



Technology Policy Institute

AI & Copyright Policy

Aspen Forum 2024 Fireside Chat

Panelist:

**Shira Perlmutter**, Register of Copyrights, US Copyright Office

Moderator:

**Joel Waldfogel**, Carlson School of Management, University of Minnesota

Monday, August 19, 2024

10:10AM-10:50AM

Event Page: <https://www.youtube.com/watch?v=j3LpzXmT08k>

**Scott Wallsten**

This next session is going to be, I think, really interesting, the Fireside Chat. I will introduce the moderator, and then he will introduce our guest. So Joel Walfogel, who is a professor at the Carlson School of Management at the University of Minnesota, will be the moderator. You may have seen him here in the past and read lots of his work on copyright and so on, and this should be a really interesting discussion. So, Joel, go ahead and introduce Shira.

**Joel Walfogel**

All right. No, thank you very much. Thank you very much for involving me. So let me begin by introducing Shira. So Shira is the Register of Copyright and the Director of the US Copyright Office. She was appointed in October 2020. She advises Congress and the executive branch agencies on copyright policy and directs the administration of important provisions of the US Copyright Act, leading a workforce of nearly 500 employees. Prior to appointment, she had served since 2012 as Chief Policy Officer and Director for International Affairs at the US Patent and Trademark Office. She previously held senior positions at the International Federation of the Phonographic Industry in London and at Time Warner in New York. She also worked at WIPO World Intellectual Property Organization in Geneva as a consultant on copyright and electronic commerce. While on the faculty at the Columbus School of Law at the Catholic University of America, from 1990 to 1995, she was a Copyright Consultant to the Clinton administration's Advisory Council on the National Information Infrastructure.

Earlier in her career, she practiced law in New York City, specializing in copyright and trademark counseling and litigation. She's a co-author of a leading casebook on international intellectual property law and policy, and has published numerous articles on copyright issues. She's also a research fellow at the Oxford Intellectual Property Research Center at Oxford University. She received an AB from Harvard and a JD from the University of Pennsylvania. So just before we get into this, I know she's in a few capacities. I first met her, I guess, at the OECD about a dozen years ago, the first time I met you, and you were working for IFPI. But you seemed very scholarly and very erudite. That was the impression I had. And I had the opportunity to work at the Copyright Office or visit as a year as a Kaminstein Fellow and work with Shira.

So it's great joy to have the opportunity to ask you questions. All right, so AI is on everybody's mind. Is there anyone else for whom the word AI evokes an image of Alan Iverson? Okay, yeah, I'm still working on that. I mean, ML, to me still means maximum likelihood and not machine learning, but I'm working on that, too. In any event, I'm pleased to welcome you. Copyright was once like the sleepy cousin of patent, you know, for people who do scholarship as well as policy. But all of a sudden, copyright is enormously interesting to everyone. And so there's a lot going on at the Copyright Office. Let's kick it off and talk about the report that you just released on artificial intelligence.

**Shira Perlmuter**

Thanks, Joel. And I will say, first of all, it's nice to be here as like an example of a targeted sectoral approach after the last panel. And I also always am a bit amused when I listen to my own bio at the quaint terms that really date me of the national information infrastructure and copyright and e-commerce, you know, all terms that we've abandoned a long time ago. We did just issue a report now about two and a half weeks ago on copyright and AI issues. And I have to say this is just the first part of a much longer, more comprehensive report. The office has been studying the full range of copyright issues that are raised by artificial intelligence, in particular generative AI.

And we realized fairly early on that there was too much to cover in just one release, or it would have been this massive volume that took maybe a couple of years to finish. And meanwhile, a lot is

happening. A lot is happening in the real world, in the boardrooms, in technology development, in Congress and the other agencies and in the courts. So we decided to do it in sections. And the first part, the one that we released two and a half weeks ago, deals with what we are calling digital replicas, and that is imitations of a real individual's likeness, image, voice, in such a way so close to reality that it looks like they are doing or saying something that they are not.

