Penguin Random House books now explicitly say ‘no’ to AI training

Book publisher Penguin Random House is putting its stance on AI training in print. The standard copyright page on both new and reprinted books will now say, “No part of this book may be used or reproduced in any manner for the purpose of training artificial intelligence technologies or systems,” according to a report from *The Bookseller* spotted by *Gizmodo*.

The clause also notes that Penguin Random House “expressly reserves this work from the text and data mining exception” in line with the European Union’s laws. *The Bookseller* says that Penguin Random House appears to be the first major publisher to account for AI on its copyright page.

What gets printed on that page might be a warning shot, but it also has little to do with actual copyright law. The amended page is sort of like Penguin Random House’s version of a robots.txt file, which websites will sometimes use to ask AI companies and others not to scrape their content. But robots.txt isn’t a legal mechanism; it’s a voluntarily-adopted norm across the web. Copyright protections exist regardless of whether the copyright page is slipped into the front of the book, and fair use and other defenses (if applicable!) also exist even if the rights holder says they do not.

The Verge contacted Penguin Random House for more information but didn’t immediately hear back.

In August, Penguin Random House published a statement saying that the publisher will “vigorously defend the intellectual

property that belongs to our authors and artists.” Not all book publishers are cautious about AI, as academic publishers like Wiley, Oxford University Press, and Taylor & Francis have already formed AI training deals.

https://www.theverge.com/2024/10/18/24273895/penguin-random-house-books-copyright-ai