Copyright + Technology Conference 2025

Panel 2

A THOUSAND TIMES NO: THE PRACTICALITIES OF OPT-OUT FOR AI TRAINING

With Adam Rendle,[1] Daan Archer,[2] Ali Sternburg,[3] and Thomas Sullivan[4]

Generative AI gives rise to many questions about how to license very large amounts of content efficiently where the need arises. One of the types of licensing schemes that has been discussed is "opt-out," in which online services use copyrighted material unless rightsholders tell them not to (e.g., YouTube), as opposed to "opt-in," in which online services require permissions from rightsholders to use their content (e.g., Spotify). This panel discusses deliberations and controversies over opt-out schemes for AI training that have taken place in the UK and EU as well as in the US, and some of the technical challenges in implementing workable opt-out schemes.

Adam Rendle: I'm Adam Rendle. I'm a partner of the law firm Taylor Wessing in London. I'm also the International Programming Committee chair [for the Copyright Society], and Bill [Rosenblatt, Copyright + Technology program chair] has got me here to talk about the UK and EU aspects of our topic on opt-outs for AI training. The concept of the panel was entirely Bill's, but I was really excited to hear him propose it because it seemed to me that the practicalities of how opt-outs for AI training work are a key piece of the puzzle as to whether opt-outs should be introduced as an exception to copyright infringement to enable AI training.

Let me introduce our panelists. We have to my immediate right, Daan Archer, the CEO of Copyright Delta. He's got a huge amount of experience in applying technologies to solve copyright problems, and he's ideally placed to show his insights on what technologies are available and what problems they face. He also used to be a diplomat, which seems to be a particularly useful skill set in navigating these challenges.

Then we have Ali Sternburg, who is the VP of information policy at the Computer and Communications Industry Association. She's got her fingers on the pulse of the US tech industry's positions on these issues and will be speaking to their views on opt-out technologies.

---

[1] Partner, Taylor Wessing, London, (last visited Nov 20, 2025) https://www.taylorwessing.com/en/people/united-kingdom/london/adam-rendle (last visited Nov 20, 2025)

[2] Copyright Delta, https://www.copyrightdelta.com/ (last visited Nov. 20, 2025).

[3] VP of information policy at the Computer and Communications Industry Association, · Computer &Communications Industry Association, https://ccianet.org/ (last visited Nov. 20, 2025).

[4] Partner, Ballard Spahr, https://www.ballardspahr.com/people/attorneys/s/sullivan-thomas, (last visited Nov 20, 2025).

And last but not least, Tom Sullivan, a partner at Ballard Spahr, who represents rights holders enforcing their copyrights. So, he knows what's at stake for them if opt-outs don't work.

Let's do a quick running order. I'm going to start with, as I mentioned, the EU and UK positions on how opt-out exceptions are working now and how they're being proposed. Tom and Ali are then going to touch on what's been discussed in the US. And then we'll collectively discuss the merits and shortcomings of the available opt-out technologies. But before that, a quick word on what the panel isn't about.

So, this is not a policy debate about which regimes provide the best outcomes for copyright holders, the tech industry, and society as a whole. We're not going to debate fair use, status quo, raw data mining exceptions, transparency, collective licensing, all that stuff. We're going to assume that an opt-out regime is the regime that is chosen, but I do accept that one's appetite to adopt these technologies and one's assessment of how good they are will depend on which side of that policy debate you're on. I also recognize that opt-out is a sort of loss for both sides of the debate.

Some developers want an unqualified TDM regime[5] without the burden of having to license or look for or respect opt-outs. And of course, rights owners want absolute control without having to go to the burden of inserting opt-outs. We're also not going to mention whether an opt-out regime defends against the Berne Convention's prohibition on formalities.[6] But Bill seems to be – based on his opening remarks – and this panel seems to be teeing up a panel on that, maybe at the next Midwinter or Annual [Meetings of the Copyright Society], if we should introduce formalities.

But we're not going to be discussing that here, and we're also not going to discuss whether the opt-out exceptions are compliant with the three-step test.[7] And we are also hopefully not going to discuss fair use[8] litigation. I know there's a lot to discuss there. So, let's jump into the EU as one of the first main jurisdictions to have exceptions for TDM. There are two. One is scientific research, which is not subject to an opt-out, but I mentioned this for completeness and because it was subject to a case I'll go on to discuss.

The one that's most relevant now is Article 4.[9] This was introduced in 2019, the Copyright Directive.[10] It debates about whether or not the EU legislator had in mind AI and gen[erative] AI, and everything that's going on. It probably didn't, and there has been some debate as to whether or not this exception applies at all. We'll come on to that shortly, but just to pull out a few key concepts from the EU's exception. The work to be suitable or allowable for text and data mining has to be lawfully accessible. You can only retain the work for as long as is necessary for the purposes of text and data mining.

---

[5] "TDM regime" refers to copyright exceptions for text and data mining, i.e., using software to analyze large bodies of text and data to, e.g., discern patterns and compute statistics.

[6] The Berne Convention for the Protection of Literary and Artistic Works, https://www.wipo.int/wipolex/en/text/283698, Article 5 (2).

[7] *Id.,* Article 9 (2).

[8] 17 U.S.C. § 107.

[9] Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market ("Directive (EU) 2019/790"), https://eur-lex.europa.eu/eli/dir/2019/790/oj, Article 4; *see also id.* Article 3.

[10] Directive (EU) 2019/790.

And then in Article 4(3) at the bottom there,[11] the exception applies in the condition that the right has not been reserved expressly – What does expressly mean? – by the right holders – Who are right holders? – in an appropriate manner – What's appropriate? What's machine-readable? – in the case of content made publicly available. All up for debate in the EU at the moment. The recital has some more detail about this.[12] It references if content has been made publicly available online, which is, I guess, what we're talking about here in the context of web scrapers, for example. It should only be considered an appropriate means if you use machine-readable means.

And then they give two examples of the metadata and terms and conditions.[13] Now, as I've said, there's been a case coming out of Germany.[14] This was a photographer. The LAION data set,[15] of which many of you may be familiar, scraped from a stock image library the photographer's photo, and put it in [its database]. He sued, and this then landed, let's just say, the first decision on TDM opt-out.[16] The first instance in Germany. It decided – and I mentioned Article 3[17] before, which is not the opt-out exception – that this kind of use did qualify as TDM, which was a big decision to recognize that works going into a data set for AI training can qualify as TDM.

But for present purposes, the *obiter* in this case, the court having found the non-commercial TDM exception applied, still went on to discuss the applicability of the commercial or the general TDM exception, which is, as I say, subject to the opt-out. It formed a very broad view of all of those issues I called out when we were looking at the exception in the first place. The opt-out we were talking about is one of those kinds of generally applicable opt-outs you see in all sorts of terms of use websites, basically prohibiting automated programs or bots accessing the website or the content on it for any purposes, including indexing, scraping, or caching.

