



AI, Licensing and the Path Forward

**A CCC Town Hall
October 12, 2023**

with

- **Prof. Daniel Gervais**
 - **Bruce Rich**
- **Carlo Scollo Lavizarri**
- **Catherine Zaller Rowland**

**For podcast release
Monday, November 27, 2023**

ZALLER ROWLAND: Hello and welcome to CCC's town hall – AI, Licensing, and the Path Forward. I am Catie Zaller Rowland, the general counsel at CCC, and I am excited to moderate today's program, where we will review the history and significance of voluntary collective licensing and the role it could play in driving innovation in science and technology, including in AI.

To talk about AI and voluntary collective licensing, I am joined by a truly expert panel. We are so lucky to have them here today. First up, we have Daniel Gervais, PhD. He is Milton R. Underwood Chair of Law at Vanderbilt University Law School, where he serves as director of the Vanderbilt Intellectual Property Program. He is a member of the American Law Institute, or ALI, where he serves as associate reporter on the Restatement of the Law, Copyright Project. He was previously a legal officer at (inaudible), the WTO, and head of section at WIPO and previously was a vice-president here at CCC. Welcome, Professor Gervais. It's so great to have you here.

GERVAIS: Thank you. Thanks for having me.

ZALLER ROWLAND: Yes, it's so wonderful. We also have Carlo Scollo Lavizarri, who is a lawyer who works in Switzerland, South Africa, England, and Wales – so truly global – specializing in intellectual property. He has almost 20 years of experience with law firms in Africa, Europe, Switzerland, and the United States. He is a strategist and problem-solver in negotiations, conflict resolution, policy, and he participates in advocacy for norm-setting and litigation on five continents. Thank you for joining us today, Carlo.



Next up, we have Bruce Rich, who is a senior partner in the international law firm of Weil, Gotshal, and Manges, where he's headed the firm's intellectual property and media litigation for more than 30 years. During his tenure, he led a transformation of the practice to address issues presented by the advent of the internet and the emergence and explosive growth of digital commerce. Right now, Bruce serves on CCC's board of directors, and he chairs the board of directors of the nonprofit education reform organization EL Education, Inc., and he's a member of the board of advisors of Dartmouth's Rockefeller Center for Public Policy.

So as you can see, we have three very expert people here today to talk about these really important issues. And without further ado, I want to jump right into it. The first thing that we wanted to talk about today was about AI. So what is AI? And before we talk about the copyright issues, to give kind of a high-level overview. Daniel, could you briefly explain what we mean when we talk about AI in today's discussion?

GERVAIS: Of course. So the term AI is often used as a synonym for machine learning. Machine learning, as the name suggests, is a process by which AI machines are given a dataset to learn from, and this learning process can be supervised or not. Sometimes, the machines just basically learn on their own – for example, just by accessing anything they can find online.

But I think for many people on the call, the most relevant form of AI is large language models, or LLMs. That's a technology made popular, of course, by ChatGPT. Those systems can do many things, including generate new text and images and other types of content. That's why they're referred to as generative AI.

Now, LLMs typically use large datasets, but it's important to note that there can be smaller-scale large language models. I think they're actually likely to flourish in the next few years in areas such as law, but also medicine, construction, engineering, and so on.

The data that is used to create LLMs and for other forms of AI very often comprised copyrighted materials, such as texts, but also music, video, images. And the process by which LLMs work is that they break up these copyrighted works into small chunks we call tokens. A token can be anything from a syllable or word or combination of words or pixels in an image. Using this dataset of tokens, the AI, the LLM, can then answer prompts by predicting – for example, in the case of text – the next best word to answer the prompt. And it can do so repeatedly. As people choose and correct the output of the machine, then the machine learns to get better at what it does. This is why AI is already able to surpass humans at more than 100 cognitive tasks, tasks that actually require what people call high cognition, including the bar exam, for example.



ZALLER ROWLAND: Thank you. It's definitely an area where we have so many different technologies and use cases, and it's good to get that foundational discussion of generally what we're talking about for the rest of this hour. So thank you very much.

It's clear that AI is multidimensional and it has many different technological aspects. What are some ways that AI intersects with copyright? Daniel?

GERVAIS: If you look at the piece that CCC put on LinkedIn that I prepared, or my checklist of issues on generative AI and IP also on LinkedIn, you'll see that there's a long list of IP and specifically copyright issues here. But for our purposes today, I think the most important point is that the process by which AI learns involves the making of a copy of the data that they are using. Making this copy has several advantages for AI companies, from preventing delays in accessing remote data, for example, but also it allows people to check the dataset, to correct or modify the dataset later if required.