And we started there with that particular set of issues because that's been the topic of a lot of attention, particularly in Congress. There's been now two different bills introduced, other discussion drafts circulated. So there's a lot of activity in this, on that angle, in that space. I will say, the other issue, well, this has been part of what we announced as a major initiative on AI about a year and a few months ago. And what we did, we started out by releasing some guidance on how to register works for copyright protection that incorporate AI generated content. And then we spent a few months really just educating ourselves, what is this new technology? What does it do? How is it being used? What are the concerns that people have? What are the challenges?

And what are the things that are exciting to them about how they can use these tools? And then we built a whole set of questions that we published in a notice of inquiry to the public, raising every issue that had been presented to us during those months of discussion, and issued a notice of inquiry in the end of August last year, and by December, which was the closing date for comments, we had over 10,000 comments. So we've been reviewing and analyzing those comments ever since. And what I find interesting, especially in light of the prior panel about regulation, is that there were only very limited calls for regulation for new legislation, and most of them focused on the digital replicas issue. And I can talk more about that later.

But most of the comments really talked about how existing principles of copyright law should apply to this new technology. And they're very divergent views. And I can summarize them a bit, but that's most of what we heard. So we will see where we go from here. But there is not that much call in the copyright space for changing the actual statute at this point in time. And, you know, just to describe the issues. So one set of issues, the digital replicas issue, second sort of basket of issues that have to do with copyright ability. So when you look at the output of generative AI, to what extent can it be protected by copyright? And then the third set of issues really have to do with liability.

When AI is trained using a data set that contains copyrighted works, and that is what's happening, there are millions of copyrighted works that are being used to train these sophisticated AI tools. To what extent does that qualify as fair use, or to what extent, and in what circumstances does it require consent or compensation, and therefore some sort of licensing mechanism? So those are, I would say, the three baskets of issues. And the third set of issues also raise a lot of other liability questions, like who should be responsible for any active infringement? Should it be the developer of the data set of the tool, the distributor of the tool, or the user of the tool? It raises questions about licensing mechanisms, it raises questions about secondary liabilities. So there's a lot of subsidiary questions, and we're looking at all of these. We expect to have the entire report completed by the end of the year.

### **Joel Walfogel**

End of the year. So Congress makes laws and courts adjudicate decisions, and private plaintiffs, you know, complain and so forth. What is the role of the copyright office in kind of shaping policy right now?

### **Shira Perlmutter**

Yeah, it's interesting, the Office, as you indicated in your introduction of my introduction, we have two major roles, and one is administering the copyright laws, and that includes registering claims to

copyright protection. But the other major role is advising on copyright law and policy, providing really impartial expertise in this area to Congress. We're located in the Library of Congress, so we're a legislative branch agency, Congress is our boss and also providing advice to other executive branch agencies and to the courts. And actually that responsibility is set out explicitly in our authorizing statute. And it's interesting because in a way it's that first role of administering the copyright system that led us to initiate the study, because about five or six years ago, we started to get applications to register copyrighted works produced by artificial intelligence. And so we had to make some determinations and we decided early on that a work that was produced entirely by AI without any human authorship, could not be protected by copyright. The copyright required requires human authorship. And so far that perspective has been upheld by the courts. And it's interesting because on that issue of copyright ability, because we have this role in examining applications for registration, the US Copyright Office has really been out in front. Most other countries don't do this. And so we are making decisions about copyright ability sooner than other countries which have to go through the court system before they get those types of decisions. So that's one area that we are involved. But then, of course, on policy issues we have this advisory role and there we often do policy studies and issue reports, make recommendations for legislation.

And in both roles we do have an influence. Often Congress takes up legislation following our recommendations. The courts often look to our decisions on copyright ability. You know, we can talk about Skidmore versus Chevron deference and what the recent Supreme Court decisions mean, but in general, what the courts say is they follow the Copyright Office analysis because we have a lot of experience looking at copyright ability questions and so therefore our opinions are worthy of some level of deference.

### **Joel Walfogel**

One question, I know some of the folks in the room are interested in sort of scholarship and research. What is the role of research, sort of social science research, in answering questions? What questions need to be answered in order for us to make good copyright policy in this era of AI?

### **Shira Perlmutter**

Yeah, I think a lot of social research questions. I mean, in a way I'd like to flip this around and ask you, Joel, because this is to a large extent your area, and one of the things we've done as part of our initiative on copyright and AI is to pull together a group of eminent economists from government and academia to discuss what are the economic implications. What research needs to be done? And we held an all-day roundtable and are currently putting together a report on the research agenda that came out of that roundtable. And Joel, I have to thank you because Joel was an active participant. So maybe I'll suggest a few things and you might want to add.

