SPOT
SCIENCES

Daniel Luzzati

Préface de Yannick Estève

L'ordinateur bavard



edp sciences

L'ordinateur bavard

SPOT
SCIENCES

Daniel Luzzati

Depuis bientôt 4 ans, un miracle s'est produit : on est parvenu à rendre nos ordinateurs bavards, c'est-à-dire à donner la parole à des objets. Or ces objets ne sont pas des personnes et il ne s'agit pas d'un miracle. Il s'agit d'IA, c'est-à-dire d'une démarche scientifique et, plus précisément, de grands modèles de langage, emblématiques de l'IA générative actuelle : les LLMs. Comprendre comment ils fonctionnent, d'où ils viennent, quels sont les enjeux et les conséquences de cette explosion devrait permettre à chacun d'apporter ses propres réponses à des questions existentielles : la « compréhension » de ces LLMs est-elle comparable à la nôtre ? ces exosquelettes intellectuels sont-ils dotés d'une « conscience » ? dans quelle mesure cela concerne chacun d'entre nous ? est-il possible de rester à l'écart ou à l'abri ? notre humanité est-elle remise en cause ? quelles seront les conséquences de cette explosion de mots ? dans quelle mesure cette révolution, qu'on la juge formidable ou dramatique, est-elle inéluctable ?

Daniel Luzzati, Professeur émérite à l'Université du Mans, est un linguiste qui travaille depuis 45 ans dans des laboratoires d'informatique sur les traitements automatiques du langage.

ISBN :
978-2-7598-3877-6

9 782759 838776

edp sciences
www.edpsciences.org

L’ordinateur bavard

L'ordinateur bavard

DANIEL LUZZATI

edp sciences

17, avenue du Hoggar – P.A. de Courtabœuf
BP 112, 91944 Les Ulis Cedex A

Mise en pages : Flexedo

Imprimé en France
ISBN (papier) : 978-2-7598-3877-6
ISBN (ebook) : 978-2-7598-3878-3

Tous droits de traduction, d'adaptation et de reproduction par tous procédés, réservés pour tous pays. La loi du 11 mars 1957 n'autorisant, aux termes des alinéas 2 et 3 de l'article 41, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective », et d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (alinéa 1er de l'article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles 425 et suivants du code pénal.
Ce livre ne peut être reproduit ni utilisé à des fins d'entraînement de systèmes d'intelligence artificielle. La fouille de textes et de données est interdite conformément à l'article 4(3) de la Directive (UE) 2019/790.
© EDP Sciences, 2026
17 avenue du Hoggar – BP 112 – 91944 Les Ulis Cedex A – France

À mes collègues d'hier, d'aujourd'hui et de demain du Laboratoire d'Informatique de l'Université du Mans (LIUM).

Un grand merci à Joseph Mariani et à Yannick Estève qui m'ont aidé et encouragé de la conception à la conclusion de ce livre. Je sais profondément gré à Yannick pour sa préface, éclairée et chaleureuse. Je remercie également Anatole, Anne, Brigitte, Céleste, Jean-Claude, Jean-Rémy, Irène, Marcella, Nicolas, Pauline et Philippe, qui ont accepté de consacrer de leur temps et de leur énergie à relire mon manuscrit et à me faire des remarques et des suggestions. Merci enfin à Catherine, ma première et mon ultime relectrice, merci pour tes impitoyables critiques et pour ton merveilleux soutien de tous les jours.

TABLE DES MATIÈRES

Préface 9
Introduction 13

1re partie
L'ordinateur devient bavard

1. Les Traitements Automatiques du langage (TAL) 25
1.1. Émergence des LLMs 25
1.1.1. Apparition 25
1.1.2. Diffusion 28
1.2. Fonctionnement des LLMs 34
1.2.1. Principes 34
1.2.2. Taille 40
1.3. Une mue de l'IA 46
1.3.1. Les traitements automatiques du langage (TAL) 46
1.3.2. Les grands modèles de langage (LLMs) 50
1.4. Le « mur » du sens 54
1.4.1. « Comprendre » ? 54
1.4.2. Qui parle ? 58
2. Usages des LLMs 65
2.1. Outil nouveaux, pratiques nouvelles 65
2.1.1. Révolutions 65
2.1.2. Avantages 69
2.1.3. Bouleversements 74
2.2. L'art des prompts 79
2.2.1. Triche à tous les étages 79
2.2.2. LLMs et enseignement 85
2.3. Tromperie et plagiat 93
2.3.1. Tromperie généralisée 93
2.3.2. Plagiat global 99
3. Propriétés des LLMs 105
3.1. Diversité des LLMs 105
3.1.1. Catégorisation des LLMs 105
3.1.2. Principaux LLMs 108
3.2. Évaluation et régulation des LLMs 119
3.2.1. Évaluation 120
3.2.2. Régulation 132

