

Article

Ce que l'IA fait à l'écriture scientifique : ethnographie participative d'un hackathon interdisciplinaire

Nathan Ferret

ENS de Lyon

Patrice Abry

ENS de Lyon

Rémy Cazabet

Université Lyon 1

Philippe Gabriel

Avignon Université

Lucie Gournay

Université Paris Est Créteil

Jean-Philippe Magué

ENS de Lyon

Emmanuel Quemener

ENS de Lyon

Julien Thiburce

ENS de Lyon

Remerciements

Nous remercions l'ensemble des participantes et participants au hackathon pour leur implication dans la réussite de l'événement, la richesse et la qualité des échanges lors de la séance de restitution et de la table ronde qui a suivi.

Nous remercions également Martin Breuil pour son aide dans la réalisation de l'observation ethnographique et Diane Barthelemy, gestionnaire de l'IXXI, pour l'organisation logistique et pratique du hackathon.

Résumé

Cet article explore comment les usages collaboratifs des grands modèles de langage (LLM, pour *Large Language Model*) participent à transformer l'écriture scientifique, en s'appuyant sur une enquête conduite lors du hackathon « Ce que l'IA fait à la pratique de la science » (IXXI et CBPsmn, ENS Lyon, 8-9 juillet 2025). L'analyse s'ancre dans une démarche combinant ethnographie et cadres de la recherche participative (Merilhou-Goudard, 2016) : un ethnographe a co-construit, avec les personnes participantes, les protocoles d'observation et d'analyse, puis ces dernières ont activement contribué à la production et à l'interprétation des données. En mobilisant la perspective de la cognition distribuée (Hutchins, 1995), nous éclairons les dynamiques de groupe et les rapports aux modèles de langage où se joue la négociation de l'écriture scientifique entre les humains et la machine : élaboration collective des requêtes, itérations critiques sur les textes générés et validation partagée des résultats. Plutôt qu'une automatisation du processus scriptural des chercheurs et des chercheuses, les LLM redéfinissent les dynamiques de collaboration, les rapports disciplinaires et les régimes de scientificité, donnant naissance à une réflexivité médiée et co-produite par la génération automatique de texte.

Mots-clés : écriture scientifique ; grands modèles de langage (LLM) ; ethnographie participante ; collaboration interdisciplinaire

Abstract

This article explores the way large language models (LLMs) are transforming scientific writing practices. The approach adopted combines ethnography and participatory science. The ethnographic work took place during the hackathon “What AI Does to the Practice of Science” (IXXI and CBPsmn, ENS Lyon, July 8–9, 2025), and was led by an ethnographer who co-constructed data collection and analysis protocols with the participants who actively contributed to the production and the interpretation of the data. Drawing on the perspective of distributed cognition (Hutchins, 1995) as well as frameworks from participatory research (Houillier et Merilhou-Goudard, 2016), we show that every stage of the scientific writing becomes a negotiated space between humans and the machine: the collective crafting of prompts, critical iteration on generated texts, and the participatory validation of the resulting final presentation. Far from automating the process, LLMs redefine collaborative dynamics, forms of acculturation, and regimes of scientificity, giving rise to a mediated, co-produced reflexivity.

Keywords: Scientific writing practices; large language models (LLM); participatory ethnography; interdisciplinary collaboration.

Dans un contexte marqué par l'essor rapide de l'intelligence artificielle (IA), l'écriture scientifique devient un lieu d'interactions encore plus étroites entre humains et outils computationnels. L'introduction récente des grands modèles de langage (LLM), capables de générer du texte de manière autonome et contextualisée, modifie les pratiques scripturales des chercheurs et chercheuses en apportant une aide plurielle à toutes les étapes du processus de recherche (Altmäe *et al.*, 2023) — y compris pour les novices ou les locuteurs non natifs (Lin et Zhu, 2025). En dehors des questions éthiques¹, cette évolution engage une redéfinition plus profonde des compétences requises, notamment en termes de communication (Yi, 2021) : il ne s'agit plus seulement de rédiger, mais de savoir dialoguer avec une machine qui écrit. Cela suppose notamment la maîtrise de l'« ingénierie du prompt » — cet art de formuler des instructions qui permettent d'orienter la production du LLM (Marvin *et al.*, 2024).

L'ampleur et la nature de ces usages des LLM dans le champ scientifique se repèrent aujourd'hui principalement à travers les textes finalisés. Dans ces versions stabilisées des productions de recherche, les modèles sont de plus en plus mobilisés comme assistants de mise en forme, intervenant à différents niveaux du texte. Ainsi, la revue de littérature, étape centrale de l'argumentation scientifique, tend à devenir automatisée, notamment sous l'effet du contexte d'inflation des publications et d'injonctions à la productivité académique (Bouchard, 2025). Si cette assistance permet de faire l'économie d'un processus de recherche documentaire souvent jugé fastidieux, elle reste imparfaite : comme le montrent Tang, Duan et Cai (2024), les LLM produisent encore des références partielles ou sélectives, inexactes, voire fictives, lorsqu'on les sollicite pour générer des textes bibliographiques. De plus, ces générateurs tendent à reconduire des biais linguistiques, disciplinaires et géographiques, déjà présents dans leurs corpus d'entraînement (Skopec *et al.*, 2020).

Dans le cas des résumés, l'automatisation est déjà à l'œuvre à large échelle. Une étude récente sur plus de quinze millions de publications biomédicales révèle une augmentation marquée des formules textuelles typiques des LLM, et estime que près de 13 % des résumés publiés en 2024 ont été partiellement ou entièrement produits à l'aide de ces outils (Kobak *et al.*, 2025). Certains modèles sont également spécialisés dans la révision structurée du corps des articles. C'est notamment le cas de XtraGPT (Chen *et al.*, 2025), conçu pour accompagner l'amélioration itérative des différentes sections (méthode, discussion, résultats) à partir d'instructions ciblées. Des évaluations croisées par des lecteurs humains confirment que ce type d'outil peut améliorer la cohérence et la lisibilité des textes produits, sans pour autant permettre d'automatiser entièrement l'écriture (Chen *et al.*, 2025).

Mais les LLM ne se limitent pas à intervenir en aval, dans la mise en forme des textes finalisés : ils jouent désormais un rôle structurant en amont de la restitution, en s'invitant à différentes étapes du processus de recherche, de la formulation d'une problématique à la mise en place d'un protocole, de la génération d'hypothèses à l'interprétation et la discussion collective de résultats. Plusieurs revues de littérature et rapports récents documentent cette diversification des usages (Alan Turing Institute, 2022 ; Bommasani *et al.*, 2021 ; Luo *et al.*, 2025 ; The Royal Society, 2024). Les LLM y apparaissent comme des outils capables d'identifier des régularités et des relations au sein de jeux de données complexes (Paul *et al.*, 2021), de contribuer à des formes de simulation exploratoire (Sarker, 2021), voire d'agir comme ressources d'inspiration intellectuelle (Lin, 2023). Par rapport à ces travaux, notre article introduit deux inflexions principales : un ancrage empirique fondé sur l'observation directe, et un recentrage sur les pratiques d'écriture, en tant que dimension clé de construction et de transformation des savoirs en contexte numérique (Cummings, 2023).

