



**Committee:** Senate Judiciary Subcommittee on Crime & Counterterrorism  
**Event:** [Too Big to Prosecute?: Examining the AI Industry's Mass Ingestion of Copyrighted Works for AI Training](#)  
**Date:** July 16, 2025  
**Time:** 12:00 PM  
**Place:** 226 Dirksen Senate Office Building

---

***Member Toplines:***

*Chair Josh Hawley (R-MO)*<sup>1</sup>: Hawley attacked artificial intelligence (AI) companies for what he called the largest intellectual property (IP) theft in U.S. history. He criticized these companies, such as Meta and Anthropic, for intentionally stealing copyrighted material from illegal online repositories and using it to train their AI models. He claimed that they have stolen billions of pages of copyrighted works, enough to fill 22 libraries the size of the Library of Congress. Hawley called for the government to protect U.S. citizens from Big Tech's destruction of IP.

*Ranking Member Richard Durbin (D-IL)*: Durbin highlighted writers, artists, musicians, and other creators who contribute over a trillion dollars to the U.S. economy. He questioned how creators can compete with AI products that generate content—that oftentimes mimics their work—at “the push of a button.” He reiterated that AI companies are pirating millions of works from so-called shadow libraries, and argued that AI innovation and protection of IP rights are not mutually exclusive.

***Witness Toplines:***

[Maxwell Pritt, Partner, Boies Schiller Flexner LLP](#): Pritt emphasized that AI companies knowingly stole millions of copyrighted publications to gain a competitive advantage in the AI industry. He illustrated that they earn billions of dollars from pirating others' works, without paying a cent to the original authors. Pritt asserted that the decisions to engage in this mass piracy were made by the top executives of these companies, despite employees recognizing that their behavior was illegal and unethical. He stated that there is no carveout in the *Copyright Act of 1976* for AI companies to conduct mass piracy of copyrighted works.

[Michael Smith, Professor of Information Technology and Marketing, Carnegie Mellon University](#): Smith condemned the tech industry's three core arguments for piracy: it is fair use because it will not harm legal sales; it is unlikely to harm creativity; and legislative efforts to restrict piracy will stifle innovation. He argued that digital piracy harms creators by reducing their ability to earn money from their creative efforts, damages society by curtailing economic incentives for investment in creative output, and copyright enforcement has been shown to reverse these harms while still allowing businesses to innovate.

---

<sup>1</sup> The opening remarks by Chair Hawley and Ranking Member Durbin were not available at the time of this memo's distribution.

[Bhamati Viswanathan, Professor of Law, New England Law School](#): Viswanathan stated that courts have not yet reached a consensus on what should be done about the training of AI models on pirated works. However, she asserted that pirate websites are illegal and have been prosecuted by the FBI and the Department of Homeland Security. She emphasized that this theft is not a victimless crime, as basic copyright incentive structures and creators' livelihoods are destroyed. Viswanathan called for greater use of already existing licensing practices to solve this problem, demanding innovation that is fair and sustainable.

[David Baldacci, Bestselling Author](#): Baldacci called creative arts the lifeblood of the U.S. He felt as though "someone had backed up a truck to [his] imagination and stolen everything [he had] ever created." He criticized AI companies for not giving any compensation to the original creators, and warned against the flood of cheap works that will come to market and drive down profits for publishers. Baldacci maintained that books are just as transformative as the AI industry.

[Edward Lee, Professor of Law, Santa Clara University School of Law](#): Lee highlighted decisions made by Judges William Alsup and Vince Chhabria, which held that training AI models with copied works serves a highly transformative purpose, thus qualifying for the protections of the fair use doctrine, outlined in Section 107 of the *Copyright Act of 1976*. He emphasized that fair use is case-specific, and in the cases against Meta and Anthropic, the judges concluded that the plaintiffs—the original creators—did not prove the AI models had produced any outputs infringing on their copyrighted material. He noted President **Donald Trump's** executive order making U.S. development of AI a national priority, and maintained that Congress must observe how courts resolve these issues in the pending copyright lawsuits across the country.

### **Major Takeaways:**

- Hawley and Viswanathan explained that AI companies steal works from illicit online repositories through a process known as torrenting; massive amounts of data are uploaded and downloaded simultaneously, allowing AI companies to access the works while disseminating them to other parties.
  - Hawley asked if the companies ever considered compensating copyright holders. Pritt responded that they did, but resorted to piracy after realizing it would cost them hundreds of millions of dollars to pursue licensing agreements.
  - Hawley displayed documents of multiple text conversations between Meta engineers. They discussed the illegality and unethical nature of their actions, yet decided to continue pirating works to gain a competitive advantage.
  - Durbin emphasized that Meta did not compensate any copyright holders for the use of their work, and asked about the willfulness behind this action. Hawley, Pritt, and Viswanathan argued that the piracy was intentional and agreed upon by top leadership, including Meta's Mark Zuckerberg. Pritt referenced evidence that Meta deliberately went through Amazon Web Services so their actions could not be traced to Meta's servers.
- Durbin criticized Congress's decision to use Section 230 to exempt tech companies from liability, simply because the internet was seen as a transformative industry. He

asserted that Lee's position shifts the burden to the authors of creative work when a corporation steals their product and asserts fair use protections. He stated that the "thievery" of creative work provides economic benefit to those creating AI models.

- Hawley raised a similar challenge to Lee's arguments, noting that the mass theft of works benefits U.S. corporations, rather than U.S. citizens. Following Lee's support of David Sacks' beliefs, Hawley remarked that the U.S. should not let an "unelected AI czar" decide what the rights of U.S. citizens are. He declared that protecting the common rights of citizens is what makes the U.S. a nation.
- Sen. **Peter Welch** (D-VT) highlighted his Transparency and Responsibility for Artificial Intelligence Networks (TRAIN) Act ([S. 5379](#)), which would protect creators' copyrighted works if they are used to train AI models. He held that AI companies need content and do not care how they source it, despite creators' right to have their copyrighted material respected. Welch called for reforms to protect artists, stating that they contribute to the development of meaningful career paths and the United States' culture.