A

WIPO
المنظمة العالمية
للملكية الفكرية

SCCR/49/6
الأصل: بالإنكليزية
التاريخ: 23 أكتوبر 2026

اللجنة الدائمة المعنية بحق المؤلف والحقوق المجاورة

الدورة التاسعة والأربعون
جنيف، من 30 نوفمبر إلى 4 ديسمبر 2026

دراسة نطاق حول تدريب الذكاء الاصطناعي وحق المؤلف

من إعداد دانيال ج. جيرفيه، وسونغ جاي تشوي، وميليسا أومينو، ورودريغو كوبكي سالينا‌س، ووانغ تشيان

دراسة نطاق حول تدريب الذكاء الاصطناعي وحق المؤلف

إخلاء مسؤولية:

أُعدّت هذه الدراسة بتكليف من أمانة المنظمة العالمية للملكية الفكرية (الويبو). الآراء الواردة في هذه الوثيقة تعبر عن آراء المؤلفين ولا يُقصد بها أن تعكس آراء الدول الأعضاء أو الأمانة.

شكر وتقدير

دانييل ج جيرفاي المؤلف المشارك الرئيسي لهذه الدراسة. وهو أستاذ القانون الذي يشغل كرسي ميلتون ر أندروود (فخري) في كلية الحقوق بجامعة فاندربيلت وأستاذ مشارك أول في جامعة فريجي بأمستردام، وعضو منتخب في معهد القانون الأمريكي. يمكن النفاذ إلى أبحاثه عبر www.collectivemanagement.net، وwww.tripsagreement.com، وwww.thehumancause,.org.

سونغ جاي تشوي أستاذ القانون في جامعة سيجونغ. حاصل على درجة الماجستير في القانون من كلية الحقوق بجامعة كولومبيا. تركز أبحاثه على الملكية الفكرية وقانون المنافسة. وهو مؤلف مشارك لقانون حق المؤلف (2024)، وحصل على الجائزة الكبرى للتحكيم الكوري لعام 2021 بصفته أحد قادة الجيل القادم.

ميليسا أومينو مديرة مركز قانون الملكية الفكرية وتكنولوجيا المعلومات (CIPIT) بجامعة ستراتمور في نيروبي. وهي حاصلة على درجة الدكتوراه في قانون الملكية الفكرية والقانون التجاري، وتتخصص في مجالات الملكية الفكرية والذكاء الاصطناعي وحماية البيانات والحقوق الرقمية في بلدان الجنوب العالمي.

رودريغو كوبكي ساليناس أستاذ في كلية الحقوق التابعة لمؤسسة خيتوليو بارغاس في ساو باولو (GVlaw)، وحاصل على درجة الماجستير في القانون من جامعة كولومبيا. وهو عضو في المجلس الخاص بحق المؤلف التابع لنقابة المحامين البرازيلية، ويتخصص في قانون الإعلام والترفيه.

وانغ تشيان أستاذ القانون ونائب رئيس اللجنة الأكاديمية في جامعة شرق الصين للعلوم السياسية والقانون. وهو أيضاً نائب رئيس جمعية حق المؤلف في الصين، وقد شغل منصب عضو في لجنة الخبراء المعنية التنقيح الثالث لقانون حق المؤلف في الصين.

ملاحظة: استُخدمت نماذج لغوية لمراجعة بُنية الدراسة وقواعدها اللغوية ومحتواها، فضلاً عن تقديم اقتراحات تحريرية بهدف توحيد نص الدراسة الذي شارك في إعداده عدة مؤلفين.

جدول المحتويات

مقدمة 4
الجزء الأول: استخدام المحتوى المحمي في تطوير نماذج الذكاء الاصطناعي التوليدي 5
القسم ألف: الحصول على البيانات 5
1. أنماط استهلاك البيانات حسب نوع النظام 5
2. استخراج البيانات من الويب 5
3. التوليد المعزّز بالاسترجاع 6
القسم باء: استخدام البيانات في التدريب 7
1. تدريب النموذج 7
2. التحسين والتعلم المعزز 9
القسم جيم: بعد انتهاء التدريب 9
1. استخدامات أخرى للذخيرة اللغوية الخاصة بالتدريب 9
2. الاحتفاظ بنسخ في النموذج 9
3. علاقة المدخلات والمخرجات 12
الجزء الثاني: التفويض والترخيص والتعويض 13
القسم ألف: الخيارات المتاحة حالياً للمطورين للحصول على مواد مؤهلة للحماية بموجب حق المؤلف لأغراض التدريب 13
1. مسارات الاستحواذ الخمسة 13
2. الترتيبات التعاونية 14
القسم باء: العناصر الرئيسية للسياسات 15
1. الحقوق المتأثرة بتدريب نماذج الذكاء الاصطناعي التوليدي 15
(أ) حق النسخ 15
(ب) حق التكييف والمصنفات المشتقة 16
(ج) التبليغ للجمهور والإتاحة 16
(د) الحقوق المعنوية 16
(هـ) RMI وتدابير مناهضة التحايل 17
2. الترخيص الطوعي (الجماعي والفردي) 17
(أ) الترخيص الفردي والجماعي لاستخدامات الذكاء الاصطناعي 17
(ب) النماذج الرئيسية للترخيص الجماعي (الطوعي، والموسع، والإلزامي) 21
3. الترخيص المجتمعي والترخيص القائم على الإشراف للبيانات اللغوية الأولية 24
4. الاستثناءات القانونية لتنقيب النصوص والبيانات والأحكام العامة 25
5. أنظمة المكافآت 28
(أ) أنواع أنظمة المكافآت 28
(ب) طرق وأسس المكافآت 29
الجزء الثالث: الإنفاذ 30
القسم ألف: الدروس المستفادة من السوابق القضائية الحالية 31
1. مسألة النموذج كنسخة 31
2. استثناء تنقيب النصوص والبيانات 33
3. الاختصاص القضائي على التدريب خارج الإقليم 34

4. مسألة الاستبدال 35
القسم باء: إنفاذ الحقوق ضد مطوري الذكاء الاصطناعي..... 38
1. الخصائص الهيكلية لبيئة الإنفاذ..... 38
2. أوجه عدم التماثل بين الأطراف 39
القسم جيم: التحديات في إنفاذ الحقوق 39
1. الشفافية 39
2. الهجمات في استخراج البيانات، والتصفية، وتعدد معايير الاعتراض على الاستخدام..... 40
3. السوابق القضائية في مجملها..... 43
الختام..... 44

مقدمة

أُعدّت هذه الدراسة بناءً على طلب اللجنة الدائمة المعنية بحق المؤلف والحقوق المجاورة. فقد انعقدت اللجنة الدائمة في دورتها السابعة والأربعين (جنيف، الفترة من 1 إلى 5 ديسمبر 2025)، وطلبت من الأمانة إعداد دراسة حول نُهُج السياسة أو النُهُج التنظيمية المتعلقة بالعلاقة بين تدريب الذكاء الاصطناعي وحق المؤلف، والقواعد ذات الصلة، وكذلك حول الممارسات المطبَّقة بشأن الإذن والإنفاذ والتعويض عن الاستخدام. وقد سُجِّل هذا الطلب في الملخص الذي أعده رئيس تلك الدورة (الوثيقة SCCR/47/SUMMARY[1] بتاريخ 5 ديسمبر 2025، الفقرة 21).

