

Pierre Levy's Blog

Various thoughts and papers in french and english

Intelligence collective, intelligence artificielle et partage du savoir

12 July 2023 //

2

Prof. Pierre Lévy, PhD, MSRC

“Ce texte propose quelques réflexions philosophiques sur la relation triangulaire entre l’intelligence collective, l’intelligence artificielle et la noble finalité de mettre la connaissance à la portée de tous.”



La valeur du savoir vérifié à l'heure de l'IA générative

Commençons par une citation de Denny Vrandečić, un des initiateurs de Wikidata, qui avait travaillé sur le graphe de connaissances de Google comme ontologiste et qui est aujourd’hui le chef de file du projet “Abstract Wikipedia”[1] visant à rendre les données des articles de Wikipédia indépendants des langues (c’est-à-dire traductibles dans toutes les langues). Denny a déclaré dans sa communication à la *Knowledge Graph Conference* de mai 2023[2] *“In a world of infinite content, knowledge becomes valuable”*. Ce monde où les contenus sont potentiellement infinis résulte évidemment de l’usage désormais massif de l’intelligence artificielle générative. Parmi tous les problèmes que posent cette nouvelle situation, citons-en deux, particulièrement prégnants du point de vue de l’accès à la connaissance. Premièrement, malgré les usages qui sont faits des modèles génératifs pour obtenir rapidement des réponses directes, il faut réaffirmer que, contrairement à l’IA symbolique classique du XXe siècle, l’IA statistique (dite aussi neuronale) d’aujourd’hui – limitée à ses seules capacités – n’offre aucune garantie de vérité. GPT4 comme les autres modèles du même genre *ne sont pas des*

bases de connaissances. Les erreurs de fait et de raisonnement sont nombreuses et il suffit d'être spécialiste d'un domaine pour constater les faiblesses de ces IA, comme nous le faisons d'habitude lorsque nous lisons un article portant sur un de nos domaines de compétence quand il est rédigé par un journaliste pressé qui se contente de reformuler la doxa dans laquelle il baigne. Les réponses – probabilistes – de Chat GPT sont seulement *vraisemblables*. Deuxièmement, comme les modèles d'IA générative sont entraînés sur les données du Web et que celles-ci sont de plus en plus rédigées et illustrées par les modèles en question, on se trouve en présence d'un dangereux cercle vicieux, et cela d'autant plus que les travailleurs au rabais chargés d'aligner les modèles et de redresser leurs biais ou erreurs utilisent eux-mêmes des IA génératives pour accomplir leur tâche![3] Pour s'extraire de ces sables mouvants épistémologiques, il est donc nécessaire d'investir plus que jamais dans la construction de sources d'information fiables. En d'autres termes, l'explosion des usages de l'IA générative, loin de nous dispenser d'alimenter et d'utiliser Wikipédia, Wikidata et d'autres bases de connaissances vérifiées, rend l'effort d'y contribuer et le plaisir de les consulter encore plus nécessaires! Ceci dit, comme nous le verrons par la suite, l'IA neuronale a néanmoins vocation à jouer un rôle positif dans le partage du savoir.

Ce texte propose quelques réflexions philosophiques sur la relation triangulaire entre l'intelligence collective, l'intelligence artificielle et la noble finalité de mettre la connaissance à la portée de tous.

L'intelligence collective

Étroitement définis comme des moyens de résoudre des problèmes, les processus d'intelligence collective[4] se présentent sous de multiples formes, dont les plus étudiées sont les espèces statistique, délibérative et stigmergique[5].

Au début du XXe siècle, en Angleterre, le savant Francis Galton visitait une foire agricole. Un concours avait été organisé: on demandait aux huit cents participants, la plupart éleveurs, de deviner le poids d'un bœuf. Mais aucun d'eux n'avait trouvé le poids exact. Galton fit la moyenne de toutes les évaluations et trouva qu'elle était beaucoup plus proche du poids réel qu'aucune des estimations individuelles. La "sagesse" de la foule était supérieure à chacune des intelligences isolées[6]. Cette forme d'intelligence collective statistique – ou comptable – suppose que les individus ne communiquent pas entre eux et ne se coordonnent d'aucune manière. Elle fonctionne d'autant mieux que la distribution des choix ou des prédictions s'étend sur un large spectre, de sorte que les erreurs et les biais individuels se compensent. Ce type d'intelligence collective suppose – paradoxalement – l'ignorance mutuelle. Elle s'exprime dans les sondages ou dans les élections, lorsqu'il est interdit de communiquer les résultats partiels avant que tout le monde ait voté. Cette approche de l'intelligence collective statistique sans connexion entre ses membres a été notamment popularisée par James Surowiecki dans son livre « la Sagesse des Foules »[7]