It didn't mention TDM explicitly. It didn't mention training data sets, et cetera, but that was enough. The photo library from where the image was found was a non-exclusive licensee. The court said that even that type of licensee, that type of rights holder in the world of legislation, could declare an opt-out. It was in natural language in terms and conditions of the website. That was enough, even though it wasn't using a sort of robots.txt exclusion protocol. And that type of opt-out did satisfy the machine readability test. Now that case is subject to appeal.

While it was, say, against a sort of non-profit non-commercial organization, commercial organizations who are involved in the background, because clearly this is a significant case, I think there will be a hearing this side of Christmas. So, we'll get a bit more clarity from Germany on that. It will probably go to the Court of Justice, given the

---

[11] *Id.,* Article 4 (3).
[12] *Id.*
[13] Directive (EU) 2019/790, preamble, (18).
[14] Hamburg Regional Court, Germany [2024]: Robert Kneschke v. LAION e.V., Case No. 310 O 227/23, Germany.
[15] LARGE-SCALE ARTIFICIAL INTELLIGENCE OPEN NETWORK, https://laion.ai/ (last visited Nov. 20, 2025.)
[16] Unofficial English translation (generated by ChatGPT) at https://chatgptiseatingtheworld.com/2024/09/28/unofficial-english-translation-of-german-courts-decision-kneschke-v-laion-under-tdm-exception/.
[17] Directive (EU) 2019/790, Article 3.

importance of the decisions. So, I said there was some debate about whether the TDM exception applied to AI training at all. That debate was answered in the AI Act.[18]

We won't go into talking about this in any more detail. A very significant piece of legislation causing all sorts of trouble for all sorts of people in the AI ecosystem. But for present purposes, providers of general-purpose AI models, which is, for example, your big foundation models,[19] have to put in place some policy to comply with [European] Union law, including how they recognize opt-outs using, for example, state-of-the-art technologies. So, all providers of general-purpose AI models will have to be creating policies to check the status of works by reference to their opt-out status. There's now been a code of practice published in July.[20]

Various big tech signatories, the names that you will recognize.[21] It's interesting, this code of practice, particularly because I think what it does is recognize the very nascent state of opt-out technologies and their utility and their applicability. The signatories commit to employing web crawlers that follow robots.txt,[22] the kind of starting position for how these crawlers work. And then when it goes onto the other techniques available, they're very vague about what techniques are to be used and adopted. They reference good things like making sure they've been adopted broadly, they're state-of-the-art, they're technically implementable, widely adopted, agreed through an inclusive process.

But it very much begs the question about what these technologies are going to be. So, we have in the EU an exception that is dependent on the application of technology, but no agreement or consensus at the political level, or that kind of stakeholder signatory level of what opt-outs can be. There is recognition, though, of some dialogue between the signatories and the rights holders, so the rights holders can work out information about the web crawlers.

And there's also some concern that if you use some web crawlers that are employed by search engines, it's kind of all or nothing. That if you're out for AI training, you're also out for indexing and search engine [use], which also has bad implications for publishers, for example. But there's a note there to sort of decouple that. So, if you've got a search engine, compliance and rights organizations won't cause adverse effects on indexing.

Now [in] the UK, we're introducing a similar exception. Now people often ask me about Brexit, particularly over on this side of the pond. "Got any benefits of Brexit yet?" No, no, no, there aren't any. But guess my political persuasion. However, I may have found one, and it may be this. What we are doing is looking at the EU experience of how

---

[18] *The Act Texts*, EU ARTIFICIAL INTELLIGENCE ACT, https://artificialintelligenceact.eu/the-act/ (last visited Nov. 20, 2025).

[19] Foundation models are the primary large computational models for producing content in generative AI, such as OpenAI's GPT series, Meta's Llama series, and Google's BERT. they are built by training on very large bodies of material of the type being generated.

[20] *The General-Purpose AI Code of Practice*, EUROPEAN COMMISSION, https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai (last visited Nov. 20, 2025).

[21] *Signatories of Code of Practice*, EUROPEAN COMMISSION, https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai#ecl-inpage-Signatories-of-the-AI-Pact (last visited Nov. 20, 2025).

[22] Robots.txt is a file that a web publisher can place on its website; it contains commands that web crawlers (such as those for search engines and AI training) can read to determine which web pages the publisher wants to exclude from crawling. Compliance with robots.txt is strictly voluntary.

opt-outs work, and to use that to develop our possible opt-out TDM exception. We have recognized that it's not clear in the EU. It's not that practically workable. So, we are actively looking at what's going on in the EU. Having said that, there's this consultation that was published a couple of days before Christmas last year. So, it led to a lot of rushing around over that period. So, we've had the most number of responses to a consultation ever run by the UK Intellectual Property Office. It closed in February.[23]

And here we are in September, [and] not a word has been spoken about what's gone on since. Our Labour Party and government have had a few things to deal with. A bit distracted, it seems. And the latest I heard is that we're not even going to get a consultation response out until January next year. And then there'll be another however-long [period] of legislative process before we even know if anything is going to get on the books. That legislative process will be very highly contested. We've already seen some examples of that. And at the same time, the government recognizes that there's a ton of work that needs to be done to establish proper functioning opt-out technologies before you can introduce an opt-out exception.

So, in the UK at least, we're in a situation where the government recognizes the importance to UK PLC, UK economic growth, of developing conditions to enable AI training to take place, but it's basically not making any decisions at all. And even when it comes to making a decision, it will have to work through all the stuff we're going to discuss in the panel about how opt-out exceptions can work in practice. It wants to form stakeholder groups. It recognizes that we may need to have some form of regulation about opt-out technologies.

So, it's going to be another 18 months or so before any legislation comes about, by which time, god knows what sort of version of ChatGPT we'll be on by then. Things have been developing very quickly. The UK government is not moving very quickly. They haven't made a decision. In my view, they effectively have made a decision which is that the status quo is going to prevail for the meantime, and it's left to the industry to work out how to license things and how to operate opt-outs. So, that's the UK and EU summary. Tom, do you want to pick up on the US?

Thomas Sullivan: Thanks, Adam. From the US perspective, the Copyright Office at least expressed earlier this year a different view. As we'll talk about for a second, not clear that's still their view or that their view necessarily matters, but they at least expressed a different view.

So, in early 2023, the Copyright Office launched an initiative to examine AI issues sort of generally and sought comment on a number of issues.[24] One of those topics was broadly the use of copyrighted materials as training data for generative AI, and specifically for our purposes today, whether there should be an opt-in or opt-out system.[25] And how the opt-out process – assuming we went with an opt-out process – how it would work and the obstacles to establishing it, some of which Adam talked about a second ago.

---

[23] Copyright and Artificial Intelligence, GOVERNMENT OF THE UNITED KINGDOM, https://www.gov.uk/government/consultations/copyright-and-artificial-intelligence (last visited Nov. 20, 2025).

[24] *Copyright Office Launches New Artificial-Intelligence Initiative*, U.S. COPYRIGHT OFFICE, https://www.copyright.gov/newsnet/2023/1004.html (last visited Nov. 24, 2025).