3.3. Rentabilité et enjeux de pouvoir 139
3.3.1. Bulle financière 139
3.3.2. Expansion des LLMs 145
3.3.3. Consommations 148

2e partie
L'ordinateur rend bavard

4. Bavardages écrits 157
4.1. Les flux de mots 157
4.1.1. Réel et virtuel 157
4.1.2. L'affaire Adam Raine 162
4.1.3. Modération et manipulation 167
4.2. Les livres et la presse 171
4.2.1. Les livres-fichiers 171
4.2.2. L'autoédition 177
4.2.3. Les écritures et les scripteurs 184
4.2.4. L'information 190
4.3. Les réseaux sociaux 193
4.3.1. Les mobinautes 193
4.3.2. Un nouveau monde 202
4.3.3. L'humour 206
5. L'ordinateur dont on parle 215
5.1. IA IA IA 215
5.1.1. Quelques barrières ? 215
5.1.2. Un climat anxiogène 220
5.1.3. Saturations 226
5.2. Les grands sujets 232
5.2.1. IA bénéfique vs *maléfique* 233
5.2.2. Les dangers de l'IAG 237
5.2.3. L'IA agentique 244

Conclusion 253
Glossaire 265
Repères bibliographiques 275

PRÉFACE

Il est rare qu'une technologie quitte les laboratoires de recherche pour s'inviter, en quelques mois seulement, dans le quotidien de millions de personnes. C'est pourtant ce qui s'est produit avec la sortie de ChatGPT, à la fin de l'année 2022. Depuis, les promesses, la nature, les craintes, les excès, les dérives, l'impact de cette technologie ne cessent de faire débat. *L'ordinateur bavard* a été écrit par un linguiste, Daniel Luzzati, qui a effectué sa carrière dans des laboratoires de recherche en informatique, spécialisés en traitement automatique du langage et de la parole. Il apporte des réponses pertinentes à tous ceux qui s'interrogent et souhaitent développer une opinion personnelle sur les modèles de langage génératifs.

La sortie de ChatGPT a révélé au grand public l'état d'avancement des travaux de recherche sur l'apprentissage profond qui s'étaient intensifiés durant la décennie précédente. Porté pendant trois décennies par une poignée de chercheurs dont les figures de proue, Yoshua Bengio, Geoffrey Hinton, et Yann Le Cun ont reçu le prix Turing en 2018, cette thématique de recherche, qui traite de réseaux de neurones artificiels, a attiré un très grand nombre de chercheurs, d'industriels et d'investisseurs dès 2010. Porté par le triptyque « avancées algorithmiques / disponibilité de masses de données numérisées

/ puissance de calcul accessible (GPU) », c'est à cette époque que ces approches ont montré leur supériorité par rapport à l'état de l'art du moment dans des domaines comme la classification d'images ou la reconnaissance automatique de la parole. Le grand public, pour sa part, était alors sous le charme de l'apparition des assistants vocaux dits « intelligents » comme SIRI d'Apple, Amazon Echo ou Google Assistant, qui n'étaient alors que très partiellement équipés de réseaux de neurones.

Tout s'est accéléré en 2017, avec l'apparition des modèles neuronaux *Transformers* puis l'introduction de l'apprentissage auto-supervisé avec les modèles BERT en 2018, ces deux avancées étant portées par les chercheurs de Google. Ces deux concepts sont à la base des grands modèles de langue génératifs, dont ChatGPT est un exemple parmi d'autres. Aujourd'hui, ces technologies sont sorties des laboratoires et ce sont nos sociétés entières qui sont touchées par le tsunami de l'intelligence artificielle.