But, you know, ultimately the question is, how will the use of AI affect the incentive structure on which our entire copyright system is based? The idea of copyright in the US Constitution is that it is an incentive to create new works that promote the progress of science and the arts. So what are the subsidiary questions? Well, will these new types of AI, generative AI, will they enable more and different types of creativity, perhaps by more people? To what extent and in what circumstances will AI generated content substitute for human generated works? What will the impact of that be on creators livelihoods? And, you know, I was thinking about this with the panel, the economics panel this morning, talking about what kinds of jobs would be displaced.

And it seems to me to some extent, there's a division between physical jobs, like medical jobs in a hospital, and jobs that involve using your mind, which are more likely to be displaced. And there is a lot of concern in the creative community about what the impact will be for artists. We can already see it. I mean, I think there are situations where movies are being made where instead of hiring

someone to draw the backdrop, that can be done using AI, where voice actors can be replaced by AI, interpreters translators, that can be done by AI. And then to me, the really fascinating question is, will this lead to less innovation in the creative context? So I find it hard to imagine that AI, which is essentially a derivative, generative AI, is producing content that is by definition derivative. It's based on having input many things that came before and being able to draw connections and determine what is the most likely next thing that would be connected to whatever you're asking for. So how can that kind of technology possibly produce entirely new genres of art and literature and music, like impressionism or cubism or even rock and roll? I mean, this just seems unlikely. So what will the impact of that be on society and creativity going forward? And then, of course, you know, on the flip side, there's also questions about what will be the impact of different copyright determinations and licensing mechanisms on future IT development, technology development?

### **Joel Walfogel**

So those are terrific questions, which could easily keep a lot of people busy for a generation. But of course, we need to answer some of these questions sooner. What's your sense about what you can talk about how these questions are going to be answered quickly or given that they can't be answered in the usual way that maybe scholars would like to work on for 2030 years?

### **Shira Perlmutter**

Well, I hope that we will start getting some, at least initial economic analysis, and then to some extent in policy making. We're also always trying to figure out what's the most likely result. We certainly are asking for public comment and input in any way possible, even after our initial notice of inquiry. So we'll see. But it's kind of interesting now because there are many things happening on many levels. I mean, the way I always look at it is the new generation of generative AI available for use by the public is really at most two years old at this point. So we spent probably the first year figuring out what hit us. And I say we, I don't mean just the Copyright Office, I mean society. We then tried to understand the technology, how it worked, what the implications were.

We then figured out what the questions were that needed to be answered. And that has all taken place in the last year or so. And I think we're now at the stage where we'll start to see some solutions beginning to emerge. And we will see them. I mean, we're already beginning to see a number of licensing models emerging. We're seeing in various countries various regulatory models emerging. There's a lot of lawsuits pending in the court right now. I think something like 25 lawsuits involving copyright and whether ingestion in the circumstances in each case might qualify as fair use. Decisions are beginning to be made, narrowing down the claims and the cases. So things are beginning to happen.

And we are also seeing other government agencies look at aspects of these issues, notably the Patent and Trademark Office, the Federal Trade Commission, the Federal Election Commission, Federal Communication Commission. So there is a lot happening. And then, of course, bills being discussed and introduced in Congress. What I will say is that we hope that our report coming out of the study will be very helpful in a number of respects. First of all, in making recommendations to Congress, which, you know, are not always taken up, but are often influential, but also to maybe assist the courts as they look at these issues in setting out what the framework for analysis should be. You know, how do you look at these fair use questions? What are the facts that are relevant and how should they be interpreted in light of the, of past litigation? And then, of course, we're following what the courts are saying as well.

### **Joel Walfogel**

Thank you. I should mention that we do want to take questions. So we could be using the technology or just later you could raise hands and use microphones. But I wanted to ask you about two related

things. One is office modernization at the Copyright Office. But I think a related issue is creating data infrastructure, both internal and external, because, again, the patent analogy, the availability of patent data for generations now has created a big body of research that's arguably potentially helpful when it comes time to occasionally change laws. So what is the Copyright Office doing in these regards, both modernization and data? Yeah.