Une seconde forme d'intelligence collective, délibérative, repose au contraire sur la communication directe entre les membres d'une communauté. Elle résulte de l'échange des arguments et des points de vue. Face à un problème commun, elle peut converger sur un consensus ou se partager entre quelques solutions dont on a – tous ensemble – pesé le pour et contre. Pourvu que l'écoute soit au rendez-vous, chacun amène son point de vue, sa compétence particulière, qui enrichit le débat général[8]. Ce type d'intelligence collective, apparemment idéale puisqu'elle est ouverte et réflexive, est d'autant plus difficile à mettre en œuvre que la collectivité est étendue. Il faut alors établir des formes de hiérarchie et de délégation, certes indispensables mais qui troublent la transparence de l'intelligence collective à elle-même.

Je voudrais maintenant introduire une forme d'intelligence collective moins connue mais qui n'en est pas moins à l'œuvre dans nombre de sociétés animales et qui a trouvé son plus haut degré d'achèvement dans l'humanité: la communication stigmergique. L'étymologie grecque explique assez bien le sens du mot « stigmergie » : des marques (*stigma*) sont laissées dans l'environnement par l'action ou le travail (*ergon*) de membres d'une collectivité, et ces marques guident en retour – et récursivement – leurs actions[9]. Le cas classique est celui des fourmis qui laissent une traîne de phéromones sur leur passage lorsqu'elles ramènent de la nourriture à la fourmilière. L'odeur des phéromones incite d'autres fourmis à remonter leurs traces pour découvrir le butin et ramener des vivres à la ville souterraine en laissant par terre à leur tour un message parfumé. Le langage confère à l'humanité un haut degré d'intelligence collective, supérieur à celui des autres mammifères et comparable à celui des abeilles ou des fourmis. Comme d'autres espèces eusociales, nous communiquons en grande partie de manière stigmergique, mais au lieu de marquer un territoire physique au moyen de phéromones ou d'autres types de signaux visuels, sonores ou olfactifs, nous laissons des traces symboliques. Au fur et à mesure de l'évolution culturelle, les signifiants s'accumulent dans des *mémoires externes* de plus en plus perfectionnées : pierres levées, totems, paysages sculptés, monuments, architectures, signes d'écriture, archives, bibliothèques, bases de données. On peut prétendre que toute forme d'écriture qui n'est pas précisément adressée est une forme de communication stigmergique : des traces sont déposées pour une lecture à venir et font office de mémoire externe d'une communauté.

Les différents processus d'intelligence collectives qui viennent d'être évoqués, statistique (sans communication), délibératifs (à communication directe) et stigmergique (à communication indirecte) ne sont évidemment pas exclusives l'une de l'autre et peuvent fort bien se succéder ou se combiner. Par exemple, les wikipédiens se coordonnent par l'intermédiaire de bases de données communes, délibèrent *et* votent.

À l'échelle de l'espèce, l'intelligence collective humaine se situe dans la continuité de l'intelligence collective animale, mais elle est plus perfectionnée à cause du langage, des techniques et des institutions politiques, économiques, légales et autres qui nous caractérisent. La principale différence entre les intelligences collectives animale et humaine tient à la culture. Dans une dimension diachronique, notre espèce est entraînée par une vitesse d'apprentissage supérieure à celle de l'évolution biologique. Nos savoir-faire s'accumulent et se transmettent d'une génération à l'autre grâce à nos mémoires externes, au moyen de systèmes de signes, de conventions sociales et d'outils. Aucun individu ne serait « intelligent » s'il n'héritait pas des connaissances créées par les ancêtres. Dans une dimension synchronique, nous participons à une intelligence collective coordonnée où résonnent et se relancent l'architecture conceptuelle de nos mémoires communes et l'organisation sociale de nos collectivités. La définition réciproque des identités et la reconnaissance des problèmes se décident à ce méta-niveau de la culture. Au-delà des procédures utiles (stigmergique, statistique et délibérative) pour résoudre des problèmes, il existe donc une intelligence collective plus holistique, qui circonscrit les capacités cognitives d'une société.