A related issue is the removal of copyright management information, or CMI, which almost necessarily happens when copyrighted material is processed by an AI. What I mean is that either the work – the copyrighted work is copied, but the CMI is ignored, and therefore effectively removed, or the work is copied without the CMI in the first place, which as I see it also amounts to the removal. This removal is a separate source of liability, by which I mean distinct from copyright infringement.

So to summarize, first, a copy of copyrighted material is made for the AI, and this copy is not transient. It's typically a copy that will remain. Second is the issue of removal of CMI.

The next question is how much of this copying is covered by an exception to exclusive rights, including fair use in the United States, fair dealing in a number of countries, including Canada, and specific exceptions adopted in various countries around the world? Now, those exceptions usually do not apply to the removal of CMI, just to the copying.

ZALLER ROWLAND: Thank you so much. You mentioned exceptions and limitations. Carlo, would you mind telling us a little bit how they might differ internationally?

SCOLLO LAVIZZARI: Yes. Thank you, Catie. Essentially, the rights involved in protecting copyright are fairly uniform throughout the world thanks to international treaties, but the exceptions are far less homogenous. Especially in a new area such as AI, there is a wide variety of different exceptions that partially overlap with the concepts of AI, such as Daniel just explained.



I think one of the best known ones is perhaps in Europe, exceptions to text and data mining, but that is not to say that there's a perfect overlap between all forms of AI, machine learning, use of training data, verification and calibration of AI machines, and the TDM exceptions. In Europe and in the UK, there is a distinction between doing text and data mining for research purposes, be they commercial or non-commercial in the UK, and commercial users doing commercial text and data mining.

In addition, in Europe and the UK, and also in Singapore, for instance, the provenance of that data embedded in copyrighted works is relevant. So the sourcing has to be from a legal source, or the access has to be lawful. That is, again, not guaranteed in all other areas where regulation of use or reuse of copyrighted works take place.

ZALLER ROWLAND: Thank you. So it seems that around the world, we have these different kind of exceptions and limitations to be careful to look at and see what is happening in the variety of places.

So right now, we have been talking a lot about the inputs, so what has gone into these AI technologies. But I know, while it's not the main focus of our talk today, that people are interested in the outputs, too. What about the outputs? What do copyright and AI works have in common when it comes to outputs, and how does copyright law treat these outputs? With that question, I'm going to ask Daniel if you might be willing to provide some insight into how copyright works and the human authorship question.

GERVAIS: Very interesting question, Catie. So the output of an LLM is technically based upon the material it has learned from. This implies that the output could be a violation of the right of reproduction and the right that is known in the United States as the right to prepare derivative works. It's usually referred to as the right of adaptation or translation in other jurisdictions. They're not exactly the same, but pretty close.

The US Copyright Act defines derivative works as works "based upon" one or more preexisting works, but courts have limited the scope of the right so that the words based upon must be read carefully. Still, there are certainly cases where infringement is possible, and I can quickly think of two scenarios.

First, let's say that the machine is given a very specific prompt. It could then copy an identifiable work in its dataset as a basis for the output. This will depend in part on the size and variability of the dataset, and it will depend in part on the quality of the algorithm, including the quality of filters that were programmed into the LLM.

So let's test the audience's knowledge of art. If I asked a machine to produce an image of a man wearing a black bowler hat, a black suit, a white shirt, and a yellow tie, with a white



dove covering his face, the machine could pick the famous painting by Magritte for its output, because this prompt is very, very specific, even if the machine has learned from a very large dataset of paintings or art.

Now, if the dataset is smaller, infringement is more likely. For example – this is my second scenario – if the machine learns only from a dataset of, say, Jackson Pollock paintings or Andy Warhol’s works, it is likely to produce something that resembles one or more preexisting works by Pollock or Warhol. That similarity could constitute an infringement, and it could infringe both the right of reproduction and the right to prepare derivative works.

The last point I’ll make is in terms of human authorship, so far, the vast majority of countries that have taken a position on this consider that there should be some human creativity in a work for copyright protection to arise, so that a machine-produced work would not be protected by copyright. However, this does not preclude the possibility of humans and a machine working together to produce some sort of copyrightable output.

ZALLER ROWLAND: Thank you. Carlo, did you have anything you wanted to add to that?

SCOLLO LAVIZZARI: Yeah. I think when we come to that second point that Daniel mentioned, where you have collaboration between machines and humans, some jurisdictions have the concept of works generated – computer-generated works. That could be one peg. But there could also be other solutions in other jurisdictions.

The one specific question of interest to any user of AI tools is if there is intellectual copyright and copyright accruing, subsisting, in any collaborative outputs, will that be owned by the, so to speak, end user of the AI tools, or will that be somehow owned by the companies that have provided the AI tool as well?

