

Un numéro spécial de la Revue Management et Avenir

Rédacteurs associés :

Sylvie Michel et Marc Bidan

sylvie.michel@u-bordeaux.fr et marc.bidan@univ-nantes.fr

Émancipation, aliénation, délégation : Analyse critique et questionnements éthiques des systèmes d'intelligence artificielle dans les organisations

Les racines de la théorie critique nous rapprochent des travaux d'Emmanuel Kant et en particulier de son ouvrage « La Critique de la Raison Pure » (1781). Kant établit par le raisonnement ce qui peut être connu par la raison : que puis-je connaître ? Bien des années plus tard, face à la montée du nazisme, la première génération de l'École de Francfort (Adorno et Horkheimer) a revisité cette théorie critique. Cette nouvelle approche poursuit néanmoins le même objectif. Il s'agit de montrer que l'idéologie constitue le principal obstacle à l'émancipation. Horkheimer définit une théorie comme étant critique lorsqu'elle vise à libérer les êtres humains des circonstances qui les asservissent.

Plus récemment, en s'appuyant sur les travaux de Huault, Golsorkhi, et Leca (2009), la perspective critique – recentrée sur les sciences de gestion et du management et sur les organisations - met en lumière les relations asymétriques de pouvoir et les diverses formes de structures de domination présentes dans les organisations. Cette approche remet en question l'ordre établi considéré comme « naturel ». Elle examine également les discours et les idéologies sous-jacents au *statu quo*, qui génèrent une distribution asymétrique du pouvoir. En complément des approches traditionnelles, la perspective critique en management offre un regard différent – complémentaire, alternatif - sur la manière dont les options stratégiques sont influencées.

Dans la logique de ce numéro spécial, il est important de souligner que l'une des caractéristiques fondamentales des études critiques en management réside dans le fait que les chercheurs doivent adopter une approche, un positionnement ou une démarche visant à déconstruire les concepts établis et à souligner la relativité des phénomènes organisationnels.

C'est pourquoi nous pouvons la mobiliser pour aborder l'Intelligence Artificielle (IA) ou plus largement les Systèmes d'Intelligence Artificielle (SIA). Pour rappel, l'IA a été présentée et introduite dans les années 1950 par John McCarthy et Marvin Minsky aux USA dans le cadre de leurs travaux au MIT. Puis, des chercheurs comme Alan Turing, Herbert Simon et John von Neumann ont contribué au développement des intelligences qui avaient pour mission « d'imiter » le fonctionnement de l'humain. C'est ensuite dans les années 1980, que les travaux ou projets entrepreneuriaux de Geoffrey Hinton, Yann Le Cun, Luc Julia, Arthur Mensch et Yoshua Bengio ont ravivé l'intérêt pour des IA de plus en plus proches de l'humain en particulier grâce aux avancées sur les réseaux de neurones, la *machine learning* et le *deep learning*. C'est enfin vers le milieu des années 2000 et la fin des années 2010 – et avec de nombreux rapports - que les progrès à la fois de la puissance de calcul des serveurs et de la capacité de transport des données ont mis l'IA, ses algorithmes et ses outils à la portée des utilisateurs et des décideurs dans les organisations !

La définition élaborée et publiée par le groupe d'experts de haut niveau sur l'IA – celui qui fut installé par la Commission européenne - insiste clairement sur trois éléments fondamentaux et caractéristiques des SIA que l'on peut formaliser ainsi :

1. Un SIA est un ensemble de techniques (algorithmes simples ou apprenants, reconnaissance d'image, traitement du langage naturel, etc.) permettant d'analyser de très importants volumes de données, de différentes natures, structurées ou non ;
2. Un SIA interagit avec son environnement pour l'appréhender à travers des capteurs et/ou le modifier ;
3. Un SIA a la capacité d'apprentissage et d'adaptation et, finalement, d'influencer fortement les comportements et les décisions.

Plus récemment, avec la montée en puissance des LLM (*Large Language Models*) et l'avènement spectaculaire de robots conversationnels propulsés par OpenAI (ChatGPT), Google (Gemini, Bard), Meta (Llama), Mistral (MistralAI), Claude 3 (Anthropic), la question des IA génératives (IAG), de leur usage, de leur mobilisation dans les organisations et de leur éthique est devenue incontournable. Les IAG ont en effet pour mission de générer, de fabriquer, de proposer du contenu vraisemblable tel que du code, du texte, du son, des images et des vidéos en mobilisant de grands modèles de langage. Les IAG – dès lors qu'elles « génèrent » la réponse et non plus seulement « accompagnent » vers la réponse - signent-elles la fin des moteurs de

recherche ? Le contenu proposé par l'outil est-il « la » réponse recherchée ou doit-il être retravaillé, corrigé, testé par l'humain ? Le prompt ou la conversation – une dizaine de prompts et réponses successifs – piloté par l'humain est-il pertinent, cohérent, spécifique, fiable ? Quid des hallucinations, des affabulations, des raccourcis, des discriminations issues des *dataset* et de l'apprentissage machine ? Le statut du contenu est-il clairement maîtrisé – plagiat, imitation, etc. - et notamment par le décideur ultime ? Toutes ces questions doivent rapidement être intégrées dans la fabrique de la décision au sein des organisations et notamment dans le couplage humain/robot ?