### **Shira Perlmutter**

Yeah. Well, I'm glad you asked those questions. So modernization is one of our top priorities, although we've started to move from calling it modernization to calling it continuous development, because we were in a situation about five years ago where some of our systems for managing, for administering the copyright law and notably recording documents related to copyright, like transfers of rights, were very antiquated. We were still using a paper based manual system, believe it or not, five years ago. So we really were in need to move into the 21st century now that we're almost finished with the first quarter of it. So what I will say is that we've been working very hard in this for about five years. We got funded by Congress in 2019 and started development in 2020.

But I do want to start with a bit of a boast, which is that even with our current systems, we're doing quite well, and that actually we have both registration and recordation processing times down to a historic low. So I'm very proud of that. Not that related to AI, but just in general. But what we are trying to do is to build a new enterprise copyright system, or ECS, because in Washington we're addicted to acronyms. And this will make all of our various services online, interconnected, user friendly, web based. It's really going to be a huge improvement. And we've now already, as I hinted at, moved our recordation system to an online system. So almost 90% of all recordation is now done online, and processing times have gone from months to weeks, or even days sometimes. And we are digitizing our historic public records and extracting metadata from them to make them more searchable. So all of that is happening, and we're now focusing on our registration system, which is the most complex of our systems. And we intend to have a limited pilot by the end of the year for the standard application and for uploading digital deposits. So a lot is happening in that area.

### **Joel Walfogel**

I see we're starting to get some interesting questions here. Can you see it in front of us here? So, for example, and you should only answer, obviously things that you can answer, but can you expand, it says on the licensing models you're seeing other countries consider or adopt, and which do you consider or which are some people considering to be most innovative and which might have risks?

### **Shira Perlmutter**

Yeah. Licensing models for AI, I think that's still very much in the early stages, other than, I think, what we're beginning to see. And probably there are people in the audience who know more about this than I do. But for AI models that are more targeted, there is a lot of licensing going on in particular spaces. The more complicated issue is for the large language model and general AI systems, where it's less easy to just find certain tailored input that is all that you need.

But among the licensing models that are possible, so you've got the individual licensing, as I said, some of which is going on, you've got collective licensing, things like what the Copyright Clearance Center just announced it's going to start doing for AI. Which means that, right holders in a given category all get together and you can go to one place to get a license that covers all of their works. And that's something that has been happening for many years in the US, in the music space in particular. So this is something we're familiar with, but it has not yet been generalized from one sector to another for AI specifically.

And then there's some discussion of what we call compulsory licenses or statutory licenses, which

could enable licensing within an entire category of works without needing to even get consent, that it would just take place by statute. The money would be collected somewhere and paid out. And so these are all under discussion. Generally, copyright owners don't like the compulsory licenses because they would prefer to be able to negotiate terms rather than have them set by the government.

**Joel Walfogel**

I see one that's just comment that might be interesting for us to, I mean, they're all interesting, but maybe to go deeper in some of the, during the live comment period, the question goes, what were some of the most pressing concerns raised by stakeholders regarding AI and copyright? Are there some that come to mind, and how have they influenced the recommendations that the Copyright Office is preparing for Congress?

**Shira Perlmutter**

Yeah, well, in many ways, the most pressing concern was exactly the digital replicas concern, because people are seeing every day in the paper examples of how this is happening. That, first of all, you see it with celebrities. You had Taylor Swift, Tom Hanks, Drake, and the Weekend, all these people whose images are being used. Biden and Trump's images have been used, making it look like they did or said things they didn't, have political context. So there's a great awareness right now of the problem. And then, of course, for ordinary individuals, for example, you've probably all seen reported in the press, students in middle schools who are putting their classmates images in pornographic settings. I mean, there's really a lot happening that people are concerned about. So that's been number one.

The other type of concern we're really hearing a lot about is this issue about the ingestion of copyrighted works for purposes of machine learning. And look, the arguments on both sides, there are concerns on both sides. On the one hand, you have copyright owners saying, look, it just, it simply can't be fair to take all of these works that we've invested skill and training and inspiration in producing, and to take all of it and feed it into a machine and use it as the raw material for stuff that's produced on the other end and competes with us in the marketplace and may undermine our ability to make a living.