ZALLER ROWLAND: That’s a really good question and one I think we will see unfold as we move along through new technological advancements.

I want to take a minute to just welcome everyone to the program today and thank you for joining. If you do have any questions, we’ll be taking them at the end, so please put them in, and we will be sure to address them as we get to the end of the program.

So we have now kind of talked about the beginnings of our understanding and the discussion of AI. With that, I want to do a little bit of a shift and talk about another really important issue in copyright, which is voluntary collective licensing. We’re going to be focusing on what is this voluntary collective licensing? How do people license and get



permission to use works? And then we will be able to tie that together with AI and really understand kind of the symbiosis of these two different concepts.

With that, there are different approaches to how people have used copyrighted works throughout the years, and as long as there has been copyright, there have been ways to use them. Licensing is a key part of the system. With that, I wanted to ask Bruce, can you please describe the different types of licenses that we see in the copyright ecosystem?

RICH: Hi, Catie and everybody. I'd be glad to do that. Rights owners are of course free to contract directly with users one to one to create what are generally called direct licenses. Those stipulate the scope of the rights granted and the terms and conditions of those uses – in other words, again, one to one.

In addition, under US law as well as a number of foreign laws, there exist statutory licenses entitling various categories of users to gain access to and exploit in prescribed ways certain categories of copyrighted works. A couple of examples – under US law, Section 111 of the Copyright Act provides a compulsory or statutory license dealing with secondary transmissions of radio and television programs by cable and satellite systems. In turn, Section 114 of the act prescribes a statutory license dealing with the public performance of sound recordings by means of digital audio transmissions. And a third example is found in Section 115 of the Copyright Act, pertaining to the making and distributing of phonorecords for non-dramatic musical works. These provisions are actuated by perceived public policy benefits which alter the default rule that normally entitles copyright owners to exercise their own discretion whether or not to license their works and the terms under which they might do so. In practice, I should add, these statutory licenses can be quite complex both in their interpretation, in their administration, and also in their implementation.

Finally, to the ultimate point and focus of much of today's conversation, there are voluntary collective licenses, such as those offered with respect to public performances of musical works by organizations, most prominently ASCAP and BMI – many of you are familiar with those organizations – and with respect to the copying and distribution of text works by the Copyright Clearance Center. These licenses aggregate the works of many, many rightsholders into one or more forms of collective license, as the name implies. These most commonly take the form of so-called blanket or repertory licenses, which typically afford the user access to all of the works under the license for a fee that's negotiated between the user and the licensing organization. The proceeds from such licenses are then equitably distributed among the participating rightsholders based on formulas that vary, but that typically take account of the value of the respective works and the frequency with which they've been exploited by licensees.



ZALLER ROWLAND: Those are a large number – a different variety of ways to license these works and get permission. For a moment here, let’s focus on the voluntary collective licenses. How have we seen them used over the years, and how have they evolved?

RICH: They’ve evolved in response, I would argue, to market need. For example, the music license organization – the original organization, ASCAP – was formed all the way back in 1914 to bring under license random, unplanned, spontaneous, widely scattered public performances of music of the type that were later described in a notable Supreme Court ruling as resulting from “the disc jockey’s itchy finger and the bandleader’s restive baton.” It gives you an idea of, at least in that earlier era, the types of uses that it would have been cumbersome for individual rightsholders to track down, secure licenses from, enforce infringements concerning.

CCC, in turn, was formed in the late 1970s initially to address the anticipated widespread and hard-to-monitor copying enabled by then-developing photocopy technology. So the fertile soil for these licenses lies where they can offer significant transactional efficiencies both to rights owners and to users of copyrighted materials. For rightsholders, this entails situations where the costs of identifying myriad users of their copyrighted works and negotiating countless individual direct license transactions would be, frankly, prohibitive. For users of significant volumes of copyrighted works, one-stop licensing is equally transactionally efficient.

Such arrangements also can be valuable in settling disputation in still-developing areas of copyright law, particularly where fair use boundaries are uncertain. You’ve heard a preview of that in the AI setting. Rightsholders and users alike benefit when a thoughtfully conceived and reasonably priced collective license is available that can minimize costly, protracted litigation, and for users, can avoid the risk of copyright infringement.

ZALLER ROWLAND: Thank you. So it seems that it is in many different types of works and different types of use cases that voluntary collective licensing has really been at play. When we talk about them, I’d love to hear a little bit more about the different types of uses that are out there. Daniel, would you mind giving us a taste of that?

GERVAIS: I certainly agree with everything Bruce just mentioned. There are several types of collective licensing, and there are even several types of voluntary collective licensing. As Bruce mentioned, collective licensing can be based on compulsory or statutory licenses. Those are basically synonyms. They’re common in the music field.