وتتسم الدراسة بطابع وصفي ومقارن. وهب تستعرض الأساليب المعتمدة في عدد من الولايات القضائية والممارسات الآخذة في التطور، وتناقش الخيارات المقترحة دون أن توصي الدول الأعضاء باعتماد أي نهج محدد، أو أن تتخذ موقفاً من المسائل الخلافية التي لم تحسمها هذه الدول بعد.

وتتناول الدراسة تدريب نماذج الذكاء الاصطناعي التوليدي، أي استخدام المحتوى لبناء هذه النماذج وتحسينها، والمسائل المتعلقة بالإذن والإنفاذ والتعويض التي يثيرها هذا الاستخدام عندما يكون المحتوى محمياً بحق المؤلف أو الحقوق المجاورة. بيد أن التقاطع الأوسع بين الذكاء الاصطناعي وحق المؤلف يتجاوز بكثير مسألة التدريب. وهي تتناول على وجه الخصوص الوضع القانوني لنواتج نموذج منشور، سواء من حيث ما إذا كان أي ناتج معين يتعدى على المصنفات التي دُرِّب عليها النموذج، أو من حيث ما إذا كان الناتج ذاته قابلاً للحماية بموجب حق المؤلف، ووفق أي شروط. وتثير الأسئلة المتعلقة بالناتج مسائل أخرى لم تطلب اللجنة الدائمة من هذه الدراسة تناولها. ولا يُشار إليها إلا عندما تقتضي الضرورة تقديم صورة كاملة عن مسألة التدريب أو لتوضيح السياق اللازم.

وتُعرض الدراسة في ثلاثة أجزاء. يصف الجزء الأول، على مستوى تقني ووصفي، العمليات التي يُستخدم من خلالها المحتوى المحمي في تطوير نموذج ذكاء اصطناعي توليدي، ويحدد النقاط التي قد تمسّ فيها تلك العمليات الحقوق الاستئثارية المعترف بها في التشريع الوطني لحق المؤلف ومعاهدات حق المؤلف الدولية. ويتناول الجزء الثاني نطاق الحقوق المعنية ومداها، والآليات التي تُستخدم من خلالها هذه الحقوق، وهي: الترخيص الطوعي، فردي كان أو جماعي؛ والاستثناءات والتقييدات المنطبقة. ويتناول الجزء الثالث مسألة الإنفاذ ويستكشف الدروس التي يمكن استخلاصها من مجموعة السوابق القضائية التي بدأت تتشكل في عدد من الولايات القضائية. كما يتناول الطابع العابر للحدود للتدريب ومسائل الإنفاذ ذات الصلة، بما في ذلك الشفافية ومعايير التعبير عن الانسحاب حيثما انطبق ذلك. والولايات القضائية المذكورة في هذه الدراسة تذكر على سبيل التوضيح وليست شاملة. وتشمل هذه الدول البرازيل، وكندا، والصين، والاتحاد الأوروبي والعديد من دوله الأعضاء، وكينيا، وجمهورية كوريا، والمملكة المتحدة، والولايات المتحدة.

[1] انظر(ي) https://www.wipo.int/meetings/ar/doc_details.jsp?doc_id=653196

الجزء الأول: استخدام المحتوى المحمي في تطوير نماذج الذكاء الاصطناعي التوليدي

نموذج الذكاء الاصطناعي التوليدي يُطوَّر على مراحل متعاقبة، وعادةً ما يتم استخدام محتوى من جهات خارجية في كل مرحلة من هذه المراحل. وقد يخضع بعض المحتوى لحق المؤلف والحقوق المجاورة؛ وقد يكون بعضه الآخر ضمن الملك العام، إما لانقضاء مدة الحماية أو لأنه لم يكن محمياً من الأساس. والعمليات التقنية التي تُطبق على المحتوى هي عموماً نفسها بغض النظر عن وضعه القانوني. وقد تُطبَّق العمليات ذاتها بالتساوي، بما في ذلك عمليات الزحف والنسخ والترميز والتحليل الإحصائي التي سيرد شرحها أدناه، سواءً على رواية حديثة أو على نص من القرن الثامن عشر في الملك العام. ويصف هذا الجزء هذه العمليات ويبيّن، على مستوى وصفي، النقاط التي قد تمسّ فيها الحقوقَ المعترف بها في التشريع الوطني لحق المؤلف وفي المعاهدات الدولية لحق المؤلف والحقوق المجاورة.

ويمكن التمييز بين ثلاث مراحل من التدريب، وتُعرض تباعاً. يناقش القسم ألف مسألة الحصول على البيانات، أي كيفية قيام مطوِّر نموذج الذكاء الاصطناعي ببناء الذخيرة اللغوية، التي تتألف في جزء كبير منها من موضوعات قابلة للحماية، والتي ستُستخدم لتدريب النموذج. ويغطي القسم باء استخدام تلك البيانات في التدريب، بما في ذلك العملية التدريبية الرئيسية وأنواع الضبط الدقيق المختلفة التي تليها. ويتناول القسم جيم الوضع بعد انتهاء التدريب، بما في ذلك التعامل مع الذخيرة اللغوية التدريبية، ومسألة ما إذا كان النموذج المدرب ذاته يحتوي على نسخ من المصنفات[2] التي دُرِّب عليها.

القسم ألف: الحصول على البيانات

يبدأ تطوير نموذج لغوي كبير بإنشاء ذخيرة لغوية، تكون عادةً على نطاق واسع جداً، ومن مصادر متنوعة. ويحصل المطورون على بيانات التدريب بعدة طرق. قد تكون بعض البيانات مرخصة، وفي هذه الحالة قد يُتاح المحتوى مقترناً بترخيص استخدامه. في مثل هذا السيناريو، يمكن للمطوّر شراء حق النفاذ إلى مجموعة بيانات منتقاه بعناية تم تجميعها من قبل جهة خارجية، أو التفاوض على اتفاقية ترخيص مباشرة مع ناشر أو مؤسسة إخبارية أو منظمة إدارة جماعية. وإذا لم يتم منح ترخيص لاستخدام المصنفات القابلة للحماية، فيجوز أن يستند استخدامها إلى افتراض وجود استثناء أو قيد ينطبق على الحق الاستئثاري ذي الصلة. ويتمثل أحد المسارات الرئيسية للتدريب، من حيث حجم البيانات والمسائل المتعلقة بحق المؤلف التي يثيرها، في الاسترجاع المؤتمت للمواد من الويب المفتوح (الإنترنت). وقد يستخدم التدريب أيضاً البيانات الاصطناعية، أي البيانات التي يولّدها نموذج ذكاء اصطناعي حالي بدلاً من جمعها من مبدعين بشريين، لا سيما في المراحل المتأخرة من التطوير.