L'évolution culturelle a déjà franchi plusieurs seuils d'intelligence collective. En effet, les inventions de l'écriture, de l'imprimerie et des médias électroniques (enregistrement musical, téléphone, radio, télévision) ont déjà augmenté de manière irréversible nos capacités de mémoire et de communication sociale. Le surgissement d'une communication globale par l'intermédiaire de la mémoire numérique est probablement le plus grand changement social des vingt-cinq dernières années. Cette nouvelle forme de communication par lecture-écriture distribuée dans une mémoire numérique collective représente une mutation anthropologique de grande ampleur. Plongés dans le nouvel environnement numérique, nous interagissons par le moyen de la masse océanique de données qui nous rassemble. Les encyclopédistes de Wikipédia et les programmeurs de GitHub collaborent par l'intermédiaire d'une même base de données. À notre insu, chaque lien que nous créons, chaque étiquette ou hashtag apposée sur une information, chaque acte d'évaluation ou d'approbation, chaque « j'aime », chaque requête, chaque achat, chaque commentaire, chaque partage, toutes ces opérations modifient

subtilement la mémoire commune, c'est-à-dire le magma inextricable des rapports entre les données. Notre comportement en ligne émet un flux continu de messages et d'indices qui transforment la structure de la mémoire et contribuent à orienter l'attention et l'activité de nos contemporains. Nous déposons dans l'environnement virtuel des phéromones électroniques qui déterminent en boucle l'action des autres internautes et qui entraînent par-dessus le marché les neurones formels des intelligences artificielles (IA).

L'intelligence artificielle comme augmentation de l'intelligence collective

Abordons maintenant le thème de l'intelligence artificielle, mais sous l'angle – qui paraîtra peut-être insolite à quelques lecteurs – de l'intelligence collective. Les journalistes et le grand public ont tendance à classer dans "l'intelligence artificielle" les applications considérées comme avancées à l'époque où elles apparaissent. Mais quelques années plus tard ces mêmes applications, devenues banales et quotidiennes, seront le plus souvent réinterprétées comme appartenant à l'informatique ordinaire[10]. Par delà les titres apocalyptiques et les images de jeunes femmes au cerveau chromé censées illustrer "l'intelligence artificielle", nous assistons depuis le milieu du XXe siècle à un processus de réification formelle et d'extériorisation des fonctions cognitives. L'augmentation de puissance et la baisse des coûts du matériel distribuent ces fonctions cognitives objectivées dans l'ensemble de la société. Des machines interconnectées enregistrent et retrouvent de l'information, effectuent des calculs arithmétiques ou algébriques, simulent des phénomènes complexes, raisonnent logiquement, respectent des syntaxes et des systèmes de règles, extraient des formes à partir de distributions statistiques enchevêtrées... L'informatique automatise et socialise nos capacités de communication, nos facultés de mémoire, de perception, d'apprentissage, d'analyse et de synthèse.

Du fait même de son nom, l'intelligence artificielle évoque naturellement l'idée d'une intelligence autonome de la machine, qui se pose en face de l'intelligence humaine, pour la simuler ou la dépasser. Mais si nous observons les usages réels des dispositifs d'intelligence artificielle, force est de constater que, la plupart du temps, ils augmentent, assistent ou accompagnent les opérations de l'intelligence humaine. Déjà, à l'époque des systèmes experts – lors des années 80 et 90 du XXe siècle – j'observais que les savoirs critiques de spécialistes au sein d'une organisation, une fois codifiés sous forme de règles animant des bases de connaissances, pouvaient être mis à la portée des membres qui en avaient le plus besoin, répondant précisément aux situations en cours et toujours disponibles. Plutôt que d'intelligences artificielles prétendument autonomes, il s'agissait de médias de diffusion des savoir-faire pratiques, qui avaient pour principal effet d'augmenter l'intelligence collective des communautés utilisatrices[11].