But for text and images, not just in the United States, but elsewhere around the world, the system is by and large voluntary, which means that authors and publishers must opt into



the system, and users must sign a license, then, with an organization that represents this large group of rightsholders. As Bruce mentioned, this can be an exclusive license or a non-exclusive license, which is the case with CCC. Again, this varies by country, by area. In the case of text and images, the system really started with photocopying in the 1970s and evolved to digital uses later on and is possibly going to evolve into other areas as well in the future.

Around the world, the situation is more or less the same, in the sense that there is very old, established music collective organizations that go back, as Bruce said, to the early part of the 20th century. Some countries also have voluntary collective licensing of theatrical plays, for example, or art reproductions. So there's really no limit. It's really, as Bruce mentioned, a matter of transactional efficiency when a group of rightsholders, often a group that will be in several countries around the world, needs to license users of fairly large amounts of copyrighted material.

One thing I would mention is that even if it's voluntary, there are jurisdictions where there is a mechanism to set rates even when there's no agreement between the parties. It could be a specialized tribunal. Some countries have that. But for CCC licenses, the process is voluntary as I understand it, which again means that the user decides whether to sign the license or not, and if they do not, then the risk, of course, is litigation.

ZALLER ROWLAND: Throughout the world, we have these different scenarios. Where are some of the examples where the government has encouraged a voluntary collective licensing option? It seems that does happen. Bruce, would you mind giving us some information on that?

RICH: Sure. The principal example I'll give is CCC, but I first want to just mention, following on Daniel's comment, that even voluntary licenses in the United States in the case of the music collectives have – because of a topic we'll touch on in a bit, antitrust concerns – created mechanisms that are quasi-compulsory. That is, once a process is initiated between users or a user group and those organizations, a failure to reach agreement can trigger judicial remedies and judicial rights in the user, and indeed, even rate-setting mechanisms. So you have the whole spectrum of use.

But to your question directly, CCC's formation is a good example of government at least nudging. As I mentioned, at the time the current Copyright Act came into force in 1978, there was this roiling debate centered on the proper treatment of then brand-new photocopy technology. Education interests among others were keen on seeing the benefits of this new technology, and necessarily the content community felt that there was a serious risk to the unlicensed and unlimited reproductions of their works.



So the Senate and House reports accompanying the enactment of the new law came up with the constructive suggestion, why not create a neutral clearinghouse to facilitate what it termed workable clearance and licensing procedures pertaining to photocopying? They had in mind a voluntary organization dedicated, in their words, to working out means by which permissions for uses beyond fair uses can be obtained early, quickly, and at reasonable fees. That's what spawned CCC. It actually opened its doors the effective date of the Copyright Act – January 1, 1978.

ZALLER ROWLAND: Thank you, Bruce. What are some key takeaways we have about the benefits of voluntary collective licensing? How does that work in the entire ecosystem of copyright permissions and whatnot? Bruce, would you mind providing us a little background on that?

RICH: I'm sure Daniel and Carlo will want to add, but I will just flag two or three. One, we've each mentioned transactional efficiencies in markets that would otherwise probably fail as a practical matter in allowing copyright commerce to operate. CCC notably has been what we like to call on the board and management likes to call a solutions provider in those situations where the markets would absolutely otherwise freeze, where there would be frustration both on the licensing end and on the fair use side and making lawful commerce work.

Relatedly and finally, it reduces that kind of legal friction. Undoubtedly, the ability to have a license that meets folks in the middle, as it were, and that allows lawful transactions to occur reduces people running to court either to bring infringement suits or to argue fair use defenses in those suits or a combination of those.

ZALLER ROWLAND: Carlo, do you have anything you'd want to add to that?

SCOLLO LAVIZZARI: Yeah. I think especially in this area of fast-developing technology that the gradient of improvements of all AI systems is such that we haven't seen the end of it, it is important to respect, frankly, property rights as a system that organizes a market. Otherwise, downstream users have no certainty over quality of what goes into the AI tool. So I think the neat summary – I think it was said in a US submission to the Copyright Office of credit, consent, and compensation can easily be addressed through voluntary collective licensing in a many-to-many situation, which the AI technology and the use of copyright in training presents to the world.

ZALLER ROWLAND: Daniel, do you have anything you would like to add?

