



L'émergence de l'IA conversationnelle comme autorité cognitive : perspectives éducatives et éthiques à l'ère de Grok

Amélie RAOUL
Université Paul Valéry, Montpellier

INTRODUCTION

Depuis 2023, **Grok**, IA conversationnelle développée par xAI d'Elon Musk, se présente comme une alternative « moins politiquement correcte » aux modèles dominants comme ChatGPT. Son intégration à la plateforme X (anciennement Twitter), et l'aura de son créateur, lui confèrent une **légitimité idéologique particulière**, en rupture avec les discours de neutralité affichés par d'autres IA. Cette contribution interroge **l'émergence d'une autorité cognitive automatisée**, et ses effets dans l'espace public numérique. L'analyse adopte une posture critique ancrée en sciences de l'information et de la communication (SIC), attentive aux processus de médiation, aux logiques de légitimation du savoir, et aux implications éducatives de cette mutation.



03 DISCUSSION

Ce déplacement de l'autorité cognitive vers un système algorithmique non questionné soulève des enjeux majeurs :

- **Le jugement humain est progressivement désactivé** : l'habitude de consulter une IA comme Grok pour obtenir des réponses définitives **affaiblit l'exercice du doute méthodique**.
- **L'autorité algorithmique s'impose sans contestation**, d'autant plus que Grok assume ses biais idéologiques (« anti-woke »), ce qui lui permet de s'inscrire dans une posture **perçue comme plus transparente que les IA neutres de façade**.
- L'évolution de Grok, qui intègre désormais **des fonctions de recherche avec liens vers les sources**, ne modifie pas fondamentalement sa structure : **la logique de synthèse prime sur celle de transparence**, et le raisonnement algorithmique **reste opaque pour l'utilisateur moyen**.
- Le risque est celui d'une **clôture épistémique** : **le savoir devient une réponse prédéterminée**, non pas un **objet de débat, d'argumentation, ou de reconstruction contextuelle**.
- Cette autorité algorithmique, **contextualisée dans l'espace social numérique**, opère une **naturalisation de sa propre légitimité** : elle devient le point de référence par défaut, sans que soit questionnée la manière dont elle produit ses réponses.

Grok **ne peut être compris comme un outil neutre**. Son fonctionnement, sa place dans les échanges publics, et la manière dont elle est perçue construisent une **autorité cognitive hybride**, à la fois technologique, idéologique et sociale. Cette dynamique appelle une réponse éducative rigoureuse et adaptée.

05 BIBLIOGRAPHIE

BUCKINGHAM D. (2019). The media education manifesto. John Wiley & Sons.

BUTLER J., JAFFE S., BAYM N., CZERWINSKI M., IQBAL S., NOWAK K., RINTEL S., SELLEN A., VORVOREANU M., ABDULHAMID N. G., AMORES J., ANDERSEN R., AWORI K., AXMED M., BOYD D., BRAND J., BUSCHER G., CARIGNAN D., CHAN M., COLEMAN A., COUNTS S., DAEPP M., FOURNEY A., GOLDSTEIN D. G., GORDON A., HALFAKER A. L., HERNANDEZ J., HOFMAN J., LAY-FLURRIE J., LIAO V., LINDLEY S., MANIVANNAN S., MCILWAIN C., NEPAL S., NEVILLE J., NYAIRO S., O'NEILL J., POZNANSKI V., RAMOS G., RANGAN N., ROSEDALE L., ROTHSCHILD D., SAFAVI T., SARKAR A., SCOTT A., SHAH C., SHAH N. P., SHAPIRO T., SHAW R., SIMKUTE A., SUH J., SURI S., TANASE I., TANKELEVITCH L., TROY A., WAN M., WHITE R. W., YANG L., HECHT B. & TEEVAN J. (2023). Microsoft New Future of Work Report 2023.

MICROSOFT (2023) Rapport interne MSR-TR-2023-34.

DOMINIQUE C. (2015). À quoi rêvent les algorithmes. nos vies à l'heure des big data. Paris, Éditions du Seuil & La république des idées.

EUBANKS V. (2018). Automating inequality : How high-tech tools profile, police, and punish the poor. St. Martin's Press.

FLORIDI L. (2019). The logic of information : A theory of philosophy as conceptual design. Oxford University Press.

FRAU-MEIGS D. (2024). Algorithm literacy as a subset of media and information literacy : Competences and design considerations. Digital, 4(2), 512–528. DOI : 10.3390/digital4020026.

GILLESPIE T. (2014). The relevance of algorithms. Media technologies : Essays on communication, materiality, and society, 167(2014), 167.

HAIDER J. & SUNDIN O. (2022). Paradoxes of media and information literacy : The crisis of information. Taylor & Francis. DOI : 10.4324/9781003163237

KAHNEMAN D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

KARSENTI T. (2018). Intelligence artificielle en éducation : L'urgence de préparer les futurs enseignants aujourd'hui pour l'école de demain. Formation et profession, 26(3), 112–119.

LE DEUFF O. & ROUMANOS R. (2022). Enjeux définitionnels et scientifiques de la littératie algorithmique : entre mécanologie et rétro-ingénierie documentaire. tic&société, 15(2-3| 2ème semestre 2021-1er semestre 2022), 325–360. DOI : 10.4000/ticetsociete.7105.

LIVINGSTONE S. (2004). Media literacy and the challenge of new information and communication technologies. The communication review, 7(1), 3–14.

