

**United States Senate
Committee on the Judiciary
Subcommittee on Crime and Counterterrorism**

**Statement of David Baldacci
July 16, 2025**

Mark Twain once said that travel is fatal to prejudice, meaning if you meet people where they live, you find out they're just like you. However, I had no chance to leave the segregated world of Richmond, Virginia when I was growing up. But I visited the library every week and I like to think through books I traveled the world without a plane ticket or a passport. And born from my love of reading came my desire to be a writer. I worked away for decades, getting rejected over and over. But I kept going, honing my craft, remaining disciplined, taking the rejections head on and using them as motivation. And finally, I was successful. And after sixty novels under my belt, I work just as hard as I ever have. It's the American way. Work hard, play fair, stay the course and you'll make it.

I truly believed that until my son asked ChatGPT to write a plot that read like a David Baldacci novel. In about five seconds three pages came up that had elements of pretty much every book I'd ever written, including plot lines, character names, narrative, the works.

That's when I found out the AI community had taken most of my novels without permission and fed them into their machine learning system.

I truly felt like someone had backed up a truck to my imagination and stolen everything I'd ever created.

I'm aware of the argument that what AI did to me and other writers is no different than an aspiring writer reading other books and learning how to use them in original ways.

I can tell you from personal experience that is flatly wrong.

I was once such an aspiring writer. My favorite novelist in college was John Irving. I read everything that Irving wrote. None of my novels read remotely like an Irving novel. Why? Well, unlike AI, I can't remember every line that Irving wrote, every detail about his characters, and his plots. The fact is, also unlike AI, I read other writers not to copy them but because I loved their stories, I appreciate their talent, it motivated me to up my game. What AI does is take what writers produce as an incredibly valuable shortcut to teach software programs what they need to know.

And I have learned that these trillion-dollar companies didn't even buy my books. They got them off websites that have pirated works. They complain that it would be far too difficult to license the works from individual creators. So apparently it was more efficient to steal it. Trillion-dollar companies with battalions of lawyers did not have the resources to do things lawfully? I was once a trial lawyer. If I had made that argument in court I would either have been laughed out of the courtroom or held in contempt by the judge. And rightly so.

Keep in mind that copyrighted books don't simply end up in AI training datasets as part of some indiscriminate sweep of the internet. Complete books are not posted online by their copyright owners like website content, blogs, news articles, or other text. Books are unique in that they are sold as digital files that include technical protection measures against copying and downloading through online retailers like Amazon, Barnes and Noble, Kobo, and others, to be read on digital devices. So the only way for the AI companies to access free books online was through pirate websites, virtually all of them based abroad, in Russia, Ukraine, and other countries outside the reach of U.S. law enforcement. And it was not an isolated instance of one bad actor—every major large language model in commercial use today was trained on pirated books, apparently with the full knowledge and authorization of the companies' highest decision-makers. This is the largest criminal-

level¹ copyright infringement ever perpetrated in this country and it was committed by some of the wealthiest companies in the country. The pirate sites they used are among the most notorious on the web that authors and publishers have been trying to get shut down for nearly a decade without success. Why? For one, lawsuits against these sites are often pyrrhic victories since civil judgments can be extremely difficult to enforce in foreign jurisdictions. For example, one of the most notorious sites, Library Genesis, has 7.5 million books, each one representing years of labor by its author. It was slapped with a \$15 million judgment in 2017 and another \$30 million in 2024 in a separate lawsuit, but continues to operate. Another notorious site, Z-Library, was indicted in federal court in 2022. The FBI seized 240 domains used by the site and arrested two of its principals—Russian nationals—in Argentina. But the site soon came back online under new domains and its pirate book repositories were reuploaded by others.

And it wasn't just the *quantity* of books that made these sites attractive to the AI companies; equally important was the quality. If AI companies only needed words, they could have fed every dictionary in the world into their machine learning. But that was not nearly good enough because it would mean decades of additional work and hundreds of billions of dollars of additional investment. What they needed was complete, well-crafted, living, breathing stories, with characters that seemed real, plots that made sense, dialogue that appeared genuine. Humanity on the page. In sum, they needed us and our craft that we earned with the sweat of our brows and the flexing of our imaginations. And these companies have swooped in, stolen that labor in order to make enormous profits. But we, the writers, the source of all this, will receive nothing.

AI will also allow anyone, with no effort at all, to order up a novel written in the vein of an established writer. And that book can be sold saying that it reads just like a David Baldacci novel. Yes, it does read like my novels. Because it is my novel. It is my imagination.

¹ See 17 U.S.C. § 506.

People complain about cheap imported goods hurting American workers. Well, we have cheap books being created by American technology flooding the market. The Authors Guild receives reports that for many if not most new anticipated top selling books (they usually have significant advance sales), an AI-generated book that is intended to directly compete against the real book and divert sales is posted the day of or even before the release date of the real book. As AI becomes more widespread, the number of such books will only increase, forcing authors into an endless game of whack-a-mole. That will mean lower profits for publishers, and less money to spend on new, emerging writers. That hurts all of us. Online vendors now require the “author” to disclose if a book was not human created. It’s getting to the point where they will have to limit the number of books that someone can publish on a weekly or even daily basis. This is insane.

All this comes at a time when writers are already facing unprecedented hurdles in earning a living. Between 2009 and 2018, authors’ median incomes dropped 42%.² The Authors Guild’s most recent authors’ earnings survey found that the median writing-related income for full-time authors in 2022 was just over \$20,000, with only half of that from books.³ Looking at all authors, including those who reported writing part-time, the median book income was \$2,000 in 2022, and the median income from books plus other writing-related work was \$5,000.

Stemming this tide of piracy and unfair competition requires congressional action. I urge Congress to pass legislation that gives copyright owners the ability to obtain judicial orders blocking access to foreign piracy sites or online services. In addition, Congress should adopt commonsense transparency legislation requiring AI companies to disclose any unlicensed copyrighted works used in training. And to help ensure that consumers aren’t

² Authors Guild Survey Shows Drastic 42 Percent Decline in Authors Earnings in Last Decade, <https://authorsguild.org/news/authors-guild-survey-shows-drastic-42-percent-decline-in-authors-earnings-in-last-decade/>

³ Key Takeaways from the Authors Guild’s 2023 Author Income Survey, <https://authorsguild.org/news/key-takeaways-from-2023-author-income-survey/>

deceived into purchasing machine-generated outputs aimed at capitalizing on the work of human authors, Congress should require that AI-generated content be labeled as such.

More broadly, I urge Congress to consider what the AI companies' position means for copyright as a whole and the future of the creative professions in this country. Source code and elements of algorithms are also protected by copyright. I would hazard to bet that if I stole any of the AI communities' source codes or algorithms and then tried to profit off them, they would unleash a tsunami of lawsuits against me. However, if, as AI companies contend, fair use is actually my entire body of work, there is no more copyright protection for anyone. I'm sure the AI community believes that their IP should be fully protected against interlopers, and I agree with them. Thus I am deeply disappointed that they don't feel that people such as myself should enjoy the same rights and protections.

The AI community apparently believes that they are entitled to steal our work product despite it being copyrighted because what they are doing is so transformational. Well, billions of people have been transformed by books. Many significant events in human history had seminal authors and their works that rode at the head of the pack. We didn't truly emerge from the dark ages until the invention of the printing press when books became widely available. Books also teach us empathy, making the world a kinder, gentler, more meaningful place.

I'm only one man but books transformed my life, propelling me to a far better existence.

I'm sure there are aspects of AI that will also transform the world.

But if you want to bet on which side is more transformational, for all of us, I will bet on books every single time.

Thank you.