GERVAIS: Sure. You were asking about governments encouraging the formation of collective systems, and Bruce rightly mentioned the discussion surrounding the adoption of the '76



Copyright Act in the United States. But around the world, it is quite common for governments to push for the formation of collective licenses in areas where authors, publishers, and users can benefit from the transactional efficiencies there. The Nordic countries are a very good example in text, because they use a lot of English-language material that is published around the world, from the US to Australia to India and everywhere else. France just actually tabled a bill. I'm not sure it'll ever pass, because it contains so many different things. But one of its purposes would be to create a collective license for generative AI. So there is this perception in many countries that collective licensing is a solution – not a panacea, but a solution that's part of a toolbox of solutions for rightsholders and users of copyrighted material.

I think it's really important to emphasize that the one-stop-shop approach that this provides – as Bruce was mentioning, if you look at people using large amounts of copyrighted material, they're probably using material that was published, made available around the world, by hundreds or perhaps even thousands of different authors and copyright holders located in several different countries. It would be really prohibitively expensive for those users to locate and negotiate a separate license with each rightsholder. But it would also be prohibitively expensive for the rightsholders to license individual users. I think that's the reason why certain governments have pushed for collective solutions – not to mention, because we tend to forget those, the need to bridge linguistic differences, currency exchange, differences between the legal systems of each country. So there's a good set of reasons why collective licensing is part of a toolkit here.

ZALLER ROWLAND: Thank you, Daniel. Bruce, would you want to say anything more about the things you might consider when looking at voluntary collective licensing? As Daniel mentioned, it is part of the toolkit that can be used when thinking of ways to get legal access to materials for use.

RICH: I think it's important to know from long experience that while voluntary collective licenses can offer many of the advantages we've been citing, they also present potential legal risks if not managed and organized and monitored correctly. Because such licenses typically offer a set price for access to a bundle of copyrights, those arrangements can raise, and indeed have raised, both antitrust in the United States, competition law more generally, concerns overseas. The concern is over the risk that the collective licensing mechanism itself will bring about a stifling or a reduction of price and other competition between and among the various participating rights owners. So it's prudent for collective licensing organizations to engage expert counsel to help guide the development and the implementation and adaptations over time of the licenses that it offers to minimize those risks.



I would say, while it's a complex subject, as a general matter, the safest environment is one where the collective license, as Daniel mentioned, is non-exclusive, meaning that users have other options than to obtain the materials they seek, strictly speaking, through that organization. It's one among other license avenues available to the user. Examples are, as I mentioned earlier, the ability – technically, at least – to get direct licenses with rightsholders, and/or in the case of a CCC, opt for a transactional license service option, where individual transactions can occur and are priced individually by rightsholders.

ZALLER ROWLAND: Thank you, Bruce. We've had a pretty good overview right now of what the collective licensing situation is and all of the different ways that it can be effectuated throughout the world. What I want to bring together now is tie together AI, licensing and the challenges that might be there, and the opportunities. Copyright is – at least in the United States, we call it the engine of free expression, and it is there to help people build, create, and innovate. So it is good to know how they can intersect and work together. First of all, what are some of the challenges that can occur when people are trying to use copyrighted works with AI technologies? Daniel, do you mind talking a little bit about that?

GERVAIS: Yes. I will try to be brief, but it's a big question. To make the basic point, I think there are clearly three ways in which the current legal uncertainty about the scope of exceptions, including fair use, will be removed or diminished in a way that will allow machine learning, and specifically LLM technology, to progress. The three ways are legislation, court decisions, and a market-based solution.

Very quickly, national legislative initiatives can only provide guidance within the borders of one jurisdiction, obviously. Users that are relying on those will need to figure out not just the different exceptions in a different country, but in fact, deal with a patchwork of different rules if and when this happens.

If you pin your hopes on courts to find the perfect solution here, you should be very, very patient – not to mention the legal fees and the uncertainty of the outcome, including possibly very substantial statutory damages in the US. Court decisions will take years. A future Supreme Court judgment on this – we're looking at four to seven years is my best guess, and who knows if it will be the end of a process or if it will start a second cycle of people reading the tea leaves of that opinion and then trying to apply it? Add to this that even in a scenario in which a court would ultimately find that much of what AI companies have done in the US is exempt from liability for copyright infringement and CMI removal, the cross-border issue will remain. Of course, copyright holders also have to bear in mind the risk that courts will eventually shrink the scope of their rights here. So litigation, it seems to me, is far from a perfect path forward.



There's another thing that we sometimes forget – is that the AI process depends on the quality of the data that the machines learn from. To ensure that AI can really contribute to progress, we need machines learning from high-quality data. A lot of this data is content published by professionals, and those professionals are more likely to want to protect and license this content as appropriate. Put simply, all users, not just the large ones, should be able to get access to legal, high-quality data to train the machines.

