Résous ton problème avec un clin d'œil

L'IA a fourni des sources. L'une d'elles existe même.

Comment vérifier les citations et les sources générées par les outils d'intelligence artificielle



Max Paradox

L’IA a fourni des sources. L’une d’elles existe même.

Comment vérifier les citations et les sources générées par les outils d’intelligence artificielle

Max Paradox

Rozwiąż swój problem z przymrużeniem oka

ISBN: 9788384906309

ALL-IN Rafał Płonkowski



TABLE DES MATIÈRES

INTRO 5

Chapitre 1 - Une bibliographie venue du pays de l'imaginaire, ou pourquoi l'IA invente des sources 11

Chapitre 2 - Une citation entre guillemets, la vérité en vacances 21

Chapitre 3 - Le lien fonctionne, mais où mène-t-il ? Anatomie d'une fausse preuve 32

Chapitre 4 - « Mais l'IA l'a vérifié ! », ou les excuses les plus coûteuses du monde 42

Chapitre 5 - Ne cherchez pas une aiguille dans une bibliothèque, ou comment retrouver une source sans perdre la raison 54

Chapitre 6 - Quand la réponse correspond trop bien à vos attentes, ou le piège de vos propres convictions 65

Chapitre 7 - Le téléphone arabe en costume-cravate, ou comment les sources se perdent entre les personnes et les documents 76

Chapitre 8 - Un PDF peut tout contenir, ou quand l'IA lit mal de vrais documents 87

Chapitre 9 - Cinq minutes pour la vérité, ou comment réaliser un premier audit d'une réponse de l'IA 100

Chapitre 10 - De la citation à l'original, ou une enquête qui connaît une véritable conclusion 111

Chapitre 11 - Les chiffres ne mentent pas, mais leurs descriptions ont beaucoup d'imagination 122

Chapitre 12 - Construisez une bibliographie dont vous n'aurez pas à rougir 133

Chapitre 13 - Apprenez à l'IA à montrer des preuves plutôt qu'à faire bonne impression 145

Chapitre 14 - De la réponse de l'IA à une décision que l'on peut défendre 156

Chapitre 15 - Deux sources, trois vérités et une personne au bord du désespoir 169

Chapitre 16 - La source qui se met à donner des ordres, ou pourquoi il ne faut pas obéir à toutes les instructions d'un document 181

Chapitre 17 - Quand il est impossible de tout vérifier, ou l'art d'accepter honnêtement l'inachevé 193

Chapitre 18 - Une vérification qui ne s'arrête pas au clic sur « Envoyer » 205

Chapitre 19 - Votre propre radar à absurdités, ou comment ne pas retomber dans vos anciennes habitudes 218

Chapitre 20 - L'IA peut écrire avec assurance. Vous pouvez savoir d'où lui viennent ses informations 230

INTRO

Imaginez que vous prépariez un rapport important. L'échéance approche à la vitesse d'un livreur qui sonne précisément au moment où vous venez d'entrer sous la douche. Vous avez les données, vous avez les arguments, il ne vous manque que quelques sources solides. Vous lancez un outil d'intelligence artificielle et lui demandez des publications qui confirment une affirmation précise. Quelques instants plus tard, vous recevez cinq références bibliographiques impeccablement présentées. Les auteurs ont des noms, les articles ont des titres, les revues ont des numéros, et certaines publications possèdent même des identifiants DOI. L'ensemble paraît si professionnel que vous auriez presque envie d'imprimer la réponse, de la faire relier en cuir et de l'installer sur une étagère à côté de votre encyclopédie.

Il y a seulement un petit problème. Le premier article n'existe pas. Le deuxième existe, mais traite d'un tout autre sujet. Le troisième possède un véritable auteur, mais un titre inventé. Le quatrième renvoie vers une page qui ne fonctionne plus depuis longtemps. Quant au cinquième, il existe bel et bien, sauf que ses auteurs affirment à peu près le contraire de ce que l'IA vient de vous raconter.

Félicitations. Vous venez de recevoir une bibliographie dans laquelle la réalité fait une apparition exceptionnelle.

