

IA GEN COMME ORTHÈSE RÉFLEXIVE : DES CLÉS POUR LA DÉVELOPPER

Pourquoi est-il essentiel de comprendre l'illusion réflexive pour mieux échanger avec les IA Gen conversationnelles et utiliser leurs capacités réelles ? Éclairage avec Olivier Las Vergnas, chercheur en sciences de l'éducation et de la formation.

Karine Sautereau



L'IA conversationnelle non seulement écrit, mais établit un raisonnement sur cette écriture.

Mieux comprendre l'illusion réflexive¹ pour mieux dialoguer avec les IA Gen² conversationnelles et tirer parti de leurs capacités réelles, tel est l'objectif d'un article paru dans la revue *L'esprit d'Archimède*, co-écrit par un chercheur en sciences de l'éducation et de la formation, Olivier Las Vergnas, et... par Chat-GPT 5. Le texte éclaire l'illusion de réflexivité produite par les modèles de langage, capables de dialoguer d'une manière qui semble consciente et réfléchie. Bien que leurs réponses donnent l'impression d'une pensée intérieure, ces systèmes ne possèdent aucune conscience : ils se contentent de calculer des enchaînements de mots probables en imitant la forme du raisonnement humain. Cette illusion naît de notre tendance à projeter notre propre réflexivité sur des discours qui en reproduisent fidèlement la structure. Comprendre cette distinction aide à utiliser ces outils avec discernement : leur efficacité ne repose pas seulement sur leurs performances techniques, mais aussi sur la manière dont les humains formulent les questions et interprètent les réponses.

Le texte met ainsi en lumière un espace de co-élaboration "où l'humain garde la direction, mais où la machine assure la continuité formelle et la cohérence méthodologique", explicite le chercheur. Ce présent article donne à voir des clés permettant de développer cet usage de l'IA Gen en tant qu'orthèse réflexive, c'est-à-dire en dialoguant avec la machine grâce à plusieurs niveaux de boucles réflexives. Au niveau de la forme, cet article a pour particularité d'être en grande partie écrit par Chat-GPT 5, à qui l'auteur humain a demandé "de permettre de mieux se préparer à ce à quoi on peut s'attendre comme performance des modèles de langage en matière d'aide à la réflexion humaine". Une forme de mise en abyme qui constitue "une occasion de s'interroger, à partir d'un cas concret, sur [la] notion d'orthèse". Mais bien sûr, comme un modèle de langage n'est en rien introspectif ou réflexif, ce texte n'est qu'une production probabiliste, résultat d'un ensemble de pondérations lexicales qui résultent d'un long dialogue antérieur entre l'auteur humain et Chat-GPT³.

Trois niveaux de dialogue

En réponse à cette question d'Olivier Las Vergnas, Chat-GPT 5 indique qu'il y aurait trois niveaux



1. L'illusion réflexive est "l'effet, très convaincant, d'un système statistique entraîné à reproduire la cohérence du langage humain", explicite Olivier Las Vergnas.

2. "L'IA Gen « grand public » a comme son nom l'indique pour but de produire de nouveaux textes, images, sons, vidéos ou fragments de

LE CHERCHEUR



Olivier Las Vergnas est professeur des universités émérite en sciences de l'éducation et de la formation. Il participe aux activités de recherche de l'équipe "Apprenance, Formation, Digital" (ApForD) du Centre de recherches éducation et formation (Cref) de l'Université Paris Nanterre. Il est membre du Liif de Centre Inffo.

codes informatiques, au cours de dialogues avec les humains, grâce à ce que l'on appelle des prompts", précise le chercheur (Las Vergnas & al., 2025).