The bottom line as I see it is that a market-based solution makes sense here. A license, whether individual or collective, can cover several jurisdictions. In fact, it can cover almost the entire world, because as Carlo mentioned, copyright is pretty uniform worldwide in terms of the rights. Now, there's a major reason why some copyright holders and users have already negotiated deals, and it's exactly that – that it puts an end to uncertainty on possibly a worldwide basis.

Then as I mentioned earlier, users are likely to need data from different sources in many different countries, which need to get to a solution. It matters, because if users can only use data from certain parts of the world because of differences in the legal system or the absence of licenses, this could lead to absence of diversity in the dataset, for example. So for all those reasons, I think a voluntary collective solution is a good way forward for AI companies and for copyright holders.

ZALLER ROWLAND: Thank you so much. Carlo, do you want to say anything about the international implications of this?

SCOLLO LAVIZZARI: Yeah. Ultimately, it's about a market. And like in any market, you have four basic parameters. You have who is contracting with who. What is the timeframe? Is it for past training or present training? Where is the training and the use taking place? And then what is the nature of the reproductions, and are they exempt or not? These can vary. The way the law in different countries looks at these four basic questions varies. So voluntary licensing is one way to bridge these four questions that can get in the way of having an efficient marketplace for AI.

This may even be in the interest of AI developers who might say, well, we have such a beautiful machine. Let's not look behind the curtain what went into it. We don't want to pay. But if they think further about the trust and confidence the downstream users need and the ability to perhaps reuse – if you think of a summarization tool, if a downstream user wants to, with confidence, be able to use part of the AI technology to summarize for internal use, perhaps, certain state-of-the-art knowledge or certain designs in architecture – who knows – the copyright will be able to infuse confidence and fairness and ethics into that process to have a wider distribution of the AI technology, to have more confidence in its outputs and results, and more confidence in the ability to reuse those outputs.



ZALLER ROWLAND: Thank you, Carlo. Bruce, do you want to add anything?

RICH: I would just add one other complexity in this long list. In an environment with as many uncertainties about both the direction of AI commercially as well as the many legal issues that arise, a challenge for a voluntary collective licensing organization is creating enough logic and comfort in the eyes of both sets of parties necessary to the transaction – that rightsholders will feel secure that what is to them an early era of exploitation of rights will be managed effectively, efficiently, and adequately from their standpoint through a collective license organization, and in turn, users have to be comfortable that the solution and the license fee structure and everything surrounding it are fair and appropriate. I don't think it would be proper to minimize the challenges in a market that is still evolving like this in bringing both sides together with an equitable license arrangement, although the process is extremely important to pursue.

ZALLER ROWLAND: Thank you so much, Bruce. So we have all these kind of swirling issues, and we do know that some copyright owners are licensing their work for AI uses already. Carlo, I wanted to ask you if you wouldn't mind giving us some examples of these kind of use cases that you've seen. What are the kinds of things that are already happening?

SCOLLO LAVIZZARI: Right. In all areas of copyright, there are uses developing. There is some in the music field, for instance – a company that develops scores mainly based on classical music for synchronization for film music. So you have some developers that are using licensed music.

There is a French company that gave a presentation recently in Paris called Artinity. They work with many rightsholders and estates of typically very famous visual artists and are themselves – it's part of their quality assurance that they use licensed visual art to generate derivative look art that a known creative painter might have done.

Then in the field of video and movie production, you have big, giant players such as Nvidia that produces typical explosions or matters that are kind of generated through AI technology or technology that's used in the costumes of actors, etc. CGI, computer-generated animation imagery, has long been with us.

In the literary field, I think summarization tools, perhaps in the legal publishing field, no doubt are playing a role.

And in the pharma industry, I think there are a number of success stories, particularly in the structural design of proteins, where AI tools have played and continue to play a major



role. In fact, I don't think it is possible, really, to be in pharma science and research without coming into touch with machine learning and AI at this point.

ZALLER ROWLAND: Thank you so much, Carlo. Really, this shows the diversity that's out there in the uses and what is going on.

One question I wanted to ask everyone is about exceptions and limitations to copyright, which we mentioned a little bit earlier on in the program. And the question I have is are there different perspective for looking at these for AI? Obviously, that is jurisdiction-specific. How does AI work with these exceptions and limitations? Are there specific ones? Is there less fair use in different countries? That's a big question, but I thought I would start with Daniel to take a crack at that one.

