

Project Panama: When is « too much » going to be enough?

The EWC condemns the AI tech practice of « destructively scanning » millions of old or rare European books and calls upon decision makers and the whole of our European society to speak up loudly on behalf of authors, publishers, and readers on this predatory exploitation.

AI companies are literally ripping apart authors' works and turning them into toilet paper: as of today, non-European tech giants have purchased at least up to two million books through second hand and antiquarian booksellers across Europe and the world. Old, rare, and also out-of-commerce titles (which in the EU are still protected under author rights legislation but are unavailable on the market) are subject to bulk scanning and being destroyed in the process. Specifically, at the AI scan factories in the U.S., this involves cutting the spine from a book, so its pages can be fed through high-speed scanners before the remaining physical copy is discarded, shredded, and recycled into toilet paper. The digitised files are exploited for the development of large language models, foundations of e.g. Anthropic's "Claude Sonnet" or Amazon's AI software.

"For those scanning programmes, the AI tech companies are operating dedicated, industrial-scale book-scanning facilities in order to train their LLMs with books published before 2022 – to avoid feeding their systems with their own reproduced AI slop," explains Maïa Bensimon, EWC Vice President. "Rare books are destroyed. Many such books will never be found again and will exist solely as the digital property of an AI tech company."

"Non-European tech companies are plundering Europe's trove of knowledge and literature", says Sebastià Portell, Catalan writer and EWC President. "AI developers have invested neither in authors nor in the funding schemes maintained by the Member States to ensure that book culture can flourish. They are building a digital archive of our works, our knowledge and our European values, and controlling access, retrieval and use. This implies censorship, but also future business models will monetise access to our books, undermining the fundamental rights that underpin Europe, as well as the principles of decency."

"The AI company Anthropic and its secret 'Project Panama' aims to 'scan all the books in the world' in its rat race for training data instead of respecting copyright and paying fees," assumes Nina George, EWC's Policial Commissioner.

From 2024 to 2025 Anthropic spent tens of millions of dollars to acquire rare and out-of-commerce works, including from European antiquarian stores or American-based antiquarians, which are holding stock from overseas' and foreign countries – books by European authors.

Yet all we learn about this is a two-line summary in the transparency template on training data recently published by Anthropic: *A portion of the data corpus comes from acquired physical texts.*

Here, too, it is evident that Europe's decision-makers are neither capable of imagining the worst-case scenario – namely, how non-European companies might operate beyond the bounds of law and decency – nor of introducing the tools needed to enforce our rights and protect the intellectual assets of Europe's third-largest economic power.

The EWC denounces this shameless, frantic and clandestine race to acquire the collected written works of humanity. Trying to subvert copyright legislation is one thing but destroying our literary cultural heritage is another.

The EWC calls upon the European Commission, the European Parliament, and upon every European State to take a clear stance against this malicious practice.

BACKGROUND

In many EU countries, books are regarded as cultural heritage, cultural assets, or cultural goods. UNESCO, for its part, considers the destruction of cultural heritage – such as the systematic destruction of books – to be a serious breach of the Hague Convention.

Under Union law, the scan and further utilisation of purchased print books is not legal. In a written answer to a request by MEP Andreas Schwab (EPP) on the legality of *“this predatory practice (that) permanently removes rare books from secondary markets, exploiting an analogue loophole to mask data provenance,”* Commissioner Henna Virkkunen answered on 16th September 2026: *“The digitisation of books protected by copyright, including books lawfully acquired in analogue form, constitutes an act of reproduction requiring the rightsholder’s prior authorisation unless a relevant exception or limitation provided for under EU law and implemented in national law applies. As the text and data mining exceptions only apply to text and data in digital form, they do not cover the digitisation of works in physical format.”*

However, the practice of scanning books that were not originally created in digital form – but as printed publications – has been an ugly tradition by U.S.-based Tech monopolists for over 25 years:

In the early 2000s, Google launched its Google Books programme. This meant scanning millions of books in order to feed its search engine with digital copies of books. Back at that time, copyright was still meaningful, even though tech companies would fight it in court. At that time also, the easiest path to scanning was to put scan machines within libraries. Machines would scan and turn pages. This was already incredible machinery. Then the books would be put back on shelves and Google would plea, in the U.S., that the process and subsequent use of the texts was Fair Use.

Fortunately, in Europe, the unauthorised scanning in libraries stopped after a French ruling condemning Google for counterfeiting acts in 2009. It is important to be reminded of the strong differences between U.S. legislation and European legislation when it comes to intellectual property. Fair Use does not exist in Europe.

But tech companies laying hold to every sentence in books scanned from libraries was not enough: Digital communication companies then developed robots (“crawlers”) that would scan the whole web and not only publicly accessible and by the rightsholders authorised works and material. They took everything, claiming that whatever works can be found on the web are available for them to use, avoiding the obvious fact just because something can be obtained, doesn’t mean it is public property. There are, as of today, more than one hundred legal actions in the world against tech companies, now called AI tech companies and their development of large language AI models (LLMs) and further generative applications, which are built on the works of authors, artists, performers and every person ever having something posted, commented or published in the internet.

But feeding companies with digital content, including illegally copied books from shadow libraries and piracy sites, seems insufficient for AI tech.

AI tech companies have now started eating up physical books. The old Google way in libraries reveals itself as no longer adequate to satisfy their T-Rex-like appetite. They desire more. It’s time to stop them – because Enough is Enough.