[25] Artificial Intelligence and Copyright, 88 Fed. Reg. 59942, 59947 (Aug. 30, 2023)

The request for comment explicitly referenced the process in the EU.[26] It recognized that sort of we're in an international conversation here. And [it] said, "Here's what the EU is doing. Talk about this, people of the world." In response by December of 2023, the Copyright Office received over 10,000 comments.[27]

The vast majority of those were on the generative AI issue. And as you can expect, they can sort of be sorted into two large buckets. The copyright owners expressed a lot of concern about an opt-out system.[28] They worried it would be antithetical to current law, unduly burdensome, impossible to utilize after the training process had occurred for these systems and would be difficult to implement. They also pointed out that it could be difficult for an opt-out system to be effective because the metadata that attaches to their works might disappear over the process as they move from site to site or place to place. And that the owners don't necessarily control every site where their works might appear. So, you license it to party A. You now need party A to always do a good job opting out. Or a pirate site could pick it up, and you've not licensed it at all. You have no control over that. On the other side, commentators from some of the AI companies or those in favor of the technological innovation point said it could be beneficial to allow greater development of these systems and that it could be sort of the only effective way to do it.[29] So, there were a range of views.

So, in May, as I assume at least a number of people in this room know, the Copyright Office issued a somewhat sort of extraordinary draft version of its report related to generative AI.[30] The Copyright Office had earlier released part one and part two in sort of the normal, "This is our final version of part one. This is our final version of part two," in July 2024 and January 2025.[31] So, May 9, the Office released this prepublication version of part three and it included an analysis on opt-out.[32] It analyzed the various approaches, including the EU, and ultimately took the side of the copyright owners, finding it was inconsistent with sort of a basic principle that you need consent to use someone's intellectual property.[33] But they acknowledged that Congress could consider some sort of exception for AI training in the future and that the ability to opt out could preserve some

---

[26] *Id. at* 59947 n.47

[27] *Copyright and Artificial Intelligence Part 3: Generative AI Training (Pre-Publication Version)*, U.S. COPYRIGHT OFFICE, at 2 https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf ("USCO AI Training Report") (last visited Nov. 20, 2025.)

[28] *Id.* at 83, 101-102.

[29] *Id.* at 83, 101-102.

[30] *Id.* at 102-03.

[31] *Copyright and Artificial Intelligence Part 1: Digital Replicas*, U.S. COPYRIGHT OFFICE, https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf (last visited Nov. 20, 2025); *Copyright and Artificial Intelligence Part 2: Copyrightability Report*, U.S. COPYRIGHT OFFICE, https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf (last visited Nov. 20, 2025).

[32] USCO AI Training Report at 101-103, 105.

[33] *Id.* at 107 ("[W]e conclude that several stages in the development of generative AI involve using copyrighted works in ways that implicate the owners' exclusive rights").

of the copyright owners' rights.[34] But they had concerns about the effectiveness and availability of the ability to opt out.

So, you might say, "Okay, the Copyright Office has weighed in." But as I said, this is sort of an extraordinary report. The report says, "We're releasing this in response to congressional inquiries and expressions of interest from stakeholders. A final version will be published in the near future without any substantive changes expected in the analysis or conclusions."[35] All of which could be true, but it was highly unusual.

The publication came the day after the firing of the Librarian of Congress,[36] who appoints the Register of Copyrights, the director of the Copyright Office.[37] The Register was purportedly fired the next day, May 10.[38] So, I think it is fair to speculate that the changes in leadership or the imminent change in leadership potentially led to the publication of this prepublication version of the report. No final report has been issued. It's been three months since May. That seems a little bit different than in the near future, but maybe they have a different definition than I do. And it's unclear, of course, whether there will be substantive changes if and when a final report is issued.

As most people know, it's not even entirely clear what's going to happen with the leadership of the Copyright Office. There's a lawsuit going on right now about whether President Trump can fire the leadership of the Office.[39] So far, a lawsuit has not resulted in a reversion to the prior leadership, but that's going to have to play out in the process.[40] And then assuming that the firing was appropriate, it's entirely unclear who President Trump is going to put in charge of the Copyright Office and what that person's interests or beliefs will be.

The other thing, of course, is that the Copyright Office's view in general is just a view. It's a highly informed view. It's a view that carries some weight, but ultimately, anything that would happen in the US would have to happen through legislation on this

---

[34] *Id.* at 101-103, 105.

[35] *Copyright and Artificial Intelligence*, U.S. COPYRIGHT OFFICE, https://www.copyright.gov/ai/ (last visited Nov. 20, 2025).

[36] Librarian of Congress Carla Hayden was fired the evening of May 8, 2025. Sueng Min Kim, Zeke Miller and Lisa Mascaro, Trump Fires Librarian of Congress Carla Hayden, ASSOCIATED PRESS, https://apnews.com/article/donald-trump-library-of-congress-carla-hayden-20a1862ce6d2e0d51a84a37b264ce2ef (last visited Nov. 24, 2025). The pre-publication report was issued on May 9, 2025. Copyright and Artificial Intelligence, U.S. COPYRIGHT OFFICE, https://www.copyright.gov/ai/ (last visited Nov. 20, 2025).

[37] 17 U.S.C. § 701(a).

[38] Blake Brittain, Trump fires head of U.S. Copyright Office, RUETERS, https://www.reuters.com/legal/government/trump-fires-head-us-copyright-office-2025-05-12/ (last visited Nov. 20, 2025).

[39] *Perlmutter v. Blanche*, No. 25 Civ. 5285 (D.D.C.).

[40] The D.C. Circuit subsequently granted an emergency injunction preventing Acting Librarian of Congress Todd Blanche and others from "interfering with [Shira Perlmutter's] service as Register of Copyrights and Director of the U.S. Copyright Office pending further order of the court." *Perlmutter v. Blanche*, No. 25-5285, 2025 WL 2627965 (D.C. Cir. Sept. 10, 2025). Blanche has sought a stay of that injunction in the Supreme Court, *see Blanche v. Perlmutter*, No. 25A (U.S.), which remains pending. As of November 24, 2025, Register Perlmutter has resumed her duties.

issue, probably. And that's going to be written by Congress, not the Copyright Office. So, we'll have to see.

Adam Rendle: So, we've got a lack of clarity in the US for much more political reasons than I think we've got a lack of clarity in the UK. We've got an EU exception, which itself is unclear on its face and is subject to a German decision which is obiter and subject to appeal. I don't know. Right, Ali, have you got any clarity you could bring here?

Ali Sternburg: I don't know if I would put it the same way that there's a lack of clarity. I think that there's extraordinary circumstances around this prepublication report and the leadership of this part of the government, but as Tom was just saying, it's Congress that is tasked with legislating and creating the laws. We also haven't discussed the courts, as well, but I'll try to stay on topic. I do want to talk about fair use and formalities in the courts and all the things I'm not supposed to talk about, but I don't think there's a lack of clarity because I think that existing law in the U.S. is working.