Le plus inquiétant, pourtant, n'est pas que l'intelligence artificielle soit capable d'inventer une source. C'est la manière incroyablement convaincante dont elle peut la présenter. Aucun avertissement ne s'affiche : « Attention, les noms ci-dessous pourraient être le fruit de l'imagination

littéraire d'un processeur. » Vous n'entendez aucun toussotement nerveux. Vous ne voyez aucune rougeur numérique monter aux joues de la machine. Vous recevez simplement une réponse calme, structurée, qui donne l'impression que son auteur revient de la bibliothèque après avoir passé trois jours à éplucher consciencieusement la littérature scientifique.

En réalité, un modèle de langage peut produire un texte qui ressemble parfaitement à une bibliographie correcte sans disposer d'éléments suffisants pour confirmer l'authenticité de chacune de ses références. Cette distinction est essentielle. Savoir rédiger une note bibliographique crédible ne signifie pas avoir retrouvé la publication, l'avoir lue et avoir vérifié qu'elle confirme réellement l'affirmation avancée. De la même manière, savoir dessiner une carte au trésor ne garantit pas qu'un coffre rempli d'or soit enterré sous le palmier indiqué. Parfois, on n'y trouve que la facture des vacances.

Et pourtant, utiliser l'IA est vraiment pratique. Elle permet d'organiser rapidement ses idées, de comprendre des sujets complexes, de préparer un premier brouillon ou d'identifier des pistes pour approfondir une recherche. Le problème commence lorsque nous confondons commodité et preuve. Une réponse paraît convaincante, alors nous cessons de nous demander d'où vient l'information. Nous recevons un lien, alors nous supposons qu'il mène vers un contenu fiable. Nous voyons des guillemets, alors nous croyons que les paroles citées ont réellement été prononcées.

C'est ainsi qu'on peut publier une citation que personne n'a jamais formulée, invoquer une étude inexistante ou transmettre à ses collègues une information erronée accompagnée de la mention « selon les experts ». Les

experts découvrent ensuite qu'ils ont participé à une étude dont ils n'avaient même jamais entendu parler. Ils sont probablement aussi surpris qu'une licorne découvrant son inscription officielle au registre des animaux d'élevage.

Ce livre n'a pas été écrit pour vous convaincre que l'intelligence artificielle ne sert à rien. Ce serait à peu près aussi raisonnable que de jeter une calculatrice parce que quelqu'un y a saisi les mauvais chiffres. L'objectif est d'apprendre à distinguer une aide à la recherche d'informations d'une véritable confirmation de leur exactitude. Vous pouvez utiliser intensivement l'IA, gagner du temps et conserver malgré tout le contrôle sur ce que vous acceptez finalement comme un fait.

Pour commencer, il faut comprendre une vérité quelque peu dérangeante : les sources erronées ne sont pas toutes entièrement inventées. Si chaque fausse référence renvoyait vers un article intitulé « Étude extrêmement importante que je viens tout juste d'inventer », la vie serait considérablement plus simple. Mais vous rencontrerez plus souvent des erreurs subtiles. Une publication authentique est attribuée au mauvais auteur. Un véritable rapport se voit affublé d'une année de publication incorrecte. Une paraphrase est présentée comme une citation mot pour mot. Un lien mène à une page parfaitement authentique, mais son contenu ne confirme absolument pas la conclusion avancée.

Voilà pourquoi demander simplement « Cette source existe-t-elle ? » ne suffit pas. Il faut également déterminer si le document indiqué est le bon, si son contenu a été rapporté fidèlement, si son auteur se prononce réellement sur la question concernée et si la conclusion tirée ne dépasse pas ce que permettent d'affirmer les preuves disponibles. Ce sont plusieurs opérations distinctes, même

si, dans l'urgence, il est tentant de les ranger toutes dans un même tiroir étiqueté « Vérifié ». Le tiroir paraît bien ordonné. Son contenu, pas nécessairement.

Dans la pratique, nous considérerons les réponses de l'IA comme un point de départ pour vérifier les informations, et non comme un certificat automatique de vérité. Vous apprendrez à reconnaître les signaux d'alerte dans les bibliographies, à retrouver les documents originaux, à contrôler les noms des auteurs et les dates, à vérifier les identifiants des publications et à replacer les citations dans leur véritable contexte. Vous découvrirez également pourquoi des différences apparemment minuscules entre « l'étude a démontré », « les auteurs suggèrent » et « une observation a été faite dans certaines conditions » peuvent transformer complètement le sens d'une information.

