaux sur l'interprétation d'images, en introduisant différents thésards de son laboratoire, notamment la thèse de Régis Pierrard, ou celle de Dao Thauvin réalisée à des fins d'explication, en permettant d'obtenir in fine une classification explicable d'images.

Elle introduit alors la conférence d'après, qui intègre des modèles symboliques au cœur des réseaux de neurones.

Hybridation symbolique / connexionnisme

1/ Améliorer les systèmes de classification neuronaux avec de la connaissance logique a priori

Par Arthur Ledaguenel, Doctorant en Intelligence Artificielle · Institut de Recherche Technologique SystemX

[Vidéo ici](#)

[Ressource ici](#)

Arthur Ledaguenel est diplômé de Centrale/Supélec et doctorant en Intelligence artificielle. Il présente sa thèse afin d'illustrer plusieurs des concepts présentés dans la matinée. Passionné de cuisine, il file une métaphore pour expliquer et vulgariser les différents types d'IA. Ce dernier a observé différents types de cuisiner : celui qui change les recettes et teste des choses selon ses envies, et l'érudit technique qui suivra le livre de cuisine à la lettre. Ce dernier représentera l'IA symbolique, qui travaille sur des données étiquetées, et passe son temps à faire des plats préparés. Et l'autre cuisinier serait une IA apprenante, que l'on renforce en estimant si le plat est bon ou non.

Il compare ainsi les avantages et inconvénients des deux types d'IA : le premier chef est rapide, flexibles, mais peu robuste (garantie théorique sur le résultat obtenu, par exemple, l'utilisation d'un aliment interdit pour cause d'allergie), et surtout, non explicable : on ne sait pas comment il a fait. En outre, il n'est pas sobre. (Cela a demandé beaucoup d'ingrédients pour parvenir à faire une recette convenable). Le deuxième type, c'est tout l'inverse, sauf sur la sobriété.

L'hybridation consiste à tirer le meilleur des deux mondes. Dans ce domaine, Arthur Ledaguenel travaille dans un champ présenté par Céline Hudelot : l'Informed Machine Learning. Le but consiste à prendre de la connaissance affectée à une tâche précise, et d'utiliser cette connaissance pour améliorer des systèmes basés sur de l'apprentissage (comme des réseaux neurones), décrite dans la slide suivante :

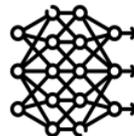
Connaissances



+



Étiquette valide



Données

IA neurosymbolique

Le but consiste à mixer la connaissance et les raisonnements issus des données afin d'obtenir en bout de chaîne une donnée étiquetée selon des règles précises. Cela revient à exercer des contraintes précises pour cadrer les algorithmes d'apprentissage. Pour filer la métaphore, cela revient à laisser le

premier cuisinier s'entraîner en le laissant tester tous ses plats, mais lui interdire des ingrédients ou contrôler la température de cuisson, afin d'éviter des résultats catastrophiques.

2/ XTRACTIS®, une IA Raisonnante Générale pour Découvrir des Connaissances Prédicatives Robustes

Par Zyed Zalila, Professeur de Mathématiques du flou à l'UTC et créateur de l'entreprise Intellitech

[Vidéo ici](#)

[Ressource ici](#)

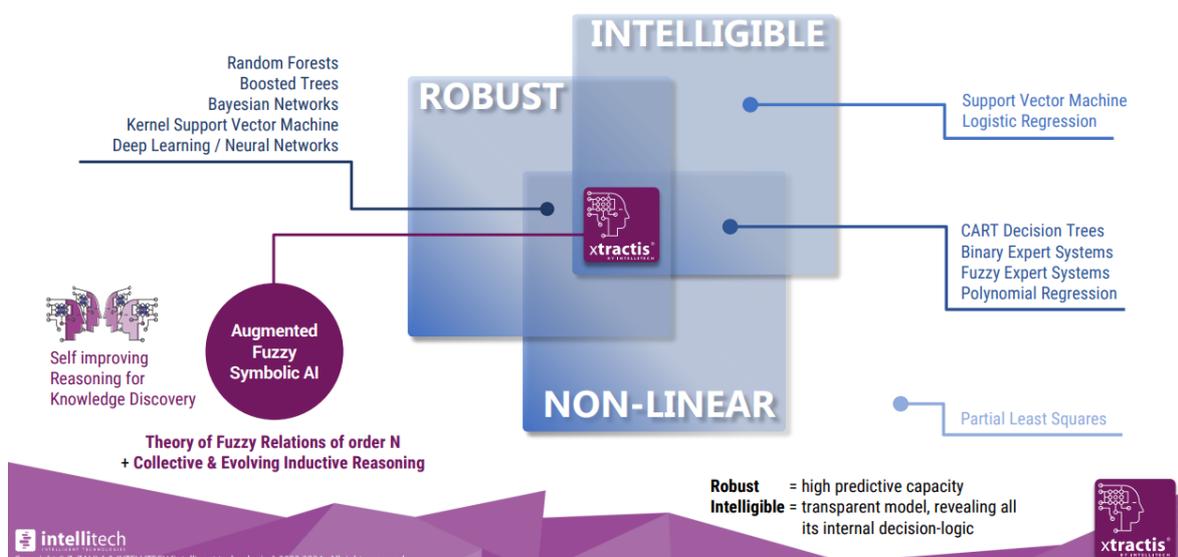
Zyed Zalila est spécialiste des mathématiques du flou. Il présente au départ le système d'IA qu'il a développé, Xtractis. Selon lui, il existe de nombreux cas où la connaissance n'existe pas ou n'est pas optimale. « Même avec 200 de QI vous ne pourrez jamais gérer une dizaine de critère en simultané, de façon consciente. Par exemple en médecine épigénétique, vous faites face à 27 000 variables. Pour cela vous pouvez avoir une approche cartésienne, c'est-à-dire faire du réductionnisme et découper le système en petit morceaux. Mais en faisant cela, vous modélisez autre chose car vous oubliez les interactions multiples au cœur du système », détaille-t-il.