3. Le dialogue entre Olivier Las Vergnas et Chat-GPT 5 qui a généré cet article est entièrement disponible en accès libre à l'adresse : <https://nakala.fr/10.34847/nkl.3c2atugj>

d'organisation du dialogue entre l'humain et la machine : "discursif", "processuel" et "méta-réflexif". Au premier niveau, basique, la machine génère un discours cohérent en enchaînant les idées selon les probabilités linguistiques les plus plausibles. Elle répond ainsi directement à la demande. Au second niveau, "processuel" ou "réflexif", la machine explicite son raisonnement, organise sa réponse en étapes et commente la manière dont elle traite la question. Enfin, au troisième niveau, "méta-réflexif", le dialogue devient lui-même l'objet de l'échange. L'humain et la machine discutent du cadre de la conversation, ajustent les règles d'interaction et redéfinissent ensemble les objectifs poursuivis. Chat-GPT 5 précise que "ces trois plans ne sont pas séparés, mais emboîtés : chaque niveau supérieur prend le précédent pour objet, comme une boucle qui se replie sur elle-même".

Deux déclencheurs langagiers

Toujours selon l'écrit de Chat-GPT, la conversation avec la machine peut être un instrument de programmation. Il a d'ailleurs identifié deux dynamiques : lexicale et procédurale. La dynamique lexicale est enclenchée lorsque le vocabulaire de l'utilisateur se concentre davantage sur la méthode, la structure ou la réflexivité (comme "processus", "niveau", "diagramme" ou "protocole"), le modèle identifie statistiquement ce registre sémantique et ajuste en conséquence son style de réponse. Quant à la dynamique procédurale, elle est lancée lorsque certains termes ou formes de questionnement (comparaisons, demandes de formalisation, etc.) activent des proximités internes qui influencent la manière dont le modèle structure sa réponse, par exemple en passant d'un texte linéaire à un tableau ou à un schéma. Ainsi, "le langage de l'humain agit comme un levier de pilotage de la forme et du niveau du discours de la machine", formule Chat-GPT.

Installation d'une boucle réflexive

Le dialogue avec l'IA Gen engendre un mécanisme d'interaction en boucle entre l'humain et la machine. De fait, lorsque cette dernière reformule ou explicite ses réponses, elle influence la manière dont l'utilisateur se positionne et formule ses questions. L'utilisateur s'adapte alors à ce nouveau registre, ce qui conduit la machine à ajuster et approfondir à son tour son discours.

"De boucle en boucle, la frontière entre simulation et co-réflexion s'estompe : le dialogue devient un système dynamique d'ajustement mutuel", fait remarquer Chat-GPT.

Exercice de méta-communication

Vu ainsi, ce mécanisme d'interaction avec l'IA Gen conversationnelle n'est pas un piège, mais une opportunité. Apprendre à dialoguer avec une IA Gen consiste à choisir consciemment le niveau d'échange souhaité : 1/ obtenir des informations, 2/ analyser des démarches ou 3/ co-élaborer un cadre méthodologique ou critique. Ainsi, "chaque interaction devient un exercice de méta-communication : un apprentissage de la pensée sur la pensée, où l'humain reste le véritable pilote de la réflexivité", explicite Olivier Las Vergnas. Il s'agit donc pour l'utilisateur de maîtriser le passage d'une commande, telle que "résume ce texte", à une régulation, comme "explique comment tu le résumes", ou bien même d'une régulation à une contractualisation : "à l'avenir, résume toujours selon ce protocole". En postface à ce texte écrit par Chat-GPT, Olivier Las Vergnas donne une vue d'ensemble des voies par lesquels un humain peut influencer sur les réponses du modèle de langage avec qui il dialogue, notamment grâce à l'usage de "mots déclencheurs" et de la "mémoire personnalisée persistante". Ainsi, "cette capacité à « changer de boucle » est au cœur de la pratique réflexive", souligne le chercheur. ●

POUR ALLER PLUS LOIN

- Las Vergnas, O. (2025). *Peut-on dépasser l'illusion réflexive dans nos rapports avec une IA ? Un article co-écrit avec Chat-GPT 5.* *L'esprit d'Archimède*, 2025(14), 30-34.
- Las Vergnas, O., Jeunesse, C. et Adinda, D. (2025). *De la délégation à l'amplification : quand l'IA Gen reconfigure nos façons d'apprendre.* *Action publique. Recherche et pratiques*, 25(2), 23-36.
- Adinda, D., Blandin, B., Jeunesse, C., & Las Vergnas, O. (2025). *Digital, IA et formation des adultes : synthèse de treize ans de recherche internationale.* *Savoirs*, 68(2), 13-118.