1. أنماط استهلاك البيانات حسب نوع النظام

لا تقتصر حاجة النظام إلى البيانات عادةً على عملية التدريب الأصلية: إذ تُستخدم البيانات في إعادة التدريب والضبط الدقيق، وفي تحليل النصوص والبيانات الذي يُجرى على مكتبة بحثية مُحتفظ بها، وفي التوليد المعزّز بالاسترجاع (RAG[3]، ولتوليد المعاينات والمقتطفات. وتختلف هذه الاستخدامات من ناحية كمية المواد التي تتطلبها ومن ناحية سهولة تتبع مصادرها. ويميل نوع النظام إلى التأثير على أنماط الاستهلاك: فأنظمة توليد المحتوى الإعلامي وروبوتات المحادثة التي تعتمد على نماذج لغوية كبيرة تستهلك البيانات في معظمها من خلال الاستيعاب التدريبي؛ أما الأنظمة القابلة للتخصيص التي تطوّرها المؤسسات بالاستناد إلى نموذج لغوي (أو أساسي)، فتضيف ذخيرة لغوية خاصة بمجال معين، وغالباً ما يكون لديها مكتبة بحثية يتم الرجوع إليها عند الاستخدام؛ في حين تتطلب محركات البحث الحوارية إمكانية النفاذ إلى المصادر المتاحة عبر الإنترنت في وقت التوليد. ويؤثر المكان الذي يقع فيه نظام ما ضمن هذا النطاق على مجموعة الحقوق الاستئثارية التي ينطوي عليها استخدامه للبيانات، وعلى الهيكل المقابل لأي ترخيص. وسيتم تناول هذا الموضوع بالتفصيل في الجزء الثاني، القسم باء.2.

2. استخراج البيانات من الويب

يُعد استخراج البيانات من الويب نهجاً شائعاً يستخدمه مطورو النماذج الأساسية لبناء مجموعات البيانات التدريبية الخاصة بهم. غير أن هذه المصطلحات لم تُعتمد عالمياً. فبعض المصادر تميّز بين "الزحف على الويب" (التنقل المؤتمت عبر مواقع الويب بواسطة برنامج يتتبع الروابط من صفحة إلى أخرى) و"استخراج البيانات من الويب" (الاستخراج الانتقائي لمواد محددة من صفحة سبق استرجاعها). وفي الممارسة العملية، قد يتم الجمع بين هذين النشاطين: حيث ينشر مطور برنامج زحف يتتبع الروابط عبر الويب، ويحدد في كل صفحة المحتوى النصي أو المرئي أو أي محتوى آخر يُراد الاحتفاظ به ويستخرجه. وغالباً ما يكون الناتج نسخة من هذا المحتوى، ويتم تخزينه عادةً بطريقة منظمة على البنية التحتية الخاصة بالمطوّر.

---

[2] ولتبسيط النص، ينبغي فهم مصطلح "المصنفات" على أنه يشمل موضوعات الحقوق المجاورة، حيثما كان ذلك مناسبًا.

[3] يُعرف التوليد المعزز بالاسترجاع، الذي يُختصر عادةً بعبارة RAG، بمزيد من التفصيل في القسم 3 من هذا الجزء.

وهناك أربع سمات ذات صلة بالمناقشة الحالية. أولاً، تنطوي هذه التقنية على إجراء نُسخ. قد يتم الاحتفاظ بالنسخ على خادم أو في ذاكرة تخزين مؤقت، أو معالجتها في الذاكرة قبل إجراء أي تحويل آخر. قد تنطوي هذه العمليات على حق النسخ على النحو المعترف به في المادة 9(1) من اتفاقية برن، وفي المادة 1(4) من معاهدة الويبو بشأن حق المؤلف، وفي الأحكام المقابلة في التشريعات الوطنية. يؤكد البيان المتفق عليه بشأن المادة 1(4) من معاهدة الويبو بشأن حق المؤلف أن تخزين مصنف محمي في شكل رقمي يُعد نسخاً، وأن حق النسخ والاستثناءات منه ينطبقان بشكل كامل في البيئة الرقمية.[4] أما ما إذا كان الطابع المؤقت لبعض النُسخ، أو انطباق استثناء مثل ذلك المتعلق بالنسخ العرضي والعابر، يغيّر من التحليل، فهذه مسألة سيتم تناولها لاحقاً في هذه الدراسة. وتتضمن المادة 7 من معاهدة الويبو بشأن الأداء والتسجيل الصوتي، والمادة نفسها في معاهدة بيجين بشأن الأداء السمعي البصري، حقاً يتعلق بالاستنساخ المباشر أو غير المباشر "بأي طريقة أو بأي شكل كان".

وثانياً، قد يكون الاسترجاع غير انتقائي. فبرنامج الزحف لا يميّز دائماً بين المحتوى الموجود في الملك العام، والمحتوى المتاح بموجب ترخيص مفتوح، والمحتوى المتاح دون أي بيان ترخيص على الإطلاق، والمحتوى المحمي المتاح وفقاً لشروط الخدمة أو محمياً بأيٍّ من تدابير الحماية التكنولوجية. على سبيل المثال، تلتزم بعض برامج الزحف، وليس جميعها، بتعليمات ملف robots.txt.[5] ويُجرى هذا النوع من التمييز أحياناً بعد وقوع الفعل، من خلال تصفية الذخيرة اللغوية أو إزالة المتكرر منها، لكن فعل النسخ الأولي لا يكون دائماً انتقائياً.