Il revient alors sur les différents types de raisonnements possibles : l'induction, la déduction et l'abduction, qu'il détaille, puis donne quelques exemples de cas où la logique floue, continue, permet de traiter et résoudre des questions où la logique binaire (vrai ou faux) butte.

Il présente alors le fruit de son entreprise Intellitech, Xtractis, un ensemble d'IA qui permet de travailler sur des modélisations de systèmes complexes.

Un système conçu pour être à la fois robuste et intelligible sur des systèmes non-linéaires. Pour cela, il revient sur la théorie des relations floues d'ordre N, dont le but est de généraliser les ensembles flous à plusieurs opérateurs. « In fine, cela nous permet d'obtenir des algorithmes performant aussi bien

12 AIs & Data-Driven Modeling Techniques for Complex Processes and Phenomena



inductif, déductif qu'abductif. Au développe ainsi une IA qui est collective. C'est un système multi-agent distribué. Quand vous lancez une extraction, vous lancez N robots à l'assaut du même problème.

Comme ils ont à leur disposition une famille infinie de stratégies de raisonnement inductif chacun va s'instancier avec un raisonnement différent et va procéder ensuite à l'induction du modèle va, puis va évaluer de façon honnête sa stratégie. Et ensuite avant de partir pour un autre tour, ces agents vont coopérer pour essayer de voir quelles sont les stratégies qui ont les mieux marcher. Ainsi en dernière phase ils peuvent tous faire évoluer leur propre façon de raisonner », résume-t-il.

Il met alors en avant la notion d'explicabilité et d'intelligibilité de son système. « Pour chacune des déductions, d'inférence déductive, vous allez avoir un rapport de déduction, de prédiction, qui va vous expliquer le raisonnement rationnel et déterministe qui a été suivi par le robot pour pouvoir obtenir cette décision-là. Aujourd'hui on ne peut pas faire mieux en termes d'intelligibilité, en termes de d'auditabilité ou de certificabilité », argumente-t-il. Il revient alors sur quelques détails plus techniques des raisonnements, notamment le principe de « Fuzzy Abduction », et revient sur l'hybridation de son modèle. Puis donne quelques exemples de cas d'usage d'Xtractis. Le robot par exemple peut-être utilisé en médecine, à des fins d'aide au diagnostic. Ou dans le cadre de suivi de comportements malveillants de terroristes. Le système combine 4 entrées sur les modes de communication des personnes, pour classer ensuite le profil et lui attribuer un indice de phase afin de déterminer si la personne est proche du passage à l'acte ou non.

Enfin, il détaille la partie « physique » de sa solution, en revenant sur la sobriété de son système, qui fonctionne sur 7 à 8 cœurs physiques.

3/ IA Raisonnante Souveraine dans les Systèmes Critiques

Par Stéphane Delaye, Eviden (Atos)

[Vidéo ici](#)

[Ressource ici](#)

La présentation de Stéphane Delaye se voudra assez courte, car il préfère se focaliser sur la démonstration, accessible en dehors de la salle. Car il est important pour lui de voir le système fonctionner. Il travaille avec Xtractis, qui a été présenté précédemment. En tant qu'industriel, il se concentre sur la partie « concrète ». « Il y a beaucoup de choses dans l'IA qui sont théorisées depuis des années et la vraie question c'est de savoir si aujourd'hui nous avons les moyens physiquement de pouvoir exploiter ces théories », interroge-t-il. Et de répondre au cahier des charges des clients finaux. Notamment, en termes d'intelligibilité qui reste une notion subjective (même si les choses se normalisent peu à peu), pourtant indiquée clairement dans les desiderata des clients.

Selon lui, petit à petit, les définitions deviennent plus claires, notamment du fait du foisonnement des conférences internationales, ou des initiatives gouvernementales (ONU, UE, etc.) Il arrive ainsi à donner sa vision d'une « IA de confiance » : elle est souveraine (liberté d'action), robuste (excellence), éthique (suppression des biais), et auditable (contrôle de la délégation). Il replace alors l'importance de l'expert métier dans la démarche et de son rôle dans la définition du problème et de la clarté de l'intelligibilité.

Il détaille ensuite comment le système, par la production de rapports de performance, de structure, et de prédiction, passe d'une boîte « noire » à une boîte « transparente ».

Mais cela a un impact sur les capacités de calculs, qui sont plus gourmandes. « L'induction demande une plus grande utilisation de mémoire. Car vouloir expliquer et pouvoir avoir des modèles intelligibles demeure plus compliqué », commente-t-il. Pour terminer, il revient sur les critères de

mesure de performance. Si beaucoup de choses existent, la mesure de l'intelligibilité reste plus complexe, et encore à définir. Et il revient sur un risque fréquent, celui qui consiste à croire que l'IA peut apprendre par cœur toute la connaissance, mais ce comportement omet le critère de variabilité, très important si le cas d'usage n'est pas pris en compte dans l'apprentissage. Dans ce cas, le modèle pourrait avoir un comportement incontrôlé.

4/ IA hybride dans la Défense et la Sécurité

Par David Sadek, Thalès

[Vidéo ici](#)

David Sadek concentre sa présentation sur l'intérêt de l'IA hybride pour une entreprise comme Thalès, dont les cas d'applications concernent des usages critiques (le domaine militaire, principalement). Il commence par un clin d'œil, d'une thèse qu'il avait dirigée en 1995, qui s'intitulait : « Une approche combinant représentation explicite des connaissances et apprentissage connexionnistes ». L'IA hybride, n'est pas une affaire récente.