GERVAIS: Thanks, Catie. Yes, there are specific AI exceptions in several countries – Europe, Japan – well, Europe's not a country, but a group of countries – Japan, Singapore, Switzerland. In all those cases, though, I think courts will need to clarify the scope of the exception. I don't want to debate fair use, which is the main US exception, but I would say this as at least a point to consider. I've heard many times that the Second Circuit opinion in the Google Books case basically provides a full answer here. I'm not sure I agree with that, because generative AI in particular is not just about giving access to snippets of material as Google Books does. Machines are processing copyrighted material in a different way. They're extracted meaning that's embedded in the specific expression of the works that they're learning from. Actually, I would call this semantic extraction. How much of this can or should be covered by an exception like fair use, especially when the material has been used to create material that can compete with the material the machine learned from, is not a slam dunk, if you'll allow me to use this expression. But fair use is admittedly a contentious issue.

Let me put it this way. One way or the other, there will always be a limit to the exception that a national law contains. To put this perhaps slightly differently, no exception is unlimited. If an exception was unlimited, there would be no right left. Obviously, there will be some limit. In fact, international law – the TRIPS agreement managed by the World Trade Organization contains something called a three-step test which limits the flexibility of national governments about adopting exceptions to copyright law.

So overall, those exceptions and limitations are obviously an important part of the copyright framework, but they cannot do all the work here. Not to mention the delays and the uncertainties in waiting for either legislative changes or judicial changes interpreting those legislative changes.

ZALLER ROWLAND: Thank you. Carlo, do you have anything you'd like to add to that?



SCOLLO LAVIZZARI: Yeah. Outside the US and fair use, the situation is spotty. There are basically two kinds of rules seemingly evolving, one on the specific exceptions that Daniel referred to that deal with aspects of copyright and give fairly specific, limited, pigeonholed types of exceptions, and then there are more broad laws, such as transparency rules. Or in China, I think around the 15th of August or the 15th of September – I can't remember now – a new 10-point set of guidelines has come into force which again demands respect for copyright and transparency of what works have actually been used. So I think there are these two broad frameworks. And then there is most recently the French bill that Daniel also already mentioned. Right now, I think one would have to say spotty. Definitely not enough to run a significantly, high-quality AI development company based purely on the shoestring of one or two exceptions.

ZALLER ROWLAND: Thank you, Carlo. I think this is a good time to tie the two issues together a little bit about voluntary collective licensing and how can we efficiently use copyrighted work in conjunction with AI? Is voluntary collective licensing something that we should be exploring and that can help with this situation? If so, how? And I will open it up to whoever would like to answer that first. Go ahead, Bruce.

RICH: Let me just jump in quickly. I think a number of the themes here have been developed over the past 45 or 50 minutes. We have obviously a lot of fascinating questions that will keep both legal practitioners and legal scholars busy for generations to come, me excluded. I'm past that immediate phase.

But as several of us have pointed out here, you can't let markets languish, point one. Point two – while fair use issues on the US side are important, even critical, there is no AI exception to copyright law any more than there was an internet exception to copyright law. The information yearns to be free arguments, while emotionally attractive, simply don't reflect the reality that copyright is meant to be a flexible doctrine over time, with fair use being the safety valve to make sure there's the right balance between the exploitation of works that encourage authors to create and the ability to enrich society by not enlarging that monopoly beyond what it can be.

Experience has shown that voluntary collective licensing can serve a really constructive role in bridging a gap which as Daniel points out, if left unresolved, would be years and years of litigation, retarding the exploitation of markets. So done right, it appears to be a very, very productive solution – not to the exclusion of other exploitation, not to the exclusion of areas of fair use, but trying to strike that proper balance to make markets work.



ZALLER ROWLAND: Thank you so much, Bruce. Carlo, Daniel, do you have anything you'd like to add to that?

SCOLLO LAVIZZARI: Yeah, I think as much as AI has a copyright deficit, it's also true that the debate is informed by how big is the deficit that copyright holders have with AI? To the extent that there are no pragmatic solutions in the market of offering licenses widely and conveniently in situations adapted to technological change and in many-to-many situations and across geographies, that foments the call for more and more exceptions among governments. So I think this is not a sort of in vitro experiment. It's alive and developing technology. Rightsholders should also consider this aspect when deciding on their options regarding their property and their copyrighted works.

ZALLER ROWLAND: Thank you so much, Carlo and Bruce. It's been a great discussion that we've had. We've got several really good questions coming in, so I wanted to take a moment so that we can answer them. The first one I wanted to talk about is about the AI Act in the EU. We have a question about what is the status of it? What is the impact with copyright? For that one, I will turn to Carlo and see if you could provide us some advice.