وثالثاً، إمكانية إطلاع أصحاب الحقوق على هذا النشاط محدودة. فقد لا يعلم صاحب حقوق شيئاً عن أن مصنفه قد تم استخراجه، أو عن برنامج الزحف الذي تولى الأمر، أو وقت حدوث ذلك، ولا عن الاستخدام اللاحق لهذا المصنف. وتتطلب التزامات الكشف التي بدأت تظهر في بعض القوانين الوطنية، ولا سيما في الاتحاد الأوروبي بموجب قانون الذكاء الاصطناعي لعام 2024، من مقدمي نماذج الذكاء الاصطناعي ذات الأغراض العامة نشر ملخص "مفصل بما فيه الكفاية" للمحتوى المستخدم في التدريب.[6] لا يزال من المبكر تقييم الأثر العملي لتلك الالتزامات. إن مسألة كيفية التوفيق بين الكشف واحترام المصالح التجارية المشروعة للمطورين هي في حد ذاتها موضوع نقاش مستمر.[7]

رابعاً، قد ينطوي إنتاج واستخدام نسخ من الأعمال المحمية لأغراض التدريب على إزالة أو تغيير معلومات إدارة الحقوق، والتحايل على التدابير التكنولوجية الفعالة التي يستخدمها أصحاب الحقوق في سياق ممارسة حقوقهم بموجب معاهدة الويبو بشأن حق المؤلف، ومعاهدة الويبو بشأن الأداء والتسجيل الصوتي، ومعاهدة بيجين.

3. التوليد المعزّز بالاسترجاع

تقنية منفصلة لكنها ذات صلة تُعرف باسم "التوليد المعزَّز بالاسترجاع" (RAG).[8] وهي طريقة يستخدمها نظام ذكاء اصطناعي للإجابة عن استفسار ما من خلال الرجوع أولاً إلى مجموعة منفصلة من الوثائق تم تجميعها مسبقاً، بدلاً من الاعتماد فقط على المواد التي تم تدريبه عليها. ومن المقارنات المفيدة على ذلك الفرق بين شخص يجيب عن سؤال من ذاكرته، والشخص ذاته حين يتوقف لبرهة ليبحث عن شيء ما في رفّ مراجع موضوع بجوار مكتبه. وعند ورود استفسار ما، يبحث النظام في تلك المجموعة، ويستخرج المقاطع الأكثر صلة بالسؤال، ويضعها بجانب الاستفسار، ويستخدمها لصياغة إجابته. ويتم إنشاء المجموعة مسبقاً

---

[4] معاهدة الويبو بشأن حق المؤلف (20 ديسمبر 1996)، البيان المتفق عليه بشأن المادة 1(4): "ينطبق حق النسخ، على النحو المنصوص عليه في المادة 9 من اتفاقية برن، والاستثناءات المسموح بها بموجبها، تطبيقاً كاملاً في البيئة الرقمية، ولا سيما على استخدام المصنفات في شكل رقمي. ومن المفهوم أن تخزين مصنف محمي رقمي الشكل في وسيط إلكتروني يُشكِّل استنساخاً بالمعنى الوارد في المادة 9 من اتفاقية برن".

[5] بروتوكول استبعاد الروبوتات، الموحد بموجب الوثيقة IETF RFC 9309 (2022)، هو بمثابة اتفاقية يضع بموجبها أحد مشغلي مواقع الويب ملفاً نصياً عادياً (robots.txt) في جذر موقع ما للإشارة إلى أي برامج زحف مؤتمتة يُسمح لها بالنفاذ إلى أجزاء معينة من الموقع. الالتزام بهذا البروتوكول طوعي؛ ولا يتضمن البروتوكول أي آلية لإنفاذه.

[6] اللائحة (الاتحاد الأوروبي) رقم 1689/2024 الصادرة عن البرلمان الأوروبي ومجلس الاتحاد الأوروبي بتاريخ 13 يونيو 2024، التي تحدد قواعد موحدة بشأن الذكاء الاصطناعي (قانون الذكاء الاصطناعي)، المادة 53(1)(د) (التزام مقدمي نماذج الذكاء الاصطناعي ذات الأغراض العامة "بإعداد وإتاحة ملخص مفصل بما فيه الكفاية للجمهور حول المحتوى المستخدم لتدريب نموذج الذكاء الاصطناعي ذي الأغراض العامة، وفقاً لنموذج يوفره مكتب الذكاء الاصطناعي").

[7] ومع ذلك يظهر في بعض الأحيان أصل المواد المستخرجة في سلوك النموذج الناتج: فقد أعادت الإصدارات المبكرة من أحد نماذج توليد الصور البارزة إنتاج شكل يمكن التعرف عليه لعلامة مائية لوكالة للصور الفوتوغرافية الجاهزة، وهو ما يشير إلى أنه تم استيعاب كمية كبيرة من صور تلك الوكالة دون إذن. وقد حظيت هذه الظاهرة بتغطية واسعة، كما تم تناولها في الدعوى القضائية التي نوقشت في الجزء الثالث، القسم ألف.1 (Getty Images (US) Inc. ضد Stability AI Ltd.). ويوضح ذلك أنه، على الرغم من أن عملية استخراج البيانات من مصنف معين تكون عادةً غير مرئية لصاحب الحقوق، فإن النموذج المدرب قد يكشف هو نفسه، من خلال نواتجه، عن استخدام مواد محمية في عملية التدريب.

[8] انظر(ي) الملاحظة 2 أعلاه.

من خلال إنشاء نسخ من الوثائق وتخزينها، وغالباً ما يتم نسخ المقاطع ذات الصلة مرة أخرى إلى النظام في اللحظة التي يتم فيها توليد الإجابة، وهي نقاط مهمة بالنسبة لتحليل حق المؤلف المبيّن أدناه.

وفي نظام RAG، يتم الاحتفاظ بذخيرة لغوية خارجية، تُفهرس عادةً في قاعدة بيانات بعد تحويل كل وثيقة أو جزء من وثيقة إلى ما يُعرف "بالتضمين"[9]. وعندما يتلقى النظام استفساراً من مستخدم، يتم تضمينه واستخدامه لاستخراج الأجزاء الأكثر صلة من الذخيرة اللغوية. ثم يجمع النظام المطالبة التي يتم إدخالها إلى نموذج اللغة، دامجاً استفسار المستخدم مع الأجزاء المسترجعة، والتي يستخدمها النموذج لتوليد الناتج بناءً على تدريبه والمواد المسترجعة.

ولأغراض هذه الدراسة، يُوصف التوليد المعزّز بالاسترجاع أحياناً بأنه يُغير طبيعة المسائل المتعلقة بحق المؤلف التي تنشأ عن تدريب نماذج اللغة. والسبب في ذلك هو أن النموذج ذاته لا يتعلم من الذخيرة اللغوية، لأن هذه الذخيرة اللغوية تُحفظ بشكل منفصل ولا يُرجع إليها إلا في وقت الاستدلال. ويتطلب إنشاء الذخيرة اللغوية المفهرسة معالجة مسبقة، وفي بنية تقليدية لتوليد معزّز بالاسترجاع، يتطلب الأمر تخزين وثائق المصدر أو الأجزاء المكونة لها.[10] ويؤدي تحويل كل وثيقة إلى "تضمين" إلى إنتاج تمثيل رقمي مشتق من الوثيقة، على الرغم من أن التضمين، على عكس النسخة المصدر المخزّنة، لا يمكن عادةً إعادة تحويله إلى النص الأصلي. وفي مرحلة التشغيل (الاستدلال)، تُنسخ الأجزاء ذات الصلة إلى نافذة سياق النموذج، وقد تُنسخ كلياً أو في جزء جوهري منها، في الناتج. وتمسّ هذه الخطوات حق النسخ؛ وقد تمس حق التكييف (المادة 12 من اتفاقية برن) وفقًا للقانون المعمول به وخصائص العرض. ويُناقش لاحقاً إمكانية انطباق هذه الحقوق.