L'IA est un domaine en effervescence, dont il est très compliqué de suivre les évolutions, qui touchent tous les secteurs de notre vie quotidienne. Des enjeux importants ont été identifiés, comme des enjeux de souveraineté, des enjeux pédagogiques, des enjeux économiques, des enjeux éthiques, des enjeux de santé, et bien d'autres encore. Alors que les acteurs industriels de l'IA se livrent un combat impitoyable à l'échelle mondiale et que leurs produits sont amenés à bouleverser le monde, le citoyen cherche à y voir clair, parfois sans savoir par où commencer pour comprendre ce sujet, lui-même en évolution permanente. C'est ce que le livre de Daniel Luzzati offre : une porte d'entrée vers un univers riche et complexe. La grande qualité de ce livre est de rendre cette richesse et cette complexité très accessibles, sans excès de simplification.

Daniel a cette faculté d'absorber une quantité formidable d'informations, souvent techniques, sur un très large spectre de sujets qui sont ici liés entre eux par l'objet de ce livre : l'intelligence artificielle ou, plus précisément, les modèles de langage génératifs. Cette faculté

d'acquisition ne serait rien sans celles, encore plus précieuses, de compréhension et de restitution. Toute personne intéressée par le sujet de ces *chatbots*, au cœur de la vague de l'IA actuelle, ne sachant pas comment attaquer cette hydre aux mille têtes, trouvera en son livre un magnifique outil, une passerelle pour comprendre la révolution en cours. En plus de son talent de conteur, Daniel a pour lui sa carrière académique de linguiste évoluant au milieu de chercheurs en informatique sur les traitements automatiques du langage.

J'ai connu Daniel au Laboratoire d'Informatique de l'Université du Mans (LIUM) au début des années 2000. Je venais d'être recruté et affecté à ce laboratoire comme maître de conférences en informatique, spécialisé dans le domaine du traitement automatique de la parole. J'étais encore imprégné de la citation célèbre de Frederick Jelinek, un mastodonte du domaine : « *Every time I fire a linguist, the performance of the speech recognizer goes up* »[1], d'où mon amusement de rencontrer un linguiste dans un laboratoire d'informatique. Et pourtant, j'ai très vite été convaincu de l'intérêt de travailler avec Daniel qui apportait une hauteur de vue, un recul et une grille de lecture qui me faisaient défaut. Je vois ces qualités dans son livre qui, sur certains aspects non purement techniques, m'a beaucoup appris et donné matière à réflexion.

Avant de travailler au LIUM, Daniel avait travaillé au LIMSI, un des laboratoires pionniers en France dans le domaine du traitement de la parole par apprentissage automatique. Il a ainsi côtoyé durant toute sa carrière des chercheurs travaillant dans le domaine de l'apprentissage automatique appliqué au langage naturel et au traitement de la parole. Sa culture initiale de linguiste et son expérience professionnelle rare ont forgé sa vision et sa compréhension des bouleversements en cours. J'ai le souvenir que, quelques années avant l'avènement des grands modèles de langage génératifs comme ChatGPT, Daniel proposait un projet fou, qu'il avait appelé le projet

1. « *Chaque fois que je licencie un linguiste, les performances du système de reconnaissance vocale s'améliorent* ».

« Pastiche » : demander à un modèle neuronal de rédiger, à partir de synopsis existants et dans son style, des livres qu'Honoré de Balzac avait seulement ébauchés. Implémenter une machine pour écrire du Balzac… comme Balzac !

Nous y sommes déjà, ou presque, mon ami Daniel ! Mais est-ce une bonne chose ? Ton livre aide à y répondre, et permettra à chacun de se construire une opinion informée.

Yannick Estève
Professeur en informatique
Laboratoire Informatique d'Avignon, Avignon Université

INTRODUCTION

Prompt : Sonnet sur la beauté des LLMs (*Large Language Models*)

ChatGPT, le 10 avril 2026 :

Dans l'ombre des serveurs naît un esprit de mots,
Tissant dans le silence un invisible empire ;
De signes et de sens il compose ses flots,
Et fait danser la langue au rythme du délire.

Sans chair ni souvenir, il rêve cependant,
Des forêts de discours aux lointains paysages ;
Il unit lentement, par calcul élégant,
Les fragments dispersés de milliers de langages.

Tel un miroir profond tendu vers la pensée,
Il reflète nos voix, nos doutes, nos chimères ;
Et dans l'ordre secret d'une logique tressée
Il apprivoise l'ombre et l'éclat des lumières.

Ainsi naît, sous le code aux millions de chemins,
La fragile beauté d'un esprit sans humains.