Terrain, méthodologie et problématique

Si l'exploration des usages scientifiques des LLM repose majoritairement sur des analyses de corpus ou des métasynthèses, les observations en situation demeurent rares. Cela tient notamment à la nature souvent solitaire et peu visible du travail d'écriture scientifique, mais aussi à la dimension socialement sensible de l'usage de ces outils. Dans de nombreux contextes professionnels, recourir à un grand modèle de langage est perçu comme un signe de paresse, d'incompétence, ou de dépendance intellectuelle, et peut même susciter des soupçons de tricherie (Reif, *et al.*, 2025). Cette disqualification est d'autant plus marquée dans le champ scientifique, où la figure doxique de « *l'homo academicus* » comme être autonome, incarnant un idéal d'autocréation et de maîtrise individuelle du savoir (Bourdieu, 1984) s'oppose à l'idée de dépendance et substitution partielle de la « créativité » par un outil technique.

En replaçant les usages des LLM dans les différentes phases scripturales de la recherche, nous les abordons sous l'angle du rôle épistémique de l'écriture. Écrire ne consiste pas seulement à consigner ou à transmettre des résultats, mais bien à penser différemment. Autrement dit, en tant que technologie cognitive, l'écriture ne se contente pas d'enregistrer une pensée déjà formée : elle la rend possible, visible, manipulable (Goody, 1977 ; 1986), fonctionnant comme un « échafaudage sémiotique » qui structure et supporte le développement de la pensée (Cobley et Stjernfelt, 2015). Elle infléchit les formes de raisonnement, de catégorisation et de planification, et se situe ainsi au

cœur de la production de connaissances scientifiques : elle permet d'énoncer des problèmes et de formuler des hypothèses, participant au travail heuristique de construction de l'objet (Grogan, 2020) ; d'ordonner les résultats et de tester des enchaînements logiques en formalisant les étapes de l'analyse (Bazerman, 1988) ; et de stabiliser et coordonner les collectifs de recherche, grâce à des inscriptions textuelles qui servent de supports partagés pour l'action scientifique (Latour et Woolgar, 1986). Pour en prendre la pleine mesure, les transformations scientifiques induites par les LLM doivent donc être saisies à partir de celles qu'ils opèrent sur l'écriture elle-même.

Afin d'ainsi aborder l'écriture scientifique assistée par LLM à chaque étape de la recherche, nous nous inscrivons dans la lignée des travaux sur la cognition distribuée, lesquels conçoivent les processus cognitifs comme co-construits entre individus, artefacts, environnements et relations sociales (Hutchins, 1995). Nous explorons ces dynamiques à partir d'un terrain concret : le hackathon scientifique « *Ce que l'IA fait à la pratique de la science* »², organisé à l'École normale supérieure de Lyon les 8 et 9 juillet 2025 par l'Institut Rhône-alpin des Systèmes Complexes (IXXI³) et le Centre Blaise Pascal de Simulation et Modélisation Numérique (CBPsmn⁴). Cet événement avait pour objectif d'explorer les inscriptions possibles des LLM dans la diversité des pratiques scientifiques, en invitant les participants et participantes à expérimenter, de manière collaborative et interdisciplinaire, sur des projets de recherche en mobilisant ces outils. Il ne s'agissait pas de rédiger des articles complets — le temps imparti ne le permettant pas — mais d'expérimenter l'usage des LLM à différentes étapes clés : construction d'objets, analyse de données et formulation de résultats (textuels et graphiques) présentés collectivement en fin d'événement. Durant deux jours, 32 personnes (les six membres de l'organisation et 26 volontaires, dont un tiers de femmes), provenant de disciplines couvrant tout le spectre depuis les sciences exactes et expérimentales jusqu'aux sciences humaines et sociales, se sont ainsi engagées dans une démarche scientifique accélérée autour de cinq jeux de données fournis par le comité d'organisation ou proposés par les personnes participantes. Les LLM étaient mobilisés librement, comme outils de génération d'idées, de traitement de données et de rédaction. Cette situation, à la fois encadrée et ouverte, offrait un point d'entrée privilégié pour observer des usages scientifiques des LLM en train de se faire, au plus près des pratiques d'écriture, des interactions langagières et des ajustements collectifs.

Conformément aux principes de la science participative (Houillier et Merilhou-Goudard, 2016 ; *Ethnographie et Recherches Participatives*, 2025), la méthode ethnographique mise en œuvre repose sur une co-construction étroite entre l'ethnographe et les personnes étudiées. Un chercheur était spécifiquement chargé de l'observation détaillée de l'événement, tandis que les personnes

participantes, en tant qu'expertes de leur propre pratique, ont activement pris part à la production et à l'interprétation des données — récits, traces d'usage, retours réflexifs — et ont collectivement validé les analyses et les résultats produits. Cette démarche participative permet d'enrichir de manière significative le matériau recueilli. Elle augmente la profondeur et la diversité des données en intégrant les savoirs situés et les préoccupations propres aux participant·es (Chaveau, 2024), produit des connaissances immédiatement mobilisables par la communauté concernée (Eyraud, 2024), et contribue à réduire l'asymétrie entre la recherche et les sujets d'étude. En plaçant la réflexivité au cœur du dispositif, elle favorise une transparence méthodologique accrue et redéfinit le rôle de l'ethnographe, non plus comme simple observation, mais comme facilitation et médiation au sein d'une communauté de savoirs, instaurant finalement un « partenariat épistémique ».

À partir de ce cas et de cette démarche, nous proposons de considérer les LLM comme des dispositifs sociotechniques dont le fonctionnement ne réside pas uniquement dans une capacité computationnelle autonome, mais dépend avant tout des forces sociales, culturelles et organisationnelles qui les investissent, orientent leur usage et leur confèrent une place spécifique dans un collectif donné (Cardon, 2015). En retour, ces dispositifs participent à la production de nouveaux rapports sociaux : le contexte singulier du hackathon — temps court, équipes interdisciplinaires, forte intensité collaborative, et légitimation explicite de l'usage des LLM — permet de rendre visibles des dynamiques habituellement plus diffuses. Il met en lumière à la fois la plasticité sociale des LLM — leur capacité à s'adapter à des langages disciplinaires, des tâches et des attentes variées — et la manière dont les collectifs scientifiques ajustent leurs pratiques pour intégrer un agent non humain dans le processus d'écriture.

Ce terrain rend ainsi observables des situations de négociation concrète : chercheurs et chercheuses y testent, discutent ou rejettent les propositions du modèle, et réorganisent leurs échanges humains autour de cette nouvelle entité scripturale. Il offre un laboratoire précieux pour analyser la manière dont les pratiques d'écriture se transforment au contact de l'IA — et avec elles, le lien entre production de textes, production de savoirs, et production de collectifs.

En suivant cette ligne, nous structurons notre propos en deux temps. Dans un premier temps, nous examinerons la manière dont le LLM s'intègre aux dynamiques de groupe en mettant en lumière les formes d'écriture collaborative qu'il rend possibles. Nous montrons comment la présence de cet agent non humain reconfigure les interactions entre les membres d'un collectif, et participe à renouveler la pratique de l'interdisciplinarité. Dans un second temps, nous nous intéresserons au statut ambivalent du texte généré par les LLM, tel qu'il se construit tout au long du processus

scientifique. Entre outil de relance, source d'erreur, brouillon provisoire ou point de départ, ce texte est constamment mis à l'épreuve par les chercheurs et chercheuses. Nous montrerons alors comment cette mise en doute permanente structure la trajectoire de recherche et d'écriture scientifiques : elle appelle des stratégies de contournement des limites du modèle, une vigilance accrue vis-à-vis de ses effets discursifs, et fait finalement émerger une « intelligence artificielle assistée » (Basso Fossali, 2025) — un renversement de perspective où l'assistance en recherche vient plus de l'humain que de la machine.