Dans la phase actuelle du développement de l'IA, le rôle de l'expert est joué par les foules qui produisent les données et le rôle de l'ingénieur cognitif qui codifie le savoir est joué par les réseaux neuronaux. Au lieu de demander à des linguistes comment traduire ou à des auteurs reconnus comment produire un texte, les modèles statistiques interrogent à leur insu les multitudes de rédacteurs anonymisés du web et ils en extraient automatiquement des patterns de patterns qu'aucun programmeur humain n'aurait pu tirer au clair. Conditionnés par leur entraînement, les algorithmes peuvent alors reconnaître et reproduire des données correspondant aux formes apprises. Mais parce qu'ils ont abstrait des structures plutôt que de tout enregistrer, les voici capables de conceptualiser correctement des formes (d'image, de textes, de musique, de code...) qu'ils n'ont jamais rencontrées et de produire une infinité d'arrangements symboliques nouveaux. C'est pourquoi l'on parle d'intelligence artificielle *générative*. L'IA neuronale synthétise et mobilise la mémoire

commune. Bien loin d'être autonome, elle prolonge et amplifie l'intelligence collective. Des millions d'utilisateurs contribuent au perfectionnement des modèles en leur posant des questions et en commentant les réponses qu'ils en reçoivent. On peut prendre l'exemple de Midjourney, dont les utilisateurs s'échangent leurs consignes (*prompts*) et améliorent constamment leurs compétences en IA. Les serveurs Discord de Midjourney sont aujourd'hui les plus peuplés de la planète, avec plus d'un million d'utilisateurs. Une nouvelle intelligence collective stigmergique émerge de la fusion des médias sociaux, de l'IA et des communautés de créateurs.

L'IA contemporaine fonctionne ainsi comme le conduit d'une boucle de rétroaction entre la mémoire numérique commune et les productions individuelles qui l'exploitent et s'accumulent à leur tour dans les centres de données. Derrière la machine il faut entrevoir l'intelligence collective qu'elle réifie et mobilise.

Le partage du savoir : vers une intelligence collective neurosymbolique

L'intelligence collective aujourd'hui supportée par l'intelligence artificielle n'est encore que partielle. En effet, l'utilisation des données de l'Internet pour entraîner les modèles mobilise les intelligences collectives stigmergique (la boucle de rétroaction entre les comportements individuels et la mémoire commune) et statistique (l'apprentissage neuronal). Au début des années 2020, la connexion et le renforcement mutuel de ces deux formes d'intelligence collective par les nouveaux dispositifs d'intelligence artificielle a provoqué un choc intellectuel – et des émotions fortes – chez ceux qui en ont aperçu la puissance. Mais une intelligence collective délibérative et réflexive manque encore à l'appel. A l'échelle où nous nous situons, cette intelligence collective délibérative doit porter sur l'organisation des données, c'est-à-dire sur la structure conceptuelle de la mémoire, inévitablement couplée aux pratiques des communautés. Comment faire en sorte que les réseaux de concepts qui informent la mémoire numérique puissent faire l'objet d'une conversation ouverte, transparente, attentive aux conséquences de ses choix? Le Web sémantique et son empilement de standards (XML, RDF, OWL, SPARQL) a certes établi une interopérabilité de formats, mais non pas l'interopérabilité proprement sémantique – celle des architectures de concepts – dont nous avons besoin. Les géants du web ont leurs graphes de connaissances, mais ces derniers sont malheureusement privés et secrets. Wikidata propose un exemple de graphe de connaissance ouvert, mais il est encore bien difficile à explorer et utiliser quotidiennement par le grand public. Il se présente de plus comme *une* ontologie, celle de l'encyclopédie Wikipédia, alors qu'il faudrait mettre en harmonie et en dialogue la multitude des ontologies qui émergent de pratiques aussi diverses que l'on voudra.