Cela ne signifie pas que vous deviez devenir détective bibliographique professionnel. Vous n'avez besoin ni d'un bureau tapissé de livres, ni d'une loupe, ni d'une musique dramatique chaque fois que vous ouvrez un nouvel onglet de votre navigateur. Vous avez besoin d'une méthode permettant de savoir ce qu'il faut vérifier en premier, où chercher une confirmation et à quel moment les preuves disponibles sont insuffisantes. Sinon, vous risquez de consacrer autant de temps à vérifier le titre d'un article sans grande importance qu'à contrôler une information dont dépend une décision essentielle.

Toutes les situations n'exigent d'ailleurs pas le même degré de rigueur. Si vous préparez une note informelle pour votre usage personnel, vous pouvez accepter une plus grande part d'incertitude, à condition de ne pas la dissimuler. En revanche, si vous rédigez une publication scientifique, un support pédagogique, un document

professionnel ou un texte concernant la santé, le droit ou les finances, les conséquences d'une erreur peuvent être beaucoup plus sérieuses. Dans ces circonstances, vérifier les sources n'est pas une précaution excessive. Cela fait partie de votre responsabilité envers les personnes auxquelles vous transmettez l'information.

Une autre compétence importante consistera à savoir reconnaître que vous n'avez pas réussi à confirmer quelque chose. Cela paraît évident, mais les outils modernes nous ont habitués à obtenir des réponses immédiates. Une case vide dans une bibliographie ressemble presque à un échec personnel. Pourtant, affirmer honnêtement « Je n'ai trouvé aucune confirmation fiable de cette citation » a bien plus de valeur qu'une référence parfaitement formatée qui ne mène nulle part. L'absence de preuve ne démontre pas toujours qu'une affirmation est fausse, mais elle ne permet certainement pas de la présenter comme vérifiée.

Avec le temps, vous commencerez à comprendre que la compétence la plus importante ne consiste pas à prendre l'IA en flagrant délit d'erreur. Elle consiste à poser les bonnes questions et à distinguer une réponse d'une preuve. Un modèle peut vous aider à trouver une piste, à identifier des publications possibles, à proposer des mots-clés de recherche et à organiser les résultats. Mais la confirmation définitive nécessite de consulter les documents appropriés, de préférence auprès de leur éditeur d'origine ou d'une autre source fiable qui les met à disposition.

Si vous avez déjà copié une référence bibliographique sans ouvrir le document correspondant, cru une citation simplement parce qu'elle semblait intelligente ou considéré qu'un lien suffisait à prouver une affirmation, vous êtes au bon endroit. Nous n'allons pas organiser le procès public de

votre navigateur. Nous allons apprendre quelque chose de beaucoup plus utile : profiter des possibilités de l'IA sans lui abandonner le pouvoir de décider de ce qui est vrai.

Car l'intelligence artificielle peut produire en quelques secondes une liste de dix sources. Votre mission n'est pas de vous émerveiller devant la longueur de cette liste. Votre mission est de savoir combien de ces dix références méritent réellement de figurer dans votre texte.

Et s'il s'avère qu'une seule existe, au moins saurez-vous laquelle.

Chapitre 1 - Une bibliographie venue du pays de l'imaginaire, ou pourquoi l'IA invente des sources

Lundi, 9 h 17. Dans une salle de réunion, vous présentez les résultats d'une analyse préparée la veille au soir. Tout se passe à merveille. Les graphiques sont lisibles, les conclusions raisonnables et la présentation comporte même une diapositive consacrée aux sources, ce qui, dans l'univers des présentations d'entreprise, passe parfois pour une marque de distinction aristocratique. À un moment donné, quelqu'un vous demande d'où provient l'information concernant l'efficacité de la solution présentée. Sans la moindre hésitation, vous citez une publication scientifique que l'intelligence artificielle vous a proposée la veille.

Votre interlocuteur saisit le titre de l'article dans un moteur de recherche. Rien. Il vérifie le nom de l'auteur. Il trouve bien un chercheur, mais celui-ci travaille dans un domaine complètement différent. Il essaie de retrouver la revue scientifique. Cette dernière existe effectivement, sauf que le numéro indiqué ne contient aucune publication ressemblant à celle que vous venez de mentionner. Le silence s'installe dans la salle. Même le vidéoprojecteur, qui produisait jusque-là un bruit évoquant le décollage d'un petit avion, semble soudain fonctionner plus discrètement.

