Commentary

Title:

"Open AI's Legal Battles and the Penguin Effect"

Author(s):

James Boudreau

March 2024

Artificial intelligence is all over the news these days, with OpenAI, the company behind ChatGPT, at the forefront. Just two months after its public launch in late 2022, ChatGPT became the fastest-growing consumer app in history, with over 100 million active users.[1] To adapt a movie quote from a bygone era, however, with great popularity often comes great legal responsibility. Over the past year, OpenAI and its parent company, Microsoft, have been hit with several lawsuits related to the use of personal[2] and copyrighted[3] data from the internet to train their popular chatbot, including one from *The New York Times*.[4] Whatever the final outcomes of those specific cases, the legal questions they bring to the country's (and the world's) attention matter not just for OpenAI itself but also for the future of the artificial intelligence industry and perhaps the whole economy.

Technological advances are one of the most important factors for sustained increases in economic growth,[5] but the trick with new technologies is that they often require new rules. New technologies can infringe on the rights of other parties in ways that are not anticipated until those innovations actually exist, which means one "penguin" must jump into the water first. Once that happens, new rules can be constructed around the experience of the first mover, but someone has to make the initial plunge for new industry guidelines to be established. The question is, who wants to take the risk? The courts may want to take the incentives facing first movers like OpenAI—and future firms in similar situations—into consideration when rendering their judgments.

The Penguin Effect

In economic theory, the term "penguin effect" was first coined by Farrell and Saloner (1986) in the context of technology adoption with competing standards.[6] Think of a firm choosing between Windows or Mac OS for its employees or households deciding between Betamax or VHS (more on that specific example below). To fully understand which option is better requires some experience with one or both options, but these products also feature what economists call *network effects*, meaning one option is more valuable when a greater number of other users choose it. As a result, potential users have an incentive to delay their choice until someone else goes first; that way they'll have a better idea of which option will catch on. "Penguins who must enter the water to find food often delay doing so because they fear the presence of predators. Each would prefer some other penguin to test the waters first."[7]

In addition to causing delays in adoption and widespread use, the penguin effect can also lead to the adoption of a *less risky* standard rather than the superior standard, as early users face more uncertainty. Subsequent users might then engage in herd behavior as the initial choice

---

[1] https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/
[2] https://www.cnn.com/2023/06/28/tech/openai-chatgpt-microsoft-data-sued/index.html
[3] https://www.cnn.com/2023/07/10/tech/sarah-silverman-openai-meta-lawsuit
[4] https://www.npr.org/2023/12/27/1221821750/new-york-times-sues-chatgpt-openai-microsoft-for-copyright-infringement
[5] https://www.nobelprize.org/prizes/economic-sciences/1987/press-release/
[6] Farrel and Saloner (1986), "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation," *The American Economic Review*.
[7] The quote is directly from Farrel and Saloner (1986, p. 943), but it's worth noting that it's unclear whether penguins really behave this way. A quick internet search shows that other fields in academia—like crowd physiology—also use the term as a metaphor, albeit with alternative meanings in some cases, but as a zoological phenomenon the penguin effect may be a myth.

becomes more common, and a suboptimal standard can end up dominating the market.[8] Again, the examples of Windows vs. Mac OS and Betamax vs. VHS are relevant comparisons as cases of new technology adoption by end users. But similar effects may exist for the *producers* of new technology as well, particularly when multiple firms compete to develop a technology that may infringe on existing rights.

The threat of costly legal battles provides an incentive for firms to delay entry into markets where such issues are present, as waiting can give them the opportunity to observe the first mover. Once any firm enters, its experience provides a valuable public good for its competitors in the form of realized experience with the relevant legal guidelines, but (at least) one firm must be confident enough to overcome the absence of that information in the first place. In addition to potentially delaying new technology, then, the expected legal environment could influence which firm enters first and in turn which standard is ultimately adopted.

The Role of Courts

When it comes to new technologies that involve potential copyright infringements, as does the case of OpenAI and ChatGPT, a model by Shahshahani (2018) provides additional guidance for the role of courts.[9] The key insight of the model is that a court's decision on any one case is not necessarily the end of the story. Rather, when a new technology requires some reconsideration of existing rules, the court's decision is merely the first stage of a sequential game. The court decides whether or not some infringement has occurred based on existing laws and precedents, but after that step, interest groups representing both sides have the opportunity to lobby Congress, who then ultimately has the ability to legislate new rules into existence. The court should, therefore, take those later stages of the game into account when making their decisions in the first stage.

In the final stage of the game, we assume legislators will choose new policy settings based on the lobbying efforts of interest groups in the second stage. This may (perhaps heroically) be interpreted as Congress balancing the public's interests as proxied by lobbying activity, but what really matters is that both the interests in favor of the new technology and those feeling their rights infringed upon have the opportunity to spend as much as they believe their cause is worth. If both sides are unconstrained in terms of spending, the resulting policy legislation should reflect both the value of allowing new technology to function and any costs it imposes on existing rights holders. What courts need to consider is thus whether their decisions have the potential to restrict either party's efforts in the game's second stage.

