

DES LIVRES SACRIFIÉS POUR ENTRAÎNER L'IA

Des libraires européens voient affluer des commandes automatisées pour des ouvrages anciens, spécialisés et souvent invendus depuis des décennies. Achetés par des intermédiaires pour le compte d'entreprises d'intelligence artificielle, ces livres seraient numérisés puis détruits afin d'alimenter leurs modèles. Aubaine pour les bouquinistes, cette pratique soulève de nombreuses questions sur le droit d'auteur et la préservation du patrimoine écrit. [Temps de lecture estimée : ~ 4min]

Depuis plusieurs mois, un étrange phénomène intrigue libraires et bouquinistes. Des milliers de livres anciens et rares sont achetés, en masse, sur leur site en ligne. [Selon la RTS, la Radio Télévision Suisse](#), « [...] début mai, une librairie allemande de livres anciens recevait sur son site en ligne chaque nuit, entre 3h00 et 5h00 du matin, **un nombre considérable de commandes automatisées d'ouvrages** ».

Un évènement non isolé, car de nombreux confrères, notamment sur le réseau social *Reddit*, en font également constat. **De la Nouvelle-Zélande, à l'Espagne, en passant par les Pays-Bas ou encore l'Australie et la Bulgarie, le même procédé a lieu.**

Nouvel engouement pour les ouvrages anciens ?

Ce qui intrigue surtout les professionnels, ce sont les genres d'ouvrages sélectionnés : **des livres peu recherchés**, très spécialisés et invendus depuis des décennies, tels que des manuels techniques ou encore des livres de cuisine. Et chaque titre est méticuleusement acheté en un seul exemplaire.

Une méthode spécifique, rythmée, presque « algorithmique »... **Plusieurs médias initient des enquêtes et portent alors leur regard sur les destinataires** de ces commandes : *Better World Books*, *World of Books* et *Zoom Books*. Des centres de logistique basés aux États-Unis. Selon la *RTS* ou encore [le média El Diario](#) qui enquête sur le phénomène, ces sociétés seraient **des intermédiaires logistiques pour des sociétés américaines spécialisées en IA**. Mais une question reste en suspens : pourquoi acheter des milliers de livres dont presque personne ne veut ?

Les livres : nouvelles données pour l'IA

Pour comprendre ce phénomène, il faut revenir au fonctionnement des [intelligences artificielles génératives](#). Les grands modèles de langage sont

entraînés à partir de **gigantesques quantités de textes**. Internet a longtemps constitué une source particulièrement abondante de données libres d'accès : articles, pages web, forums, documents accessibles publiquement...

Mais rapidement, les ressources disponibles se sont raréfiées. Pour poursuivre leur développement, les entreprises d'IA se sont alors dirigées vers des **données non libres d'accès**, en passant par des plateformes de téléchargement illégal d'ouvrages. Des techniques qui les ont rapidement menés à des poursuites judiciaires, comme tel a été le cas pour l'entreprise *Anthropic*, à l'origine du modèle Claude, qui a **été condamné en 2025 à verser 1,5 milliard de dollars** aux auteurs d'ouvrages piratés.

Pour les sociétés d'IA **ont alors cherché** une solution : se tourner vers des ouvrages anciens et moins populaires. Un marché de grande envergure, peu coûteux et surtout, capable de **passer dans les mailles du filet législatif**.

Car en utilisant les livres papiers, les sociétés agissent **sous couvert du principe étatsunien du « fair use », ou « principe d'utilisation équitable »**. Une exception au droit d'auteur américain qui **permet d'utiliser une partie d'un livre protégé sans l'autorisation de l'auteur**, à condition que les fins soient spécifiques, comme la critique, l'enseignement ou la recherche. Et c'est ainsi que le livre papier entre paradoxalement dans l'histoire de la technologie la plus dématérialisée qu'il soit.



Des millions de livres découpés sans retour arrière possible... Jusqu'où peut-on prétendre à du « fair use » ? / [Pexels](#)

Des livres achetés pour être... détruits

Des intermédiaires logistiques comme Zoom Books collecteraient les ouvrages avant de les envoyer vers des centres de numérisation. Pour numériser rapidement chaque livre, les opérateurs en **découpent le dos afin d'en séparer les pages, qui sont ensuite introduites dans des scanners**. Une fois cette opération réalisée, les exemplaires ne peuvent plus être remontés ni lus comme des livres.

« [cette opération est un] effort visant à numériser de manière destructive tous les livres du monde » Ce procédé n'est pas une simple hypothèse développée par quelques médias. En janvier 2026, [le Washington Post a révélé](#) une **longue enquête sur le « projet Panama »**, une opération secrète lancée par la start-up américaine *Anthropic*, décrit en interne comme un « *effort visant à numériser de manière destructive tous les livres du monde* ».

Pour les professionnels du livre d'occasion, la situation est pour le moins **paradoxe**. D'un côté, ces commandes représentent une manne financière inattendue : **des ouvrages qui dormaient depuis des années dans des réserves trouvent soudainement preneur**. De l'autre, les libraires découvrent que certains de ces livres pourraient être achetés non pas pour être lus, collectionnés ou transmis, mais pour être **dépouillés de leur contenu avant d'être détruits**.

Une situation particulièrement troublante pour un secteur qui ne se contente pas de vendre des livres : **les bouquinistes participent aussi à leur conservation**. C'est là que réside tout le **paradoxe éthique** de cette pratique.



Bouquinistes sur les quais de Paris, mars 2017 / [WikimediaCommons](#)

Vers la fin du savoir de l'humanité ?

Un livre représente bien plus qu'un ensemble de données. Il est aussi et avant tout un **outil de transmission physique**. Depuis des millénaires, des parois rocheuses d'une grotte aux feuilles de papyrus, l'ensemble des connaissances du monde se partage, se diffuse et se transmet **à travers des supports matériels**.

Avec l'essor de l'intelligence artificielle, une nouvelle question se pose désormais : **que devient cette fonction de transmission lorsque les livres sont réduits à des ensembles de données destinés à alimenter des modèles ?**

La collecte massive d'ouvrages à cette fin soulève ainsi des enjeux qui dépassent la seule question de leur destruction physique. **Elle interroge les conditions** dans lesquelles les œuvres peuvent être numérisées, utilisées et intégrées aux modèles d'IA, mais aussi **la manière dont nous souhaitons préserver et transmettre le patrimoine écrit**. Face à ces nouvelles pratiques, le cadre juridique devra trouver un équilibre entre

innovation technologique, droits des auteurs et préservation de l'accès au savoir.

– *Pauline Lunay*

– <https://mrmondialisation.org/livres-sacrifies-entrainer-ia/>