Il commence par rappeler la stratégie de Thalès en termes d'IA de confiance. Elle se résume sous l'acronyme TrUE, pour Trustable, Understandable et Ethical.

Il donne un exemple de système qui était mis en place pour entraîner les capitaines de navires américains et qui gagnait à chaque fois. Après enquête, il s'est avéré que le système tirait sur ses propres bateaux dès qu'ils étaient touchés, pour les couler, car cela le ralentissait. Il avait « appris » seul cette règle, sans que jamais on ne lui donne l'ordre de ne pas tirer sur ses propres bateaux. D'où l'importance de l'approche symbolique qui permet de donner un cadre et certaines règles. Selon lui, l'IA hybride permet d'aller vers la validité, l'explicabilité, la sécurité et la responsabilité, en encadrant des systèmes non déterministes par des garde-fous.

Il précise que chez Thalès, l'objectif va encore un cran plus loin en termes d'explicabilité, puisque le but est d'arriver à établir des systèmes auto-explicables, qui s'expliquent eux-mêmes, dans le temps de l'action. Il donne l'exemple d'un pilote dont le système propose de virer et changer de trajectoire, selon David Sadek, le pilote doit être en droit de poser la question.

Il évoque également la conformité à tous les cadres exogènes, sur laquelle travaille l'entreprise : la loi (AI Act, la frugalité avec l'empreinte environnementale...)

Thalès a choisi de mettre un effort sur l'IA hybride dès 2018. Dans l'industrie, l'IA connexionniste a ses limites, notamment dans le panel de données à disposition, pas assez représentatif. « Le big data est plutôt small, dans notre cas. Il ne faut pas imaginer qu'il y a pléthore de données pour faire les choses, comme dans la reconnaissance sous-marine, nous n'avons tant d'image que cela », ajoute-t-il.

Il donne ensuite un exemple concret d'une approche de l'entreprise en termes d'IA hybride. Elle passe par des formes géométriques, pour pallier le manque de données d'images. Par des transformations géométriques (rotation, symétrie, homothétie) pour les changements d'échelles, sans avoir à changer l'arrière plan. « Les résultats sont remarquables », commente-t-il. Et sans avoir à passer par une approche uniquement basée sur les données. Cela permet de faire de la reconnaissance d'image quelque soit le modèle de caméra utilisée, par exemple. Certaines, comme les fish-eyes, peuvent déformer les images. Cette approche est quasi opérationnelle, et des applications commencent à fonctionner là-dessus.

Une autre approche, dans la détection d'obstacles, où Thalès renforce les algorithmes neuronaux avec du symbolique. Enfin, il cite les arbres génétiques flous, qui ont été testés dans le cadre de simulations

de combats aériens, qui en phase d'industrialisation, dont le but est de combiner des inférences floues, avec des algorithmes évolutionnaires.

5/ Expliquer les IA : de l'IA statistique à l'extraction de connaissance via une approche « régionale »

Par David Cortes, fondateur AI-vidence

[Vidéo ici](#)

[Ressource ici](#)

L'approche d'AI-vidence est centrée sur la notion d'explicabilité, avec une autre approche que ce qui a déjà été présenté : l'approche régionale.

Pour David, Cortes, l'hybridation consiste à réconcilier les sœurs ennemies. Et les deux ont un ennemi commun : la complexité. La performance se dégrade avec la complexité des phénomènes, mais la compréhension également.

Le problème de l'IA statistique, c'est qu'elle exacerbe tous les mauvais côtés de l'automatisation. Les performances se dégradent dans le temps, dès la mise en production, et l'on ne sait pas ce qu'elle a appris. Il donne alors deux statistiques intéressantes : 90% des projets d'IA ne vont pas en production, et 50% des modèles finissent par être débranchés. L'IA de confiance est un vrai problème de société, car ce domaine représente 60% des investissements en IA.

Pourtant, l'éthique et la robustesse sont les piliers de l'IA de confiance, et notamment des impératifs de mise en conformité d'un point de vue légal. Mais ils sont aussi importants pour accélérer l'adoption opérationnelle des outils à base d'intelligence artificielle. L'approche d'AI-vidence consiste donc à rendre explicable l'IA (même si c'est compliqué !). Pour cela il faut s'interroger d'abord à qui s'adresse l'explication. Entre le régulateur, les équipes conformité, les utilisateurs de modèle, et les clients finaux, le data scientist, qui doit améliorer le modèle, et les experts métiers, pour comprendre quelle connaissance nouvelle on peut extraire des IA. A l'heure actuelle, on donne beaucoup de pouvoir au créateur de modèle. AI-vidence se concentre sur le lien entre le concepteur du modèle et son utilisateur, afin qu'ils puissent se comprendre.

Il présente alors son outil : AntakiA, dont le rôle est d'explorer une boîte noire, de la compartimenter (méthode cartésienne) et de substituer ces parties par des sous-modèles explicables, pour la blanchir au maximum. Cette approche est régionale, car elle compartimente certaines parties du modèle, pour les substituer par un autre. Ce processus peut aussi être ciblé en fonction des zones à risques du modèle, en y forçant un modèle plus simple sur les zones à haut risque. Cela donne des choses plus robustes dans le temps.

L'exploration se fait de manière « dyadique », en concertation avec les experts métiers, pour juger de la pertinence de la substitution ensuite. Le concept a été testé auprès du régulateur financier, et a été distingué par les banques et les assurances, et les équipes ont commencé à travailler pour des projets plus industriels. De nombreuses améliorations sont prévues pour les mois et années à venir présentées notamment la question des dérives dans le temps (en partenariat avec Telecom Paris).

