



HAL
open science

ALGORITHMES ET LIBRE ARBITRE, QUEL AVENIR POUR NOTRE LIBERTE ?

Cécile Godé

► **To cite this version:**

Cécile Godé. ALGORITHMES ET LIBRE ARBITRE, QUEL AVENIR POUR NOTRE LIBERTE ?. 23ième rencontres économiques d'Aix-en-Provence, Jul 2023, Aix en Provence, France. hal-04123097

HAL Id: hal-04123097

<https://amu.hal.science/hal-04123097>

Submitted on 9 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LES RENCONTRES ÉCONOMIQUES D'AIX-EN-PROVENCE

LES 7, 8 ET 9 JUILLET 2023

SESSION 29 : ALGORITHMES ET LIBRE ARBITRE, QUEL AVENIR POUR NOTRE LIBERTÉ ?

Cécile Godé, Professeure agrégée des universités en sciences de gestion et du management

Aix-Marseille Université, CERGAM

NOTE DE CADRAGE

Les algorithmes, ces suites finies et ordonnées d'instructions destinées à résoudre un problème, sont omniprésents au quotidien. C'est particulièrement le cas des algorithmes d'intelligence artificielle (AIA), capables d'automatiser des tâches cognitives complexes dévolues aux seuls experts il y a encore peu. Intégrés aux réseaux sociaux, plateformes et autres applications logicielles, les AIA analysent des corpus massifs de données pour produire des recommandations, prédire les crimes ou encore préconiser un traitement médical. Développés par l'homme, ils suscitent de fortes attentes de la part d'utilisateurs qui s'en emparent largement tant leur efficacité exécutoire est importante. Légitimement, ces mêmes utilisateurs s'inquiètent des risques de manipulation, de dépendance et de réduction du libre-arbitre. Cette session met en débat les capacités à la fois habilitantes et contraignantes des AIA en interrogeant leurs usages et pensant leurs impacts sur les modèles de gouvernance organisationnelle et sociale.

CONTRIBUTION

Entre opportunités et menaces : les Algorithmes d'Intelligence Artificielle en débat

Les algorithmes permettent de résoudre des problèmes du quotidien. La recette de cuisine décrite pas à pas sur l'écran d'un robot cuiseur multifonctions, la trajectoire calculée par GPS ou encore le placement des ordres de marché par négociation haute fréquence sont autant d'algorithmes couramment utilisés pour prendre en charge des tâches d'ordre domestique et professionnel. Parmi les algorithmes, une catégorie spécifique, les Algorithmes d'Intelligence Artificielles (AIA), suscite de vifs et essentiels débats, en particulier depuis qu'Open AI a ouvert au public l'accès à ChatGPT 3.5, fin novembre 2022. C'est dans cette actualité du moment que la session 29 « Algorithmes et libre arbitre, quel avenir pour notre liberté ? » souhaite s'inscrire, en soumettant à discussion les capacités à la fois habilitantes et contraignantes des AIA et en interrogeant plus largement leurs impacts sur les modèles de gouvernance organisationnelle et politique.

De l'algorithme à l'intelligence artificielle

Les algorithmes peuvent être appréhendés comme des suites finies et ordonnées d'instructions destinées à résoudre un problème. Ainsi définis, ils renvoient notamment à une vision de la connaissance décomposable en règles, à partir desquelles un moteur d'inférence peut déduire des propositions. Les systèmes experts, dont Newell et Simon (1972) ont produit les bases théoriques, entrent dans cette perspective, dite symbolique (Cardon et al., 2018). Une autre approche, connexionniste, considère que toutes les connaissances ne sont pas explicables ni décomposables. Les

contributions développent alors des méthodes de réseaux de neurones, optimisées à partir d'algorithmes d'apprentissage dédiés (voir par exemple Rumelhart et al., 1986 ; LeCun, 2016).

Les Algorithmes dit d'Intelligence Artificielle (AIA ; Benbya et al., 2021) permettent « à des machines d'accomplir des tâches et de résoudre des problèmes normalement réservés aux humains » (LeCun, 2016). Dans une approche symbolique, les AIA abordent des problèmes prédéterminés à partir de règles connues, alors que d'un point de vue connexionniste, ils sont entraînaables et auto-apprenants (Pascal et Godé, 2023). L'AIA apprend lui-même à partir des données d'apprentissage qui lui sont fournies. C'est cet apprentissage automatique supervisé qui permet par exemple à l'IA de reconnaître un visage parmi des milliers d'autres ou de conduire une voiture. Les récents progrès de l'IA permettent également d'associer des réseaux neuronaux multi-couches à des algorithmes d'apprentissage profond. Dans ce cadre, ce sont les règles du jeu qui sont fournies à l'algorithme d'IA, qui va ensuite générer ses propres données d'apprentissage.