القسم باء: استخدام البيانات في التدريب

يتطلب تدريب نموذج الذكاء الاصطناعي التوليدي إجراء سلسلة من العمليات التي تُوصف عادةً بأنها تتألف من أربع مراحل: التدريب المسبق، والضبط الدقيق تحت الإشراف، ونمذجة المكافأة، والتعلم المعزز.[11] وتُعد المرحلة الأولى من هذه المراحل بشكل عام المرحلة الأكثر تطلباً، سواء من حيث التكلفة الحاسوبية أو من حيث حجم البيانات المستهلكة؛ أما المراحل الثلاث اللاحقة، فتوصف مجتمعةً في أغلب الأحيان بأنها "الضبط الدقيق" أو، في بعض الأحيان، بأنها "مرحلة ما بعد التدريب". ويصاحب التدريبَ نشاطٌ منفصل لكنه وثيق الصلة به، وهو تقييم النموذج وتحسينه، ويستمر في مرحلة نشر النموذج.

ولأغراض حق المؤلف، تتمثل الملاحظة الرئيسية في أن كل مرحلة من هذه المراحل تنطوي عادةً على تخزين الذخيرة اللغوية الأساسية بأكملها أو جزء منها، ومعالجتها. ولا يتم في مسار التدريب الاطلاع على البيانات بصورة عابرة ثم التخلص منها؛ بل يتم تخزين البيانات وتحميلها وتحويلها واستخدامها لتحديث النموذج. وحيث تتضمن البيانات مصنفات محمية، فإن تلك العمليات تنطوي على نسخ أو تحويلات ذات صلة بتحليل النسخ الذي تم تناوله في الجزء الثاني.

1. تدريب النموذج

تُعد عملية التدريب الرئيسية، التي تُعرف عموماً باسم "التدريب المسبق"، المرحلة التي يكتسب فيها النموذج الجزء الأكبر من معرفته باللغة (في حالة نموذج لغوي) أو بالصور والصوت والفيديو (في حالة نموذج متعدد الوسائط). ويجري التدريب المسبق

---

[9] وتمثّل تضمينات الكلمات كل رمز بمتجه كثيف عالي الأبعاد من القيم الحقيقية. وتُوصف التضمينات بأنها "سياقية" عندما يتوقف التمثيل المسند إلى رمز معين على الرموز المحيطة به. انظر(ي) جوهر المسألة، الحاشية 10 أعلاه، الجزء الثاني، جيم.

[10] من الناحية التقنية، ينبغي التمييز بين تضمينات المقاطع المستخدمة في البحث المتجه والنص المصدر الذي تُولَّد منه. يعدُّ التضمين تمثيلاً رقمياً كثيفاً ينتجه جهاز تشفير؛ وهو ليس نسخة مطابقة من النص، ولا يمكن عادةً عكسه بمجرد فك تشفير المتجه. ومع ذلك، يمكن لتقنيات "انعكاس التضمين" المتخصصة أن تعيد بناء بعض مقاطع المصدر القصيرة نسبياً، أو بناء جميعها في ظروف تجريبية معيّنة، انطلاقاً من تضميناتها. وفي نظام توليد معزّز بالاسترجاع التقليدي، يتم الاحتفاظ أيضاً بوثائق أو أجزاء المصدر في شكل نص، إما في مخزن المتجهات ذاته أو في مخزن وثائق منفصل مرتبط بالمتجهات. تُستخدم التضمينات لتحديد المقاطع ذات الصلة؛ ثم يُسترجَع النص المرتبط بها، وليس متجه التضمين، وتزويد نموذج اللغة به كجزء من سياق الإدخال الخاص به. انظر(ي): جون إكزافيير موريس وآخرون، " (Almost) Text Embeddings Reveal As Much As Text"، وقائع. مؤتمر 2023 حول الأساليب التجريبية في معالجة اللغة الطبيعية () (2023EMNLP). للاطلاع على وصف تقني لمسار الاسترجاع، انظر(ي):
https://infohub.delltechnologies.com/en-us/p/demystifying-on-device-intelligent-search-using-rag-architecture/

[11] يستند الوصف القائم على أربع مراحل والمستخدم هنا إلى الأدبيات التقنية، ويُناقش بمزيد من التفصيل في المقال بعنوان The Heart of the Matter: Copyright, AI Training, and LLMs 71 J Cop Soc'y of the USA 482 (2024) (يُشار إليه فيما يلي بعبارة "Heart of the Matter")، ومتاح على الرابط https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4963711. يُعرض مسار مماثل مكون من أربع مراحل (جمع البيانات، والمعالجة المسبقة للبيانات، وتدريب النموذج، وتقييم النموذج وتحسينه) في "دليل وزارة الثقافة والرياضة والسياحة واللجنة الكورية لحق المؤلف بشأن قابلية تطبيق مبدأ الاستخدام المنصف على تدريب نماذج الذكاء الاصطناعي التوليدي" (فبراير 2026) (يُشار إليه فيما يلي "بدليل الاستخدام المنصف الكوري")، في 10.

عبر عدة خطوات تقنية، تُنفَّذ كل منها على نسخة مخزَّنة من الذخيرة اللغوية. ويخضع المحتوى النصي أولاً للترميز، أي تحويل الكلمات وغيرها من التسلسلات إلى سلسلة من الرموز العددية المستمدة من مفردات ثابتة.[12] وتعمل أدوات التقطيع إلى رموز الحديثة عموماً على مستوى أجزاء الكلمات لا الكلمات الكاملة.[13] وتكتسب هذه الخطوة أهميتها في التحليل الحالي لأنها، في حالة أدوات التقطيع على مستوى البايت التي تستخدمها النماذج المعاصرة الرئيسية، تجري دون فقدان وقابلة للعكس: فالنص الذي تُرمِّزه أداة تقطيع من هذا النوع ثم يُفكُّ ترميزه يعطي النص الأصلي، ومن ثم فإن الذخيرة اللغوية المقطَّعة إلى رموز ليست ملخصاً للمصنفات ولا تجريداً لها، بل إعادة ترميز لها في صورة أخرى.