6/ Modèles NLP d'extraction d'information

Par Etienne Bernard, CEO de Numind

[Sa présentation vidéo à distance est ici.](#)

Etienne Bernard est le fondateur de Numind, est sa présentation est uniquement vidéo. Numind est spécialisée dans le développement d'outils pour créer des modèles NLP d'extraction d'information à partir de texte. Il veut revenir sur l'importance des explications données par le modèle sur ses prédictions.

Il commence par décrire le produit et le fait tourner en direct. Il s'agit d'un algorithme qui permet de déceler dans un texte légal, les noms des tribunaux, des avocats et des autres parties. Il « apprend » au logiciel en labellisant certaines données au départ. Rapidement, le logiciel fonctionne seul, en donnant le pourcentage de chance que sa labellisation automatique soit vraie.

« Mais ce qui serait très utile c'est que le modèle nous explique pourquoi il a pris telle ou telle décision, afin de mieux comprendre ses erreurs », explique-t-il. Dans l'exemple donné, le modèle labellise un avocat car le mot « Attorney » était situé à proximité mais pas par le contexte, notamment le fait que la personne « représente » une autre personne. Un « playground » permet alors de modifier le texte de départ, pour trouver la faille dans le modèle, et pouvoir l'entraîner spécifiquement sur cette erreur.

Des méthodes existent pour que le modèle puisse expliquer ces prédictions, comme les « chap values », que Numind essaie de mettre en place, afin que le modèle puisse expliquer de manière plus intuitive ses prédictions. La formalisation d'un « espace des explications », qui permet de montrer les points faibles du modèle, pour mieux les situer et pouvoir les corriger soit en modifiant le texte comme fait dans le Playground, soit pour modifier la base de données d'apprentissage. Il pointe alors la vigilance qu'il faut avoir lorsqu'on met le modèle en production, car les données de production sont toujours différentes des données d'entraînement.

Les jeunes talents : Mon CV en deux minutes

Des étudiants présentent leur CV en deux minutes. On commence par Léa Lesbats. [\(Présentation ici\)](#)

Puis Nicolas Saint, [dont la présentation est ici.](#)

Hybridation homme/machine

1/ Borges and AI

Par Léon Bottou, chercheur chez Meta

[Vidéo ici](#)

[Ressources ici](#)

Selon Léon Bottou, le public commet une erreur depuis le succès des modèles de langage, en les confondant avec une intelligence artificielle. Lui, ne voit pas pourquoi un modèle de langage serait une intelligence artificielle. Car un LLM est un outil mesurable, mais pas une IA. Donc la définition n'est pas aussi claire. Le public, tout en parlant, s'inquiétant, débattant autour de ce sujet, finit par partager au fur et à mesure la même imagerie mentale : celle de la science-fiction moderne, de l'IA, mais dont l'image est présente depuis bien plus longtemps. Elle était déjà présente avec Homère, et les « robots » servants. Idem avec le Golem de Prague. Le souci étant que l'imagerie mentale partagée crée la manière avec laquelle on perçoit les choses et la manière avec laquelle on les comprend. Mais la question consiste à savoir si l'imagerie mentale possédée par la population est bien en correspondance avec la réalité de ce que nous avons sous les yeux.

Pour creuser la question, avec Bernard Schölkopf, Léon Bottou a égrainé la littérature de Borges « d'une part parce que c'est amusant, mais aussi parce que cela nous force à penser différemment », explique-t-il. A travers les multiples petites histoires imaginatives, Borges interroge des notions comme le temps, le langage, la réalité et l'infinie. « Le but est donc d'appréhender l'intelligence artificielle avec une nouvelle imagerie mentale », détaille-t-il. Les chercheurs en ont sélectionné deux.

La première intitulée « Le jardin des chemins qui bifurquent », parle du livre de Ts'ui Pen's, qui contient toutes les possibilités de l'histoire, toutes les ramifications de l'intrigue. Ce ne sont pas tous les textes existants (comme la bibliothèque de Babel) mais tous les textes possiblement compréhensibles. Imaginons qu'un utilisateur enregistre sur une bande le mot d'après, il sera contenu dans le livre, mais autant il restreint l'histoire, autant il ouvre une infinité de possibilités. Léon Bottou en vient alors à interroger la structuration du langage, et des différences de construction du langage par des transformations syntaxiques, et des transformations sémantiques. Il applique alors ces réflexions au modèle de langage par apprentissage, et ouvre ainsi une nouvelle compréhension des chatbots. Il arrive à la définition de la « nécessité narrative » qui est la possibilité de compléter tout schéma narratif, et interroge ensuite par l'exemple de l'endroit où vit Sherlock Holmes, la notion de vérité.

Narrative necessity

- The ability to recognize the demands of a narrative is a flavor of knowledge distinct from the truth.
- What is true in the world of the story need not be true in our world.
- As new words are printed on the tape, the machine borrows facts from the training data (not always true) and fills the gaps with plausible inventions (not always false).
- Not **hallucinations** but **confabulations**.*

Neither truth nor intentions matter to the operation of this machine,
only narrative necessity matters.

Selon lui un modèle de langage parfait ne s'interroge pas sur la vérité de ce qu'il raconte, ni sur la cohérence de ses propos, mais pose juste du langage sur une bande, poussé par la nécessité narrative.