Vous venez de découvrir que votre bibliographie possède davantage d'imagination que le service création.

La situation est particulièrement désagréable, car l'erreur ne ressemblait pas à une erreur. Aucune faute de frappe, aucun mot incongru, aucune phrase suspecte. La référence

comportait un auteur, une date, un titre, le nom d'une revue et un numéro de volume. Tous les éléments étaient exactement à leur place. Il ne manquait qu'un minuscule détail : une véritable publication que quelqu'un aurait pu consulter.

Il est important de comprendre pourquoi de telles situations se produisent. Sans cette compréhension, on risque d'adopter l'une de deux positions extrêmes. La première consiste à considérer l'intelligence artificielle comme un bibliothécaire génial auquel il faudrait accorder une confiance absolue. La seconde la présente comme un escroc numérique qui chercherait délibérément à nous ridiculiser. Ces deux visions ont un certain panache, mais aucune ne nous aide à utiliser correctement cet outil.

Un modèle de langage n'a pas besoin d'avoir réellement l'intention de vous tromper pour produire une information fausse. Il peut simplement générer une réponse correspondant aux structures et aux régularités apprises pendant son entraînement, même si cette réponse n'est pas suffisamment rattachée à un document réel. C'est toute la différence entre une réponse dont la présentation est correcte et une réponse dont le contenu est exact.

Quand une référence élégante se fait passer pour une preuve

Imaginez un exercice consistant à compléter la phrase : « Selon des recherches publiées dans la revue… » Une personne qui maîtrise la langue et connaît les principes élémentaires de la rédaction scientifique peut proposer une suite parfaitement naturelle. Cela ne signifie pourtant pas qu'elle vient de consulter les archives de l'éditeur concerné. De la même manière, un modèle peut produire des références bibliographiques très convaincantes en utilisant

les relations qu'il a apprises entre les noms des chercheurs, les titres, les disciplines et les conventions de présentation des publications.

Selon sa conception et les fonctionnalités dont il dispose, un outil peut également effectuer des recherches sur Internet, consulter des documents ou interroger des bases de données externes. Ces possibilités facilitent la vérification, mais ne garantissent pas à elles seules l'exactitude de la réponse finale. Le modèle peut mal interpréter un document trouvé, mélanger les informations de deux publications ou attribuer à une source authentique une conclusion que ses auteurs n'ont jamais formulée.

Le problème comporte donc plusieurs niveaux. Parfois, la source n'existe pas. Parfois, elle existe, mais sa description est erronée. Dans d'autres cas, toutes les informations bibliographiques sont correctes et l'erreur concerne le sens du texte. Cette dernière situation est particulièrement trompeuse, car une simple recherche destinée à vérifier l'existence de la publication ne permet pas de détecter le problème.

Prenons un exemple fictif. Vous demandez à l'IA des études concernant les effets des courtes pauses sur la concentration dans le travail de bureau. Vous recevez un titre comme « The Impact of Microbreaks on Cognitive Performance in Office Environments », accompagné de noms d'auteurs, d'une année de publication et du nom d'une revue de psychologie. Tout paraît extrêmement plausible. Le sujet correspond à la discipline, le titre possède le sérieux académique nécessaire et les noms des auteurs donnent l'impression que leurs propriétaires participent régulièrement à des conférences scientifiques.

Cela ne signifie absolument pas que l'article mentionné existe. Le titre utilisé ici est une illustration, pas une référence bibliographique confirmée. C'est précisément ainsi qu'il faut considérer toute proposition de source non vérifiée : comme une information nécessitant un contrôle, même lorsqu'elle présente toutes les apparences du professionnalisme.

Curieusement, plus une réponse est détaillée, plus elle peut donner une impression de fiabilité. Un numéro de page, une date de publication et un identifiant DOI ressemblent à des tampons administratifs. En les voyant, nous supposons presque inconsciemment que quelqu'un a déjà effectué les vérifications nécessaires. Pourtant, la précision peut n'être que de la précision. Une erreur extrêmement détaillée reste une erreur, simplement mieux habillée.