If the court's decision has no bearing on the ability of either party to lobby, in an important sense whichever party they initially side with doesn't matter. As long as both sides are able to convey the costs and benefits of the new technology in the lobbying stage, Congress can choose the optimal policy mix; whether the technology is initially determined to be in violation of the current rules is irrelevant. This is very similar to another result from economic theory known as the Coase theorem. Named after Nobel laureate Ronald Coase, the idea is that two parties can reach an efficient agreement over the use of a resource regardless of who initially has the property

---

[8] Choi (1997), "Herd Behavior, the "Penguin Effect," and the Suppression of Informational Diffusion: An Analysis of Informational Externalities and Payoff Interdependency," *The RAND Journal of Economics*.

[9] Shahshahani (2018), "The Role of Courts in Technology Policy," *The Journal of Law & Economics*.

rights for it to begin with, so long as there are no costs to bargaining.[10] In this case, the court's decision determines the initial state of rights, and the lobbying process that informs legislators is analogous to the bargaining process that ultimately determines how resources are used. As long as the court's decision does not interfere with that process, it doesn't matter.

Sometimes, however, a court's decision will impact the future capabilities of interested parties, specifically when a ruling against a firm implementing new technology has negative financial consequences. Any penalties suffered as a result of the initial court decision could reduce or even eliminate a firm's opportunity to fight for its interests in the second stage of the game, thereby influencing the timing of entry decisions and the eventual policy mix implemented by Congress. Shahshahani (2018), therefore, concludes that courts should lean toward favoring whichever party is more likely to be constrained in terms of future lobbying efforts in the event of a negative decision, which is usually the fledgling firm bringing new tech to market. That way the lobbying stage can proceed unimpeded, and the resulting policy legislation is more likely to reflect the value of the new technology in question. In fact, in the press release announcing Coase's Nobel prize in 1991, the committee explicitly recognizes this corollary to the Coase theorem in more general terms:

> *"Coase found that courts probably try to distribute the rights among the parties so as to realize the solution which would have been the outcome of an agreement, if such an agreement had been possible. The underlying idea is that this is a natural and rational way for a court to reason if it is more intent on setting a precedent to generate expedient incentives for the future than solving a particular dispute."*[11]

Previous Cases of Disruption

Though artificial intelligence brings its own unique challenges in terms of potential rights infringements (in addition to other potential concerns like the threat of apocalypse), it is certainly not the first time technology has disrupted the status quo, leading to legal challenges and eventually new policy. The following three examples all hinged primarily on copyright infringements, but each is instructive in its own way in terms of the court's decision and the eventual outcomes for the industries involved.

Copiers. They may not seem like fancy technology anymore, but before Xerox machines existed (roughly before 1959), if anyone wanted their own copy of written material, they had few (if any) options other than to buy the whole thing—even if they were only interested in one section of it. With the wonders of photocopying, however, anyone could go to a library and make their own copies of whatever they were interested in for a small fraction of the cost. Publishers weren't happy about that, of course, and it turns out academics printing copies of journal articles were what really set them off. In 1973, a publisher by the name of Williams & Wilkins Co. sued the National

---

[10] The article often credited as originating the Coase theorem is Coase (1960), "The Problem of Social Cost," *Journal of Law & Economics*. For a more specific example involving radio stations bargaining over who has control over contested frequencies, see Coase (1959), "The Federal Communications Commission," *The Journal of Law & Economics.*

[11] https://www.nobelprize.org/prizes/economic-sciences/1991/press-release/

Institutes of Health and the National Library of Medicine, claiming that the unauthorized photocopying of medical journal articles violated their copyright. The case eventually reached the Supreme Court as *Williams & Wilkins Co. v. United States*[12] (1975), where a divided court ended up ruling 4-4 (Justice Harry Blackmun recused himself), affirming the lower court's ruling that the photocopying involved constituted "fair use" of copyrighted materials.[13]

It was close, but the court allowed the new technology's use to continue, and in doing so it also set a major precedent for the "fair use" doctrine, which allows the limited use of copyrighted works without permission. Though the term "fair use" had been used prior to *Williams & Wilkins Co.*, it and several similar decisions in cases also related to copiers—along with related lobbying efforts—led Congress to formalize the term in the Copyright Act of 1976.[14] That piece of legislation, which is still the basis of copyright law today, paved the way for other new technologies to make other new uses of previously copyrighted materials. But not always quietly.