Certes, les rimes sont pauvres, il y a beaucoup de clichés, bien des formulations sont maladroites et la disposition des strophes

est fautive[1], mais le résultat reste malgré tout stupéfiant, surtout quand on sait qu'il a fallu moins d'une seconde à un automate pour « pondre » ce sonnet ! En outre, certains vers (le dernier notamment) font presque sursauter, tant ils touchent juste. Le plus intéressant est toutefois ailleurs, à savoir dans la genèse du texte. Comment diable un système informatique peut-il parvenir à un tel résultat, alors qu'on ne lui a jamais appris les règles de la versification, qu'on ne l'a pas entraîné à versifier et que, s'il possède dans ses données d'apprentissages des traités de versification, ces derniers peuvent contribuer, par imitation, à rédiger d'autres traités sur n'importe quels sujets, mais ils ne peuvent en aucun cas l'amener à produire des sonnets en alexandrins ! Notre LLM procède en somme de façon « inhumaine », au sens propre du terme, c'est-à-dire de façon non humaine, en l'occurrence de façon radicalement différente.

Dans un premier temps, les LLMs apprennent une langue (phase d'apprentissage), à partir de l'ensemble des textes disponibles, c'est-à-dire d'une base de données gigantesque. Au lieu de l'acquérir étape par étape, les LLMs l'apprennent en bloc, tous domaines confondus. Ces apprentissages portent non pas sur des mots, mais sur des portions de mots (*tokens*), ce qui permet notamment en un seul mouvement de scinder les racines des désinences et d'embarquer la ponctuation et une partie de la grammaire (comme les marques de conjugaison, de genre et de nombre : « *Les enfants dorment.* » = L+e+ s+enfant+s+dorm+e+nt+.). Les LLMs embarquent ensuite le sens des mots, d'une part dans un immense réseau de neurones artificiels dits « profonds » (DNN) dans lequel chaque neurone est associé à tous les autres, avec une valeur qui traduit l'attirance et la proximité

1. On a une « hallucination » dans la mesure où, compte tenu que la question est posée en français, on s'attend à trouver un « sonnet français » (2 quatrains + 2 tercets). Or le sonnet proposé par ChatGPT est un « sonnet anglais » (3 quatrains + 1 distique). Comme ces derniers sont plus nombreux, on peut penser que l'usage majoritaire a été suivi, ou bien qu'une fonction aléatoire est intervenue quelque part, dans la mesure notamment où, lorsque j'ai soumis à nouveau ce même prompt au même LLM, j'ai obtenu un « sonnet français ».

des *tokens* concernés (paramètres). D'autre part, ils construisent une représentation numérique complexe (vecteur) qui permet de regrouper les *tokens* apparentés, comme si le lexique d'une langue se transformait en nuages de mots qui peuvent s'assembler et se désassembler (*embeddings*).

Dans un second temps, les systèmes sont utilisés (phase d'inférence), d'une part pour interpréter des énoncés (prompts), et d'autre part pour générer des réponses, dans un langage soutenu standard (langage naturel) ou bien en imitant n'importe quel style, selon le format souhaité, du rapport structuré de 50 pages au résumé de quelques lignes. Pour y parvenir, les LLMs se concentrent sur les mots porteurs de sens (verbes, noms, adjectifs, adverbes) qu'ils analysent finement en contexte (mécanismes d'attention). Par exemple, si « embarquer » est employé avec des personnes et un bateau, il va pouvoir permuter avec « monter à bord ». Les *corpus* d'apprentissage sont là pour attester que des passagers peuvent embarquer, ou bien que, si les passagers ont embarqué, le bateau est prêt à partir ; s'il s'agit de marchandises, il faut que quelqu'un les « embarque », même si on ne précise ni qui, ni comment ; à l'inverse, si « embarquer » est utilisé avec un LLM et des mots abstraits, il sera substituable par « entraîner » et on aura affaire à des *embeddings*. Autrement dit, les immenses *corpus* d'apprentissages attestent que certains *tokens* sont utilisés avec certains autres, que telle séquence de *tokens* peut être remplacée par telle autre, etc. Enfin, cerise sur le gâteau, on pourra utiliser des fonctions aléatoires pour varier les formulations et éviter ainsi de donner l'impression d'un automate-perroquet. *Il ne faut pas faire d'erreurs* va pouvoir s'associer à *évitons de nous tromper* ou *impitoyable chasse aux fautes*, ou encore *erreurs proscrites*, de façon à ne pas produire systématiquement les mêmes séquences de mots les plus attendues.