Ainsi, aujourd'hui, l'IA continue d'évoluer et ses applications se sont considérablement accrues, suscitant à la fois de nombreuses opportunités et des controverses (Véry & Cailluet, 2019 ; Benbya et al., 2021 ; Fuhrer, 2023).

Au travers de cet appel à article, les deux coordinateurs, le comité scientifique et le comité de lecture seront donc particulièrement sensibles à des propositions d'articles critiques oscillant autour des trois thèmes principaux – mais non exhaustifs - suivants :

- Le développement de perspectives critiques sur les transformations induites par l'IA au sein de nos organisations et de nos sociétés soulève des questions fondamentales. L'IA a le potentiel de remodeler le travail et son organisation, remettant en question le sens même du travail. Ces transformations touchent tous les secteurs : santé, éducation (Quinio & Bidan), tourisme, communication, défense, recherche, etc. Nous pouvons nous demander quelles sont les implications des IA dans l'ensemble de ces domaines (Chintalapati & Pandey, 2022 ; Dwivedi et al., 2023). Comment ces technologies sont-elles élaborées et intégrées dans les pratiques institutionnelles ? L'exemple de la crise - principalement liée à la différence d'idéologie avec ses partenaires - autour de l'éviction/réintégration de Sam Altman illustre cette problématique. Quelle est l'idéologie sous-jacente à l'implémentation des systèmes d'IA dans les organisations (Vesa & Tienari, 2022) ? Quels aspects de l'IA relèvent du mythe ou de l'idéologie dans l'imaginaire collectif (Benbya et al., 2020) ? L'IA va-t-elle redéfinir les relations de pouvoir au sein des entreprises, le management des hommes (Heyder et al., 2023) ? Va-t-elle modifier la répartition du pouvoir ou renforcer les asymétries existantes (Monod et al.) ? Assisterons-nous à un *statu quo* ?

• Les interrogations sur l'identification, la caractérisation et la compréhension des défis éthiques posés par l'IA (Michel et al., 2022). Il convient de se poser la question de savoir si les dilemmes éthiques posés par l'IA sont nouveaux ou similaires à ceux des technologies du numérique (Hamet et Michel, 2018). Cela signifie qu'il est strictement nécessaire dans un premier temps d'identifier ces grandes questions éthiques liées à l'IA (Tsamados et al., 2021 ; Zacklad & Rouvroy, 2022 ; Giarmoleo et al., 2024). Déjà des problématiques concernant la surveillance, la vie privée, la propriété intellectuelle, les biais et préjugés, la transparence, la question de la « vérité », ont été mises à jour. Il convient aussi de ne pas négliger l'impact environnemental et énergétique du développement des IA et des IAG (Kirkpatrick, 2023 ; Bidan et al., 2024). Ces questionnements semblent d'autant plus complexes qu'ils sont interdépendants et intrinsèquement liés (Rowe et al, 2020 ; Michel et al., 2023).

• L'impact des tentatives de régulation que ce soit au niveau local, national ou international qui fleurissent depuis 5 ans (Adams-Prassl, 2022 ; Fournier-Tombs, 2021 ; Commission IA, 2024). Au niveau organisationnel, les chartes déontologiques, les codes éthiques et autres « catalogue de bonnes pratiques » se multiplient. La question qui s'impose est donc : sont-ils efficaces ? Les comités d'éthique sont-ils performants ? Au niveau national et international, quels sont les intérêts et enjeux des acteurs impliqués ? Comment se définissent les rapports de pouvoir entre les divers groupes engagés ? Qui sont les innovateurs, les développeurs, les régulateurs (pour citer une célèbre formule) ? Quelles sont les lacunes de ces réglementations et comment les combler ?

Calendrier à retenir :

- *29 mai 2024 : parution de l'appel à article*
- *15 décembre 2024 : soumission des V1*
- *1 février 2024 : retour des évaluations*
 - *1 Avril 2025 : soumission des V2*
 - *15 mai 2025 : retour des évaluations*
 - *15 juin 2025 : soumission des V3*
- *1 juillet 2025 : parution du numéro spécial RMA/Systèmes d'IA*

Bibliographie indicative

Adams-Prassl, J. (2022). Regulating algorithms at work: Lessons for a 'European approach to artificial intelligence'. *European Labour Law Journal*, 13(1), 30-50.