Il n'est d'ailleurs pas nécessaire d'inventer une revue scientifique pour provoquer des ennuis. Il suffit d'associer le nom authentique d'un éditeur prestigieux à un article inexistant. L'élément véritable donne à l'ensemble une apparence de sérieux. C'est un peu comme si quelqu'un vous communiquait une adresse d'appartement fictive dans un immeuble qui existe réellement. Le bâtiment est là, la rue correspond, mais derrière la porte indiquée, il n'y a qu'un mur.

Pourquoi nous nous laissons convaincre si facilement

Nos difficultés ne viennent pas uniquement de la technologie. L'autre moitié du problème réside dans notre manière d'évaluer les informations. Lorsqu'un texte est fluide, bien organisé et rédigé sur un ton assuré, nous avons davantage tendance à le trouver compétent. C'est particulièrement vrai lorsqu'il traite d'un domaine que nous ne maîtrisons pas suffisamment.

Le fait que nous demandions généralement de l'aide à l'IA précisément pour gagner du temps n'arrange rien. Nous avons besoin d'un document rapidement, et la bibliographie reçue répond parfaitement à nos attentes. Vérifier les références représente un travail supplémentaire. Ouvrir les publications, contrôler les auteurs et comparer les citations semble beaucoup moins séduisant que de cliquer sur le bouton « Copier ».

À cela s'ajoute une tentation plus subtile : celle de confirmer notre propre position. Si l'outil nous présente cinq sources soutenant notre thèse, nous éprouvons une certaine satisfaction. S'il nous proposait cinq publications remettant cette même thèse en question, nous commencerions probablement à poser des questions supplémentaires. Voilà pourquoi il faut se montrer particulièrement attentif aux sources qui affirment exactement ce que nous espérions entendre.

Imaginons, par exemple, que vous prépariez des arguments en faveur d'une nouvelle méthode d'organisation du travail. Vous demandez : « Quelles études prouvent que cette méthode augmente la productivité ? » La question elle-même suppose déjà que les preuves recherchées existent. Le modèle peut répondre conformément à cette attente plutôt que de remettre en question son fondement.

Il serait plus judicieux de demander : « Quelles publications fiables étudient l'efficacité de cette méthode ? Présente les résultats favorables comme les résultats défavorables, et signale clairement les informations que tu ne peux pas confirmer. » Une telle formulation ne garantit pas l'absence d'erreurs, mais elle modifie l'orientation de la recherche. Vous cessez de commander des preuves destinées à soutenir une conclusion préparée à l'avance.

Cette précaution reste importante même lorsque vous utilisez un outil qui possède réellement un accès à un moteur de recherche. Internet contient une quantité considérable de documents de qualité très variable. Le simple fait de trouver une page ne signifie pas que l'affirmation qu'elle contient est vraie. Rechercher une information et évaluer sa fiabilité sont deux opérations distinctes.

Les quatre visages d'une fausse source

Dans la pratique, il est utile de reconnaître quatre situations fondamentales, car chacune exige une réaction différente.

La première est la fiction bibliographique complète. La publication, ses auteurs ou la revue mentionnée ne correspondent à aucun document réel. Vous pouvez passer une demi-heure à chercher un article qui n'existe que dans une réponse générée. Si, après une vérification raisonnable dans les catalogues et les bases de données, vous ne trouvez aucune confirmation, ne l'inscrivez pas dans votre bibliographie comme s'il s'agissait d'un fait établi. Vous n'avez pas besoin de démontrer que le document n'a jamais existé dans toute l'histoire de l'humanité. Il suffit de reconnaître honnêtement que vous n'avez pas réussi à le vérifier.

La deuxième situation correspond au mélange d'informations. L'article existe, mais l'année, le titre ou les auteurs indiqués sont incorrects. Il arrive qu'un modèle compose une référence bibliographique à partir de fragments provenant de plusieurs publications authentiques. Pour reconnaître cette erreur, il faut comparer l'ensemble de la référence avec la notice bibliographique originale. Il est déconseillé de corriger les différents éléments à l'intuition,

car on risque de créer accidentellement une nouvelle publication hybride.

La troisième forme concerne les citations déformées. L'auteur a réellement écrit une phrase, mais l'IA en modifie quelques mots, supprime une réserve importante ou fusionne deux passages en une seule déclaration. Le sens peut rester proche, mais un texte placé entre guillemets doit correspondre à l'original. Si vous souhaitez présenter l'idée d'un auteur avec vos propres mots, utilisez une paraphrase en indiquant correctement la source. Ne déguisez pas une paraphrase en citation littérale.