ثم يتم تمثيل الرموز في شكل تضمينات.[14] والنهج الذي هيمن على المجال منذ عام 2017، وهو المحول (الحرف "T" في الاختصار GPT الذي يرمز إلى المحول التوليدي مسبق التدريب)، فيُعالج تلك التمثيلات عبر آلية تُعرف بالانتباه.[15] أما وصف التضمين، والنموذج المدرب نفسه، اللذين ينطويان على درجة أكبر من التحويل، فيتم تناولهما في القسم C.2 أدناه

بمجرد تهيئة النموذج، يُدرَّب بتعريضه مراراً لتسلسل من الرموز المستمدة من الذخيرة اللغوية. ويتنبأ النموذج بكل رمز على التوالي استناداً إلى الرموز التي تسبقه، ويُعدَّل على زيادات صغيرة كثيرة حتى تقترب تنبؤاته من النص المرصود؛ وبذلك لا يُرجَع إلى الذخيرة اللغوية مرة واحدة، بل تُستخدم، رمزاً برمز، طوال عملية التدريب. وبعد معالجة كم هائل من الرموز المستمدة من الذخيرة اللغوية، اكتسب النموذج توزيعاً احتمالياً ضخماً على فضاء واسع من تسلسلات الرموز الممكنة. وهذا التوزيع، المرمَّز في معلمات النموذج، هو ما يتيح للنموذج توليد نص متماسك استجابةً لمطالبة.

وتشكِّل العلاقة بين بيانات التدريب والنموذج المدرَّب موضوع نقاش كبير. ويتمثل أحد الآراء في أنه نظراً لأن النموذج المدرّب لا يحتوي إلا على المعلمات وليس على بيانات التدريب الأصلية، فلا تبقى أي نسخ من المصنفات المحمية بعد اكتمال عملية التدريب.[16] كما أن المعلمات تصف توزيعاً احتمالياً على فضاء متجهي هائل الحجم. ومع أن هذا التوزيع يتيح للنموذج التعميم وتوليد تسلسلات جديدة تماماً، فإنه قد يسند احتمالات مرتفعة ومركَّزة على نحو غير متناسب إلى تسلسلات محددة ودقيقة تعرَّض لها مراراً في أثناء التدريب. ولما كان التوزيع مرجَّحاً بشدة نحو مسارات التدريب المحدودة هذه، فيمكن في حالات كثيرة استحثاث النموذج على إعادة إنتاجها حرفياً أو شبه حرفياً. وقد وُثِّقت هذه الظاهرة على نطاق واسع في الأدبيات التقنية، ويرد بحثها بمزيد من التفصيل في القسم جيم أدناه. وكما يرد توضيحه أدناه، ثمة فرق بين هذه الظاهرة (التي يشار إليها غالباً بـ"الحفظ") ومسألة ما إذا كان ثمة "استنساخ" داخل النموذج لجميع المصنفات المستخدمة في التدريب.

وبالنسبة للنماذج التي تُنتج صورًا أو ملفات صوتية أو مقاطع فيديو، تُعرف النماذج الأكثر شيوعاً باسم نماذج "الانتشار". وهي تتعلم من خلال إضافة ضوضاء تدريجيًّا إلى أمثلة التدريب، ثم تتعلم عكس تلك العملية بهدف إنتاج أمثلة جديدة من الضوضاء[17]. وهذا يختلف عن نماذج اللغة (النصية)، لكن في كلتا الحالتين تُنتج عملية التدريب تمثيلاً متطوراً لبيانات المدخلات.

---

[12] بشأن الترميز، انظر(ي) جوهر المسألة، الحاشية 10 أعلاه، الجزء الثاني، باء. يشير الدليل الكوري للاستخدام المنصف إلى الترميز باعتباره أحد الإجراءات التقنية لمعالجة البيانات المسبقة، إلى جانب التنقية والتوحيد. انظر(ي) الدليل الكوري للاستخدام المنصف، الحاشية رقم 10 أعلاه، الصفحة 10.

[13] وتُبنى المفردات بالبدء من الأحرف المفردة ثم دمج الأزواج المتجاورة الأكثر تواتراً تدريجياً في وحدات أكبر، بحيث تُمثَّل الكلمات الشائعة برمز واحد، في حين تُقسَّم الكلمات الأقل شيوعاً إلى عدة رموز. وتستخدم شركة OpenAI ترميز أزواج البايتات في نماذج المحول التوليدي مسبق التدريب الخاص بها، كما يستخدمه عدد من كبار المطورين الآخرين. انظر(ي) جوهر المسألة، الحاشية رقم 10 أعلاه، في الجزء الثاني، باء.

[14] وبشكل أكثر تحديدًا، يتم تعيينها في فضاء متجه عالي الأبعاد، حيث يُمثَّل كل منها على أنه تضمين.

[15] وتتيح آلية الانتباه للنموذج، عند حساب تمثيل رمز معين، ترجيح صلة كل رمز آخر في التسلسل، وهو ما يمكّنه من التقاط الاعتماديات عبر مقاطع نصية طويلة. أشيش فاسواني وآخرون، "Attention Is All You Need"، المؤتمر الحادي والثلاثون. أنظمة معالجة المعلومات العصبية () (2017NeurIPS). تشكِّل بنية المحول الأساس الذي تستند إليه النماذج اللغوية الكبيرة المعاصرة الرئيسية.

[16] انظر(ي) جوهر المسألة، الحاشية 10 أعلاه، الجزء الثاني، جيم.

[17] بشأن أصول نماذج النشر، انظر(ي): جاشا سول-ديكشتاين وآخرون، " Deep Unsupervised Learning using Nonequilibrium Thermodynamics"، وقائع المؤتمر الدولي الثاني والثلاثون للتعلّم الآلي () (2015ICML)؛ جوناثان هو، وأجاي جاين، وبيتر أبييل، " Denoising Diffusion Probabilistic Models"، المؤتمر الرابع والثلاثون. أنظمة معالجة المعلومات العصبية () (2020NeurIPS)؛ وبشأن بنية الانتشار الكامن المستخدمة في "ستيبل دفيوجن" والمطروحة في دعوى "غيتي" التي يرد بحثها في الجزء الثالث، انظر(ي): روبن روماباخ وآخرون، " High-Resolution Image Synthesis with Latent Diffusion Models"، مؤتمر IEEE/CVF المعني بالرؤية الحاسوبية والتعرف على الأنماط ((CVPR 2022). للاطلاع على مقدمة مبسطة يسهل فهمها، انظر(ي) أيضاً يانغ سونغ وستيفانو إرمون، " Generative Modeling by Estimating Gradients of the Data Distribution"، المؤتمر الثالث والثلاثون. أنظمة معالجة المعلومات العصبية () (2019NeurIPS) متاح على http://papers.neurips.cc/paper/9361-generative-modeling-by-estimating-gradients-of-the-data-distribution.pdf. لا تتمحور النقطة العامة في المناقشة الحالية حول التفاصيل الرياضية لأي فئة معينة من النماذج.