And the reason that the EU and UK are looking at these really specific exemptions around text and data mining are because they don't have fair use. And in the U.S., we have fair use, and that is precedent.[41] That enables the training of AI and these successful companies that we have in the U.S., and the benefits that they provide for the public. So, I do think that's another important copyright stakeholder. Whenever we're talking about copyright policy, a lot of times it's seen as two sides, but as I learned from my copyright professor Peter Jaszi,[42] there's the three legs of the stool. You always want to include not just the rights holders, the intermediaries and service providers, but also the public.

And that's where copyright comes from in the Constitution, when we're talking about promoting progress and always thinking about those important stakeholders as well. So, there are a lot of voluntary initiatives I think we're going to probably talk about next. And so, yeah, while there might not be anything mandatory, I think from robots.txt, the standard that's been working for a long time, to some other initiatives that are happening in the private sector or from individual companies, from vendors and service providers as well, and then some task force, IETF[43] kind of standard setters as well.

So, I think a lot of people are very engaged in this space, want to see things work, want people to be able to express their preferences, want to enable innovation and free speech, and protection of IP. And so, I think that these are not necessarily not aligned. And I think that, yeah, if changes need to be made, that's something that can be figured out. But right now, even though AI is very novel in some ways, the copyright system in the U.S. has adapted to new technologies for centuries going back. So, it's another challenge, but it's not something where we need a whole new AI law regime. I think copyright law can adapt flexibly to this new technology.

Adam Rendle: Thanks, Ali. And a very important segue, I think, as one realizes the limitations of being a lawyer and a policy maker. And realize that certainly in the UK and EU, to your point, Ali, that in a kind of fair use regime, maybe opt-out is sort of, I guess, irrelevant or less relevant. But at least in the UK and EU and other regimes that are contingent on the success of opt-out to strike the balance between the different interests that you mentioned, I think we quickly have to shift the conversation to technical

---

[41] 17 U.S.C. § 107.

[42] Emeritus Professor of Law, Washington College of Law, American University. He is also a Partner at Jaszi Butler, https://www.usefairuse.com/about.

[43] INTERNET ENGINEERING TASK FORCE, https://www.ietf.org/ (last visited Nov. 20, 2025).

standards. I think us as lawyers and policy makers can come up with various qualities of what we define to be a workable opt-out standard.

And in the UK, to note, the government has said an opt-out exception will only be introduced if we have workable opt-outs. And that, from what we're hearing, could be a very big if. And what I've done here is bring together twelve kind of qualities that a workable opt-out standard may need to meet. Maybe this is actually asking for too much. Maybe this reveals a kind of pro-rights-holder bias but just trying to aggregate various standards around machine readability. Work out if it's going to be location and or unit-based. Location being URL-related, unit-based being sort of attached into the specific content.

We've already heard about it being standardized and respected. It's got to be automated, simple, and scalable so it can work for creators from the biggest creator to the smallest. And we heard on the panel before that the interest of indie artists, for example, is going to be targeted, whether that's at a content level, different types of crawlers, different types of uses that it enables, or prohibits. Indelibility, I think, is really important. And the UK government has recognized that opt-out technologies may not be indelible, i.e., they may kind of fall off.

And if you've attached them to your content on your newspaper website, for example, as Tom mentioned, they can disappear. So, maybe being indelible is really important. And that can prohibit or apply to downstream and multi-stream use. Multi-stream, I mean different types of technology than just viewing things on a browser. The UK government actually recognized that because technology isn't indelible at the moment, there may be an obligation on technology companies to look elsewhere than the place where they're scraping it from, or to create registries where opt-outs can be centralized and recorded.

And they've got to be quick to market. They've got to be up to date as crawler technology changes. Do they have to be forward and backward looking? Does it apply to data that's already been scraped? I've mentioned the existence of a central repository. And this is something that the UK government didn't recognize. Do there have to be, or should there be legal protections for opt-out technology?

As in, should there be incentives or carrot-and-stick-type regimes for tech companies or the developers to look for and respect opt-outs? Should there be prohibitions on stripping out opt-outs from content? But as I say, we recognize the limitations of lawyers and policy makers here. So, it's a good job we've got a technologist. Daan, over to you. What are you seeing? What's out there? What are you building? What are you hoping to see?

Daan Archer: Hi. Thank you so much for the invitation. I'm Daan Archer. I've been an AI software developer for 25 years. I've also been a technology diplomat for the Dutch government in Japan. I did my MBA at MIT, trying to look to compare capitalism versus the Japanese consensus-building. And I've been in the music industry since 2016 with Howie [Singer][44] and Bill [Rosenblatt].[45]

---

[44] Adjunct Professor, Music Business, New York University. *See* in this issue, Aruni Soni, Patra Sinner, Jon Glass, and Katherine Reilly, *The Masquerade: Streaming Fraud and Copyright*, J. COPYRIGHT SOC'Y 1005 (2025).

[45] Lead of Technical Working Groups, Open Music Initiative, 2016-2018.

A year ago, we were asked by the music societies in Ireland and Finland to build an opt-out system, which resulted that we have aligned all the rights that are needed to opt out because you need to clear 300%: society, publishing, and the recording.[46] That is functional, operational, and now it can serve as an opt-in system because you have cleared all the rights. The next step with Finland has not been publicly announced, but I can mention it. It's now going to build it into a global licensing framework. So, technically, it's possible. You do need to align parties. And going back to the earlier question, I do believe that we need to get into a KYC[47] world because we work with a lot of banks like ING Bank, and they've got stablecoins,[48] and you can make global direct payments, but you need to have KYCs there.[49]

Going back to opt-out and the robots.txt, the difficulty with that is that we called 180 AI companies because that was part of what we did. Half of them didn't pick up the phone. So, we said, "Hey, we have opt-out here for you," but nobody is answering it. The other thing is with robots.txt, you actually need to integrate it into the file level, because then the file has it. But the problem is, as soon as you upload a file, like if you upload to Spotify, you have an FTP[50] and XML[51] files, but then the text is not in there. So, how do you tackle that? And I was just actually looking just now for the latest, but for example, C2PA[52] could be a good standard for this. They actually include at the file level the opt-out mechanism.

The other one I was looking at was ISCC, but they don't have it yet.[53] We've built a small data structure for that. So, technically, it's not there yet. So, there's a platform level.

---

[46] Mr. Archer refers here to three types of entities that have to provide licenses for typical online uses of recorded music: music publishers and collecting societies (for rights to musical works), and record labels (for rights to sound recordings that embody the musical works).

[47] Tom Sullivan, *What KYC is and why it matters in financial services*, PLAID, https://plaid.com/resources/banking/what-is-kyc/ (last visited Nov. 21, 2025).