Ces capacités d'apprentissage ouvrent le champ des possibles. Les AIA peuvent ainsi conseiller un programme TV, préconiser un traitement médical, prédire le comportement de consommateurs, ou encore produire des recommandations de justice. Parmi les différents types d'algorithmes d'apprentissage automatique, une nouvelle génération fait l'actualité ces derniers mois : les algorithmes d'IA génératives, qui se spécialisent dans la création de nouveaux contenus à partir de données existantes (texte, images, audio et codes) – et non plus dans les seules analyses et classifications des données et règles.

Les Algorithmes d'Intelligence Artificielle : saisir les opportunités, gérer les menaces

Les AIA suscitent de fortes attentes de la part des utilisateurs qui s'en emparent largement au quotidien tant leur efficacité exécutoire est importante. Légitimement, ces mêmes acteurs s'inquiètent cependant des risques associés à leurs usages : les résultats produits peuvent en effet contenir des biais algorithmiques d'envergure – issus des biais cognitifs et de données – qui rejaillissent sur les préconisations produites. Les risques liés au respect de la vie privée méritent également l'attention au regard des volumes massifs de données, parfois personnelles, collectées et traitées. Enfin, les utilisateurs interrogent le risque de dépendance aux AIA, qui viendraient décider à leur place et réduire leurs marges de manœuvre. Plus globalement, c'est la question de l'exercice du libre arbitre qui est posée.

Par exemple, le manque de transparence des AIA est aujourd'hui discuté. Ces derniers souffrent en effet d'un certain niveau d'opacité qui peut faire douter les utilisateurs et freiner leur adoption. Faut-il ouvrir la boîte noire des AIA pour rendre explicables leurs préconisations et favoriser transparence et confiance ? Se développe récemment une approche allant dans ce sens : appelée intelligence artificielle explicable (Explainable Artificial Intelligence – XAI), elle vise le développement de modèles et méta-modèles permettant de comprendre comment et pourquoi l'AIA conclut au résultat proposé (voir par exemple, Schoenborn et al., 2019). L'explicabilité des modèles d'AIA adresse un des enjeux scientifiques majeurs autour des AIA, dont il est important de se saisir.

D'autres questionnements peuvent porter sur les contextes organisationnels et sociétaux dans lesquels les AIA fonctionnent et sont exploités. C'est par exemple le cas des débats autour des cadres de gouvernance à mettre en œuvre pour réguler les menaces potentielles de l'algocratie (voir par exemple, Bersini, 2023). C'est alors la question du type de société démocratique à construire et promouvoir avec les AIA qui est posée (Ménissier, 2022)

Références

- Benbya, H., Pachidi, S., & Jarvenpaa, S. (2021). Special issue editorial: Artificial intelligence in organizations: Implications for information systems research. *Journal of the Association for Information Systems*, 22(2). <https://doi.org/10.17705/1jais.00662>
- Bersini, H. (2023). *Algocratie : Allons-nous donner le pouvoir aux algorithmes ?* De Boeck Sup.
- Cardon, D., Cointet, J., & Mazières, A. (2018). La revanche des neurones : l'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*, 211, 173-220. <https://doi.org/10.3917/res.211.0173>
- LeCun, Y. (2016). Les enjeux de la recherche en intelligence artificielle. *Interstices*. <https://interstices.info/les-enjeux-de-la-recherche-en-intelligence-artificielle/>
- Pascal, A., & Godé, C. (2023). Les compétences transversales à l'heure de l'intelligence artificielles, in Zerbib, R. et Martinache, G. (coord.), *Formation : comment relever le défi des compétences transversales ?* Editions EMS, à paraître.
- Ménissier, T. (2022). L'IA, un artefact technologique porteur de promesses d'amélioration et riche de ses zones d'ombre, *Quaderni*, 105 | 2022, 9-18.
- Newell, A., & Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Rumelhart, D.E., Hinton, G., & Williams, R.J. (1986). Learning representations by back-propagating errors, *Nature*, 323, 533-536. <https://doi.org/10.1038/323533a0>
- Schoenborn, J.M., et Althoff, K.D. (2019). Recent Trends in XAI: A Broad Overview on current Approaches, Methodologies and Interactions. In *Case-Based Reasoning for the Explanation of intelligent systems (XCBR) Workshop*, vol. 2567. <https://ceur-ws.org/Vol-2567/paper5.pdf>