2. التحسين والتعلم المعزز

تتولى المراحل اللاحقة للتدريب المسبق تكييف النموذج المدرَّب مسبقاً مع مهام معينة، وحالات استخدام معينة، ومعايير سلوك معينة. وتتباين المصطلحات من مطور إلى آخر، لكن التسلسل يشمل الضبط الدقيق الخاضع للإشراف، ونمذجة المكافأة، والتعلم المعزز.[18]

ويتضمن الضبط الدقيق الخاضع للإشراف مواصلة التدريب على مجموعة بيانات أصغر ومنتقاة تتألف من أزواج المدخلات والمخرجات من النوع الذي يُتوقع أن ينتجه النموذج. ومجموعة البيانات هذه أصغر بكثير من الذخيرة اللغوية المستخدمة في التدريب المسبق، لكنها تُختار لجودتها، وغالباً ما تكون من إعداد البشر أو تحريرهم. ويُحمَّل كل مثال من أمثلة الضبط الدقيق، شأنه شأن كل مثال من أمثلة التدريب المسبق، ويُستخدم لتحديث معلمات النموذج.

وتتضمن نمذجة المكافأة إنشاء نموذج مكافأة منفصل، يقيّم المخرجات المرشحة الصادرة عن النموذج الرئيسي بحسب مدى استحسانها. ويُدرَّب نموذج المكافأة نفسه على مقارنات من إعداد البشر بين المخرجات المرشحة.

ويستخدم التعلم المعزز من ردود الفعل البشرية، وهو تقنية واسعة الاستخدام لمواءمة نماذج الذكاء الاصطناعي التوليدي مع التفضيلات البشرية، نموذج المكافأة لتوجيه المزيد من التحديثات للنموذج الرئيسي.

وتتطلب كل مرحلة من هذه المراحل تخزين مواد التدريب ومعالجتها. وأحجام هذه المواد أصغر من الأحجام اللازمة للتدريب المسبق، لكن مواد التدريب ذاتها قد تتضمن مصنفات محمية، لا سيما حيثما تستند المقارنات التي يقدمها المعلّقون البشريون إلى هذه المصنفات أو تتضمنها. والتركيز على جودة المادة لا على كميتها لا يغيّر التحليل من منظور حق المؤلف.

ويتضمن تقييم النموذج وتحسينه، اللذان يجريان بالتوازي مع مراحل التدريب ويستمران بعد النشر، استخدام مجموعات بيانات التقييم لتقدير أداء النموذج وتحديد المجالات التي تلزم فيها مواصلة التدريب. وقد تتضمن مجموعات بيانات التقييم ذاتها مصنفات محمية، لا سيما حيثما أُنشئت المقاييس المرجعية انطلاقاً من مجموعات بيانات اختبارية متاحة للجمهور.

القسم جيم: بعد انتهاء التدريب

تُطرح ثلاثة أسئلة محددة بمجرد انتهاء التدريب. الأول هو ما إذا كان يُحتفظ بمجموعة النصوص، ومن الذي يحتفظ بها، ولأي أغراض لاحقة. والثاني هو ما إذا كان النموذج يحتوي على نسخ من المصنفات التي دُرب عليها، وبأي معنى. والثالث هو ما إذا كانت المخرجات تعيد إنتاج المصنفات التي استُخدمت في التدريب، أو تُكيفها، أو تتفاعل معها بأي شكل آخر.

1. استخدامات أخرى للذخيرة اللغوية الخاصة بالتدريب

وتظل الذخيرة اللغوية المستخدمة في التدريب ذات صلة بعد انتهاء عملية التدريب. وقد يحتفظ بها المطور لعدة أغراض: تقييم الإصدارات المستقبلية من النموذج، أو دعم مواصلة الضبط الدقيق على مجموعات فرعية معينة، أو إتاحة إجراء أبحاث جارية، أو إتاحتها لجهات خارجية بموجب ترخيص أو بغيره. وقد تُدمج مع مواد جُمعت حديثاً لتكوين ذخيرة لغوية أوسع للجيل التالي من النماذج. وقد تُباع أو تُشارك (على سبيل المثال، مع المنظمات المنتسبة). وفي بعض الحالات، تُتاح مشتقات الذخيرة اللغوية، بما في ذلك المجموعات الفرعية المرشحة أو المنقحة بإزالة التكرارات، في الملك العام بموجب تراخيص متساهلة.

وقد يتضمن كل استخدام من هذه الاستخدامات اللاحقة فعلاً من أفعال الاستنساخ والتوزيع والإتاحة بالمعنى المقصود في القواعد الوطنية والدولية ذات الصلة. بعبارة أخرى، قد يستمر استخدام ذخيرة لغوية حققت غرضها التدريبي لسنوات بعد ذلك، وتتوافق القواعد التي تحكم استخدامها اللاحق مع تلك التي كانت تحكم تجميعها في البداية.

2. الاحتفاظ بنسخ في النموذج

يناقش هذا الجزء أحد أكثر الأسئلة إلحاحاً من الناحية التقنية، وهو ما إذا كان النموذج المدرَّب ذاته يحتوي على "نسخ" من المصنفات التي دُرِّب عليها، وهو ما يميِّز الذكاء الاصطناعي التوليدي عن التكنولوجيات السابقة. وكانت المواقف التي اتخذها مطورو الذكاء الاصطناعي من جهة وأصحاب الحقوق من جهة أخرى متعارضة تعارضاً حاداً حتى وقت قريب.[19] وقد ذهب مطورو

---

[18] انظر(ي) جوهر المسألة، الحاشية 10 أعلاه، الجزء الثاني، دال. وتُستخدم المصطلحية "التدريب المسبق" و"الضبط الدقيق الخاضع للإشراف" و"نمذجة المكافأة" و"التعلم المعزز" على نطاق واسع بين مطوري الذكاء الاصطناعي.