[48] Cryptocurrencies whose values are tied to fiat currencies such as the U.S. dollar or commodities such as gold. *See, e.g.,* Nathan Reiff, Understanding Gold-Backed Cryptocurrencies: Benefits and Challenges, INVESTOPEDIA, https://www.investopedia.com/tech/goldpegged-vs-usdpegged-cryptocurrencies/ (last visited Nov. 21, 2025).

[49] KYC regulations are common in the banking industry. *See, e.g.,* James Chen, *Know Your Client (KYC): Key Requirements and Compliance for Financial Services*, INVESTOPEDIA, https://www.investopedia.com/terms/k/knowyourclient.asp (last visited Nov. 21, 2025).

[50] i.e., You upload a file using the standard File Transfer Protocol (FTP). *See File Transfer Protocol,* INTERNET ENGINEERING TASK FORCE, https://datatracker.ietf.org/doc/html/rfc959 (last visited Nov. 21, 2025).

[51] eXtensible Markup Language (XML), a widely-used standard language for describing structured data. *See Extensible Markup Language (XML) 1.0 (Fifth Edition)*, WORLD WIDE WEB CONSORTIUM, https://www.w3.org/TR/xml/ (last visited Nov. 21, 2025).

[52] Coalition for Content Provenance and Authenticity, a standards initiative for storing and tracking information about how and by whom an item of content was created and modified. *See* COALITION FOR CONTENT PROVENANCE AND AUTHENTICITY, https://c2pa.org/ (last visited Nov. 21, 2025).

[53] International Standard Content Code, a digital content identifier standard. *See* INTERNATIONAL STANDARD CONTENT CODE, https://iscc.codes/ (last visited Nov. 21, 2025).

There's an opt-out. You can say, "Hey, which ISWCs,[54] which ISRCs[55] can I train on and which not?" So, the ultimate technical solution at a far level is not there yet. But yes, opt-out, it is available.

Adam Rendle: I would just push a little bit on the file levels. This is the song file that gets shared with the DSP[56] or could be a movie file that gets shared with a DSP in Netflix, for example. What is it that creates a limitation such that it isn't available at that file level?

Daan Archer: Yeah, this is a quote from a good friend of mine who was part of the HTML in the 1990s. And he said, "The oldest standards that have won are the easiest lightweight standards." The XML didn't win; JSON[57] won. PDF has no metadata. MP3 has no metadata. So, all the standards without metadata are easy to use. So, sadly, today we live in a world where the easiest standard won without metadata. It's easy.

Adam Rendle: So, robots.txt seems to be the frontrunner, and certainly it's been recognized as such by the U.K. government. And I think Google, for example, in its submission to the U.K. government, based its position on robots.txt.

Daan Archer: If I can add to that. So, your society, you represent ISWCs. So, what if the robots.txt on your website doesn't link all the way to the ISRCs on the sound recording? So, you need to look there. That's the thing. So, that's what we did with the opt-out. We connect the ISWCs with the ISRCs and the music files.[58] But you don't get that from your robots.txt. That's the thing.

Adam Rendle: For those in the audience who aren't as au fait with acronyms in the music industry, do you just want to explain what you mean there?

Daan Archer: Yes. So sorry. So sorry. ISWC, international standard work code. It is for the writer and for the composition. And then ISRC is the international standard [identifier] for the recording. So, basically, every composition, every recording, they need to be glued together. It's not there yet. Often, distributors, they only upload the recording information. Often, not even the publishing information, so this is where a lot of the black boxes are. So, if you want to do a proper opt-out, you say you need to opt out the full 200%. So, we thought this was the U.K.'s case to now finally start connecting all those items. So, now we have a system in place that connects all those.

Adam Rendle: Will you talk a little bit about how that system operates to the extent you can without giving away any trade secrets?

---

[54] International Standard Works Code, a standard identifier for creative works, most commonly used with musical works. *See* ISWC NETWORK, https://www.iswc.org/ (last visited Nov. 21, 2025).

[55] International Standard Recording Code, a standard identifier for sound recordings. *See* INTERNATIONAL STANDARD RECORDING CODE, https://isrc.ifpi.org/ (last visited Nov. 21, 2025).

[56] Digital service provider, e.g., Spotify or Netflix.

[57] JavaScript Object Notation, another machine-readable language for describing structured data. *See ECMA-404 The JSON Data Interchange Standard*, JSON, https://www.json.org/json-en.html (last visited Nov. 21, 2025) JSON is often preferred over XML because it is more syntactically compact and easier to read. *See, e.g., What's the Difference Between JSON and XML?*, AMAZON WEB SERVICES, https://aws.amazon.com/compare/the-difference-between-json-xml/ (last visited Nov. 21, 2025).

[58] *See e.g.,* Everything You Need to Know About ISWCs, SONGTRUST, https://help.songtrust.com/knowledge/everything-you-need-to-know-about-iswc-codes (last visited Nov. 21, 2025).

Daan Archer: No, no, no, no. So, we're huge fans of compliance and data privacy. So, we built a system that allows everybody to host their own data, but we've connected the splits[59] and the equity splits in between on those identifiers. So, it's a single API which could have like the robot. So, it's a single API that AI companies can talk to and others.

And behind the scenes, it connects all those percentages and all those parties, but they're all storing the data in their own locations. So, it's data sovereign, data privacy, but you have a single API. A one-stop shop. You have to make it easy. So, we work with a lot of lawyers and bankers. Yeah. Because I believe you have to be compliant in this if we want to tackle this problem once and for all.

Adam Rendle: If there is a one-stop shop, is that one-stop shop attached to where the works are available, or is the shop somewhere else? And I ask that because it seems to me that one easy way or, sorry, very important way of making opt-outs work is that they are expressed at the point of crawl or the point of consumption.

That if a crawl is originating out to find these works, wherever they may be, on the authorized paid-for services, or as Tom mentioned in the outreach, is in an unlicensed place. If the opt-out isn't expressed or reserved alongside the work, then you've already kind of lost the ability to opt out. In the model you're describing, is the opt-out attached to the work, or is it somewhere else that the scraper, that the aggregator will then have to go and check?

Daan Archer: Very good point.

Adam Rendle: Thanks.

Daan Archer: So, we've connected it to the endpoints. So, Spotify endpoints, YouTube endpoints, as in, "Hey, this endpoint, don't train on that because the rights [owner] associated to this has opted out." But the ideal case is at all these endpoints on Spotify, on YouTube, on all those, you should have the opt-outs robots.txt there. So, participation of the platforms.

Adam Rendle: And by endpoint, you mean the points at which the consumption happens?

Daan Archer: Yeah. So, it's the public address. So, for me, if I were to be a black hat hacker, it's really easy to download the internet. That's what Google did, but then it was hard. Now it's easy. So, all your collective knowledge together with that power is very powerful.

Adam Rendle: Ali, do you want to chip in?

Ali Sternburg: Sure. So, I have a few other things I wanted to react to. So, yeah, when we're thinking about expressing these preferences around opt-out, I think it's a good point about when you do that. I like to compare when we're thinking about AI to what could a human do. Can a human unlearn and forget? Would you have to completely start from scratch? How much would that cost? What environmental impact would that have on the trees, as we were talking about the other day? Both the compliance cost, but also that would make the dataset and the tool, and all the information on the internet, would make the model worse.