I. La coécriture interdisciplinaire avec un LLM : médiations, ajustements et inventions collectives

L'interdisciplinarité est aujourd'hui souvent présentée comme une condition nécessaire pour exploiter pleinement le potentiel des grands modèles de langage, que ce soit pour croiser des savoirs, ouvrir de nouvelles hypothèses ou améliorer l'interprétabilité des opérations de la machine (Cartwright, 2023). Mais comment cette interdisciplinarité se fabrique-t-elle concrètement dans le cadre d'un usage collaboratif des LLM? Plutôt qu'être des facilitateurs transparents, ces outils imposent une série de contraintes techniques, discursives et organisationnelles qui obligent les membres d'un collectif à ajuster leurs modes d'interaction. L'écriture collaborative interdisciplinaire qui en résulte est le produit de ces ajustements croisés — à la fois avec la machine et entre les humains — dans une dynamique de médiations partagées.

1. Le LLM comme problème commun : coopérer pour surmonter les contraintes d'usages

L'utilisation d'un LLM en contexte scientifique pose une série de problèmes d'ordre technique et organisationnel, qui nécessitent d'emblée une coopération académique spécifique. Ces contraintes participent à la structuration du collectif de travail en devenant des objets concrets de coordination partagée.

Au plan infrastructurel et financier, tout d'abord. L'ensemble des modèles utilisés durant le hackathon s'appuyait sur l'infrastructure informatique du CBPsmn, grâce à l'environnement de déploiement SIDUS (*Single Instance Distributing Universal System*) (Quemener et Corvellec, 2013). Ce système permettait l'accès local à une série de modèles de langage en accès libre via le gestionnaire

Ollama (llama4, mistral-large, Deepseek, Qwen, etc.), tout en offrant la possibilité d'en installer d'autres au besoin. Ce choix d'infrastructure locale — hébergée sur le réseau de l'ENS — incarnait un enjeu fort : garantir une relative indépendance vis-à-vis des plateformes commerciales comme OpenAI, et poser les bases d'une autonomie de la recherche publique en matière d'outillage IA. Mais cette autonomie a un coût : les modèles accessibles localement étaient plus modestes en taille et en performance, et leur déploiement nécessitait un matériel spécialisé — des machines dont le coût d'acquisition est élevé (plusieurs dizaines de milliers d'euros).

En contrepartie, l'accès aux modèles via une interface Python ouvrait des possibilités d'usage avancé : appel aux modèles par interface de programmation (API), paramétrage fin (température, format de réponse), structuration contrôlée des sorties, ou encore automatisation de tâches répétitives — comme interroger un texte paragraphe par paragraphe dans une boucle plutôt que le soumettre d'un bloc. Ces usages, plus proches de « l'ingénierie du *prompt* » (Marvin *et al.*, 2024) que du simple dialogue en interface avec un agent conversationnel, exigeaient toutefois une maîtrise de l'écriture du code, en particulier en langage Python ; une compétence dont dépendait l'autonomie de chaque équipe vis-à-vis de la machine.

Sur le plan communicationnel, l'usage des LLM introduit une contrainte moins immédiatement visible, mais tout aussi structurante : par leur architecture même, ces outils favorisent une interaction individualisée avec un utilisateur unique. Pensés dans une logique de communication dyadique, les LLM s'inscrivent dans un paradigme d'échange personnalisé souvent centré sur une personne, qui devient l'interlocutrice principale du modèle. Cette individualisation peut renforcer la fluidité de l'interaction : elle favorise une continuité dans les instructions, une plus grande précision des réponses, et une rétroaction plus adaptée (Shin, 2021 ; Zamora, 2017). En ce sens, elle agit comme un lubrifiant cognitif de l'échange scriptural, en optimisant la qualité perçue de la réponse et l'engagement conversationnel.

Mais, en contexte collectif, cette efficacité dyadique a un effet paradoxal. En personnalisant le lien au modèle, elle tend à fragmenter la dynamique de groupe. Lors du hackathon, plusieurs équipes ont rapidement été confrontées à ce phénomène : l'un des membres se retrouvait en position d'interlocuteur principal avec le LLM, concentrant les interactions sur un seul « fil » d'échange. D'autres membres développaient souvent parallèlement des usages plus autonomes, lançant leurs propres requêtes sur des instances distinctes. Cette multiplicité de dialogues individuels, en dehors de la sphère partagée, pouvait nuire à la co-construction collective du raisonnement. Le LLM agit ici comme un dispositif individualisant de l'écriture scientifique, dont l'usage spontané tend à atomiser

la communauté discursive. Ce phénomène renvoie à ce que Hutchins (1995) décrivait comme un déplacement de la cognition collective vers une cognition distribuée mal synchronisée. Il rappelle aussi les observations de Suchman (2007) sur les interfaces conversationnelles, qui, bien qu'interactives, imposent un cadre relationnel asymétrique et difficilement mutualisable en groupe.

Ces contraintes techniques et organisationnelles ont dès lors induit une forme de coopération humaine spécifique, fondée sur des ajustements, des détournements et des « bricolages » au sens de Certeau (1980) par lesquels les personnes participantes ont adapté les outils à leur contexte d'usage. Cette coopération s'est d'abord manifestée à l'échelle de l'événement, où une coordination technique assurée par l'ingénieur système du CBPsmn, s'est révélée indispensable. En charge de la préparation de l'infrastructure, de la mise à disposition des modèles et du suivi de leur bon fonctionnement tout au long du hackathon, il a joué un rôle pivot dans l'opérationnalisation même de l'expérimentation. Sa présence a permis d'anticiper les problèmes de connexion, de répondre aux incidents techniques et d'ajuster les modèles aux contraintes des projets en cours, assurant ainsi un cadre stable pour l'expérimentation collective. Elle a également eu lieu en amont, l'équipe d'organisation ayant dû mettre à disposition des tutoriels techniques afin d'aider les participants et participantes à installer l'environnement nécessaire à l'accès aux modèles locaux.

À l'échelle des équipes, la coopération reposait sur la présence d'au moins une personne maîtrisant le langage Python, indispensable pour interagir efficacement avec les modèles via l'interface mise à disposition (*Visual Studio Code*). Cette compétence technique conférait à ces membres un statut privilégié dans les échanges avec la machine, centralisant l'accès au modèle et orientant les dynamiques d'échange au sein du collectif. La coordination passait alors par des dispositifs concrets de partage et de visualisation du texte — écriture au tableau, affichage sur écran, usage de papier — permettant de rendre lisibles les instructions comme les réponses générées, et d'ouvrir leur interprétation à la discussion collective. Dans ce contexte, l'écriture humaine, à la main ou projetée, jouait un rôle essentiel, permettant de rendre visible le texte en circulation, de spatialiser les échanges, et de soutenir une co-élaboration autour des propositions du modèle.