SCOLLO LAVIZZARI: Yes, I think the EU under French Commissioner Breton has taken strides in coming up with an AI Act, initially not really concerned with intellectual property or copyright, but with the quality and the tiered safety aspect for letting loose AI technology on the general public in a safe and regulated manner. But as the large language models, the foundation models, have become part of everybody's consciousness in the last eight to nine months, the AI Act has clauses added to it to require foundation models to be transparent in what copyright-protected works have been or are being used to train or calibrate these models. In fact, there are broadly calls now also to more widely accept that this transparency is needed, even though fully respecting that the trade secret laws that are equally important, it is entirely possible to respond to these transparency rules, as AI developers undoubtedly need to engage in indexing and the normalization of those reproductions that are made.

So it's a rather late graft on, but perhaps similar to the GDPR, where Europe in a way has shown the way, I think in large measure, to other parts of the globe, I think here again, credit is due to Commissioner Breton and the European legislature saying we need some mechanism to have AI evolve in a human-centered manner.

ZALLER ROWLAND: Thank you, Carlo. I was actually going to piggyback on your question, because you mentioned GDPR, and there was a question in the comments about compliance risks. What do you do in terms of confidentiality of inputs, privacy infringement when uploading training datasets – if you might have a moment to address that, that would be great.



SCOLLO LAVIZZARI: Yeah, by all means, copyright aren't the only legal frameworks that AI is subject to. It hasn't been catapulted into a legal nothing. All of these are issues. I would compare this to another arcane area. Perhaps some of you – because the audience is very, very expert, I can tell – might have known or heard of the Nagoya Protocol of protection of genetic resources and provenance – very prevalent in pharma, in food, health companies. Nowadays, if you are a perfume producer – let's say, Dior – when you go to a chemical company, you will in your contracts require the chemical company to give you a statement what the provenance is of any compounds that have been included in any generated perfume.

In the same way, I think more and more of these contracts between downstream users of AI tools and AI producers, developers, will have to have compliance statements that there was no sensitive personal information used without consent, that any clinical information has been anonymized, that copyrighted works have in fact been licensed, or otherwise it's a non-infringing development exercise. I think we will see more and more compliance contracts between developers and downstream users. That, I think, will also be an added pressure to arrive at collective licensing models upstream.

ZALLER ROWLAND: Thank you, Carlo. We have another question that implicates open source and open access work. That is the question of what is the situation with using these open access works, which is a very complex issue. But I will turn to Daniel to address that in the first instance.

GERVAIS: Yeah, very briefly, two things. First, there's a misconception that everything that is available online is free of copyright somehow. In certain discussions, at least, I've heard that. That is obviously not the case. So there's copyright protection in material that's found online. There are terms of use. There's a license. Perhaps there's some legal exception that allows some use, but it doesn't mean that it is free of copyright.

The second about open source is a similar misconception – that there is no copyright on open source. It's actually exactly the opposite. The reason that it can be open source is precisely that there is copyright, and because there is copyright, when you license open source, you can require the user to share what they do with the open source material. That's possible because there is copyright. If there was no copyright, if this was a public domain work, then anybody could appropriate it, make something with it, and then keep it as their copyrighted work with no obligation to share. So these two misconceptions really need to be dispelled.

ZALLER ROWLAND: Thank you, Daniel. We have, I think, time for one final question, and unfortunately the time to answer the question is not enough to handle the incredible



breadth of it. But the question is about the cases. There are cases that are out there in the United States dealing with copyright and AI. I was just counting them up last week. We're up to nine that deal with just the input part, not to mention the human authorship aspect of things. In two minutes, Daniel, can you give a rundown of the general takeaway from the cases and what the situation is? These are all US cases, but also there was a companion to one of them in the UK.

GERVAIS: Sure. Well, on human authorship, the Copyright Office of the United States and the District Court in Washington both said that you need human authorship, but I believe the notice of appeal was filed today. So we'll hear from the Court of Appeals on that one. I hope that they will confirm the ruling from the lower court.

But in terms of infringement, most of the cases are class-actions, so there'll be debates about the class. In terms of copyright law, the claims are generally about violations of the right of reproduction in taking the works, the copyrighted material of the plaintiffs, to train AI. And then there are additional claims in most of these lawsuits about copyright management information being removed, sometimes about the output – so some of the lawsuits also say the output of the generative AI system was infringing. And then a whole bunch of typically state law claims – anything from unfair competition to unjust enrichment. There's many of them. I have a long list. But it'll be interesting to see what the courts make of fair use in that context. That will take several years. We may end up with different appellate opinions in different circuits. So that's why I'm saying four to seven years before the Supreme Court issues its first opinion, which may not be its last on this, is maybe even optimistic.

ZALLER ROWLAND: Thank you so much, Daniel. So we've had a really, really exciting conversation today that's been really fascinating. I think I have learned things, and I hope that everyone in the audience has taken away some good points about AI and the relationship with copyright. With that, thank you for coming, and have a great rest of your day.

END OF FILE