

Generative AI and SDOs

Michael Lindsay
Dorsey & Whitney LLP



WORLD STANDARDS WEEK, 2023

Objectives of SDOs

- **High-quality standards**
 - An SDO seeks to develop and publish high-quality standards that, if broadly adopted, will achieve their purpose
 - The desired degree of quality may vary with purpose for standard and may be satisfied by more than one potential approach
- **Broad adoption of standards**
 - An SDO seeks broad adoption of the standards that it develops
 - Broad adoption maximizes the social value of the standard

Generative AI and SDO Objectives: Quality

- Is generative AI currently capable of producing, consistently and reliably, the kind of quality that an SDO requires?
- Problem of “hallucination”
- Problem of “opinion”

Texas judge bans filings solely created by AI after ChatGPT made up cases

AI Hallucinations in the Courtroom: A Wake-Up Call for the Legal Profession

A brief overview of how generative AI works, the issues in professional responsibility that are raised by the use of the generative AI, and some considerations for navigating a dynamic technological landscape.

June 7, 2023, 11:00 AM CDT

OpenAI Hit With First Defamation Suit Over ChatGPT Hallucination



Isaiah Poritz
Legal Reporter



Generative AI and “Hallucination”

- Lawyers’ brief cited at least six cases, in proper citation format, *that did not exist at all*
- “In researching and drafting court submissions, good lawyers appropriately obtain assistance from junior lawyers, law students, contract lawyers, legal encyclopedias and databases such as Westlaw and LexisNexis.”
- “Technological advances are commonplace and there is nothing inherently improper about using a reliable artificial intelligence tool for assistance. But existing rules impose a gatekeeping role on attorneys to ensure the accuracy of their filings.”
- The attorneys at issue “abandoned their responsibilities when they submitted non-existent judicial opinions with fake quotes and citations created by the artificial intelligence tool ChatGPT, then continued to stand by the fake opinions after judicial orders called their existence into question.”

Mata v. Avianca, Inc., Opinion & Order on Sanctions, Case 1:22-cv-01461, Dkt 54 (S.D.N.Y. Jun. 22, 2023)

Generative AI and Opinion

The Washington Post
Democracy Dies in Darkness

**Amazon's Alexa has been claiming the
2020 election was stolen**

Generative AI and SDO Objectives: Broad Adoptability

- Broad adoptability is advanced if SDO can obtain copyright on its standards
 - Copyright law protects “original works of authorship”
 - Includes compilations and derivative works
 - You should have attended the IBR panel!
- Broad adoptability requires that SDO be able to publish free of copyright infringement claims
 - SDOs adopt copyright policy to effectuate this objective
- Is a Generative AI compilation copyrightable?
- Do Generative AI-assisted contributions infringe copyrights?

Copyright Law

- “Works are original when they are independently created by a human author and have a minimal degree of creativity. Independent creation simply means that you create it yourself, without copying.”

[U.S. Copyright Office, What is Copyright?](#)

- “To qualify as a work of ‘authorship’ a work must be created by a human being.”
- Copyright Office “will not register works produced by a machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author.”
 - Is the work “basically one of human authorship, with the computer [or other device] merely being an assisting instrument,” or are “the traditional elements of authorship in the work (literary, artistic, or musical expression or elements of selection, arrangement, etc.) ... actually conceived and executed not by man but by a machine.”

Compendium of U.S. Copyright Office Practices § 313.2 (Jan. 28, 2021),
<https://www.copyright.gov/comp3/chap300/ch300-copyrightable-authorship.pdf>

Generative AI and Copyright



A Recent Entrance to Paradise

- “Plaintiff develops and owns computer programs he describes as having artificial intelligence capable of generating original pieces of visual art, akin to the output of a human artist.”
- “Copyright has never stretched so far ... as to protect works generated by new forms of technology operating absent any guiding human hand”

Thaler v. Perlmutter, Mem. Op., No. 22-1564 (D.D.C. Aug. 18, 2023)



[Animal-made art - Wikipedia](#)

Generative AI as Copyright Infringement

- In *Thompson Reuters v. Ross*, district court recently held that generative AI did copying full text of Westlaw headnotes
- A jury will determine
 - Whether the copied material was original authorship (vs. rehash of uncopyrightable judicial opinions)
 - Whether the product resulting from the copying was substantially similar to the copied work
 - Whether the copying constituted “fair use”

Thomson Reuters Enterprise Centre GmbH v. Ross Intelligence Inc., No. 1:20-cv-613 (D. Del. Sep. 25, 2023)

Thompson Reuters v. Ross

- “Deciding whether the public’s interest is better served by protecting a creator or a copier is perilous, and an uncomfortable position for a court. Copyright tries to encourage creative expression by protecting both.”
- “Here, we run into a hotly debated question: Is it in the public benefit to allow AI to be trained with copyrighted material? The value of any given AI is likely to be reflected in the traditional factors: How transformative is it? Can the public use it for free? Does it discourage other creators by swallowing up their markets?”
- “So an independent evaluation of the benefits of AI is unlikely to be useful yet, even though both the potential benefits and risks are huge. Suffice it to say, each side presents a plausible and powerful account of the public benefit that would result from ruling for it. So a jury must decide the fourth factor—and the ultimate conclusion on fair use.”

*Thomson Reuters Enterprise Centre GmbH v. Ross
Intelligence Inc.*, No. 1:20-cv-613 (D. Del. Sep. 25, 2023)

Generative AI as Copyright Infringement

- Authors lawsuit against OpenAI

- “OpenAI has admitted that, of all sources and content types that can be used to train the GPT models, written works, plays and articles are valuable training material because they offer the best examples of high-quality, long form writing and contain long stretches of contiguous text, which allows the generative model to learn to condition on long-range information.... Among the content OpenAI has scraped from the internet to construct its training datasets are Plaintiffs’ copyrighted works.”

Chabon v. OpenAI Inc., 1st Am. Complaint ¶¶ 37, 40, Dkt 11, No. 23-cv-04625 (N.D. Cal. Oct. 5, 2023) (cleaned up)

- Authors lawsuit against Meta

- “A large language model is ‘trained’ by copying massive amounts of text from various sources and feeding these copies into the model. This corpus of input material is called the training dataset. During training, the large language model copies each piece of text in the training dataset and extracts expressive information from it. ... Much of the material in Meta’s training dataset, however, comes from copyrighted works—including works written by Plaintiffs—that were copied by Meta without consent, without credit, and without compensation.”

Chabon v. Meta Platforms Inc., 1st Am. Complaint ¶¶ 26-27, Dkt 16, No. 23-cv-04663 (N.D. Cal. Oct. 5, 2023) (cleaned up)