Un des enjeux centraux de cette collaboration humaine résidait alors dans l'élaboration collective des consignes au modèle. Leur écriture faisait systématiquement l'objet de discussions orales, souvent structurées autour d'un ou d'une membre qui tapait pendant que les autres formulaient des commentaires, proposaient des reformulations ou débattaient du contenu. Ces échanges portaient à la fois sur des dimensions scientifiques (choix des concepts, clarification d'une démarche, cadrage analytique) et sur des dimensions plus discursives, comme le registre d'adresse, la langue ou l'usage

de marques de politesse. Ce travail de formulation faisait émerger une voix collective, construite à travers les ajustements successifs d'un texte destiné à la machine.

Il ne s'agissait pas nécessairement de parler au LLM comme à un humain, mais la formulation des instructions tendait parfois à s'en rapprocher. Même si le modèle était majoritairement perçu comme un outil et que les instructions gardaient un statut d'instructions pour « faire faire », l'élaboration collective faisait émerger une forme d'anthropomorphisation pragmatique, où le LLM était ponctuellement traité comme un interlocuteur, capable de comprendre un propos, de l'interpréter et d'en tirer des conséquences non souhaitées à l'origine. Ce glissement est par exemple devenu manifeste lorsqu'un participant a proposé à son collègue de reformuler son explication directement « pour le modèle », en utilisant l'interface vocale : « redis-lui comme tu viens de me l'expliquer à moi », comme si l'explication pouvait être transposée à un quasi-collègue. Ce flou statutaire du LLM — entre automate, interface et agent collaboratif — structurait la manière dont les chercheurs et chercheuses s'accordaient sur les textes à proposer à l'outil.

Dans ce processus, les équipes ont fréquemment eu recours à plusieurs LLM en parallèle, dans une logique de bricolage stratégique. Ainsi, il arrivait que l'un des groupes interroge un modèle pour vérifier la clarté ou la compréhension d'une instruction avant de la transférer vers un autre modèle. Ce va-et-vient illustre ce que l'on peut appeler une assistance par recoupement, où la sortie d'un LLM sert à calibrer l'entrée pour un autre. De façon plus générale, tous les groupes ont spontanément mobilisé des interfaces conversationnelles externes (principalement ChatGPT ou Mistral) depuis leurs comptes personnels, gratuits ou payants. Ce recours massif aux robots conversationnels commerciaux, pourtant extérieurs à l'infrastructure déployée pour l'événement, révèle plusieurs mécanismes d'ajustement.

L'usage de ces agents permet d'échapper aux contraintes techniques associées à l'environnement Python ou à l'appel d'interface de programmation, en favorisant une interaction en langage naturel immédiatement accessible. Cela réduit le coût d'entrée, en particulier pour les non-spécialistes, et renforce l'autonomie d'usage. Par ailleurs, les interfaces conversationnelles déploient un « contexte » qui se construit au fil des échanges, qui permet d'ajuster progressivement les instructions sans devoir tout reformuler. Elles encouragent ainsi un rapport plus intuitif à la machine, centré sur la reformulation dialogique plutôt que sur l'encodage technique. Cette dynamique déplace, par-là, les compétences requises : ce n'est plus tant la maîtrise du code qui conditionne l'accès aux outils, mais une littératie discursive stricte. À cela s'ajoute la question, plus pragmatique, de la performance :

dans plusieurs cas, les réponses fournies par les modèles commerciaux ont été jugées plus cohérentes ou plus puissantes que celles générées localement.

Enfin, au fil de l'événement, on a pu observer l'émergence de pratiques visant à intégrer le LLM plus étroitement dans la dynamique collective, en enrichissant progressivement les instructions de métadonnées contextuelles. Certains groupes ont ainsi commencé à expliciter non seulement la tâche demandée, mais aussi les conditions concrètes de son exécution : nombre de personnes impliquées, répartition disciplinaire, objectifs de la session, contraintes de temps, etc. Cette forme « d'ingénierie contextuelle » (Dorotic *et al.*, 2024), qui dépasse la simple consigne, visait à simuler auprès du modèle une forme d'intelligence situationnelle, à rendre le LLM sensible à la configuration sociale dans laquelle il est mobilisé. Cette démarche traduit un effort humain explicite pour adapter le fonctionnement des modèles à une activité de recherche collaborative, en décalage avec les usages individuels pour lesquels ces outils ont été initialement conçus.

2. Le LLM comme médiateur entre les disciplines

Le hackathon observé réunissait des scientifiques issus de disciplines très diverses, tant à l'échelle de l'événement qu'au sein même des équipes. Comme on l'a vu précédemment, cette pluridisciplinarité — entre les spécialistes en informatique et les autres — s'est révélée indispensable pour mobiliser pleinement les potentialités du LLM. En retour, celui-ci a joué le rôle d'un vecteur de dialogue entre disciplines, en rendant possible un travail commun entre chercheurs et chercheuses aux horizons variés.

À l'exception d'un participant issu du monde industriel, les autres provenaient du tissu académique français, principalement régional, dont ils représentent la diversité en termes de disciplines — sciences du langage, sciences de l'éducation, sociologie, communication, humanités numériques, linguistique, traductologie, philosophie, sciences cognitives, neurosciences, biologie, chimie, physique, sciences des données, informatique —, de statuts — de la direction de recherche aux postes d'enseignement-recherche (professeur, maître de conférences), d'ingénierie (recherche et études), ainsi que les post-doctorants et post-doctorantes, doctorants et doctorantes, et étudiantes et étudiants de master —, et d'âges.

Les participants et participantes se sont répartis par groupe sur la base d'un choix volontaire, autour des cinq problèmes et jeux de données associés proposés, eux-mêmes rendant compte de la forte pluridisciplinarité de l'événement :

1 - Équipe « Récits de rêves »

(<https://github.com/IXXI-fr/Hackathon25/blob/main/données/Rêves.pdf>),

2 - Équipe « Traductions de *La Curée* (E. Zola) »

(<https://github.com/IXXI-fr/Hackathon25/blob/main/données/LaCurée.pdf>),

3 - Équipe « Visites guidées de musées »

(<https://github.com/IXXI-fr/Hackathon25/blob/main/données/VisitesGuidées.pdf>),

4 - Équipe « Comptes rendus étudiants de formation permanente »

(<https://github.com/IXXI-fr/Hackathon25/blob/main/données/ComptesRendus.pdf>),

5 - Équipe « Dynamique de la pandémie de Covid19 »

(<https://github.com/IXXI-fr/Hackathon25/blob/main/données/Covid.pdf>).

Au-delà de la recommandation d'inclure pour chaque équipe au moins une personne familière de l'usage de la programmation en python, la constitution des groupes s'est traduite à la fois par une forte interdisciplinarité et une homogénéité de statuts et d'âges :

- Le groupe 1 (une femme et six hommes) ne comptait que deux titulaires et rassemblait des compétences en physique, informatique, neurosciences, neuro-économie, sociologie, littérature et sciences cognitives.
- Le groupe 2 (trois femmes et trois hommes) ne comptait qu'une doctorante et un ingénieur en CDD pour quatre titulaires. Il rassemblait des compétences en informatique, biogéographie, linguistique contrastive, traductologie, sciences du langage et sciences cognitives.
- Le groupe 3 (deux femmes et quatre hommes) était constitué de quatre ingénieurs de recherche ou d'étude, une doctorante et un étudiant et rassemblait des compétences en sciences du langage, humanités numériques, communication, sciences des données et programmation.
- Le groupe 4 (deux femmes et quatre hommes) comptait autant de personnes permanentes que non permanentes, l'unique participant non universitaire et rassemblait des compétences en sciences cognitives, de l'éducation, du langage, des données et linguistique.
- Le groupe 5 (cinq hommes) ne comptait qu'un étudiant et rassemblait des compétences en biologie, informatique, physique, sciences des données, philosophie et sciences de l'éducation.