Adorno, T. W., & M. Horkheimer (1947). *La Dialectique de la raison*, trad. E. Kaufholz, Gallimard, Paris, 1974 ; rééd. coll. Tel, n 82.

Bawack, R. E., Wamba, S. F., Carillo, K. D. A., & Akter, S. (2022). Artificial intelligence in E-Commerce: a bibliometric study and literature review. *Electronic markets*, 32(1), 297-338.

Benbya, H., Davenport, T. H., & Pachidi, S. (2020). Artificial intelligence in organizations: Current state and future opportunities. *MIS Quarterly Executive*, 19(4).

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), 10.

Bidan, M., Ferreboeuf, H, Marraud L., & Rodhain, F. (2024) Le numérique est une industrie de plus en plus lourde Reflets de la physique, 77. P. 151-154 DOI: <https://doi.org/10.1051/refdp/202477151>

Chintalapati, S., & Pandey, S. K. (2022). Artificial intelligence in marketing: A systematic literature review. *International Journal of Market Research*, 64(1), 38-68.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642.

Fournier-Tombs, E. (2021). Towards a United Nations internal regulation for artificial intelligence. *Big Data & Society*, 8(2), 20539517211039493.

Fuhrer, C. (2023). Intelligence Artificielle : que dit la recherche récente ? Une approche combinée bibliométrique et textuelle. *Management & Avenir*, 137, 89-111.

Giarmoleo, F. V., Ferrero, I., Rocchi, M., & Pellegrini, M. M. (2024). What ethics can say on artificial intelligence: Insights from a systematic literature review. *Business and Society Review*.

Hamet, J., & Michel, S. (2018). Les questionnements éthiques en systèmes d'information. *Revue française de gestion*, (2), 99-129.

Heyder, T., Passlack, N., & Posegga, O. (2023). Ethical management of human-AI interaction: Theory development review. *The Journal of Strategic Information Systems*, 32(3), 101772.

Huault, I., Golsorkhi, D., & Leca, B. (2009). *Les études critiques en management* (p. 350). Presses universitaires de Laval.

Kant, I. (1905). *Critique de la raison pure*. E. Flammarion.

Kirkpatrick, K. (2023). The Carbon Footprint of Artificial Intelligence. *Communications of the ACM*, 66(8), 17-19.

Michel, S., Gerbaix, S., & Bidan, M. (2022, October). A plea for choosing ex ante an ethical theoretical position for a relevant response to ethical issues posed by algorithmic systems. In *2022 IEEE - 3rd International Conference on Next Generation Computing Applications*, 1-6.

Michel, S., Gerbaix, S., & Bidan, M. (2023). Questionnement éthique des systèmes algorithmiques. *RIMHE: Revue Interdisciplinaire Management, Homme (s) & Entreprise*, (1), 105-116.

Monod, E., Lissillour, R., Köster, A., & Jiayin, Q. (2023). Does AI control or support? Power shifts after AI system implementation in customer relationship management. *Journal of Decision Systems*, 32(3), 542-565.

Ngwenyama, Ojelanki and Rowe, Frantz (2024) "Should We Collaborate with AI to Conduct Literature Reviews? Changing Epistemic Values in a Flattening World," *Journal of the Association for Information Systems*, 25(1), 122-136. DOI: 10.17705/1jais.00869 Available at: <https://aisel.aisnet.org/jais/vol25/iss1/5>.

Quinio, B., & Bidan, M. (2023). ChatGPT: Un robot conversationnel peut-il enseigner?. *Management & Datascience*, 7(1).

Rowe, F., Ngwenyama O. & Richet, J.L. (2020) Contact-tracing apps and alienation in the age of COVID-19, *European Journal of Information Systems*, 29:5, 545-562, DOI: 10.1080/0960085X.2020.1803155

Tsamados, A., Aggarwal, N., Cowls, J., Morley, J., Roberts, H., Taddeo, M., & Floridi, L. (2021). The ethics of algorithms: key problems and solutions. *Ethics, governance, and policies in artificial intelligence*, 97-123.

Véry, P., & Cailluet, L. (2019). Intelligence artificielle et recherche en gestion. *Revue française de gestion*, 45(8), 119-134.

Vesa, M., & Tienari, J. (2022). Artificial intelligence and rationalized unaccountability: Ideology of the elites?. *Organization*, 29(6), 1133-1145.