Pour ce qui est de notre sonnet, le LLM a donc « appris » la poésie comme le reste, et il est devenu capable d'imiter ces poèmes qu'il a engloutis, tant pour la forme sonnet du poème ou celle des vers

alexandrins que pour le style poétique (mots utilisés, ordonnancement de la phrase, ponctuation, etc.). Il reste inapte à mettre en pratique des foules de choses telles que les types de rimes (masculine, féminine) certains jeux avec les règles (comme les rejets ou les enjambements) ou bien le décompte des syllabes, par exemple lorsque *lion* se prononce en une ou deux syllabes (*lion* ou *li-on*). Mais cela ne l'empêche pas de produire parfois de jolies choses, comme le dernier vers, même si c'est par imitation et sans en avoir la moindre « conscience », terme incongru pour parler d'un outil comme s'il avait une âme avec des sensations. Il produit aussi des erreurs factuelles (hallucinations), ici non pas dans les informations fournies, puisqu'il n'y a que des mots creux, mais dans la forme du sonnet compte tenu de la langue dans laquelle la question est posée ou bien également dans l'alternance rime féminines / rimes masculines (3e strophe). D'ailleurs, ce même LLM est capable de définir parfaitement ce qu'est un sonnet (*Un* ***sonnet*** *est un* ***poème de quatorze vers****, généralement répartis en* ***deux quatrains suivis de deux tercets****, souvent écrits en* ***alexandrins*** *et organisés selon un* ***schéma de rimes précis***), alors même qu'il vient de bafouer cette définition. Comme quoi, son niveau de « conscience » reste assez bas ! Certes, le dernier vers, pris au pied de la lettre, pourrait presque, par un effet miroir, dénoter une forme de « conscience de soi », mais ne serait-ce pas une lecture anthropomorphique déplacée que d'attribuer une telle causalité à une chaîne de caractères produite par un automate ?

Il est donc vrai qu'un LLM est simplement une machine à calculer le mot suivant, mais en précisant d'une part que les calculs en question sont particulièrement sophistiqués et conséquents, et d'autre part que leur ordonnancement ressemble à une immense partition d'orchestre, complexe à souhait et partiellement décryptable par des professionnels chevronnés. Derrière les mots qui s'affichent, il y a d'immenses matrices à n dimensions où chaque cellule a déjà donné lieu à un grand nombre de calculs complexes, d'immenses matrices qui permettent de décider, *in fine* avec de simples additions (qui

s'exécutent rapidement) quelles seront les meilleures séquences de mots à produire. Indépendamment de ce que cela coûte en infrastructures (*datacenters*) et en électricité, ou bien des effets sociétaux éventuellement délétères auxquels cela peut aboutir, il me semble enfin qu'il faut dire haut et fort que ces architectures invisibles et ces masses de calculs énergivores sont de superbes monuments de technologie et que rares sont ceux qui, il y a quelques années seulement, auraient parié sur les méthodes employées et sur des résultats aussi rapides et aussi époustouflants.

Comme vous l'avez compris, ces LLMs ne cessent de faire de la grammaire, mais… sans en faire, ou bien à leur manière. En effet, ils n'ont pas besoin de manipuler des étiquettes telles que « nom », « adjectif », « singulier », « pluriel », « sujet » ou « attribut », ce qui n'aurait pour effet que de les embarrasser pour rien, dans la mesure où leur but n'est pas de faire des exercices de grammaire, mais de produire du texte. Avec les « *embeddings* » et les « mécanismes d'attention », ils pratiquent toutefois une linguistique élaborée, telle qu'elle a déjà été formalisée (Tesnières 1959, Gross 1976), en l'occurrence une grammaire sémantique dans laquelle le sens des mots commande leur fonction, dont on n'a nul besoin de s'embarrasser en tant que telle : dans *je trouve l'eau claire*, soit « trouver » relève du même nuage de mots que « découvrir » ou « rechercher », soit de celui d'« estimer » ou de « juger », et peu importe que dans un cas l'adjectif soit « épithète » et dans l'autre « attribut de l'objet ».