On the other hand, what you're hearing from technology companies and a lot of public interest groups and user groups is, look, in order for this technology to work, it requires vast amounts of information of data to be fed into it. And given the number of works and the variety of works that need to be used to do effective training, it's simply not feasible to license every single work, and to find the owners and to license it. So the question now is, how do we achieve some sort of balance here? How do we make sure that the technology can thrive and continue to develop efficiently, effectively, productively, and at the same time make sure that human creativity also continues to thrive? And that is a really tough balance to find. And obviously there's a lot of concern on both ends.

**Joel Walfogel**

Yeah, that sounds like a hard problem.

**Shira Perlmutter**

You wish us luck?

**Joel Walfogel**

Well, I hope to help you. I wish you luck.

**Shira Perlmutter**

You wish all of us luck.

**Joel Walfogel**

Wish us all luck. So you had worked at the Patent Office and you've brought some, I think, some things from the Patent Office to the Copyright Office. You want to talk at all about any things you brought over?

**Shira Perlmutter**

Yeah, it's been interesting. I mean, I think one of the things I miss about the work at the Patent Office was looking at how all areas of IP fit together. And, you know, so that's been fascinating. And I've been doing a lot of thinking now about how the copyright issues and the patent issues involved in protecting the output of AI are similar and are different. I think the main thing that I brought with me, in addition to a couple of really good people, which I'm glad to have, has been the need to have the economic analysis function in house. And so I established a position of chief economist, hired someone very good, he's put in place a staff as well.

And Joel was kind enough to come and sort of initiate that function for us as a fellow for a year and did our first such study, because I really want to make sure that both our administrative work, our operational work, and our policy work is informed by economic research and analysis. And Joel started out with a fantastic study on the role of women in the copyright system, looking at data from our registration system from the past 40 or so years, and found that in fact there is a gap, a gender gap, that women were using the registration system less than men, although the gap was narrowing, and he might want to say more about it. And then while Joel was there, we advertised and hired the Chief Economist, Brent Lutzenhe, and it's been really interesting.

So he's doing data driven analysis of our operations, research on the economic aspects of copyright to guide policy, and also doing work to enable external researchers, including through making our data widely available, and that's something Joel was hinting at. So he's got a number of projects going on. We issued a report on the impact of Covid-19 on the copyright reliant industries. We're looking now at geographic patterns in registration, demographics of registrants. So I think all of this will be very useful. And then, of course, the AI research through the roundtable that we convened, and I do want to say we now have a website dedicated to the work of the Chief Economist. I recommend it. And also a public distribution platform for bulk data, which, as Joel indicated, was something that the Patent Office did for years and was not available in the past for copyright data.

**Joel Walfogel**

So let me just make one comment and then ask you one more question from the-- take a look. While I'm making the comment, you can look at that top question. It looks like it's a hot one.

**Shira Perlmutter**

Oh yeah.

**Joel Walfogel**

The comment I want to make is that the gender stuff in copyright, actually, women make up more than 50% of the copyright registrants for books, not for all kinds of copyrights, but for books. So it's really a remarkable context in which there's gender parity in an innovative kind of context. The other thing I just want to say is, because of what Shira has done, just as the patent data are available, the copyright data, back to '78, are now digitally available. It's a little bit cumbersome to work with, but they're out there and can be used. So I really hope that will foster research relevant to people in the

room. Okay, so let's hard question. Humans also learn from an existing body of knowledge. Why is it different for machines to learn that way as long as they don't just replicate it?

**Shira Perlmutter**

Yeah, that's an interesting point. And one that people have been talking about for a while. I'd say, you know, there's two main differences. So one is that one is a technical difference, which is just that humans do not make an exact copy of everything they see or hear, in their brains. So a reproduction is actually being made, and the machine, even if it's not kept permanently. Although there's some dispute over the extent to which there is more long term memorization in these different models. And the other main difference to me is just the speed and scale of it. Humans, yes, you're influenced by what you saw before. It affects everything that comes after, and that's the way that there's a lot of creative progress. It's a good thing. But what we're talking about here is exact copies made, immediately being able to be drawn upon in a way that is not just inspiration, it's actual use of every word or every pixel in the original.