Zacklad, M., & Rouvroy, A. (2022). L'éthique située de l'IA et ses controverses. *Revue française des sciences de l'information et de la communication*, (25).

=>AIM2024 <https://aim2024.sciencesconf.org/program?lang=fr>

.....

Emancipation, Alienation, Delegation: Critical analysis and Ethical Questioning of Artificial Intelligence Systems in Organizations

The roots of Critical Theory can be found in the works of Emmanuel Kant, especially in his book "Critique of Pure Reason" (1781). Kant argued about what can be known by reason: What

can I know? Years later, faced with the rise of Nazism, the first generation of the Frankfurt School (Adorno and Horkheimer) revisited this critical theory. This new approach aims to show that ideology is the main obstacle to emancipation. Horkheimer defines a theory as critical when it seeks to free people from circumstances that enslave them.

More recently, building on the work of Huault, Golsorkhi, and Leca (2009), the critical perspective - focused on management and organization studies - highlights asymmetrical power relations and various forms of domination structures in organizations. This approach challenges the established order that is considered "natural" and examines the discourses and ideologies that underlie the status quo and produce asymmetrical power relations. Complementing traditional approaches, the critical perspective in management offers a different - complementary, alternative - view of how strategic options are influenced by the organization.

In the logic of this special issue, it is important to emphasize that one of the fundamental characteristics of critical studies in management is that researchers must adopt an approach or stance aimed at deconstructing established concepts and highlighting the relativity of organizational phenomena.

Therefore, we can mobilize it to address Artificial Intelligence (AI) or, more broadly, Artificial Intelligence Systems (AIS). For reference, AI was introduced in the USA in the 1950s by John McCarthy and Marvin Minsky during their work at MIT. Over the years, researchers such as Alan Turing, Herbert Simon, and John von Neumann contributed to the development of intelligence that aimed to "mimic" human functioning. In the 1980s, the work or entrepreneurial projects of Geoffrey Hinton, Yann Le Cun, Luc Julia, Arthur Mensch, and Yoshua Bengio revived interest in AIs that increasingly resembled humans, especially through advances in neural networks, machine learning, and deep learning. Finally, in the late 2010s - and with many reports in the mid-2000s - advances in server computing power and data transport capacity made AI, its algorithms, and tools accessible to users and decision makers in organizations!

The definition elaborated and published by the High Level Expert Group on AI - installed by the European Commission - clearly emphasizes three fundamental and characteristic elements of AI:

1. An AI is a set of techniques (simple or learning algorithms, image recognition, natural language processing, etc.) for analyzing very large volumes of data of various types, structured or unstructured;
2. An AIS interacts with its environment to perceive it through sensors and/or to modify it;
3. An AIS has the ability to learn, adapt, and ultimately influence behaviors and decisions.

More recently, with the rise of LLMs (Large Language Models) and the spectacular emergence of conversational robots from OpenAI (ChatGPT), Google (Gemini, Bard), Meta (Llama), Mistral (MistralAI), Claude 3 (Anthropic), the question of generative AIs (GAI), their use, their mobilization in organizations, and their ethics has become unavoidable. The mission of GAI is to generate, produce, and propose plausible content such as code, text, sound, images, and video by mobilizing large language models. Since GAI "generates" the answer and not just "assists" it, does it signal the end of search engines? Is the content suggested by the tool "the" desired answer, or does it need to be revised, corrected, tested by humans? Is the prompt or conversation - a dozen successive prompts and responses - human-driven, relevant, coherent, specific, reliable? What about hallucinations, fabrications, shortcuts, discriminations arising from data sets and machine learning? Is the status of the content clearly controlled - plagiarism, imitation, etc.? - and especially by the ultimate decision maker? All these questions need to be quickly integrated into the decision making processes within organizations, especially in the human/robot coupling!

Today, AI continues to evolve and its applications have significantly increased, generating both numerous opportunities and controversies.

Through this Call for Papers (CFP), the two coordinators, the Scientific Committee, and the Editorial Board will be particularly sensitive to critical article proposals that oscillate around the following three main, but not exhaustive, themes:

- Developing critical perspectives on the transformations that AI is bringing to our organizations and societies raises fundamental questions. AI has the potential to reshape work and its organization and to question the very meaning of work. These transformations affect all sectors: health, education, tourism, communication, defense, research, etc. What are the implications of generative AI in all these domains? How are these technologies being developed and integrated into institutional practices? What is the ideology behind the implementation of AI systems in

organizations? Will AI redefine power relations within organizations, human management? Will it change the distribution of power or reinforce existing asymmetries? Will we witness a status quo? What aspects of AI are myth or ideology in the collective imagination?