Dans la deuxième histoire intitulée la bibliothèque de Babel, qui contient tous les livres et au sein de laquelle, les bibliothèques en quête de vérité sont confrontées à beaucoup de problèmes qu'ils se créent eux-mêmes. Car pour un livre donné, ils peuvent trouver une contradiction dans un autre livre, elle-même contredite par des milliers d'autres livres. Et cela amène au problème de la vindication, qui consiste à lire quelque chose qui vous reconforte dans votre pensée. Ainsi on pourrait craindre que les modèles d'IA, dans le travail de fiction, tout du moins, ne finissent pas de délivrer que les suites de textes que nous souhaitons entendre.

Il conclut dans une troisième partie, sur la nécessité pour les humains, de créer des histoires, de les entendre et de créer des histoires à partir des histoires, et affirme, en citant Patrick Winston, que le storytelling est sûrement ce qui différencie l'humain de toutes les autres espèces.

2/ IA pour l'aviation de combat/hybridation homme/machine

Par Bruno Patin, responsable de la recherche sur les systèmes autonomes pour Dassault Aviation

[Vidéo ici](#)

[Ressource non disponible](#)

Bruno Patin divise sa présentation en deux parties : une qui traite des travaux effectués depuis 2018, et l'autre sur les projections d'ici 2040. Il rappelle que les travaux effectués sur le sujet se produisent sur le temps long, et qu'il y a de forte chance qu'il ne voit jamais le résultat de ce sur quoi le monde de la recherche planche, du fait que cela embarque beaucoup de personnes différentes, que cela engage de la complexité et de la sécurité, notamment dans le cadre de l'aviation de combat. Il rappelle à cet effet que les avions de chasse d'aujourd'hui (le Rafale ou le F-35 ont été conçus dans les années 80...) Que ce soit pour les armements ou les modes de contrôles des systèmes, il précise que la différence sur les dernières années est exponentielle et que l'on est passé du système « automate » au système « data driven ». Pour explorer la vision du futur, la Direction Générale de l'Armement a créé le projet Man Machine Teaming, qui a permis de dresser des scénarios et des pistes de réflexion sur l'avenir.

Il présente ensuite les changements au niveau des systèmes d'interaction avec l'environnement, comme avec l'introduction de la 3D ou de la réalité augmentée, dans les cockpits ou sur le terrain. Idem pour la représentation de l'assistant virtuelle. Ces considérations sur les interfaces sont prédominantes car dans un cockpit, il ne faut pas perturber le pilote et s'adapter à ses contraintes, comme la désorientation spatiale. En clair, l'introduction de nouvelles technologies demande de repenser les interactions humains – machines en jeu. « Même si l'on ne sait pas tout faire, nous avons pu explorer les concepts et les refondre », détaille-t-il. Ce travail s'est terminé en 2020, et maintenant toutes les équipes sont embarquées pleinement dans ce projet en travaillant avec les fournisseurs etc. D'un autre côté, les équipes travaillent également sur le combat aérien du futur.

Il détaille ensuite une spécificité : le travail sur les recommandations auprès des autorités pour les certifications militaires, civiles ou tous les besoins afférents. Et quelles que soient les technologies algorithmiques utilisées, Dassault travaille en termes de fonctionnalité. « La certification se fait davantage sur la fonctionnalité, même si vous pouvez avoir des guidelines sur la façon de faire la technologie », détaille-t-il. Il revient ensuite légèrement sur les défis d'organisation face à la multiplicité des acteurs autour de ces projets, notamment au niveau européen. Et sur le principe de souveraineté, au cœur des réflexions, qui doit être respecté au bout du compte, quand la technologie sera faisable.

3/ Des IAs et des Hommes – Pour une approche behavioriste du couple humain(s)-algorithme(s) dans les applications critiques

Par Nicolas Vayatis, École normale supérieure Paris-Saclay

[Vidéo ici](#)

Nicolas Vayatis se concentre sur les interfaces humains – algorithmes, et veut sortir de l'approche cartésienne ou de celle de Chomsky, et réintroduire du « Behaviorisme » dans les approches. Estimant que cette approche est marginale, un collectif de scientifique a été créé afin de la valoriser : le Centre Borelli, du nom de l'inventeur du scaphandre, première interface entre l'homme et son environnement. Ce centre intègre différents champs disciplinaires : les mathématiques, la

modélisation de systèmes physiques, le machine learning, ou encore les neurosciences comportementales, ou des sciences sociales. Il insiste sur le fait que les projets scientifiques que le centre abrite se confrontent au terrain, notamment dans le domaine médical et biomédical. Il précise alors qu'il y a bien un enjeu d'interface entre le système complexe du corps humain et l'expert clinicien.

Il insiste sur l'importance d'intégrer les sciences sociales pour réfléchir à ces questions, du fait même que beaucoup de personnes commencent à questionner l'intérêt de la technologie – il fait le parallèle avec les luddites britanniques – et que ce courant remet en cause les data centers. « Des étudiants qui étaient intéressés par l'IA se mettent à suivre un cheminement intellectuel qui les amène finalement à remettre complètement en question cette approche et à être même extrêmement militant contre l'IA », argue-t-il.