- Enfin, l'équipe d'organisation (une femme et cinq hommes) était elle-même fortement interdisciplinaire, constituée d'une gestionnaire, d'un ingénieur système, d'un linguiste-informaticien, d'un sociologue, d'un physicien et d'un expert en sciences des données.

Dans ce contexte, le LLM a, dès lors, joué un rôle de médiateur entre ces disciplines et statuts, en facilitant l'explicitation des savoirs, la mise en langage des hypothèses, et la coordination des démarches analytiques — autant de fonctions qui ont permis de maintenir un espace de travail commun malgré la diversité des cultures scientifiques.

La nécessité de dialoguer avec un modèle mis en partage a d'abord conduit les personnes participantes à reformuler leurs objets et leurs approches disciplinaires dans un langage opératoire commun — une sorte de *lingua franca* façonnée par les exigences de la machine. Cette transposition a contribué à forger une grammaire partagée, où la rédaction d'une requête supposait de rendre explicites des présupposés, de clarifier une démarche ou de définir un terme technique en vue d'une bonne compréhension par le modèle. En centralisant les échanges dans un fil unique, le LLM fournissait également une mémoire accessible et stabilisait la progression collective du raisonnement. Ce centre de gravité textuel permettait de rassembler et de faire dialoguer, dans un même espace dialogique, des contributions issues de spécialités différentes.

L'exemple du groupe travaillant sur *La Curée* de Zola (groupe 2) en offre une bonne illustration. Leur objectif était d'identifier, parmi trois traductions anglaises du roman, celle qui était la plus littérale. À partir des bitextes alignés produits avec l'outil Logiterm, les deux spécialistes en traductologie ont dû opérationnaliser la notion experte de littéralité — difficilement mobilisable telle quelle par un modèle de langage — en la traduisant en indicateurs quantitatifs et qualitatifs qui puissent être formulables dans des instructions au LLM. Ce travail a nécessité une formalisation collective, rendant le concept intelligible à la fois par la machine et pour les membres non spécialistes de l'équipe. La littéralité a ainsi été discutée en termes de densité lexicale, d'écarts de longueur entre l'original et la traduction, ou encore de filiation syntaxique. Un des critères quantitatifs proposés a par exemple été énoncé ainsi : « si la traduction contient plus (ou moins) de mots que l'original, cela peut constituer un premier indice de littéralité ». Ces échanges ont été consignés sur tableau, discutés collectivement, puis traduits en requêtes structurées adressées au modèle. Ce processus illustre comment un concept disciplinaire spécialisé peut être reformulé dans un langage opératoire commun, par le biais de la médiation imposée par le fonctionnement du LLM.

Par ailleurs, les réponses produites par le LLM, formulées en langage naturel, jouaient souvent un rôle de vulgarisation entre disciplines. En rendant accessibles des contenus spécialisés à des non-spécialistes, elles facilitaient la discussion collective et les arbitrages au sein des groupes. Ce rôle s'est particulièrement manifesté dans l'équipe travaillant sur les comptes rendus d'étudiants, où l'un des débats portait sur la nature exacte de ces documents : s'agissait-il de simples descriptions factuelles ou intégraient-ils une dimension réflexive ? Cette distinction avait des conséquences directes sur la manière d'interpréter les formulations générées par le LLM. Le choix du terme à retenir pour orienter les analyses ultérieures — « description », « interprétation », « retour d'expérience », « analyse réflexive » — a ainsi fait l'objet d'échanges nourris, engageant à la fois une définition disciplinaire de l'objet et une stratégie de mise en opération compatible avec les capacités du modèle.

Ce travail d'interprétation et de formulation collective, guidé par la nécessité de produire et de lire un langage commun intelligible par la machine, a donc agi comme un vecteur de coordination interdisciplinaire. En exigeant une explicitation constante des intentions et des cadres d'analyse, le LLM rendait visibles les différences de positionnement épistémologique tout en favorisant leur articulation dans l'action.

II. Assister le LLM, dompter l'aléatoire : l'écriture scientifique comme négociation avec le texte généré

L'introduction des LLM dans le processus scientifique transforme en profondeur le rapport que les chercheurs et chercheuses entretiennent à leurs propres textes. Que signifie écrire avec un modèle qui génère du contenu à partir de données, sans que l'on saisisse précisément comment il les articule ? Fondés sur des calculs probabilistes et opérant selon des logiques en grande partie opaques, ces modèles soulèvent des questions cruciales quant au statut des textes produits : peut-on leur accorder une valeur de preuve ? Comment en juger la cohérence, la scientificité ou la reproductibilité ? Ces interrogations rejoignent les alertes formulées sur les obstacles à une intégration rigoureuse de ces technologies dans le champ scientifique, en raison de leur instabilité, de leur opacité, et du manque de transparence des entreprises qui les conçoivent (Haibe-Kains *et al.*, 2020). À partir du cas du hackathon, cette section examine ce que signifie adopter un usage véritablement *scientifique* des LLM, en analysant les postures réflexives, les discours critiques et ajustements pratiques par lesquels les personnes participantes s'en sont saisis en tant que chercheurs et chercheuses.

1. Du « fun » au « sérieux » : LLM, littératie scientifique et jugement sur la valeur des textes générés

Les chercheurs et chercheuses n'ont pas seulement mobilisé les LLM pour analyser leurs données, mais ont également dû négocier avec les textes que ces modèles produisaient, qu'il s'agisse de résultats d'analyse ou de formulations intermédiaires. L'introduction des LLM dans la recherche ne se limite alors pas à leur capacité d'ainsi produire du texte, mais transforme en profondeur les pratiques d'évaluation, de validation et de reconnaissance scientifique des productions scripturales. Ce déplacement engage ce que Lang, Drake et Olson (2006) nomment la littératie scientifique, entendue comme la capacité non seulement à comprendre les discours relatifs aux sciences et aux technologies, mais aussi à les évaluer de manière critique. Dans un contexte où les textes sont générés par un modèle probabiliste, l'enjeu devient double : il s'agit à la fois de savoir écrire avec un LLM et de savoir reconnaître dans ses productions ce qui relève — ou non — des standards de scientificité, afin de poursuivre ou de réorienter l'écriture⁵.

Durant l'événement, cette dynamique itérative entre génération, évaluation et réorientation a été omniprésente. Les participants et participantes ont mobilisé des schèmes souvent implicites pour classer les productions du modèle selon des critères de qualité, de rigueur ou de légitimité. Ces critères varient sensiblement selon les disciplines, les statuts, et les habitudes professionnelles. La composition interdisciplinaire des équipes — où peu de membres possédaient une expertise disciplinaire approfondie sur les données étudiées — accentuait le flou évaluatif concernant la qualité et les critères de scientificité, révélant ainsi des mécanismes ordinairement implicites de jugement. Dans certains cas, les textes générés étaient jugés trop « bavards », trop généraux ou trop lyriques pour être considérés comme véritablement scientifiques. L'usage du LLM dans un registre ludique, créatif, ou visuel — favorisé par le contexte même d'un hackathon comme espace d'expérimentation à faible enjeu académique — pouvait également être perçu comme éloigné des normes épistémiques attendues. Par exemple, l'équipe travaillant sur les récits de rêves, a voulu profiter d'un « temps mort » dans la démarche de recherche pour « tester » la génération d'illustrations à partir de ces récits : cette mise en image a été cadrée comme un temps de relâche, stimulée par la « curiosité » sans faire l'objet d'une réappropriation analytique, simplement « pour voir ce que ça donne » et partager un moment de convivialité autour de ces visualisations.