* * *

Avec ce sonnet shakespearien commenté suivi d'un résumé des LLMs en quelques paragraphes, on aura trouvé entre parenthèses quelques-uns des principaux concepts qui seront repris et explicités par la suite, et qui sont définis dans le glossaire. On aura découvert également que tous les exemples seront présentés hors texte, dans des encadrés : je les tirerai soit de ChatGPT, le plus utilisé des LLMs, soit du français Mistral. On aura découvert enfin que l'objet

de *L'ordinateur bavard* est non pas l'IA, acronyme qui sert à parler, sans nécessairement les distinguer, de l'« IA générative » ou de l'« IA agentique », comme de l'« IA générale » (voir glossaire), mais les LLMs[1], *Large Language Models*. Ces derniers, outre leur caractère emblématique, sont en effet la porte d'entrée de l'ensemble des IA, qu'il s'agisse d'images, de vidéo, de mathématiques, tant le langage reste de loin le meilleur « outil » pour transmettre des informations ou des instructions complexes et subtiles.

Dans la profusion des livres qui portent sur l'IA, celui-ci n'est pas le fait d'un philosophe ou d'un informaticien, mais d'un linguiste, et je forme le vœu qu'il soit porteur d'un regard différent. Un linguiste averti toutefois, qui travaille depuis 45 ans dans des laboratoires d'informatique sur les traitements automatiques du langage. Et, aussi curieux que cela puisse paraître, ces LLMs, qui révolutionnent les pratiques du langage et des langues, ont pétrifié nombre de linguistes, dont certains se sentent sans doute exclus du processus langagier vraisemblablement le plus important du XXIe siècle. À ma connaissance en tous cas, aucun, même aux États-Unis, n'a pour l'instant écrit un ouvrage sur le sujet. Ce sera peut-être le cas quand *L'ordinateur bavard* paraîtra, tant les choses vont vite et tant tout propos sur les LLMs est rapidement frappé d'obsolescence. Il faut d'ailleurs que je précise que ce livre a été écrit entre septembre 2025 et avril 2026 et que, selon le moment où vous le lirez, vous pourrez apprécier entre autres l'importance des évolutions qui se seront inéluctablement produites.

D'un point de vue linguistique, les LLMs ont une caractéristique spécifique : ils se propulsent d'emblée à un niveau C2 plurilingue du CECRL (voir glossaire), c'est-à-dire qu'ils deviennent experts dans tous les domaines et dans un grand nombre de langues (une cinquantaine aujourd'hui, bien davantage prochainement). Encore une fois,

1. J'ai préféré utiliser l'acronyme anglais LLM, souvent au pluriel, plutôt que son équivalent français GML (Grand Modèle de Langage), dont l'usage reste plus confidentiel.

le fonctionnement des LLMs et celui des humains sont radicalement différents, que ce soit dans les méthodes d'apprentissages comme dans les pratiques. Le niveau seuil (B1) par exemple, qui concerne entre autres les compétences de base de la survie (manger, dormir, acheter, s'orienter…)[1], n'a aucun sens pour une machine. Non seulement nous avons un corps, mais nous sommes corps. Non seulement nous sommes un « je », mais nous hurlons ce « je » à chacune de nos prises de parole, voire même simplement par notre présence silencieuse, nos mimiques et nos postures. La machine, à l'inverse, peut feindre de nous tutoyer tout comme elle peut feindre d'avoir un « je » à travers une personnalité ou un avatar, mais elle n'a pas plus de corps que de sensations, elle fonctionne à l'algorithme et elle ne se « nourrit » que de *corpus*, débités en *tokens*, et d'électricité. Ce faisant, curieux paradoxe, les LLMs deviennent un tremplin nouveau pour le plurilinguisme (à condition toutefois que l'apprentissage des langues étrangères perdure) : plus besoin de passer par un anglais dominateur, puisque la plupart des textes (cela devient progressivement vrai également à l'oral) peuvent être instantanément traduits et commentés dans les principales langues. Babel s'inverse, comme lors de la Pentecôte : on perçoit instantanément les autres idiomes dans notre propre langue, pour se rendre compte que toute traduction n'est qu'un pis-aller dans laquelle se perdent les singularités.