And that again isn't just about the AI developers themselves but also about all the different use cases that will be for AI in healthcare. When you're thinking about the impact and getting artificial intelligence right in all the different use cases, economic,

[59] Percent shares of royalty collection for a given musical work.

non-economic, research, and then just thinking about the U.S. perspective, the competitiveness, wanting that economic activity and benefits to be available here. And so, I think, yeah, understanding that it has such an impact on so much of the economy and so much of society. And the open web is something I did want to emphasize.

And that having opt-out tools, it impacts how they actually work and the benefits it can provide. And then, back to fair use, I did want to stress that in the U.S., the courts have found it "exceedingly," "spectacularly"[60] fair use. And so, I think this is such an important conversation to be having, and a lot of people are thinking about it. So, the EU IP Office put out a study, I think earlier this year, I believe, where they had looked at all the different opt-out tools.[61] And I think you mentioned, are they workable? I think they found that none of them were ideal or something.

That doesn't mean that there won't be, but it is really challenging to meet all those different 12 categories you had earlier, and to Daan's point about them being lightweight, easy. If it's going to be adopted, whether it's voluntary or if it's mandatory, it has to be something that is only going to be adopted if it's voluntary, if it is that easy and lightweight. And so, that is really interesting, looking at past examples that we've seen around MP3 and other standards that maybe have less of the metadata, but it makes them more interoperable and more useful, and then more adopted by internet users.

Daan Archer: Can I? So, if we look at AI from a geopolitical perspective, so we have the U.S., China, Europe, and where do we go? And nobody wants to lose the AI race. At the same time, AI is a great tool. You can mass crawl a lot of data, search through it. There's no guarantees right now because AI systems you have input, you have output. We don't know the middle. It's basically a fault in the design from the 1960s. We still have the old architectures, but a lot of people are now working on the attribution[62] in the AI systems. So, our expectation is that in the next two to three years, we'll have really good attribution within these AI systems.

It is happening. We need to go there because there's so much data out there. We need easy search tools in that data, and the implication is also important. If you look at pharmaceuticals, medical, we're going to need that. So, I would like to vouch for [the fact that] we need attribution systems that are good. It's too important both from an economical and from a safety perspective on the security. So, I hope that we go in that direction and that everybody can make good money with that, but it shouldn't be one-side only.

Adam Rendle: Well, attribution systems pick up a number of dynamics, don't they? I didn't mention this before, but one of the things the U.K. is considering is requiring

---

[60] Quoting from Judge William Alsup's decision in *Bartz v. Anthropic PBC*, 787 F. Supp. 3d 1007 (N.D. Cal. 2025).

[61] *Development of Generative Artificial Intelligence from a Copyright Perspective*, E.U. INTELLECTUAL PROPERTY OFFICE, https://www.euipo.europa.eu/en/publications/genai-from-a-copyright-perspective-2025 (last visited Dec. 11, 2025).

[62] Technology for identifying which items of training data are most influential on a particular output of a generative AI model, and in what proportion, without having access to the AI model itself. *See, e.g.,* MUSICAL AI, https://www.wearemusical.ai/ (last visited Nov. 21, 2025); SUREEL (https://www.sureel.ai (last visited Nov. 21, 2025), BRIA (https://bria.ai/technology#attribution) (last visited Nov. 21, 2025); PRORATA https://prorata.ai/ (last visited Nov. 21, 2025.)

transparency of training data. And obviously, the EU AI Act[63] requires that as well of the general purpose models, and there's a separate code of practice about transparency.[64] And from an output perspective, some monetization models for training deals depend on use. My article or whatever, my image appeared in your output; therefore, that is a monetizable action that goes into the finances of that deal.

And we have opt-outs in the mix as well, So, to solve attribution in the middle, as you say, isn't that the same sort of thinking as we're thinking about in an opt-out? And if we're talking about attribution at the output level, you need to know where the work came from, who owns it, et cetera, to be able to credit that output and give it proper citations in the former world as well. So, is what you're saying, Daan, is opt-out is the sort of start of the conversation, but any answers we can produce for opt-out will answer a whole other set of questions that are flying around AI as well?

Daan Archer: It's a good point. Because of the opt-out, we started building a system like this, and now you can have good opt-in solutions. But you need to start aligning all those rights. You have to start somewhere. But now we're going to the next phase collectively, I would say.

Adam Rendle: Ali?

Ali Sternburg: Sure. So, yeah, there have been some efforts towards transparency in the U.S. as well. I think there are a few transparency mandates that are well-intended, but often… Okay. So, that's a great transition[65] because if we bring back formalities, then we'd have a lot more of this information and data. And I guess going back to the Copyright Office, too, a lot of this information is not available about who owns what, is it digitized, is it interoperable, versus at the PTO, there's APIs and all of it is available on tools like Google Patents.[66] And so, yeah, I think that transparency that's voluntary versus mandatory transparency, I think, can lead to some useful licensing and business activities. But some of the stakeholders have gotten ahead of it. A transparency AI-related bill passed in Colorado,[67] and now they're looking towards amending it in a special session.

And I feel like they kind of got a little bit ahead of it. And so, I think there were more than 1,000 state AI bills this term.[68] So, a few have been introduced in Congress, a lot at the federal level.[69] A lot has been introduced in the states too. And so, I think it is something that there's a variety of different AI-related issues, but a lot of them do often have to do with transparency-related considerations. Yeah, so I think transparency that's optional, great. If it's mandatory, I think sometimes there can be concerns whether it's

---

[63] EU ARTIFICIAL INTELLIGENCE ACT, https://artificialintelligenceact.eu/ (last visited Nov. 21, 2025.)

[64] Jimmy Farrell and Tekla Emborg, *An Introduction to the Code of Practice for General-Purpose AI*, EU ARTIFICIAL INTELLIGENCE ACT, https://artificialintelligenceact.eu/introduction-to-code-of-practice/ (last visited Nov. 21, 2025).

[65] A brief break was taken to reveal a CLE code word: "formality."

[66] GOOGLE PATENTS, https://patents.google.com/ (last visited Nov. 21, 2025).

[67] *SB24-205 Consumer Protections for Artificial Intelligence*, COLORADO GENERAL ASSEMBLY, https://leg.colorado.gov/bills/sb24-205 (last visited Nov. 21, 2025).

[68] *Artificial Intelligence 2025 Legislation*, NATIONAL CONFERENCE OF STATE LEGISLATURES, https://www.ncsl.org/technology-and-communication/artificial-intelligence-2025-legislation (last visited Nov. 21, 2025).