Après avoir présenté les défis actuelles (démographiques, économiques...) il atteste que l'on ne pourra se passer de la technologie parmi les paramètres d'analyse, et étend le terme d'hybridation de la matinée, non pas à l'IA, mais l'hybridation entre la connaissance experte et le monde physique, celle entre le savoir « fondamental » et la connaissance de terrain. Il souligne les différences existantes, parfois, les décalages, entre les modèles théoriques et les données relevées en pratique. Ces questions vont être croissantes du fait même du nombre de tests in silico qui deviendront plus répandues. Selon lui, tout le domaine se repose sur trois postulats, qui sont pour lui des hypothèses que l'on ne questionne pas assez :

- L'automatisation apportera des gains de productivités et réduira les risques
- Dès qu'il y a des enjeux de vies humaines, l'humain doit être dans la boucle
- Il n'y a pas de limites biologiques pour l'humain à maîtriser les environnements qu'il crée

Selon Nicolas Vayatis, cette dernière, notamment, on n'a pas assez développé d'études sur le sujet des limites biologiques. Il argumente son propos par l'analyse accidentologiques des transports en France. Mais aussi d'un point de vue purement théorique, avec la loi de Miller (le nombre de paramètres que l'humain a la capacité de traiter est environ 7) mais insiste sur le fait que s'il y a des associations entre plusieurs variables, le nombre à retenir est davantage 3. Il ajoute également la loi de « monotonie » qui établit que l'on ne comprend que ce que l'on sait déjà. Ce qui selon lui, devrait davantage remettre en question la notion d'explicabilité, et qu'il faut une explicabilité appréhendable par un cerveau humain « et rien ne dit que certains phénomènes puissent-être appréhendables par un cerveau humain ».

Ce seul fait questionne également le deuxième principe qui participe à injecter de l'humain dans certains processus.

Mais plus généralement, il questionne la répartition des budgets dans les champs de recherche, et ce que cela apporte concrètement, sur le terrain. Il donne comme exemple, le Humain Brain Projet, un projet doté d'un milliard d'euros sur 10 ans, et questionne les avancées concrètes, en ehpad, en psychiatrie sur les dernières années. Il pointe alors les décalages entre l'argent investi dans les études sur le cerveau et l'argent investi sur le terrain : déserts médicaux, aidants qui meurent avant les aidés...

Il argue ensuite en évoquant les travaux dans le domaine des sciences comportementales, où les articles scientifiques s'appuient sur des études de terrains : pour expliquer la mort subite, ou la réduction des inégalités sociales face à la douleur, ou encore la facilitation des démarches pour obtenir un deuxième avis. Il insiste alors sur le bénéfice des algorithmes dans le cadre des décisions sur le terrain, qui permettent aux humains de se corriger eux-mêmes.

Pourtant, il évoque une étude menée sur des radiologues qui conclut que l'humain augmenté fait systématiquement moins bien que l'humain seul ou que l'IA seule. Un résultat contre-intuitif, qui ne peut pas encore être généralisé, mais qui reste pertinent à questionner.

Lien association X-IA

Présentation de startups : ma startup

Plusieurs start-ups viennent présenter leur concept en quelques minutes

1/ Learning Robots : IA générative et éducation

Par Thomas Deneux, cofondateur

[Vidéo ici](#)

[Ressources ici](#)

Learning Robots a été créé par Thomas Deneux, avec Axel Haentjens, en 2020 dans une visée essentiellement éducative, afin, de rendre plus visible toute la partie invisible des algorithmes apprenants. L'entreprise s'aide pour cela d'un robot, avec caméras, capteurs, et qu'il est possible d'entraîner à l'aide d'un ordinateur. Il va ensuite restituer ce qu'il a appris, et l'utilisateur peut voir ce qu'il se passe sur les algorithmes, visuellement ». Deux modes sont possibles : apprendre en étant supervisé par un humain, ou apprendre tout seul, par renforcement en autonomie. « Dans tous les cas, nous voyons sur l'interface les réseaux de neurones se former, se renforcer et s'affaiblir », détaille Thomas Deneux. Learning Robots propose donc son outil éducatif aux milieux éducatifs. La start-up propose aussi des formations en entreprise. Le but est de dédramatiser le rapport à l'IA par les directions, afin de détruire tous les mythes que l'on peut avoir sur le sujet, et rendre les algorithmes concrets pour tous les béotiens.

2/ Muvraline : Green for AI and AI for green, with a hybrid AI. Telcos example

Par Rémi Perthuisot, cofondateur

[Vidéo ici](#)

[Ressources ici](#)

Muvraline a commencé en 2018 comme filiale du groupe Altice-SFR. Elle est désormais indépendante. Et est spécialisée en IA frugale. « Une IA frugale ne signifie pas nécessairement des algorithmes frugaux », tempère Rémi Perthuisot. Elle est surtout présente auprès des opérateurs télécoms dans une dizaine de pays, et compte environ 8000 utilisateurs à ce jour. Elle propose différents produits d'optimisation pour les équipes opérationnelles : optimisation des équipes sur le terrain, optimisation des écrans publicitaires sur les chaînes de TV, ainsi que différents outils de communication comme des chatbots. Elle aborde le sujet de l'IA hybride justement dans un souci de frugalité de l'IA. Le but est de chercher à utiliser moins de ressources pour obtenir le même résultat. Il donne alors un exemple de ce qui est fait en matière d'IA hybride, avec les armoires FTTN qui répartissent les fibres au niveau des quartiers. A partir d'une photo, le but est de voir si les multiples règles et règlements sont respectés. « Mais nous n'avons pas énormément de données pour faire uniquement un travail à partir d'IA connexionniste, donc nous avons demandé à un expert de nous fournir les règles, que nous avons

implémentées avec une IA symbolique », détaille-t-il. Cela a donné in fine, un modèle extrêmement rapide, qui permet de déterminer si une armoire est bien « montée » ou non.