À l'inverse de ces usages spontanément qualifiés de « fun », d'autres configurations faisaient apparaître le LLM comme un outil « sérieux ». Lorsqu'il était étroitement encadré par des instructions

précises, appuyées sur des hypothèses clairement définies et des jeux de données rigoureusement structurés, le LLM n'apparaissait plus comme un générateur de texte flou, mais comme un instrument maîtrisé, au fonctionnement proche d'un logiciel d'analyse statistique familier. À l'inverse des images, les agrégats lexicométriques thématiques obtenus par l'équipe Rêves avec le LLM activaient des logiques de reconnaissance professionnelle spontanées : « ça a l'air pas mal du tout ça », « c'est propre, ça ressemble à une vraie » (à propos d'un graphique d'analyse factorielle généré par la machine). Dans ces cas, l'écriture scientifique n'était plus simplement confiée au LLM, mais préconstruite, « pré-mâchée », puis injectée dans un cadre formel laissant peu de place à l'aléatoire. Le modèle était alors perçu davantage comme un relais technique et objectif d'un raisonnement déjà établi, plutôt que comme un partenaire discursif facilement susceptible de produire des réponses biaisées.

Ce spectre d'usages — du « fun » au « sérieux », du génératif au calculé — révèle la pluralité des représentations de l'écriture scientifique à l'ère des LLM. Il donne à voir les seuils implicites à partir desquels un texte est reconnu comme légitime dans un contexte académique, et rend visibles les ajustements réflexifs opérés par les chercheurs et chercheuses pour évaluer ce qu'ils reconnaissent, ou non, comme scientifique dans les productions de la machine.

Ces seuils de scientificité n'étaient dès lors pas figés : ils s'inscrivaient dans le cours même de la recherche, où la frontière entre jeu exploratoire et rigueur méthodique est sans cesse redéfinie. Comme l'a montré Marilyn Strathern (2004), les pratiques scientifiques tolèrent, et même mobilisent, la co-présence de registres hétérogènes — ici, une alternance entre expérimentation ludique et validation exigeante. Dans plusieurs équipes, on a ainsi observé des phases de « relâchement » méthodologique où l'on « balance » l'ensemble des données pour voir ce qui ressort, où l'on teste la réactivité du modèle sur des requêtes inattendues, voire ironiques — par exemple, en lui demandant s'il connaît un article écrit par un membre du groupe présent. Ces moments, qui relèvent d'un usage joueur du LLM, précèdent souvent un recadrage plus strict, dans lequel les instructions sont affinées, les hypothèses resserrées, et les critères d'analyse réaffirmés. Ce jonglage avec le caractère génératif du modèle montre bien que la scientificité des textes produits n'est jamais donnée d'emblée, mais est sans cesse construite dans l'interaction avec la machine et les autres membres du collectif.

2. La machine sous surveillance : des ajustements itératifs aux capacités du LLM

Les usages du LLM observés lors du hackathon ne peuvent être dissociés du profil social et professionnel de ses utilisateurs et utilisatrices. Issus des sphères académiques supérieures, ils et elles ont mobilisé, dans leur rapport au modèle, l'ensemble des dispositions critiques acquises au cours de leur socialisation au champ scientifique. Leurs pratiques se caractérisaient par une forme d'hypervigilance, typique des usages « savants » des technologies informatiques (Granjon, 2022) : évaluation systématique des réponses générées, mise à l'épreuve des affirmations du modèle, retour récurrent aux sources de données, attention portée aux biais de représentation inhérents à ces outils. Ce rapport critique, en suscitant un scepticisme largement partagé quant à l'idée qu'un LLM puisse réellement automatiser la recherche, a directement influencé la manière d'interagir avec les modèles et d'interpréter les réponses qu'ils génèrent.

Comparés à des usagers plus ordinaires ou profanes — où le dialogue avec un LLM peut être perçu comme fluide, mais peu interrogé — ces participants et participantes se sont ainsi révélés particulièrement vigilants. Cette différence est d'autant plus marquée si l'on considère, comme le montre Benbouzid (2025) à partir d'une enquête auprès d'étudiants et étudiantes, que les publics les moins dotés en capital culturel — notamment ceux issus de milieux populaires — sont souvent moins outillés pour remettre en cause les réponses produites par ces systèmes. À l'inverse, dans le cadre du hackathon, la méfiance face aux formes rhétoriques du modèle a été explicitement formulée dès les premières heures. Les équipes se montraient attentives à son ton conversationnel, en particulier lorsqu'il prenait les accents du développement personnel ou d'une autorité bienveillante. Certaines allaient jusqu'à tester activement les limites du modèle, en variant les formulations pour évaluer la stabilité des réponses, ou en posant des questions volontairement piégeuses afin d'en sonder la robustesse. Dans le groupe 5 (Covid) par exemple, une série d'instructions à la formulation très technique a été testée afin de vérifier si le LLM pouvait maintenir un cadre bayésien d'estimation tout en effectuant un calcul explicite de la moyenne de la loi *a posteriori*. Cette requête a par la suite permis de remplacer la stratégie d'échantillonnage aléatoire coûteuse initialement proposée par le modèle.

Cette vigilance s'est déployée tout au long des différentes phases de la recherche. Au stade de la problématisation, les attitudes des personnes participantes étaient largement guidées par leurs représentations initiales des capacités du LLM. Avant même d'interagir concrètement avec l'outil, les hypothèses et les approches de travail étaient orientées par ce que les membres du groupe pensaient

que le modèle était — ou n'était pas — capable de faire. Cette projection a parfois joué un rôle de frein : dans plusieurs cas, les équipes ont d'abord cherché à reproduire, avec le LLM, des tâches déjà réalisables avec d'autres outils plus spécialisés qu'elles maîtrisaient, tels que des logiciels d'analyse statistique, de traitement de texte ou de visualisation de données. Le modèle se retrouvait ainsi mis en concurrence directe avec des solutions techniques plus robustes, dans une comparaison qui lui était souvent défavorable. Ce biais d'évaluation négative, alimenté par une forme d'hypervigilance scientifique, pouvait occulter le fait que le LLM est susceptible de produire des résultats moins « optimaux » d'un point de vue technique, mais qui relèvent de registres nouveaux, jusque-là difficiles à automatiser.

Au fil de l'expérimentation, cette posture initiale s'est progressivement infléchie. L'enchaînement d'essais, d'erreurs et de reformulations a permis une réévaluation plus fine des capacités réelles du modèle et a conduit à des usages génératifs plus poussés. Par exemple, un groupe a exploré l'utilisation des plongements lexicaux (*embeddings*) pour comparer deux ensembles de comptes rendus d'étudiants : le LLM a servi à générer des représentations vectorielles des textes afin de visualiser les différences sémantiques entre les deux séries, une approche heuristique que peu d'outils classiques permettent d'aborder avec autant de précision. Un autre groupe a utilisé le LLM pour produire des descriptions d'œuvres d'art non pas en termes techniques ou historiques, mais dans un registre subjectif et émotionnel, en s'appuyant sur des extraits de discours de visiteurs de musée intégrés dans les instructions — une manière de simuler une réception sensible d'une manière difficilement automatisable jusqu'ici.