C'est pour résoudre ce problème de l'émergence d'une intelligence collective délibérative (ou réflexive) à support numérique que j'ai inventé IEML (Information Economy MetaLanguage) : une langue artificielle pourvue d'une structure algébrique régulière, dont la sémantique est calculable, qui permet de tout dire et qui peut traduire n'importe quel réseau de concept[12]. IEML est un langage à source ouverte, qui se place dans la perspective d'une augmentation des communs de la connaissance, et dont le développement doit faire l'objet d'une gouvernance décentralisée. Aussi hétérogènes ou divers qu'ils soient, IEML projette les ontologies, graphes de connaissances, collections d'étiquettes et modèles de données sur le même système de coordonnées sémantique : un univers virtuellement infini de différences conceptuelles donnant prise aux algorithmes. IEML peut servir de *langage pivot* entre les langues naturelles, entre les humains et les machines, entre les modèles d'IA. Il va sans dire que la plupart de ses bénéficiaires n'auront pas à l'apprendre puisque les interfaces des applications, y compris l'éditeur lui-même, seront en langues naturelles ou sous

forme iconique. La face « code » d'IEML n'est destinée qu'aux ordinateurs. On peut dès lors envisager qu'une multitude de bases de connaissances aux architectures conceptuelles singulières puissent échanger des modules ontologiques et des informations grâce à l'interopérabilité sémantique assurée par ce *langage de métadonnées* commun.

Considérons maintenant l'idéal des lumières de *mettre la connaissance à la portée de tous*. Cette finalité dépasse l'objet « encyclopédie » – qui n'est finalement qu'un moyen particulier adapté aux possibilités techniques et culturelles d'une époque – pour ouvrir de vastes horizons et retentir jusque dans un avenir encore inimaginable.

Ce concept se décompose en deux exercices complémentaires: (a) celui de permettre à toutes les connaissances de s'exprimer, de s'accumuler et de communiquer ; (b) celui de faciliter l'exploration et l'appropriation des connaissances selon la gamme étendue des situations pratiques, des parcours d'apprentissage et des styles cognitifs. On voit l'affinité de cet idéal avec celui d'une intelligence collective – diamétralement opposée au "*group think*" – qui vise à maximiser simultanément la liberté créatrice et l'efficacité collaborative.

On pardonnera au philosophe que je suis l'évocation d'une utopie concrète, sans doute techniquement réalisable, mais qui – à court terme – vise d'abord à faire penser. Imaginons donc un dispositif destiné au partage des connaissances et qui tire le maximum des possibilités techniques contemporaines. Au cœur de ce dispositif évolue *un écosystème ouvert de bases de connaissances* catégorisées en IEML, qui émergent d'une multitude de communautés de recherche et de pratique. Entre ce noyau de bases de connaissances interopérables et les utilisateurs humains vivants s'interpose une *interface neuronale* (un écosystème de modèles) « *no code* » qui donne accès au contrôle, à l'alimentation, à l'exploration et à l'analyse des données. Tout se passe de manière intuitive et directe, selon les modalités sensori-motrices sélectionnées. C'est aussi par l'intermédiaire de ce giga-perceptron – un métavers immersif, social et génératif – que les collectifs échangent et discutent les modèles de données et réseaux sémantiques qui organisent leurs mémoires. En bonne gestion des connaissances, le nouveau dispositif de partage des savoirs favorise l'enregistrement des créations, accompagne les parcours d'apprentissage et présente les informations utiles aux acteurs engagés dans leurs pratiques.

Pour ce qui est commun, chaque base de connaissance – personnelle ou collective – affiche son univers de discours, ses données et ses statistiques, aussi transparente aux algorithmes qu'elle l'est aux regards humains. Mais pour ce qui est privé, notre dispositif de partage des connaissances assure la *souveraineté* pratique et légale des individus et des groupes sur les données qu'ils produisent et qu'ils ne divulguent qu'aux acteurs choisis.

L'augmentation décisive de la dimension délibérative de l'intelligence collective grâce à l'utilisation d'un langage de métadonnées commun a des effets multiplicateurs sur les intelligences collectives statistique et stigmergique déjà à l'œuvre aujourd'hui. Une nouvelle infrastructure neurosymbolique plonge l'intelligence collective du futur dans l'univers explorable émanant de ses propres activités cognitives. Il faut cependant bien distinguer l'intelligence collective qui anime les personnes et les collectivités humaines vivantes des extensions mécaniques et des représentations médiatiques qui l'augmentent. Ne faisons pas une idole de l'intelligence artificielle.