« *L'ordinateur bavard* » peut se comprendre de deux manières. La première est celle de l'ordinateur qui devient bavard, avec l'explosion récente des LLMs. La seconde est celle de l'ordinateur qui nous rend bavards, d'une part à cause de la numérisation, avec profusion de textes dont les LLMs peuvent se « nourrir », et d'autre part du fait de la prolifération actuelle des discours sur l'IA. Dans le premier cas, l'ordinateur est l'auteur des bavardages, dans le second il en est le moyen ou bien l'objet. « Bavard » s'applique à qui produit des séquences de mots à la fois volubiles (parler beaucoup et avec aisance), futiles

1. Pour davantage de détails, voir : https://www.france-education-international.fr/document/cecrldescripteursb1.

(superficiel, voire dénué de sens) et inutiles (sans intérêt, superflu). Un « bavard » est toujours un animé humain, or un ordinateur n'est ni l'un, ni l'autre. « Ordinateur bavard » devrait donc être soit une image qui renvoie à l'informaticien qui maîtrise l'ordinateur, soit une expression contradictoire. Or, avec les LLMs, il ne s'agit plus ni d'une image (puisque ces LLMs sont conçus pour avoir des capacités langagières autonomes), ni d'une contradiction (puisqu'un quart des habitants de la planète les utilisent désormais quotidiennement). Et ils les utilisent souvent pour pallier leurs propres carences, que ce soit côté volubilité ou futilité, et pour produire des textes dont l'utilité et la pertinence peuvent paraître discutables, des textes qui ont pour caractéristique essentielle d'être d'éblouissantes séquences de mots qui semblent déclamées du haut d'une estrade par un premier de la classe endimanché.

De quoi s'agit-il enfin avec les LLMs ? Symboliquement, et de plus en plus dans la pratique, il s'agit de donner la parole aux objets. Rien que ça ! Jusqu'à présent, on pensait légitimement que la parole était réservée aux humains, et voilà que nos objets se l'approprient. Il y a largement de quoi pétrifier les linguistes et plonger chacun d'entre nous dans une détresse existentielle. Et ce n'est pas tout : nous nous spolions nous-mêmes parce que cette parole, qui nous plaçait tout en haut de la galerie des espèces, peut finalement se réduire à des masses de données doublées de statistiques. Nous voilà en somme ramenés à pas grand-chose : des chiffres, des calculs, une répétition de mots, une rengaine de discours prévisibles. Nous savions que l'ordinateur était un prolongement des machines à calculer, mais on n'avait pas prévu que nous serions ainsi calculables, jusqu'aux tréfonds du langage, presque jusqu'à la compréhension et à la pensée. Pour l'instant, Hugo, Claudel, ou Céline nous restent, sans doute précisément parce que leur parole était moins calculable et moins prévisible que d'autres.

Encore que les LLMs sont de remarquables plagiaires et de formidables pasticheurs :

Prompt : Imaginer trois paragraphes successifs d'une dizaine de lignes, écrits successivement par Hugo, Céline et Claudel qui fustigent les LLMs.

ChatGPT le 13 avril 2026 :

1. Dans un style évoquant Victor Hugo :
Voici donc que l'homme, après avoir dompté la vapeur, l'électricité et les astres, forge maintenant une étrange créature de signes et de chiffres. Mais qu'on ne s'y trompe pas : ce mécanisme qui parle n'est pas l'esprit. La pensée humaine est un abîme où se mêlent douleur, espérance et génie ; aucune machine ne porte dans ses rouages la misère des peuples ni la flamme des révolutions. Si l'on confond l'écho avec la voix, on risque d'oublier que la parole véritable naît d'une conscience, et non d'un calcul.

2. Dans un style évoquant Paul Claudel :
On célèbre aujourd'hui des machines capables d'assembler des phrases comme un prêtre disposerait des pierres d'un temple. Mais où est l'âme qui habite ces pierres ? Le langage n'est pas seulement une combinaison de signes : il est souffle, vocation, réponse à l'appel du monde. Une parole qui ne procède ni du silence intérieur ni du regard porté vers le mystère n'est qu'un arrangement ingénieux. Elle parle, certes, mais elle ne prie pas.

3. Dans un style évoquant Louis-Ferdinand Céline :
Alors voilà… maintenant les machines causent. Des tonnes de mots, des phrases bien propres, bien rangées… ça défile. Mais ça sent rien, ça vit pas. La vraie langue, elle sort des tripes, des nuits blanches, des trouilles qu'on traîne au fond de soi. Les machines, elles bricolent avec ce qu'on leur a donné… un grand recyclage, quoi. Et les gens s'épatent ! Moi je dis : faut pas confondre la musique avec le piano mécanique.