D'une façon générale, pour toutes les équipes, des effets de surprise sont apparus à plusieurs reprises, notamment lorsque le LLM générait des représentations graphiques ou des reformulations synthétiques à la forme très « scientifique », qui donnaient à voir un résultat impressionnant au premier abord. Mais ces emballements initiaux étaient immédiatement suivis d'un retour réflexif : dans le groupe travaillant sur les données Covid par exemple, un graphique produit automatiquement à partir des données disponibles a d'abord suscité l'enthousiasme, avant que l'équipe ne remarque que la méthode statistique utilisée, bien que fonctionnelle, n'était pas optimale. Ces phases d'excitation temporaire, suivies d'un recadrage critique, témoignent d'un processus d'apprentissage collectif, où la recherche avance par la prise en compte de plus en plus juste des capacités réelles des modèles.

Conclusion

Le lien d'écriture au modèle est apparu, tout au long du hackathon, comme une phase critique du travail scientifique : un moment où se concentrent le raisonnement collectif, le dialogue interdisciplinaire et la coopération de groupe. Cette écriture, qu'elle soit adressée à un LLM ou issue de ses réponses, relève autant d'une opération technique de transmission d'informations que d'un exercice de réflexivité épistémique et méthodologique. Elle oblige en effet les chercheurs et chercheuses à formuler, expliciter et cadrer ce qu'ils et elles attendent — non seulement de la machine, mais aussi de leur propre démarche collective.

Ce geste s'inscrit ainsi dans la longue tradition de l'écriture comme technique de mise à distance de soi (Lahire, 2021). Par la mise en symboles normés, le texte permet de revenir sur ses propres opérations de pensée. Ici, cette réflexivité est médiée par la machine : d'un côté, les instructions servent à orienter, contraindre, clarifier ; de l'autre, les textes générés sont lus comme des miroirs déformants, dans lesquels les scientifiques cherchent à reconnaître leurs intentions, leurs cadres conceptuels, leurs exigences disciplinaires.

Dans ce cadre, déléguer au LLM des tâches d'écriture aussi centrales que la formulation, l'analyse ou la synthèse ne signifie pas automatiser le travail scientifique, mais en reconfigurer les modalités, en impliquant autant de travail humain qu'auparavant. Ce détour par une machine faillible, en permanence mise en doute, rend ainsi visible le rôle central de l'écriture humaine dans la production de connaissances par l'IA, et ouvre potentiellement la voie à une réflexivité plus aiguisée sur sa place renouvelée dans la recherche.

Notes de Fin

1. Ces questions sont largement débattues dans la sphère sociale et scientifique suite à l'apparition massive d'articles générés par LLM et les contre-offensives éthiques et algorithmiques récentes (voir, par exemple, la lettre à l'éditeur de Pal *et al.*, 2024 ; ou la publication d'un guide de bonnes pratiques par l'INSERM, Atlan *et al.*, 2025).

2. La présentation de l'événement est disponible à l'adresse suivante :

<https://www.ixxi.fr/agenda/evenements/hackathon-grands-modele-de-langage>

3. <http://www.ixxi.fr>

4. <http://www.cbp.ens-lyon.fr/doku.php>

5. Cette transformation touche également le contrôle par les pairs : des études estiment que 7 à 17 % des évaluations d'article en informatique sont déjà partiellement générées par LLM (Liang *et al.*, 2024). Celles-ci se caractérisent notamment par leur superficialité (mentions rares de sections spécifiques, manque de références précises) et posent la question du maintien de l'autonomie du champ scientifique dans la définition collective de ses normes épistémiques (Hosseini et Horbach, 2023).

Références

- Alan Turing Institute. (2022). *Review of digital research infrastructure requirements for AI*. https://www.turing.ac.uk/sites/default/files/2022-09/ukri-requirements-report_final_edits.pdf
- Altmäe, S., Sola-Leyva, A., et Salumets, A. (2023). Artificial intelligence in scientific writing: A friend or a foe? *Reproductive BioMedicine Online*, 47(1), 3–9. <https://doi.org/10.1016/j.rbmo.2023.04.009>
- Atlas, H., Bourgain, C., Chneiweiss, H., Eisinger, F., Vidal, C., *et al.* (2025). *Recommandations de bonnes pratiques suite à l'analyse des questions éthiques soulevées par l'utilisation de l'Intelligence Artificielle dans la recherche à l'Inserm*. [Guidedebonnespratiquesdel'IntelligenceArtificielleàl'Inserm/Février 2025,2025. inserm-04975393](https://www.inserm.fr/fr/guide-de-bonnes-pratiques-del-intelligence-artificielle-a-l-inserm-fevrier-2025)
- Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. University of Wisconsin Press.
- Basso Fossali, P. (2025). Rationalités correctives et intelligence artificielle assistée : Les doubles contraintes des humanités numériques. *Semiotica*, 2025(262), 71–109. <https://doi.org/10.1515/sem-2024-0189>
- Benbouzid, B. (2025). *Écrire à l'université à l'heure des IA génératives : égalité instrumentale, inégalité structurelle*. AOC. <https://aoc.media/analyse/2025/05/07/ecrire-a-luniversite-a-lheure-des-ia-generatives-egalite-instrumentale-inegalite-structurelle-2-2/>
- Bommasani, R., Hudson, D. A., Adeli, *et al.* (2021). *On the opportunities and risks of foundation models* (arXiv:2108.07258). arXiv. <https://doi.org/10.48550/arXiv.2108.07258>
- Bouchard, A. (2025). #WorkInProgress : IA générative et outils de recherche de littérature académique. *URFISTinfo*. <https://hal.archives-ouvertes.fr/hal-04960003>

- Bourdieu, P. (1984). *Homo academicus*. Les Éditions de Minuit.
- Calenda. (2025). *Actes du colloque Ethnographie et recherches participatives*.
<https://calenda.org/1088968>
- Cardon, D. (2015). *À quoi rêvent les algorithmes ? Nos vies à l'heure du big data*. Seuil.
- Cartwright, N. (2023). *Getting serious about statistics: Scientific method and the reproducibility crisis*. Cambridge University Press.
- Certeau, M. (de) (1980). *L'invention du quotidien*. Union générale d'éditions.
- Chauveau, H. (2024). *Accompagner l'acculturation entre acteurs associatifs et apprentis chercheurs : l'expérience de la Boutique des Sciences de Lyon en tant que tiers-veilleur*. Actes du colloque Ethnographie et recherches participatives. https://ethno-rech-part.sciencesconf.org/data/pages/Actes_colloque_Ethnographie_et_recherches_participatives.pdf
- Chen, N., HuiKai, A. L., Wu, J. et al. (2025). *XtraGPT: LLMs for human-AI collaboration on controllable academic paper revision* (arXiv:2505.11336). arXiv.
<https://doi.org/10.48550/arXiv.2505.11336>
- Cobley, P. & Stjernfelt, F. (2015). Scaffolding Development and the Human Condition. *Biosemiotics*, 8 (2), 291-304. <https://doi.org/10.1007/s12304-015-9238-z>
- Cummings, L. (2023). Writing processes in the digital age: A networked interpretation. In O. Kruse, C. Rapp, C. M. Anson, K. Benetos, E. Cotos, A. Devitt, & A. Shibani (Eds.), *Digital writing technologies in higher education: Theory, research, and practice* (pp. 485–497). Springer.
https://doi.org/10.1007/978-3-031-36033-6_30
- Dorotic, M., Stagno, E., & Warlop, L. (2024). AI on the street: Context-dependent responses to artificial intelligence. *International Journal of Research in Marketing*, 41(1), 113–137.
<https://doi.org/10.1016/j.ijresmar.2023.08.010>
- Ethnographie et Recherches Participatives. (2025). *Recherche participative en santé ou urbanisme : ethnographes et citoyens collectent et analysent ensemble des récits, cartographies, pratiques pour produire des solutions adaptées aux territoires*. Université Lyon 2. <https://www.univ-lyon2.fr/recherche/agenda-scientifique/ethnographie-et-recherches-participatives>
- Eyraud, B. (2024). *L'auto-ethnographie coopérative et coresponsable dans la pratique et la gouvernance d'une démarche de recherche citoyenne : L'exemple de la démarche Capdroits*. Actes du colloque Ethnographie et recherches participatives. <https://ethno-rech>