Citant Ibn Roshd (l'Averroès des latins), Dante écrit au chapitre I, 3 de sa Monarchie : "Le terme extrême proposé à la puissance de l'humanité est la puissance, ou vertu, intellectuelle. Et parce que cette puissance ne peut, d'un seul coup, se réduire toute entière en acte par le moyen d'un seul homme ou d'une communauté particulière, il est nécessaire qu'il règne dans le genre humain une multitude par le moyen de laquelle soit mise en acte cette puissance toute entière." Que cette multitude devienne transparente à elle-même dans le nouveau médium algorithmique et nous serons passés de la fourmilière à la cité.

- [1] https://meta.wikimedia.org/wiki/Abstract_Wikipedia consulté le 8 juillet 2023
- [2] <https://www.knowledgetech.com/> consulté le 8 juillet 2023
- [3] https://www.theregister.com/2023/06/16/crowd_workers_bots_ai_training consulté le 8 juillet 2023
- [4] Lévy, Pierre. *L'Intelligence collective. Pour une anthropologie du cyberspace*. Paris: La Découverte, 1994.
- [5] Voir par exemple : Baltzersen, Rolf. *Cultural-Historical Perspectives on Collective Intelligence: Patterns in Problem Solving and Innovation*. Cambridge, Mass: Cambridge University Press, 2022.
- [6] Galton, Francis 1907, *Vox populi*, *Nature*, 75, 450-451.
- [7] Surowiecki, James. *The Wisdom of Crowds*. Doubleday, 2004
- [8] Voir par exemple : Zara, Olivier. *Le chef parle toujours en dernier: Manifeste de l'intelligence collective*, Axiopole, 2021, et Mulgan, Geoff. *Big Mind. How Collective Intelligence Can Change Our World*. Princeton: Princeton University Press, 2017.
- [9] Heylighen, Francis. "Stigmergy as a Universal Coordination Mechanism I: Definition and Components." *Cognitive Systems Research* 38 (2016): 4–13. <https://doi.org/10.1016/j.cogsys.2015.12.002>
- Heylighen, Francis. "Stigmergy as a Universal Coordination Mechanism II: Varieties and Evolution." *Cognitive Systems Research*, 2016, 50–59. <https://doi.org/10.1016/j.cogsys.2015.12.007>
- [10] Lévy, Pierre. "Pour un changement de paradigme en Intelligence artificielle", *Giornale di Filosofia* (Roma) numéro spécial sur *Technology and Constructive Critical Thought*, 15 décembre 2021
- [11] Lévy, Pierre. "Les systèmes à base de connaissance comme médias de transmission de l'expertise", *Intellectica* numéro spécial "Expertise et sciences cognitives", ed. Violaine Prince. 1991. p. 187 – 219.
- [12] Lévy, Pierre. *La Sphère Sémantique. Computation, Cognition, Économie de l'information*. Paris-London: Hermès-Lavoisier, 2011.
- Id. « Le calcul sémantique avec IEML », *Humanités numériques*, à paraître dans le numéro d'hiver 2023. <https://journals.openedition.org/revuehn/>
 Tags artificial intelligence, categorization, Collective intelligence, IEML
 Categories French, Semantic Sphere

2 Comments

William THEAUX

16 July 2023 at 1:51 pm

J'ai fait référence à votre page et travaux à <https://youtu.be/JoEObv7TavM> ainsi que dans son descriptif <https://www.lasainteethique.org/2023/htm/libemedic.htm#indiv> . Ce sont aussi vos travaux qui m'avaient encouragé à soutenir en 2018 <https://youtu.be/Hru4AwlDDi4> où des politiciens de l'auditoire avaient soudain compris que je parlais de 'cyberdémocratie' ! Ils émirent un sourd grondement (comme j'en avais l'habitude ^_^). Merci pour vos remarquables conceptions.

Reply

Pierre Lévy

16 July 2023 at 3:58 pm



Reply

Blog at WordPress.com.

Back to the top