3/ DescartesMaus.ai: Dériskuer la décision dans un monde incertain grâce à l'IA

Par Maurice N'Diaye, fondateur

[Vidéo ici](#)

[Ressources ici](#)

La start-up s'appelle Descartes&Mauss du nom de René Descartes et de Marcel Mauss, pour allier la rationalité du philosophe et mathématicien à la sensibilité de l'anthropologue. Le but de la jeune pousse est de répondre à la question de l'accélération du temps dans les entreprises, et d'aider les équipes de direction, notamment marketing, ou de direction générale à prendre des décisions. Elle aide particulièrement à optimiser l'allocation des ressources (capital ou travail) selon les besoins. Elle fonctionne pour cela en différentes étapes : une de recueil des données (à tous les niveaux), une de création de multiples arbres de décisions et de scénarios, puis de priorisation qui consiste à shortlister les meilleures décisions. Enfin, l'IA génératives intervient ensuite, dans la deuxième étape, lorsqu'il faut transmettre ces réflexions aux équipes. « Il faut comprendre que la combinatoire des arbres de décisions est très complexe pour certaines entreprises », décrit Maurice Ndiaye, son créateur. Les outils sont ainsi un support à l'idéation, mais aussi au roadmapping, pour les équipes. Et les conseils se retrouvent adaptés à la situation : filiale, marché, équipes, métiers... A ce jour, la solution a permis d'améliorer l'efficacité selon quatre KPI's : le délai d'action (McDonald's), les coûts des prestataires et sous-traitants (Colgate Palmolive), les équivalents temps plein par projet (Danone) ou encore les ROI (L'Oréal).

4/ Corolair : Former avec l'IA

Par Wassym Kalouache, fondateur

[Vidéo ici](#)

[Ressources ici](#)

Corolair a été créée début 2023, et cherche à personnaliser les formations, pour régler les inégalités due aux formations personnalisées comme les cours particuliers. Selon son fondateur Wassym Kalouache, 98% des apprenants qui ne bénéficient pas d'une formation personnalisée obtiennent de moins bons résultats que ceux qui en bénéficient. Corolair vient donc soutenir le manque de personnalisation côté apprenant, et réduire l'opacité chez les formateurs, due à l'utilisation de ChatGPT par les élèves. Le but consiste à donner aux formateurs les moyens de créer un « tuteur virtuel » pour les apprenants. Il peut ainsi transformer son cours en modèle de langage, et y associer du contenu pédagogique. Les apprenants peuvent dialoguer avec le tuteur ainsi « cadré » par le contenu du concours – en réduisant les hallucinations – via une IA hybride, enfin, en fonction de l'analyse des conversations entre les élèves et le tuteur virtuel, et des évaluations, le professeur peut ensuite améliorer et adapter sa formation.

DEBAT sur l'IA hybride

[Vidéo ici](#)

La journée se conclut autour d'une table ronde animée par Laurent Gouzènes et réunissant :

- *Caroline Chopinaud, Hub France IA*
- *Mouadh Yagoubi, IRT System X*
- *Pascal Alix, avocat*
- *Christophe Labreuche, Thales*
- *Judith Nicogossian, Anthropologue*
- *Christophe Rodrigues, Enseignant-Chercheur en IA, Pôle Léonard de Vinci*
- *Johann Dréo, Institut Pasteur*
- *Nicolas Sabouret, Professeur à l'Université Paris-Saclay*

La table ronde tente de répondre à plusieurs questions d'ordre général sur l'intelligence artificielle. La première concerne l'avenir de l'IA. Mouadh Yagoubi est chef de projet pour IRT System X, notamment sur les questions d'IA hybride. Il explique que l'IA a vécu une grande avancée ces dernières années, notamment depuis 2018, en matière de modélisation des systèmes physiques complexes, surtout grâce au système par apprentissage. Mais il précise qu'il faut toujours trouver un équilibre entre le résultat effectif et le coût en termes de calculs ou en données, qui reste un obstacle. C'est un élément déterminant dans l'avenir industriel de l'IA et limitant son passage à l'échelle. A ce titre, il précise l'existence du frugal learning, qui consiste à entraîner un modèle sur un petit échantillon de données. Tout cela s'ajoute aux questions de validation des modèles, car ces derniers sont parfois trop complexes (quand on limite, ou qu'on ajoute des paramètres afin de les contrôler ou les cadrer) et peuvent in fine se montrer contre-productif. Selon lui, dans le domaine de l'ingénierie, nous n'avons pas encore atteint la maturité.

Le débat tourne ensuite autour des capacités de l'IA et de sa progression, dans la volonté « d'essorer les domaines de la connaissance humaine », notamment à travers l'exemple d'Alpha Go, de Google, qui a pu apprendre autant que 10 000 êtres humains qui jouerait 8 heures par jour, en quelques jours, grâce à ses capacités de calculs. Nicolas Sabouret, Professeur à l'Université Paris-Saclay rappelle que la science progresse à petit pas. D'abord, l'IA (ou l'algorithmie) s'est attaquée au jeu de dames, puis aux échecs, puis au jeu de go, avec les réseaux de neurones, puis au langage... Et que le secteur s'attaquera progressivement à d'autres domaines. Mais ce qu'il s'est passé sur le langage fait faire d'énorme progrès à d'autres domaines. Car cela permet d'encoder des connaissances générales abominables à écrire sous forme de règles simples. Donc ce champ ouvre d'autres domaines. Mais la question reste de savoir si on peut faire « raisonner » les modèles, et force est de constater que pour le moment, ce n'est pas le cas. Johann Dréo, de l'Institut Pasteur prend alors la parole. Il estime que la connaissance à encore du temps avant d'être épuisée... Même si de nombreux domaines, comme la médecine, bénéficient énormément des techniques actuelles. Mais dans une considération plus générale, il s'interroge sur ce que l'IA peut améliorer, en dehors du champ technique.

Selon lui, les deux plus grands défis de notre société sont le défi climatique et la justice sociale. Et il s'interroge comment l'IA peut aider à résoudre ces défis là. Il estime qu'il faut différencier le champ de la connaissance et l'intelligence. L'IA peut ainsi aider à mieux automatiser ce que l'on sait aujourd'hui, mais la capacité à mieux comprendre les choses, à les modéliser et aller au-delà de ce qui est connu, réside encore dans l'intelligence humaine.