La quatrième situation est la plus sournoise : la source existe, la citation est exacte, mais elle ne confirme pas l'affirmation avancée. Une publication décrit une association entre deux variables, et l'IA la présente comme une preuve de causalité. Une étude porte sur un groupe restreint et particulier, mais la réponse formule une conclusion concernant l'ensemble de la population. Les auteurs évoquent une possibilité, et le résumé généré annonce une certitude.

Vous pouvez donc disposer d'un véritable article, d'un lien correct et malgré tout défendre un argument erroné. La bibliographie est acquittée, mais la conclusion vient de prendre la fuite.

Un premier filtre avant de commencer vos recherches

Vous n'avez pas besoin de procéder immédiatement à une analyse exhaustive de chaque publication. Dans le travail quotidien, un premier filtre permet de déterminer quelles références méritent une vérification approfondie. Commencez par lire la réponse de l'IA en considérant qu'il

s’agit d’une proposition, et non d’un travail de recherche déjà terminé.

Prêtez attention aux sources présentées sans informations bibliographiques complètes, à celles qui semblent correspondre de manière suspecte à votre question ou à celles qui accompagnent des affirmations très catégoriques sans indiquer de passages précis du document. Ces caractéristiques ne prouvent pas qu’une source soit fausse, mais elles justifient un contrôle. De même, l’absence de nom d’auteur ou de date ne signifie pas nécessairement qu’un document manque de fiabilité, car toutes les ressources pertinentes ne suivent pas la structure d’une publication scientifique.

Pour chaque référence importante, essayez de répondre à trois questions. Suis-je capable de retrouver le document indépendamment de la réponse de l’IA ? Sa description bibliographique correspond-elle au document réel ? Est-ce que je sais précisément quel passage confirme mon affirmation ? Si vous êtes bloqué dès la première question, ne passez pas automatiquement à la troisième en faisant comme si tout allait bien.

Supposons que vous ayez reçu cinq sources pour une présentation. Vous en retrouvez rapidement deux dans des catalogues fiables. Une troisième comporte une année incorrecte, mais vous parvenez à identifier le document original. Les deux dernières restent introuvables. Cela ne signifie pas que vous devez jeter tout votre travail. Vous pouvez conserver les références confirmées, corriger les informations de la troisième et supprimer les deux références non vérifiées. L’essentiel est de ne pas considérer les cinq comme également fiables.

Si une source joue un rôle déterminant dans une décision importante, confirmer son existence ne constituera que la première étape. Il faudra également examiner son contenu et les limites de la publication. En revanche, lorsqu'il s'agit d'un document secondaire, vous pouvez simplement renoncer à une référence incertaine au lieu de consacrer plusieurs heures à la retrouver.

Il est également utile de conserver une trace de vos vérifications. Inutile de construire un système compliqué. Une courte note comme « confirmé sur le site de l'éditeur », « texte intégral introuvable » ou « la source ne soutient pas la conclusion » suffit souvent pour comprendre, une semaine plus tard, pourquoi une publication a été utilisée ou écartée.

L'IA comme assistante, pas comme témoin principal

Le changement le plus utile consiste à distinguer deux rôles. L'intelligence artificielle peut être remarquablement efficace pour préparer des questions, organiser des notions et élaborer une stratégie de recherche. Elle ne devrait cependant pas devenir automatiquement le témoin ultime de la véracité des informations qu'elle a elle-même produites.

Lorsque vous demandez : « Les sources que tu m'as données sont-elles authentiques ? », vous pouvez recevoir une nouvelle réponse très convaincante, qui nécessitera elle aussi une vérification. Ce n'est pas un contrôle indépendant. L'outil peut corriger son erreur, mais il peut également maintenir une inexactitude antérieure. La simple affirmation « j'ai vérifié » ne remplace pas la possibilité d'indiquer ce qui a effectivement été contrôlé et où.

Il est préférable d'utiliser l'IA pour préparer une liste des éléments à confirmer, puis de comparer ceux-ci aux

documents originaux. Vous pouvez également lui demander d'identifier les limites de sa propre réponse. Vous bénéficierez ainsi d'une aide pour organiser votre travail sans transférer l'ensemble de la responsabilité au système.

