LEGAL UPDATE

January 25, 2024

"ChatGPT, the New York Times, and the KFTC: A Legal Revolution"?

The legal war over AI has begun. At the end of December 2023, the New York Times ("NYT") filed a lawsuit claiming that OpenAI "steals" NYT reporting to train its generative AI, ChatGPT. Without sources for the latest news, ChatGPT's capabilities would be severely limited. Almost the same day, the Korean Association of Newspapers ("KAN") filed a similar complaint about Naver's hangul-based generative AI, Hyperclova X. These and related cases will define the legal landscape for content creators and generative AI. It is unclear what the outcome will be.

New York Times v. Open AI/Microsoft

The New York Times has accused OpenAI and Microsoft of copyright infringement related to the training and use of ChatGPT.[1] ChatGPT is a "Large Language Model generative AI," ("LLM AI") that can understand and respond to questions from the user.

To create an LLM AI, a developer must copy large amounts of text to "train" the LLM AI. This text can come from newspapers, books, blogs, and even internet cafes.

ChatGPT can respond to questions about the news in part because it was trained with NYT articles. As a result, users can learn the news by asking ChatGPT, without paying for a subscription to the NYT. The NYT claims that constitutes copyright infringement of its original reporting.

Korean Association of Newspapers' KFTC opinion

On essentially the same day, KAN filed an opinion with the Korean Fair Trade Commission ("KFTC") raising similar concerns about Naver.

ChatGPT can understand and respond to user questions in Korean. However, Naver's hangul based LLM AI, "Hyperclova X," was trained with 6,500 times more hangul material than ChatGPT. Therefore, just as Naver surpassed Google, Yahoo, and Bing in Korean search engines, Hyperclova X may have a significant impact on the Korean AI market.

Copyright infringement: "transformative use" v. "market substitution"

Under Korean and United States copyright law, people are allowed to use copyrighted material for "fair use" purposes, without paying a royalty.[2]

"Fair use" includes "transformative uses" of copyrighted material that "add something new, with a further purpose or different character, altering the [original] with new expression, meaning or message."[3] On the other hand, uses that create a "market substitute" for the original work usually constitute copyright infringement.[4]

---

1   Michael M. Grynbaum and Ryan Mac, *The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work*, The New York Times (Jan. 19, 2024), https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

2   *See,* U.S. 1976 Copyright Act § 107; Korea Copyright Act Art. 35-5; KAN opinion (citing US cases as persuasive authority).

3   Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 579, 114 S. Ct. 1164, 1171, 127 L. Ed. 2d 500 (1994).

4   Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 528, 555, 143 S. Ct. 1258, 1290, 215 L. Ed. 2d 473 (2023).

"Transformative use": Google's scanning of millions of books

For example, in 2004, Google began scanning the entire text of millions of copyrighted books that previously existed only in hard copy.[5] Google used these scans to create an online database that allows users to do full-text searches for words and phrases, without reading the entire book. Google did not pay a licensing fee to the copyright owners.

In a case cited in the KAN's KFTC opinion, the US court held:

> "that the creation of a full-text searchable database is a quintessentially transformative use. . . [T]he result of a word search is different in purpose, character, expression, meaning, and message from the page (and the book) from which it is drawn."[6]

Copying an entire book is normally copyright infringement, because reading the copy "substitutes" for a potential sale of the book. In the Google case, however, "the full-text search function does <u>not</u> serve as a substitute for the books [themselves.]"[7] Because Google's massive copying was a "transformative use," it was "fair use," and Google was not required to pay any licensing fees.

"Market substitution": Andy Warhol's artistic copying of a photograph

On the other hand, the United States Supreme Court held last year that art that <u>appears</u> to be transformative can infringe on copyrights if it functions as a "market substitute" for the original.[8]

Andy Warhol has been called "the very embodiment of transformative copying," in both the artistic and legal senses.[9] Andy Warhol created "Orange Prince" by making a silkscreen copy of a photograph of the musician Prince that appeared in a magazine. Warhol licensed "Orange Prince" to appear in another magazine. The original photographer sued.

[10]

In a case cited by the KAN, the Court explained:

> "the use of an original work to achieve a purpose that is the same as, or highly similar to, that of the original work is more likely to substitute for, or supplant the work [on the market.]"[11]

---

5 Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 96 (2d Cir. 2014).

6 Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 97 (2d Cir. 2014).

7 *Id*. at 99-100 (emphasis added)

8 *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 561, 143 S. Ct. 1258, 1293, 215 L. Ed. 2d 473 (2023); *see* KAN opinion at III.2 (citing this case).

9 *Id*. at 592 (Kagan, J. dissenting) (citing Supreme Court opinions using Warhol's work as examples of "transformative use.").

10 *Id*. at fig. 6.

11 *Id*. at 528 (quotation omitted).

The original photograph was included in a magazine story about Prince. A different magazine paid Warhol to use "Orange Prince" for another Prince story. Therefore, the Court held that "even if the two [images] were not perfect substitutes," they "shared the same objectives," and Warhol's estate was liable for copyright infringement.[12]

LLM AI's use of news: "market substitution" or "transformative use"?