**Joel Walfogel**

I don't know if pixel is bit byte. I don't know.

**Shira Perlmutter**

But it's a very good question, and I think it is incumbent on us to think about, from a policy perspective, what these differences are and why they should or shouldn't matter.

**Joel Walfogel**

So I think if we are to stay on schedule, we should probably be wrapping up. Where are that powers be? There's one. Am I right, Scott, should we kind of wrap it up? Okay, okay. Well, is there anything else you'd like to talk about? That was our last question. Oh, okay. Burning question.

**Speaker from the Audience**

I have a burning question. So, you know, the Copyright Office says there has to be a human involved. But one of the things that's happening is you big large language models are made by teams, and then the large language models, algorithms are getting more powerful. And so when you think about ten years from now, there will be no human that doesn't use algorithms to enhance their ideation, execution, implementation. So this notion that humans are the ones that have to have copyright, it feels to me like that is a losing battle, that it's a salmon swimming upstream, and that over time, you're not going to be able to just copyright humans because they're going to be less and less a part of the winning strategy versus algorithm assistance.

**Shira Perlmutter**

Well, a couple of thoughts in response to that. So, first of all, it's going to be almost never that a human isn't involved in some way in AI generated content. And in fact, the early test case, which was brought by Dr. Steven Thaler, which is the one that we rejected the registration, and that's been upheld by the courts, he claimed that the work was entirely generated by AI. But of course, humans always have to provide the prompts, and they're going to be involved in various ways in the output. So the difficult issue, which is what we're going to be addressing in the next part of our report, is when does a human contribute enough that there is, that the human is, well, the test is really, is the human controlling and providing elements of expression sufficient to constitute copyright in what the output is?

And to a large extent, a lot of what's done with these AI tools now is exactly that, using them as tools like you would use a camera. And how do you distinguish that from when the AI itself is determining what the output looks like? And it's a very difficult question, but it is very similar to the question that

we've dealt with in copyright law for a long time of distinguishing between idea and expression. One is the human contributing the idea and telling the machine what to do, and then it goes into the black box of what the tool does, and something comes out that the machine has determined. You know, you may be asking, you know, a more philosophical question, which is just, you know, why do we still have a human creativity requirement, a human authorship requirement?

And I just would say this is, first of all in the Constitution, which talks about giving authors, which we interpret to mean human authors rights, just as it talks about giving human inventors rights to their patentable inventions. But also the copyright system in the US is based very much on this idea of incentives. So the ultimate question is you can incentivize humans to produce, but you can't really incentivize a machine. You could incentivize a human to program a machine. And that's another question. Do you need copyright and the output to incentivize humans to program the machines to do the work? And that's one of the questions in our study, too.

**Joel Walfogel**

So let me ask you one final question that's inspired by something we didn't talk about, but this last question down here about regulatory arbitrage. Do you see risks in regulatory arbitrage with respect to different fair use rules, potential for different fair use rules across jurisdictions?

**Shira Perlmutter**

Yeah, this is a fascinating question. One of the things that I'm spending a lot of time doing is talking to counterparts in other countries. Everyone is looking for this balance, right? They all have creative communities. They all want to encourage the development of exciting new technology. They're looking for a balance. But there also is always in the back of other countries minds, to the extent they have collective minds, this idea that look, the United States has gotten out in front with these big, sophisticated AI systems, and so how do we compete? How are we going to do that? And there's already some suggestion. You know, the EU AI Act has a provision in it that says you cannot put an AI tool in the market in the EU if it was trained in ways that are inconsistent with EU law.

And so they're trying to control, to put their stamp on how AI is developed and trained in other countries in the world. So I think this is going to be an issue. I think there is a risk and regulatory arbitrage. I think we are likely to see some level of sort of de facto harmonization as countries talk to each other and understand, policymakers talk to each other and understand, why they're making the decisions they are and what they're testing and how it's working. But there is also this risk at the same time that we need to control for, and maybe ultimately there will be some need for some international treaty work, whether it's through trade agreements or the World Intellectual Property Organization. But of course, we're nowhere near ready for that at this point.

**Joel Walfogel**

All right, well, thank you so very much.

**Shira Perlmutter**

Thank you.