Il n'est pas nécessaire de traiter chaque phrase générée comme un suspect dans une enquête criminelle. Il suffit d'adapter votre niveau de rigueur aux conséquences possibles d'une erreur. Une information utilisée comme simple inspiration n'a pas le même poids qu'une citation dans un mémoire universitaire ou qu'un élément servant à prendre une décision commerciale. C'est précisément dans ces situations importantes qu'il vous faut davantage qu'une référence élégamment présentée.

Ainsi, la prochaine fois que vous recevrez une bibliographie ressemblant au résultat d'un mois de travail effectué par toute une équipe de chercheurs, ne la rejetez pas automatiquement. Mais ne l'acceptez pas non plus sans vérification. Considérez-la comme une liste d'adresses potentielles auxquelles il pourrait être intéressant de frapper.

Car même la bibliographie la mieux présentée du monde ne vous sera d'aucune utilité si la moitié de ses auteurs n'habitent que dans l'imagination du modèle.

Chapitre 2 - Une citation entre guillemets, la vérité en vacances

Les guillemets comptent parmi les plus petits signes typographiques auxquels les êtres humains accordent un pouvoir étonnamment considérable. Il suffit de placer quelques mots entre deux paires de signes, d'ajouter le nom d'une personnalité célèbre et, soudain, une phrase ordinaire prend l'allure d'une déclaration historique. Si une date, le titre d'un livre ou le nom d'une conférence apparaît à côté, sa crédibilité semble encore augmenter. On cesse alors de se demander si quelqu'un a réellement prononcé ces paroles pour commencer à réfléchir à la meilleure manière de les utiliser dans une présentation.

Un jour, vous préparez un article consacré à la prise de décision. Vous demandez à l'IA une courte citation d'un penseur célèbre soulignant l'importance de l'esprit critique. Vous obtenez une phrase si pertinente que vous auriez envie de l'imprimer sur une affiche au-dessus de votre bureau. La formulation possède toute la gravité nécessaire, l'auteur est connu et le message correspond parfaitement à votre texte.

Mais lorsque vous essayez de retrouver la citation dans la publication originale, un problème apparaît. Le moteur de recherche affiche des centaines de pages répétant exactement les mêmes mots, mais aucune ne précise où ils ont été prononcés ou publiés pour la première fois. Un site mentionne un livre, un autre un discours, un troisième une date qui ne correspond à aucun événement confirmé. Le quatrième cite le premier, le cinquième cite le quatrième, et le sixième vous assure que cette phrase est « bien connue ».

Vous venez d'assister à la naissance d'un orchestre numérique dont tous les musiciens jouent à partir de la même partition erronée.

Le problème des fausses citations est plus ancien que l'intelligence artificielle. Depuis longtemps, les humains attribuent à des personnalités connues des paroles qui semblent suffisamment intelligentes, amusantes ou inspirantes. Internet a facilité leur diffusion, et les outils d'IA peuvent accélérer encore la création de nouvelles variantes. Cependant, toutes les citations attribuées à tort ne sont pas nées de l'activité d'un modèle de langage. Certaines erreurs figuraient déjà dans les documents disponibles bien avant son entraînement.

C'est précisément pour cette raison que la vérification des citations exige davantage de précision qu'une simple question sur le nom de l'auteur. Il faut déterminer quels mots ont réellement été prononcés, dans quel contexte, où ils ont été consignés et si la personne désignée en est effectivement l'auteur. Ce n'est qu'à partir de ce moment que les guillemets commencent à remplir correctement leur fonction.

Quand une paraphrase enfile le costume de quelqu'un d'autre

Le problème le plus fréquent ne consiste pas nécessairement à inventer une déclaration de toutes pièces. Parfois, l'IA reconnaît le sens général d'une pensée et le reformule à sa manière avant de présenter le résultat comme une citation littérale. On obtient ainsi une phrase fidèle à une certaine idée, mais infidèle à la formulation réelle du texte.

Supposons que l'auteur d'une publication écrive : « Dans le groupe de participants étudié, les pauses régulières

étaient associées à une diminution de la fatigue déclarée. » L'IA peut transformer ce passage en une formulation beaucoup plus percutante : « Les pauses régulières réduisent la fatigue des salariés. » À première vue, les deux phrases parlent d'un sujet similaire. Pourtant, la différence est considérable.

