EVENT TITLE: Generative Artificial Intelligence, Copyright, and Authorship

EVENT DESCRIPTION: What are the implications of the meteoric rise of generative AI technologies on authors' writings, the landscape of publishing, and copyright law? In this presentation, Authors Alliance staff present a high-level overview of the ways in which generative AI programs affect working authors, serving both as powerful creative tools and potential disruptions. Drawing on our expertise in the field of copyright and AI, we explain how laws affecting authors accommodate generative AI technologies and answer audience questions about how authors can harness their potential.

EVENT CATEGORY: Publishing, Editing, and Technology

Event Organizer and Moderator
Rachel Brooke, Authors Alliance: Rachel is a senior staff attorney at Authors Alliance, and is based in the San Francisco Bay Area. She has spent several years in a variety of different roles at Authors Alliance, and currently focuses on policy work, including on some of the difficult questions around copyright and generative artificial intelligence. Before becoming an attorney, Rachel spent several years working at a small literary agency in New York.

Event Participant
Dave Hansen, Authors Alliance: Dave is the executive director of Authors Alliance, based in Chapel Hill, NC. He is trained as a lawyer and librarian, and has extensive experience helping authors navigate copyright and information law issues. Prior to Authors Alliance he was responsible for leading publishing and copyright support for academic authors at Duke University.

Moderator Opening Remarks and Housekeeping Announcements:
Welcome to Authors Alliance’s presentation on Generative Artificial Intelligence, copyright, and authorship. We’re really excited to be here with you all today to talk about this topic.

Authors Alliance is a nonprofit organization that exists to advance the interest of authors who want to serve the public good by sharing their creations broadly. We have over 2,500 members—many of our members are academic nonfiction authors, but we welcome authors of all stripes who share our values, and we count journalists, poets, narrative nonfiction authors, fanfiction authors, and academic researchers among our members. Founded in 2014 in Berkeley, California, we advocate for policies that keep knowledge and culture accessible and discoverable as well as creating educational resources for authors to help guide them through legal issues related to authorship. On the advocacy side, we write amicus briefs, make policy submissions to the Copyright Office and Congress, and otherwise speak out on behalf of policies that advance our members’ interests. On the education side, we have a series of book length guides for authors on legal issues related to writing, on topics like fair use, negotiating publication contracts, permissions clearance, open access publishing, and more.
While Authors Alliance has expressed views on many of the issues we'll speak about today, we're going to try to keep things relatively neutral and explain how copyright and generative AI interact in today's policy debates, the ongoing open questions and discussions, in addition to how we are thinking about these questions. We want to focus on practical takeaways for authors who are interested, curious, or apprehensive about how the rise of generative AI is affecting, or will affect, authors.

We'll structure the presentation as follows: Dave and I will give a brief overview of generative AI, how U.S. copyright law has developed (or not developed) to address it. Then, we'll drill down on how all of this affects authors, focusing on practical takeaways. I'll pose a discussion question, and Dave and I will talk through the answer. Then we'll open it up for questions from all of you.

Introduction

- Generative AI 10,000 foot overview (Dave)
  - How models/systems work
  - Timeline (e.g., from Google Books (?) through popularization of ChatGPT)
  - Vocabulary (outputs, inputs, model vs. system)
  - Text Data Mining and its relationship with AI LLMs
  - Applications for authors/scholars/researchers

- Generative AI and copyright (Rachel)
  - Copyright protects original creative expression from exploitation by others, and exists to incentivize authors and other creators to create new works
  - Zarya of the Dawn copyright registration letter that ignited current debates
  - Copyright Office listening sessions
  - Ongoing class action lawsuits

Moderator Questions:

Three main legal questions about copyright and generative AI, and how we think about the answers:

Are “outputs” of generative AI systems protected by copyright?

- In general, no. The Copyright Office has taken this position.
- Copyright’s “human authorship requirement” and its background
- The difficulty here is for creators who want to use generative AI tools as part of their authorship but also want to rely on the protection

Is the use of copyrighted material to train generative AI models fair use?

- We think so, but it’s important to acknowledge that this is one of the more controversial parts of conversations around generative AI and copyright. We’ll talk through Authors Alliance’s view, but want to acknowledge that the issue isn’t settled
• Fair use case law, and similarities and differences between prior cases (e.g., Google Books, HathiTrust, iParadigms) and current LLM cases. Licensing problems (cost, practicality, bias…)

What happens when an output looks strikingly similar to an input the model is trained on?

• Outputs can “infringe” copyrights.
• In GenAI, this is often called memorization
• Rare, but happens: the Snoopy problem
• We think copyright already has the tools it needs to deal with this: substantial similarity test for infringement

Practical takeaways for authors
What should authors who want to use generative AI programs as part of their authorship be aware of when it comes to copyright?

• Copyright registration for works that are “AI-assisted,” “AI-generated” can be complicated: check Copyright Office guidance, webinars
• Essentially, it’s important to disclose the use of generative AI in copyright registration applications
• Also be aware that the output itself is not protected by copyright

What practical advice do you have for authors who are interested in using generative AI tools as part of their creative process?

• Think about what you might need help with: we have heard lots of examples of authors who have used generative AI in all sorts of ways as part of writing and publishing, to save time so they can work their day jobs and still find time to write
• Practices and policies are still evolving when it comes to publishing works that were created using generative AI in some way—both in self-publishing world (e.g. Amazon rules) and conventional publishing

What about advice for authors who don’t want their work to be used as part of AI training?
• Currently, many of the largests AI orgs are building or have established opt-out mechanisms that you can use.
• US law is still being worked out. In the EU, the regulatory framework is a little more clear. The new AI Act treats AI work for commercial purposes as covered by its existing directive on TDM, which provides for an opt-out for commercial applications (but no opt out for nonprofit research uses).
• Important to think about how opt-out affects your work and its relevance. Do you want people who are increasingly using AI as a quasi-search function to discover and learn about your work?