VCRs. As copiers did before them, videotape recorders (or VTRs, as they are referred to in court transcripts from that era) allowed individuals to make their own copies of copyrighted materials. In place of publishers, this time it was movie studios that were unhappy about the copying, and in what is now often referred to as "The Betamax Case,"[15] Universal City Studios and Walt Disney Productions sued Sony Corp., maker of the Betamax VTR, for copyright infringement. Again, the court ultimately ruled in favor of the new technology. In *Sony Corp. of America v. Universal City Studios, Inc.*, the court ruled that the use of VTRs to "time shift"—meaning viewers could record programs to watch at a later time and/or date—was legitimate, and thus "fair," use.

Somewhat ironically, but fitting with the Coasean model, the studios themselves may have benefitted as much as (if not more than) viewers from the resulting legal landscape, as revenues from the prerecorded video market would soon far exceed those of theatrical releases.[16] In fact, evidence has also shown that journal publishers were not ultimately harmed due to copiers, either, as they were able to price discriminate by charging higher fees for institutional memberships and increasing the number of different journals they offered.[17] In both examples, the relevant judicial decisions served to begin the process of forming new guidelines, but the immediate consequences were not the end. Both producers and consumers ultimately benefitted under the new rules, irrespective of the initial rights environment. (It is unclear whether legal struggles ended up being a significant factor in the demise of Betamax).

File sharing. The internet and digitization meant change was inevitable for a number of industries, including motion pictures and various forms of print media, but it is the music industry that remains most (in)famous for its struggles against file sharing. The industry's battle against the peer-to-peer filesharing platform known as Napster marked a cataclysmic shift in how people

---

[12] *Williams & Wilkins Co. v. United States*, 420 U.S. 376 (1975).

[13] For details, see Shaw (1972), "Williams & Wilkins Co. v. US," *American Libraries*.

[14] https://www.copyright.gov/title17/

[15] Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417 (1984).

[16] Liebowitz (2004), "The Elusive Symbiosis: The Impact of Radio on the Record Industry," *Review of Economic Research on Copyright Issues*.

[17] Liebowitz (1985), "Copying and Indirect Appropriability: Photocopying of Journals," *Journal of Political Economy*.

would consume digital media for the foreseeable future. While copiers and VCRs allowed consumers to enjoy facsimiles of copyrighted materials in different ways than they had before (reading outside the library, watching at later times), filesharing went further, allowing seemingly unlimited, free access to copyrighted materials.

As the most popular filesharing platform, Napster faced several copyright infringement lawsuits, perhaps most famously one from an aged version of Metallica.[18] The death knell was a suit representing 18 different record companies from the Recording Industry Association of America as Napster was officially found liable for copyright infringement.[19] This would also, eventually, mean the death of most other peer-to-peer file sharing services.

The court's treatment of Napster remains somewhat contentious, as it could have ruled more in line with the *Sony Corp.* case and the fair use doctrine. The ruling also runs counter to the advice from economic theory that the court should side with the more budget-constrained party, as Napster clearly was. In line with the previously discussed theories, however, once the rules were spelled out, the firms waiting (e.g. Spotify, which was already in its early stages) were able to take advantage of the new landscape. Napster was the penguin that jumped first and paid the price, but it seems the interests of consumers and producers were represented in the later stages of the game. Listeners now have cheaper, more easily available music, and empirical studies have (so far) found no evidence of a decline in the *quality* of music since 1999 (Waldfogel, 2012; 2017).[20] In fact, though recording labels remain below their mid-to-late 90's heights in terms of revenue, there have been dramatic increases in the numbers of new musical works, movies, books, and television shows created since the advent of digitization (Waldfogel, 2017).

What Now?

While the allegations against OpenAI are unique in their own ways, particularly those that involve the use of personal data, the rapid proliferation of "large language models" like the one used for ChatGPT suggests they're here to stay. Future lawmakers will have to clarify a variety of issues related to how those programs use data from the internet, and possibly even how the programs themselves are used in practice. Firms like Microsoft, Google, Meta, and others have deep pockets for the lobbying stage, but other politically powerful groups such as labor unions (e.g., SAG-AFTRA) may emerge as competing interests. The good news is that eventually, new guidelines will emerge, and hopefully, as in the past, they will be based on the economic interests involved.

Whether OpenAI itself ends up a casualty for jumping in first remains to be seen. Will it end up like Napster, devoured by legal sharks? Will ChatGPT end up the Betamax of artificial intelligence? Whichever way the courts lean, it's important to recognize that their decisions will inform future creators of new technology as to whether it's better to ask forgiveness or permission.

---

[18] *Metallica, et al. v. Napster, Inc.* (2000).

[19] *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004 (9th Cir. 2001).

[20] Waldfogel (2012), "Copyright Protection, Technological Change, and the Quality of New Products: Evidence from Recorded Music since Napster, " *The Journal of Law & Economics*.
Waldfogel (2017), "How Digitization Has Created a Golden Age of Music, Movies, Books, and Television," *The Journal of Economic Perspectives*.