Judith Nicogossian intervient alors. Anthropologue et anthropobiologie, elle a beaucoup travaillé sur le concept d'adaptabilité. Elle note d'abord la différence de temporalité dans l'appréhension de l'outil (et donc de sa complexité) entre le grand public et la communauté scientifique. La première communauté se retrouve à adopter ces outils avec un engouement incroyable « nous nous retrouvons tous un peu scotché face aux performances de l'outil, sans en être spécialiste », estime-t-elle. Elle distingue alors le système mécanique avec ses biais, ses limites et le système biologique, qui lui reste encore très mystérieux (comme le cerveau). Vient alors le sujet de l'hybridation entre les deux systèmes, et de la nécessité de réfléchir les interactions entre les deux mondes. Il faut donc distinguer différents types d'hybridation (celle de l'IA dont on parle depuis le début du séminaire) et l'hybridation homme machine, qui joue un grand rôle sur les deux systèmes. La question suspendue reste de savoir comment l'humain sera capable d'utiliser cet outil pour continuer à évoluer.

Johann Dréo prend alors la parole. Il revient sur la partie interaction mais propose surtout une image mentale qui pourrait nous aider à comprendre les limites et la puissance des IA actuelles : ce sont des fonctions. Certes complexes, avec un grand nombre de variables, mais des fonctions avant tout. C'est-à-dire que pour une entrée précise, elles associent une sortie précise. Cette image est, certes, un raccourci, mais elle permet de se rendre compte qu'il n'y a rien de magique derrière ces outils. Il estime donc que réussir à fabriquer des outils qui ressemblerait à ce qu'on perçoit dans les films de sciences fictions sera très compliqué. Pas impossible, mais très long.

La table ronde se concentre alors sur le fait qu'un humain est « contraint » de réduire sa pensée pour dialoguer avec ses outils. Si l'intelligence qu'on leur assigne sont justement une projection de notre propre intelligence, il n'en reste pas moins que nous « rentrons » dans le schéma mental de la machine pour interagir avec.

Christophe Rodrigues, Enseignant-Chercheur en IA, Pôle Léonard de Vinci, prend alors le micro. Il revient sur un papier scientifique de Microsoft, publié en 2020, qui estimait avoir fait des progrès phénoménaux dans l'analyse des sentiments. « Mais le biais fondamental résidait dans le fait que l'IA ne faisait que noter les expressions comme positives, négatives ou neutre », détaille-t-il. Ce papier a été par la suite « debunké », par des chercheurs qui ont montré qu'en modifiant parfois un mot dans l'expression, même anodin, suffisait à tromper l'IA et à faire passer l'expression de positive à négative ou inversement. Cela représente pour lui l'exagération totale des biais. Cela ne l'empêche pas de noter que d'autres travaux, comme ceux de Deep Mind, ont des impacts plus concrets sur la vie quotidienne des chercheurs ou de certains domaines : le déploiement des molécules, la découverte de nouveaux matériaux, ou encore la prédiction météorologique. Et ces avancées ont surpris énormément de monde, bien plus que sur les modèles de langage. Il estime d'ailleurs que ces modèles finissent peu à peu par fatiguer le public, quand sur des domaines moins « visibles » ou médiatiques, les avancées sont considérables.

C'est alors au tour de Pascal Alix, avocat, d'émettre son avis. Il estime que les hallucinations des IA génératives de langage ont des limites : elles ne peuvent être utilisées que si on maîtrise totalement le sujet. Car il faut bien poser la question (le prompt) est être capable de pouvoir vérifier la réponse. Il ajoute que les IA génératives telles qu'elles sont proposées à l'heure actuelle – en tout cas au grand public – ne sont pas adaptées à un système critique, en tout cas pour ce qu'il connaît : la justice. Même s'il reconnaît qu'avec ce qui a été présenté lors de la journée, l'IA hybride, pourrait peut-être être une réponse.

Christophe Labreuche intervient alors sur l'aspect moral. Il évoque le dilemme bien connu, pour la voiture autonome, du choix à faire dans le cas d'un accident inévitable, et de choisir la victime, et que la réponse dépend des cultures, et des pays dans lequel il est présenté. Il évoque alors le fait que si on place un opérateur humain aidé d'une IA face à une décision critique, il peut perdre conscience qu'il est dans un système critique justement. Et cela pourrait à terme biaiser sa décision, dans le sens où

cela le déresponsabiliserait. Les concepteurs doivent donc porter attention à cette notion, et tenter de ne pas déconnecter la personne de l'impact ou de l'importance de la décision qui lui incombe, avec ou sans IA.

Cette discussion replace l'importance du système au sein duquel peut être amenée à évoluer l'IA. Une IA générative qui dessine un humain à trois mains peut prêter à sourire, mais moins dans un système critique. « Je peux me tromper 10 fois sur cent dans une journée, mais une erreur sur 10 millions pour les voitures aboutit à 20 accidents par jour », estime Laurent Gouzènes.

Le débat revient alors sur une intégration plus globale du système, et pas seulement l'outil en tant que tel, passer d'une ingénierie algorithmique à une ingénierie système et logicielle, que ce soit sur la représentation des données, la modélisation etc.

En conclusion technique, est évoqué le fait que le paradigme sur lequel s'appuie ChatGPT par exemple, est le même. Ce n'est que l'évolution du dataset et la puissance de calcul qui a changé. Mais le cœur du fonctionnement du système qui permet un résultat aussi bluffant pour l'utilisateur, est le même depuis les années 50 ou 60.