It is unclear how US courts and the KFTC will rule on the newspapers' claims.

On the one hand, the NYT and KAN argue that ChatGPT and Hyperclova X have created a "market substitute" for their original reporting.

•   LLM AIs answer questions about the news by analyzing copyrighted articles written by human reporters.

•   Because users can learn the same news through LLM AIs, they arguably "serve the same informative purposes as the original" news reporting.[13]

It is true that LLM AI is not a "perfect substitute" for the original articles. However, the newspapers can reasonably argue that LLM AIs "steal" monetizable traffic from their websites and violate their copyrights.[14]

On the other hand, the LLM AIs may "transform" how users *interact* with the news. In the Google case, readers could always find particular text by carefully reading the entire book themselves. However, Google's unlicensed scanning allowed readers to find the text in seconds, with a full-text search. Similarly, LLM AIs arguably provide a different way to access the information in copyrighted articles. For example, people who want to learn about the NYT v. OpenAI case can browse the NYT website or use traditional search engines to find and read articles on the subject. ChatGPT, however, allows users to access the answers more directly, by simply asking a question.[15]
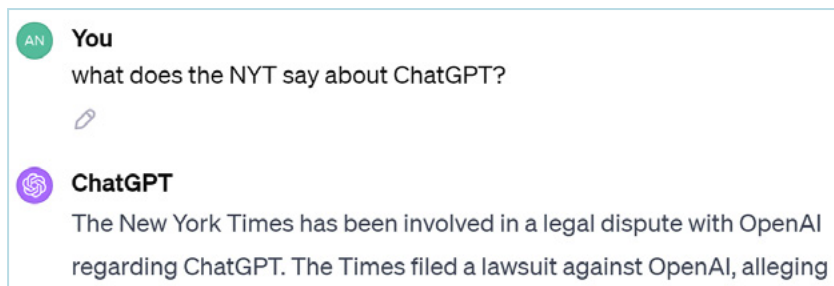
Therefore, OpenAI and Naver can argue that LLM AIs "transform" how users interact with the knowledge conveyed in news articles, and so constitute "fair use."[16]

Other AI lawsuits

Generative AI is already disrupting our legal understanding of rights beyond "hot news."[17]

---

12 *Id*. at 536.

13 Complaint at ¶108.

14 *Warhol* at 536.

15 ChatGPT and Hyperclova X both make mistakes in their responses. Worse, LLM AIs are known to "hallucinate," by referencing facts and cases that simply do not exist. Nonetheless, both products continue to increase their reliability.

16 *Cf*. Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 99 (2d Cir. 2014) (explaining that book reviews, which often quote and summarize the original, generally do not constitute copyright infringement).

17 International News Serv. v. Associated Press, 248 U.S. 215 (1918).

- Biometric data: An AI company that creates avatars based on users' actual faces was sued for allegedly violating privacy laws on biometric data.[18]

- Likeness and publicity: A "deep fake" AI creator was sued for unlicensed use of celebrities' likenesses to create fake images that are often embarrassing.[19]

- Photographs and images: Getty Images sued an AI image generation company for using its photographs without permission.[20]

- Books: Authors like George R.R. Martin ("왕좌의 게임") have sued OpenAI for using their books to train ChatGPT.[21]

What should Korean companies do?

An International Monetary Fund report released in January 2024 shows that AI will soon disrupt 40% of jobs worldwide.[22] This number rises to 60% of jobs for advanced economies like South Korea. AI will affect all industries, including the practice of law.

The law of AI is changing every day. For AI developers, it is not sufficient to wait and respond to lawsuits after they are filed. AI developers should seek legal advice to help them anticipate legal changes and plan their businesses accordingly.

For content creators, AI will change not only the legal definitions of their rights, but also the nature of the market as a whole. In the short-term, Korean content creators should consider what rights they may have related to generative AI, and how to protect those rights. Options could include filing actions in US, Korean, or European courts, or joining class action lawsuits in the US.

---

18 Flora v. Prisma Labs, Inc., 2023 WL 5061955 (N.D.Cal., 2023).

19 Young v. NeoCortext, Inc., 2023 WL 6166975 (C.D.Cal., 2023).

20 Getty Images (US), Inc. v. Stability AI, Inc., U.S. District Court for the District of Delaware, No. 1:23-cv-00135. Date Filed: Feb. 3, 2023.

21 Authors Guild v. OpenAI Inc., 1:23-cv-08292, (S.D.N.Y.). Date Filed: Sept. 19, 2023.

22 Mauro Cazzaniga et al., "Gen-AI: Artificial Intelligence and the Future of Work", INTERNATIONAL MONETARY FUND Staff Discussion Notes No. 2024/001, https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379.

Related Professionals

Andrew Chongseh Kim
Senior Foreign Attorney

T 82.2.3404.0473
E andrew.kim@bkl.co.kr

Tony DongWook Kang
Partner

T 82.2.3404.0158
E tony.dw.kang@bkl.co.kr