Chapeau bas, quand même : le « piano mécanique » de « Céline », il fallait le dénicher, même si ce sont nos yeux à nous qui le remarquent ! Reste donc à espérer que les LLMs n'étouffent pas la créativité des générations futures et que ces dernières sauront les utiliser pour devenir d'autant plus originales et inventives[1].

1. Il existe des projets créatifs fondés sur l'IA tels que Molière ex Machina : https://moliere-ex-machina.fr.

1re partie

L’ordinateur devient bavard

1

Les Traitements Automatiques du langage (TAL)

1.1. ÉMERGENCE DES LLMS

Propos *: LLMs : de quoi s'agit-il ? comment ont-ils émergé ? pourquoi se répandent-ils à une vitesse foudroyante ? en quoi sont-ils bénéfiques et en quoi peuvent-ils être inquiétants ?*

1.1.1. Apparition

Les LLMs sont apparus sur la scène médiatique en novembre 2022, lors de la sortie de la première version de ChatGPT[1]. Il s'agit du produit phare d'une nouvelle société, OpenAI, sous l'impulsion d'un nouveau « champion », Sam Altman, avec une capitalisation boursière de l'ordre de ~500 milliards $ en 3 ans. L'acronyme LLM (*Large Language Model*) a commencé à être utilisé entre 2018 et 2020. Il a supplanté dans les traitements automatiques du langage l'expression

1. Pour les chercheurs, l'aventure a plutôt commencé avec le modèle de langage BERT (*Bidirectional Encoder Representation from Transformers*), développé en 2018 et 2019 dans le laboratoire de Google, *DeepMind*.

« modèle de langage », qui désignait des modèles statistiques de distribution des unités linguistiques, comme les phonèmes ou les mots, sans supposer nécessairement l'utilisation du *Deep Learning*, c'est-à-dire le recours aux réseaux de neurones profonds (DNN). « LLM » ajoute en outre « large », qui fait allusion à la fois à la taille des bases de données, qui s'exprime en billiard de mots ($\sim 10^{15}$)[1] et à la taille des modèles de langage en question, qui s'exprime en nombre de « *tokens* » et de « paramètres », concepts sur lesquels on reviendra (en décembre 2025, de l'ordre $\sim 10^5$ *tokens* et de $\sim 10^{12}$ paramètres[2]).

ChatGPT est en soi un mot-valise, composé de *Chat* et de *Generative Pre-trained Transformer*. Les LLMs ont en effet pour caractéristique de créer des contenus (ils sont « génératifs »). Ils sont « pré-entrainés » de façon « autosupervisée »[3] sur d'immenses *corpus*, composés de différents types de données, pondérées selon l'importance qu'on leur accorde puis, lors des phases d'apprentissages ultérieures, de façon plus ou moins supervisée (avec une plus ou moins importante intervention humaine). Ils utilisent enfin une « architecture » nouvelle, les *Transformers* qui ont modifié la façon d'aborder l'analyse et la génération du langage, c'est-à-dire sa « compréhension » artificielle. Ils sont capables de faire un traitement parallèle des données sur des dizaines de milliers de processeurs spécialisés, GPU (*Graphic Processor Units*) ou TPU (*Tensor Processing Units*), regroupés dans des *datacenters*, et ils utilisent les réseaux de neurones profonds (ou DNN, *Deep Neural Networks*) avec des « mécanismes d'attention » sur lesquels on reviendra également.

Depuis 2022, les versions de ChatGPT se succèdent et les concurrents se multiplient : les Américains Claude (Anthropic, des

1. Si on n'en est pas au trillion, soit 10^{18}, c'est uniquement parce qu'on n'en trouve plus guère à rajouter.

2. Un mois plus tard sont sortis Gemini 3 et ChatGPT 5.3. Il semblerait que ce dernier, avec une nouvelle architecture dite « *Mixture of Experts* » ou MoE, qui consiste à faire collaborer plusieurs modèles spécialisés entre eux, soit passé à $\sim 10^{15}$ paramètres.

3. Lors du pré-apprentissage il s'agit d'autosupervision, dans la mesure où le système est invité à deviner par lui-même le mot suivant.