

La Relève et La Peste

Pour entraîner l'IA, les géants de la tech achètent des millions de livres avant de les détruire

<https://lareleveetlapeste.fr/pour-entraîner-lia-les-geants-de-la-tech-achètent-des-millions-de-livres-avant-de-les-détruire/>

Cette destruction massive de livres résonne comme une version moderne de l'autodafé. Si l'autodafé historique procédait d'une volonté dogmatique, religieuse ou politique visant à censurer et anéantir

la pensée par le feu, la numérisation destructive obéitelle à une implacable rationalité capitaliste et productiviste. Dans les deux cas, une forme de totalitarisme dénué de bon sens est à l'œuvre.

Pour nourrir leurs algorithmes en données textuelles, les géants de la technologie se lancent dans l'achat massif d'ouvrages imprimés. Mais derrière cette ruée vers le papier se cache une réalité cynique : après avoir été numérisés à la chaîne, des millions de livres sont détruits. Le constat est posé : les livres imprimés sont souvent les mieux écrits et les plus fiables, puisqu'ils sont édités. Pour entraîner une IA de qualité, il faut donc des textes de qualité. Partant de ce constat, les géants de l'IA se tournent aujourd'hui vers nos bibliothèques pour nourrir leurs algorithmes.

Le « Project Panama »

On pourrait croire à un scénario dystopique d'Orwell, mais c'est pourtant la réalité. Le Project Panama est un projet secret lancé par l'entreprise Anthropic en 2024, consistant à acheter et à détruire des millions de livres anciens imprimés afin de constituer une bibliothèque numérique géante pour entraîner ses modèles d'intelligence artificielle.

Décrit par le Washington Post comme un « *effort pour numériser de manière destructive tous les livres du monde* », le projet a fait sauter de nouvelles digues dans l'univers de l'IA. En l'espace d'un an, l'entreprise Anthropic a dépensé plusieurs dizaines de millions de dollars pour acquérir des ouvrages par lots de dizaines de milliers, conscients qu'ils seraient détruits.

Cette ruée vers le papier bouleverse profondément le marché de

l'occasion. Depuis 2024, plusieurs libraires spécialisés constatent une hausse inhabituelle et marquée des commandes. Comme le témoigne un libraire anonyme interrogé par 404 Media en juillet 2026 : « En avril, je suis soudainement passé de pas plus de 20 livres vendus par semaine à des centaines, et je suis presque certain que les clients sont des laboratoires d'IA. » Les acheteurs recherchent des textes rares, dans de multiples langues et sur des sujets variés, transformant le patrimoine littéraire mondial en nourriture pour IA.

Autodafé 2.0

L'objectif ? Produire toujours plus, et toujours plus vite. Les prestataires mandatés par la tech utilisent une méthode radicale appelée la *numérisation destructive* : la reliure de chaque livre est tranchée net pour désolidariser l'ensemble des pages. Les feuilles volantes sont ensuite introduites dans des scanners industriels à grande vitesse afin d'être converties en copies numériques exploitables par les algorithmes. Une fois l'opération terminée, les exemplaires réduits en miettes sont directement envoyés au recyclage. En un an, plusieurs millions de livres ont été détruits.

Cette destruction massive de livres résonne comme une version moderne de l'autodafé. Si l'autodafé historique procédait d'une volonté dogmatique, religieuse ou politique visant à censurer et anéantir la pensée par le feu, la numérisation destructive obéitelle à une implacable rationalité capitaliste et productiviste. Dans les deux cas, une forme de totalitarisme dénué de bon sens est à l'œuvre.

Les géants de la tech détruisent le livre physique mais l'intelligence artificielle n'offre aucune garantie quant à l'exactitude de la transcription, de la restitution ou de la compréhension subtile des textes ainsi digérés. Une fois de plus, l'IA peut ainsi modeler et modifier le réel comme bon lui semble. Le livre en tant qu'objet culturel authentique est définitivement anéanti pour ne laisser place qu'à une simple modélisation probabiliste de la pensée.

La bataille culturelle

Des actions en justice d'envergure, à l'image des class actions menées par des collectifs d'auteurs, attaquent frontalement le pillage des œuvres. Les syndicats et associations professionnelles à l'instar de la Society of Authors et de la Publishers Association montent vigoureusement au front

pour briser l'omerta.

Portée par l'Authors Guild et des écrivains au Tribunal fédéral de San Francisco (Californie), l'action collective Bartz c. Anthropic a démarré en 2025, stoppant net le Project Panama. Elle aboutit en juillet 2026 d'un accord historique de 1,5 milliard de dollars pour sanctionner l'exploitation illégale d'œuvres protégées. C'est officiellement la plus grande affaire de violation de droits d'auteur de l'histoire des États Unis. Comme le résume Mary Rasenberger, directrice générale de l'organisation :

« Nous espérons que cela envoie un message fort à l'industrie de l'IA : il y a de lourdes conséquences lorsque l'on pille les œuvres des auteurs pour entraîner ses intelligences artificielles. »

Malgré cette affaire monstre, les laboratoires d'IA cherchent à explorer toutes les zones grises juridiques pour trouver de nouvelles sources de données. La bataille juridique continue d'ailleurs sur d'autres terrains avec d'autres entreprises (OpenAI, Meta, etc.) Cet accord marque un coup d'arrêt pour l'utilisation directe de bibliothèques pirates et de méthodes clandestines comme le Project Panama mais la question de l'entraînement des IA sans l'accord explicite des créateurs – et sans destruction massives de livres – reste loin d'être réglée.