[69] *AI Legislation Tracker*, AMERICAN ACTION FORUM, https://www.americanactionforum.org/list-of-proposed-ai-bills-table/ (last visited Nov. 21, 2025).

about privacy or trade secrets, or First Amendment-protected speech. And so, I think yeah, a lot of the bills that have been proposed so far have had some problems, but I do think that there are unique challenges here that a lot of stakeholders care about getting right. And so, even if that hasn't happened yet, I think these are problems that people are trying to solve with policy, with legislation, with private sector tools, et cetera.

Thomas Sullivan: A tricky thing about attribution is you have to make sure the data coming in, the attribution data coming in, is not garbage data. I think from a rights holder's perspective, there are sites all over the internet that are taking my client's content. Not any AI companies, just random websites that are tomsnews.com. And tomsnews.com is full of articles from Politico, and they don't write anything. If tomsnews.com opts in or doesn't opt out, and all that data goes in there, that doesn't do us any good.

So, I think it needs to be a functional opt-out and a functional attribution that actually attaches to the actual people that wrote it. And I think my concern is that I don't know how we solve that for an internet that's been saturated with data or works that were not created by the person who is claiming them. I think that's a fundamental issue with opt-out, that there's all these sites that don't opt out, even though my client has opted out and actually owns it. I'm not sure how that gets solved without something that could track like the music licenses to a real individual work-by-work basis, which, when you're talking about text, is daunting.

Ali Sternburg: And I think understanding the volume. So, [regarding] Politico, I don't know how that works. If I'm meeting with people on the Hill and trying to explain it's not just the copyrighted stuff, that's good, professional. Everything is automatically copyrighted upon fixation, such as the scale of what's out there. Yeah, it is competing with all types of stuff too. But yeah, understanding just the scale of information online, unless it's facts. Obviously, there are exceptions, but yeah, there's a lot of information on the internet.

Daan Archer: If I can add to Thomas's point. Not to pull in Russia, but apparently, there's a lot of troll farms right now that are adding bad information online, knowing that AI will train on it. And that's what we use as knowledge. So, we're moving in that direction.

Ali Sternburg: Yeah. Even if they're trying to just go after, "Take this, whatever AI," that has effects beyond that too.

Daan Archer: So, how about if we could go two directions? One of them is the open web; anything is possible. Then we have verified information, because then you can have an end-to-end system. Indeed, if you want to go away from "garbage in, garbage out," so then on the "garbage in" side, you need to have it at the starting point. And it was mentioned early on the KYC. So, this is also why CISAC,[70] the umbrella organization of the music industry, the music societies, three months ago they've made it mandatory for

[70] INTERNATIONAL CONFEDERATION OF SOCIETIES OF AUTHORS AND COMPOSERS, https://www.cisac.org/ (last visited Nov. 21, 2025).

all societies in the world to implement KYC.[71] Sadly, KYC is already being hacked as well, so you can now do AI fraudulent selfies.

So, the EU now has a system in place where you can actually take [information from] the NFC chip [in] your passport.[72] We worked with the Ministry of Interior in Holland so you can connect the NFC chip off your passport to a KYC.[73] So, all that's being tackled so you can have a verified system and an open system, anything goes. And then you can connect the input with the training set and the output. So, it is possible to make a track and trace there. It's doable, and we have to because once you start applying this to medical data, pharma, and all those. By the way, to them, the music industry is very interesting because it's like, with all respect to the music industry, it's the lowest risk of legally compliant domains.

So, as soon as you start taking AI to medical and privacy, music is a great test case with respect to the music industry. Yeah, so to go away from "garbage in, garbage out," we need a verified system, and you need better data structures for this. I'm not pitching it at [Copyright] Delta, but we're building one with the societies. And more will do that. And we need to have a good track and trace in place, in my perspective.

Adam Rendle: So, I think what we're touching on is that – just to conclude before we go to Q&A – that we talk a lot about opt-out standards and their benefits or limitations, but it's difficult to focus solely on that piece of the jigsaw puzzle without thinking about the rest of the ecosystem. You've done well not to mention GRD, the global repertoire database,[74] and all the challenges that faced. You're talking about a lot of different touchpoints in the ecosystem that all depend on data, accurate data, and data that's actually attached to the works. And to get that right is really important for lots of different reasons across the whole ecosystem, as I say.

So, if we're going to try and get a workable opt-out, we can either try and sort of boil the ocean and solve everything in play, which would be amazing, or we try and do something in the meantime to improve opt-out. But in the meantime, from a policy perspective, it kind of puts a challenge to legislatures to make that policy decision we weren't going to talk about.

But the policy decision of well if everything is contingent on opt-out but opt-out is limited, is fundamentally limited, it will never really truly work to protect rights holders and to provide the right kind of accessibility and usability for tech, I think the message for policy makers from this is that you're going to have to make a policy decision based on an imperfect and inevitably imperfect set of opt-out technology.

---

[71] *CISAC Annual Report 2025*, INTERNATIONAL CONFEDERATION OF SOCIETIES OF AUTHORS AND COMPOSERS, https://iprs.org/wp-content/uploads/AG25-0354_CISAC_Annual_Report_2025_2025-05-22_EN.pdf (last visited Nov. 21, 2025).

[72] *See, e.g., Chips in: Why it's time to tap into NFC-enabled identity verification,* IDNOW, https://www.idnow.io/blog/chips-nfc-enabled-identity-verification/ (last visited Dec. 11, 2025).

[73] *See, e.g., Netherlands Coverage*, IDNOW, https://idenfo.com/netherlands-coverage/ (last visited Dec. 11, 2025).

[74] Global Repertoire Database (GRD) was a project to build a global comprehensive database of music rights and rights ownership. It began in 2010 and was abandoned in 2014. *See The Failure Of The Global Repertoire Database,* HYPEBOT, https://www.hypebot.com/hypebot/2015/08/the-failure-of-the-global-repertoire-database-effort-draft.html (last visited Nov. 21, 2025).

And that makes your policy decision way harder because you can't just rely on a totally functioning robots.txt or whatever exclusion. Very helpful to conclude on that note.

Participant 1 [Simon Locke[75]:] My name is Simon Locke, I'm the CEO of Tauth.[76] And we're actually implementing C2PA standards.So, I've contained authentication, as you I think said, that there are kind of technologies actually out there. My question is, if we talk about opt-out, it's a layer of C2PA, right? You can build that, and you're building in robust watermarks. So, it seems to solve a lot of problems, and in a way, it seems like a joined-up thinking approach to addressing a number of things at the same time. The question is, is there something that you kind of see as either the advantage of going down the C2PA route, and what are the disadvantages from this broader part of the conversation? Thank you.

Daan Archer: Absolutely. We need, at a file level, these verifications, like Canon, all the camera companies in Japan. So, if you take a photo, they're already implementing C2PA. They have proof this photo was taken by this hardware because with fake news and fake video, you don't know what's real anymore. You need to put in track and trace there. Yes, definitely. And the C2PA is a very interesting initiative. It's Adobe, Intel, Universal, all the big ones are involved. And they really want to get into the file level. Difficulties there, what happens when you make a copy and a copy and a copy? And they're tackling that as well, but it's still digital files.