METZGER M. J. & FLANAGIN A. J. (2013). Credibility and trust of information in on-line environments : The use of cognitive heuristics. Journal of Pragmatics, 59, 210–220. Biases and constraints in communication : Argumentation, persuasion and manipulation, DOI : https://doi.org/10.1016/j.pragma.2013.07.012.

UNESCO (2023). Guidelines for the governance of digital platforms : safeguarding freedom of expression and access to information through a multi-stakeholder approach. Paris, France : UNESCO. Foreword by the Director-General of UNESCO, Audrey Azoulay. Includes bibliography. Available in multiple languages. Licensed under CC BY-SA 3.0 IGO.

ZAWACKI-RICHTER O., MARIN V. I., BOND M. & GOUVERNEUR F. (2019). Systematic re-view of research on artificial intelligence applications in higher education—where are the educa- tors? International journal of educational technology in higher education, 16(1), 1–27. DOI : 10.1186/s41239-019-0171-0

01 CONTEXTE ET METHODOLOGIE

Posture théorique

ancrée en SIC, à l'intersection de la philosophie de la technique, de la sociologie des usages numériques et de la critique des médiations algorithmiques.

Méthode

étude de cas centrée sur Grok, analysée dans sa double dimension discursive et technopolitique.

Corpus

interactions entre Grok et les utilisateurs sur X, analyse de ses modalités de réponse, observation des effets de validation sociale dans les threads.

Références théoriques principales

- Autorité cognitive (Gillespie, 2014)
- Littératie algorithmique (Le Deuff & Roumanos, 2022 ; Frau-Meigs, 2024)
- Editorialisation automatisée (Dominique, 2015 ; Haider & Sundin, 2022)

PHASE 1
Revue de littérature sur l'autorité cognitive, la littératie algorithmique et les médiations numériques.

PHASE 2
Étude de la réception discursive (commentaires, reprises, interactions sociales).

PHASE 3
proposition d'un cadre pédagogique à partir du concept de pédagogie de l'incertitude algorithmique.

Phases de l'étude exploratoire

02 RÉSULTATS/PLAN D'ACTION PROJET

1. Émergence d'une autorité cognitive algorithmique

L'étude de Grok met en évidence **l'émergence d'une autorité cognitive automatisée**, qui ne repose plus sur des figures humaines (enseignants, chercheurs, journalistes), mais sur un **dispositif algorithmique** intégré à l'espace public numérique.

Grok fonctionne comme un acteur discursif autonome, doté d'une autorité symbolique. Pour certains usagers, il ne s'agit plus d'un outil d'appui à la recherche, mais d'un **médiateur investi d'un pouvoir de validation** (Gillespie, 2014).

L'IA produit des **réponses synthétiques, immédiates, et souvent univoques**, qui se substituent à la délibération. Le jugement humain est **supplanté par la décision computationnelle**.

2. Un imaginaire de neutralité renforcée et Transformation de la délibération

Ce phénomène est accentué par un **imaginaire collectif de la neutralité algorithmique**, paradoxalement renforcé dans le cas de Grok, **dont les biais sont à la fois revendiqués et perçus comme plus sincères**.

La **structure même des réponses** produites par Grok opère une **clôture discursive** : les usagers ne questionnent plus, ils invoquent. La vérité n'est plus construite, elle est **délivrée par l'algorithme**.

Dans un contexte de **surcharge informationnelle** et d'**urgence cognitive** (Metzger & Flanagin, 2013), les usagers cherchent des raccourcis. Grok **répond à ce besoin par son ton affirmatif, sa fluidité et sa disponibilité**, autant de critères qui deviennent **heuristiques implicites de crédibilité**.

3. Parallèlement, l'étude met en évidence une performativité sociotechnique de l'IA au sein d'un espace de sociabilité :

- **L'ancrage de Grok dans la plateforme X** accentue sa performativité : ses réponses sont **reprises, commentées, partagées dans les fils de discussion**, s'insérant dans une dynamique de **validation sociale instantanée**.
- **L'interaction n'a plus lieu uniquement entre l'humain et la machine**, mais **dans un espace de sociabilité numérique** où l'IA agit comme **arbitre implicite du vrai et du faux**, souvent sans que ses critères de jugement ne soient interrogés.

Plan d'action pour la suite de l'étude :

1. Problématique : Comment l'intégration d'une IA générative comme Grok dans un espace de sociabilité numérique (X) redéfinit-elle les modalités de production, de diffusion et de validation du savoir ?

2. Hypothèses

H1. Grok est perçu comme une instance cognitive crédible en raison de sa performativité discursive, plus que de la véracité de son contenu.

H2. La dynamique sociale de la plateforme X renforce la légitimité de ses énoncés sans remise en question épistémologique.

H3. L'autorité cognitive de Grok transforme les comportements informationnels en favorisant une posture de consommation plutôt que de délibération.

3. Méthodologie

Analyse de contenu : Corpus de publications sur X intégrant des réponses de Grok (qualitative + quantitative).

Entretiens semi-directifs : Usagers ayant interagi avec Grok (journalistes, étudiants, enseignants, citoyens).

Expérimentation contrôlée : Comparaison de la réception de contenus similaires émis par Grok vs. experts humains.

Analyse linguistique : Étude de la forme des réponses de Grok (style, modalisation, usage des sources).