part.sciencesconf.org/data/pages/Actes_colloque_Ethnographie_et_recherches_participatives.pdf

- Goody, J. (1977). *The domestication of the savage mind*. Cambridge University Press.
- Goody, J. (1986). *The logic of writing and the organization of society*. Cambridge University Press.
- Granjon, F. (2022). *Classes populaires et usages de l'informatique connectée : Des inégalités sociales-numériques*. Presses des Mines.
- Grogan, K. E. (2020). Writing science: What makes scientific writing hard and how to make it easier. *The Bulletin of the Ecological Society of America*, 102(1), e01800.
<https://doi.org/10.1002/bes2.1800>
- Haibe-Kains, B., Adam, G.A., Hosny, A. et al. (2020). Transparency and reproducibility in artificial intelligence. *Nature*, 586, E14–E16. <https://doi.org/10.1038/s41586-020-2766-y>
- Hosseini, M. & Horbach, S.P.J.M. (2023) Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. *Res Integr Peer Rev* 8, 4. <https://doi.org/10.1186/s41073-023-00133-5>
- Houillier, F., et Merilhou-Goudard J-B., (2016). *Les sciences participatives en France : États des lieux, bonnes pratiques et recommandations, rapport*.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge. MIT Press.
- Kobak, D., González-Márquez, R., Horvát, E.-Á., et Lause, J. (2025). *Delving into LLM-assisted writing in biomedical publications through excess vocabulary*. arXiv.
<https://doi.org/10.48550/arXiv.2406.07016>
- Lahire, B. (2021). *Culture écrite et inégalités scolaires. Sociologie de l'« échec scolaire » à l'école primaire*. Presses universitaires de Lyon.
- Lang, M., Drake, S., et Olson, J. (2006). Discourse and the new didactics of scientific literacy. *Journal of Curriculum Studies*, 38(2), 177–188. <https://doi.org/10.1080/00220270500122539>
- Latour, B., et Woolgar, S. (1986). *Laboratory life: The construction of scientific facts* (2nd ed.). Princeton University Press.
- Liang, W., Izzo, Z., Zhang, Y., Lepp, H., Cao, H., et al. (2024). Monitoring AI-Modified Content at Scale : A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews. arXiv:2403.07183.
<https://doi.org/10.48550/arXiv.2403.07183>
- Lin, Z. (2023). Why and how to embrace AI such as ChatGPT in your academic life. *Royal Society Open Science*, 10, 230658 <https://doi.org/10.1098/rsos.230658>

- Lin, C. W., et Zhu, W. (2025). Divergent llm adoption and heterogeneous convergence paths in research writing. arXiv:2504.13629.
- Luo, Z., Yang, Z., Xu, Z., Yang, W., et Du, X. (2025). *LLM4SR: A survey on large language models for scientific research*. arXiv:2501.04306. <https://doi.org/10.48550/arXiv.2501.04306>
- Marvin, G., Hellen, N., Jjingo, D., et Nakatumba-Nabende, J. (2024). Prompt Engineering in Large Language Models. In I. J. Jacob, S. Piramuthu, & P. Falkowski-Gilski (Éds.), *Data Intelligence and Cognitive Informatics*, 387-402. Springer Nature. https://doi.org/10.1007/978-981-99-7962-2_30
- Pal, S., Bhattacharya, M., Islam, M. A., et Chakraborty, C. (2024). AI-enabled ChatGPT or LLM: a new algorithm is required for plagiarism-free scientific writing. *International Journal of Surgery*, 110(2), 1329-1330.
- Paul, D., Sanap, G., Shenoy, S., Kalyane, D., Kalia, K., et Tekade, R. K. (2021). Artificial intelligence in drug discovery and development. *Drug Discovery Today*, 26(1), 80–93. <https://doi.org/10.1016/j.drudis.2020.10.010>
- Quemener, E., et Corvellec, M. (2013). Sidus—the solution for extreme deduplication of an operating system. *Linux Journal*, 2013(235), 3–9. <https://dl.acm.org/doi/abs/10.5555/2555789.2555792>
- Reif, J. A., Larrick, R. P., et Soll, J. B. (2025). Evidence of a social evaluation penalty for using AI. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 122(19). <https://doi.org/10.1073/pnas.2426766122>
- Sarker, I. H. (2021). Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, 2, 420. <https://doi.org/10.1007/s42979-021-00815-1>
- Skopec, M., Issa, H., Reed, J., et Harris, M. (2020). *The role of geographic bias in knowledge diffusion: A systematic review and narrative synthesis*. *Research Integrity and Peer Review*, 5, 1–14. <https://doi.org/10.1186/s41073-019-0088-0>
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human–Computer Studies*, 146, 102551. <https://doi.org/10.1016/j.ijhcs.2020.102551>
- Strathern, M. (2004). *Partial connections*. Rowman & Littlefield Publishers.
- Suchman, L. A. (2007). *Human-machine reconfigurations: Plans and situated actions* (2nd ed.). Cambridge University Press.

Tang, X., Duan, X., et Cai, Z. G. (2024). *Large language models for automated literature review: An evaluation of reference generation, abstract writing, and review composition* (arXiv 2412.13612).

arXiv. <https://doi.org/10.48550/arXiv.2412.13612>

The Royal Society. (2024). *Science in the age of AI: How artificial intelligence is changing the nature and method of scientific research*. [https://royalsociety.org/-/media/policy/projects/science-](https://royalsociety.org/-/media/policy/projects/science-in-the-age-of-ai/science-in-the-age-of-ai-report.pdf)

[in-the-age-of-ai/science-in-the-age-of-ai-report.pdf](https://royalsociety.org/-/media/policy/projects/science-in-the-age-of-ai/science-in-the-age-of-ai-report.pdf)

Yi, Y. (2021). Establishing the concept of AI literacy: Focusing on competence and purpose. *JAHR--European Journal of Bioethics*, 12 (2), 353–368.

Zamora, J. (2017). I'm sorry, Dave, I'm afraid I can't do that: Chatbot perception and expectations. In *Proceedings of the 5th International Conference on Human Agent Interaction*, 253–256.

<https://doi.org/10.1145/3125739.3125766>