La première décrit une association observée au sein d'un groupe déterminé. La seconde suggère une relation causale générale. La restriction concernant les personnes étudiées a disparu, le caractère déclaré de la fatigue n'est plus précisé, et une formulation prudente a été remplacée par une affirmation catégorique.

Si vous placez la deuxième phrase entre guillemets et la signez du nom de l'auteur, vous lui attribuerez des mots qu'il n'a pas prononcés dans cet exemple. Plus grave encore, vous modifierez le sens de son propos. Ce n'est plus une simple retouche stylistique innocente.

Gardez à l'esprit qu'une citation littérale et une paraphrase remplissent des fonctions différentes. La citation sert à reproduire fidèlement des mots précis. La paraphrase permet de présenter la pensée d'autrui avec ses propres mots, tout en respectant son sens et en indiquant correctement la source. Dans les deux cas, il faut comprendre le texte original, mais seule la citation littérale exige de reproduire exactement la formulation employée.

L'intelligence artificielle peut vous aider à préparer une paraphrase, à condition que vous disposiez du texte source et soyez en mesure d'en évaluer l'exactitude. En revanche, lorsque vous demandez une citation précise, vous avez besoin d'un passage qui figure réellement dans le document. Le modèle ne devrait pas remplacer les mots manquants par

ses propres inventions, même s'il le fait avec une élégance littéraire remarquable.

Les citations traduites présentent une difficulté supplémentaire. Une traduction française d'une déclaration anglaise n'est pas une reproduction littérale des mots originaux dans la même langue. Vous pouvez la citer comme une traduction, mais vous devez en signaler clairement la nature. Si vous utilisez une traduction publiée, il est préférable d'en indiquer la source. Si vous traduisez vous-même, ne présentez pas votre version comme l'unique formulation officiellement autorisée.

Cette distinction est particulièrement importante dans les textes scientifiques et professionnels, où un seul mot peut modifier la portée d'une affirmation. Le terme anglais « may » ne signifie pas la même chose que « must », et « associated with » ne veut pas automatiquement dire « caused by ». En cherchant à rendre une phrase plus fluide, on peut discrètement transformer une conclusion prudente en certitude absolue.

Sept cents pages répètent la même chose. Mais qui a vérifié ?

L'une des caractéristiques les plus trompeuses d'Internet est sa capacité à reproduire une information des centaines de fois sans augmenter le nombre de preuves indépendantes disponibles. Une citation peut apparaître sur des centaines de sites, dans des milliers de publications sur les réseaux sociaux et dans des dizaines de présentations, alors que tous ces contenus proviennent finalement d'un seul message jamais vérifié.

Imaginez que quelqu'un publie une phrase percutante attribuée à un scientifique célèbre. Une autre personne la

copie dans un article, une troisième l'intègre à une illustration, tandis qu'une quatrième l'ajoute à une collection de citations inspirantes en ligne. Quelques années plus tard, le moteur de recherche renvoie une multitude de résultats. À force de répétition, l'information finit par paraître familière, et cette familiarité est parfois confondue avec la vérité.

Il ne suffit donc pas de compter le nombre de pages contenant les mots recherchés. Il faut déterminer si au moins l'une d'entre elles renvoie à une source primaire fiable.

Lorsqu'il s'agit d'une déclaration provenant d'un livre, cette source peut être une édition précise accompagnée d'un numéro de page ou d'un autre emplacement clairement identifiable dans le texte. Pour un discours public, un enregistrement officiel ou une transcription fiable sera particulièrement utile. Dans le cas d'un article scientifique, mieux vaut accéder au texte de la publication elle-même plutôt que de s'arrêter à un commentaire rédigé par une autre personne.

Les documents originaux ne sont pas toujours accessibles. Certains discours anciens n'ont jamais été enregistrés, des documents peuvent être conservés dans des archives et certaines déclarations peuvent n'avoir survécu que grâce aux témoignages de personnes présentes. L'évaluation exige alors davantage de prudence. Cela ne signifie pas automatiquement que la citation est fausse, mais votre degré de certitude doit correspondre à la qualité des preuves disponibles.

Une règle simple peut vous aider : répéter plusieurs fois la même information ne remplace pas une confirmation indépendante. Si dix sites renvoient tous vers un seul blog,