Participant 2 [Gili Karev:] Hi. I'm Gili Karev. I'm a litigator at Klaris Law.[77] So, I had a question about the sort of recourse that we're looking at. I know there's this unsettled question right now about robots.txt and whether it can qualify as an access control. And under the DMCA,[78] there are various other sorts of CFAA violations that may or may not arise depending on how certain data is accessed.[79] But I guess my question is, even if we were able to implement effective opt-out measures and even if there was an effective track and control measure, what exactly is the thinking behind how violations of those measures are enforced?

Because it is so easy to ignore a robots.txt directive by programming a crawler to either not read the directive or to just circumvent it. So, it's all sort of well and nice to have methodologies by which we can control and identify infringement and violations of opt-out measures, but what's the thinking behind the legal ramifications and the incentivization structure on that end for entities to actually comply? Thanks.

---

[75] Co-Founder and CEO, Tauth Labs. *See generally* Simon Erskine Locke, *How to Protect Your Brand and Audiences from Fake Press Releases, Imposter Content and Deepfakes*, TAUTH, https://tauth.io/blog/how-to-protect-your-brand-and-audiences-from-fake-press-releases-imposter-content-and-deepfakes (last visited Nov. 21, 2025).

[76] TAUTH, https://tauth.io/ (last visited Nov. 21, 2025).

[77] Associate, Klaris Law. *See Gili Karev*, KLARIS LAW, https://www.klarislaw.com/team/gili-karev (last visited Nov. 21, 2025).

[78] In this case, 17 U.S.C. § 1201, the provision of the Digital Millennium Copyright Act (DMCA) that makes it a violation of copyright law to circumvent (crack) a content protection scheme, e.g., digital rights management (DRM).

[79] 17 U.S.C. § 1201(g) describes acts of encryption research that are exempt from the law in *id*. § 1201(g)(2)(D) states that any such acts must not violate the Computer Fraud and Abuse Act (CFAA, 18 U.S.C. § 1030). § 1201(j)(2) contains a similar provision regarding the exception for security testing.

Adam Rendle: Should I take that from sort of the U.K. political perspective and then maybe go over to you guys to the U.S.? It was in my view very revealing that in the UK consultation document, I think also supplied in the EU, there was no consideration of anything if just asked.[80] Which is to say, well, we're going to put opt-out as a sort of scale to balance things on. You've got opt-out as the fulcrum. There was no sense in which we were going to look at how you can strengthen that fulcrum to make the balance actually work.

So, I was doing a response to the government for a client, and we were filling all sorts of ideas about how you try to make the opt-out as a concept as strong and robust as possible in terms of requiring AI companies to look at, to respect, this repository idea that I mentioned. If there is an opt-out and it's ignored, does that go to an additional damages regime? I know in the U.K. and EU we don't have the kind of $150,000 worth of damages type of regime that we do in the U.S.[81] But do you introduce something like that that enhances your damages exposure if you're not properly respecting opt-outs?

Yes, we've heard a lot about how opt-out technologies are fallible, but do we still require an obligation on the scrapers, the model developers, to go and look to do the best they possibly can? There is no discussion of that whatsoever. All are possibly reliant on the fact that, well, this magical unicorn technology may exist. And I think what we've heard very clearly is that it doesn't. So, if opt-out is to be the fulcrum, then my view is that legislatures would have to work out the answers to your questions so that we can come back in a year or two, and we can ask them again, and we'll have something to say. From the US?

Thomas Sullivan: No, I mean from the perspective of the rights holders, I agree. I think if the opt-out system doesn't have any teeth, then it doesn't really work. It's a system you effectively can't opt out of, because I think once the data is in the training set, figuring out how to pull it out, you can't unmake that.

And I think the technology companies reasonably are going to say, "I can't do that," or "I won't do that because it would be unreasonable in this perspective." We had this discussion about the environmental costs of having to constantly start anew and create a new thing. We're not going to do that. But there needs to be some real, serious, beyond just "you infringed" teeth, or they're just going to blow right past it.

Ali Sternburg: Can I respond real quick? I mean, it makes me think of DMCA, where it's not mandatory – or Section 512 – to participate.[82] There's benefits and burdens on all sides, and then you're not automatically liable if you lose. So, there are still layers of secondary liability. But I think there should not be statutory damages for a lot of other cases too. But that's a whole other panel; would love to discuss that more.

Bill Rosenblatt: I want to read a question from one of our virtual attendees, which is from Geoffrey Wilson.[83] "Any thoughts about the new framework developed by Creative

---

[80] *Copyright and Artificial Intelligence*, GOVERNMENT OF THE UNITED KINGDOM, https://www.gov.uk/government/consultations/copyright-and-artificial-intelligence/copyright-and-artificial-intelligence (last visited Nov. 21, 2025).

[81] 17 U.S.C. § 504(c).

[82] 17 U.S.C. § 512 (limitations of liability for service providers that are available if they take steps to become eligible for them).

[83] IP Manager, MIT Open Learning. *See Geoff Wilson*, MIT OPEN LEARNING, https://openlearning.mit.edu/about/our-team/geoff-wilson (last visited Nov. 21, 2025).

Commons related to preference signals for training data and how this might relate to any new opt-out standards?"

Ali Sternburg: I think Creative Commons is a really interesting model.[84] For those that aren't familiar, if copyright is "all rights reserved," Creative Commons is "some rights reserved." And they also have some models that are "no rights reserved," dedicated to the public domain. [Under "some rights reserved," Creative Commons] provides tools that allow creators to say, "You can use my work, just give me attribution or only non-commercial uses."[85] And so, I think AI is something that they're looking at and then tackling too.[86] And I think they are uniquely positioned to have some kind of layer where people can express their preferences. I think they're still kind of figuring it out, but I think they are a really important organization. I use Creative Commons licenses on my own works through SoundCloud and things too.[87] So, yeah, I think they have a lot to balance, and I think there's a lot of anxiety, skepticism, just different views on AI in the public. And so, I think there's going to be a lot. I don't know how long it's going to take to find something that everyone is okay with or everyone is the least not okay with. But yeah, I think Creative Commons is a great stakeholder to have in this space too.

---

[84] CREATIVE COMMONS, https://creativecommons.org/ (last visited Nov. 21, 2025).
[85] *About CC Licenses*, CREATIVE COMMONS, https://creativecommons.org/share-your-work/cclicenses/ (last visited Nov. 21, 2025).
[86] *Understanding CC Licenses*, CREATIVE COMMONS, https://creativecommons.org/2025/05/15/understanding-cc-licenses-and-ai-training-a-legal-primer/ (last visited Nov. 21, 2025).
[87] SoundCloud allows musical artists to select Creative Commons licenses for music being submitted to the service. *See Choosing a license for your track*, SOUNDCLOUD, https://help.soundcloud.com/hc/en-us/articles/115003566468-Choosing-a-license-for-your-track (last visited Nov. 21, 2025).