[19] ولا يقتصر هذا التباين على المواقف في الدعاوى القضائية، بل يمتد إلى أولويات التصميم. وخلصت إحدى دراسات دلفي متعددة المراحل، التي شملت مختصين في أخلاقيات الذكاء الاصطناعي، وخبراء قانونيين، وممارسين في مجال الذكاء الاصطناعي الموسيقي، وفنانين، وأُجريت للتحقق من صحة إطار للتصميم المسؤول لأنظمة الموسيقى التوليدية، إلى وجود توافق ذي توجه "حمائي" بين الفنانين والخبراء القانونيين بشأن مسائل سيادة البيانات، في مقابل تركيز الممارسين على الجدوى العملية؛ وخلصت الدراسة إلى أن هذه الأطر تؤدي وظيفتها على أفضل وجه عندما تعمل بوصفها طبقة ترجمة بين أولويات أصحاب المصلحة المتباينة، بما في ذلك المفاضلة بين شفافية النماذج ومصالح الملكية الفكرية،

الذكاء الاصطناعي عموماً إلى أن النموذج المدرَّب يحتوي على معلمات فقط وليس المصنفات ذاتها، وأنه لا يُحتفظ بأي نسخ لأغراض قانون حق المؤلف. وقد ذهب أصحاب الحقوق إلى أن النموذج، في الواقع، نسخة من الذخيرة اللغوية بأكملها. وتشير الأدلة التجريبية التي تراكمت في العامين الماضيين إلى أن كلا الموقفين يميل إلى المبالغة في تبسيط العلاقة بين الحفظ والنموذج. علاوة على ذلك، أياً كانت نتيجة هذا النزاع، فإنها لا تؤثر في الأفعال السابقة: فجمع الذخيرة اللغوية واستخدامها في التدريب ينطويان على عمليات استنساخ لا يتوقف توصيفها القانوني على ما إذا كان النموذج المكتمل ذاته يحتوي على نسخ.

ويمكن استخدام نوعين من الأدلة التقنية في التحليل. الحفظ هو خاصية من خصائص النموذج المدرب، تتيح استرجاع تسلسلات تدريب محددة في معلمات النموذج بدقة كافية. ونظراً إلى أنه لا يمكن فحص البنية الداخلية لأحد النماذج مباشرة بسهولة، فإن الحفظ يُدرس من خلال الاستخراج، أي القدرة التي يثبتها النموذج، في ظل مطالبات مناسبة، على إعادة إنتاج تسلسلات من بيانات تدريبه حرفياً أو شبه حرفياً. ثم يُستخدم الاسترجاع كدليل على الحفظ. ولا يثبت إخفاق إجراء استخراج بعينه أن مصنفاً ما لم يُحفظ. وطبقت دراسة منهجية، أجراها كوبر وآخرون، تقنية استخراج احتمالية على 200 كتاب عبر أربعة عشر نموذجاً لغوياً مفتوح الأوزان، وخلصت إلى أن مدى الحفظ قد تفاوت تفاوتاً ملحوظاً بحسب النموذج وبحسب الكتاب على حد سواء.[20] ومعظم النماذج الكبرى لم تحفظ معظم الكتب، كلياً أو جزئياً. ومع ذلك، عندما كان الحفظ يحدث، كان يمكن أن يكون كبيراً: فقد أفادت الدراسة بأن أحد النماذج المنتشرة على نطاق واسع قد حفظ رواية هاري بوتر الأولى ورواية 1984 لجورج أورويل شبه كاملتين؛ وفي حالة الأولى، بينت الدراسة أنه يمكن توليد الكتاب كله بصورة شبه حرفية.

ومع ذلك، لم تسلم هذه الإثباتات التجريبية للحفظ من الاعتراض. وقد ردّ مطورو الذكاء الاصطناعي بأن تقنيات المطالبات اللازمة لاستحثاث النسخ الحرفي هي تقنيات "خصومية" لا تمثل طريقة استخدام النظم عملياً، وقد تشكّل ذاتها تلاعباً بالمنتج أو إساءة استخدام له تخرج عن نطاق الاستخدام العادي (وعن شروط الخدمة في بعض الحالات). في الدعوى القضائية التي رفعتها صحيفة نيويورك تايمز ضد شركة OpenAI، تقدّم المدّعى عليه بطلب رد دعاوى التعدي المباشر على حق المؤلف جزئياً، استناداً إلى أن المدّعي استخدم مطالبات وصفها المدّعى عليه بأنها معقّدة ومتلاعَب بها لتصنيع الاجترار الذي يشكو منه.[21] ورفضت المحكمة الجزئية للمنطقة الجنوبية من نيويورك في أبريل 2025 رد الدعاوى على هذا الأساس، ومنذ ذلك الحين انتقل النزاع، في مرحلة الكشف عن الأدلة، من مسألة ما إذا كان يمكن تصنيع "الاجترار" إلى مدى تواتر حدوثه عفوياً في الممارسة العملية، وبلغ ذروته في قرار يلزم بتقديم عينة تمثيلية من 20 مليون سجل من سجلات تفاعل المستخدمين الفعلية.[22] ويرد بحث الدعوى وأهمية الأمر بمزيد من التفصيل في الجزء الثالث، ألف 4، من هذه الدراسة.

وباختصار، تزيد النتائج التجريبية كلا الموقفين المستقطبين تعقيداً: فالنموذج المدرَّب ليس، من حيث المبدأ العام، "نسخة" من ذخيرته اللغوية التدريبية، ولكنه ليس أيضاً "بمنأى" عنها. وسيختلف النطاق الدقيق باختلاف النموذج.

بدلاً من أن تكون أدوات لفرض توافق قسري. ج دي بيراردينيس وآخرون، " RAIM: Operationalising Trustworthy AI in Music Generation"، IEEE Transactions on Technology and Society، عدد خاص بعنوان الابتكار الأخلاقي بشأن تكنولوجيا الموسيقى https://casbahlab.github.io/raim-framework/ تم النفاذ في 25 أغسطس 2026.

[20] فيدير كوبر، وآرون جوكاسلان، وأحمد أحمد، وآمي ب سايفرت، وكريستوفر دي سا، ومارك أ ليملي، ودانيال إ هو، وبيرسي لیانغ، " Extracting Memorized Pieces of (Copyrighted) Books from Open-Weight Language Models"، arXiv:2505.12546 (2025)، إصدار 30 نوفمبر 2025 (ونُشر أيضاً بوصفه ورقة عمل في القانون العام بجامعة ستانفورد / ورقة بحثية لكلية الحقوق بجامعة وست فرجينيا رقم 2025-005). وتستحق صياغة المؤلفين للمسألة (في الصفحة 1) أن تُقتبس كاملة: "كثيراً ما يطرح المدّعون والمدّعى عليهم في دعاوى حق المؤلف المتعلقة بالذكاء الاصطناعي التوليدي ادعاءات تعميمية ومتعارضة بشأن مدى حفظ النماذج اللغوية الكبيرة (LLMs) للتعبير المحمي للمدّعين في ذاكرتها. واستناداً إلى كل من التعلم الآلي وقانون حق المؤلف معاً، نبيّن أن هذه المواقف المستقطبة تبالغ كثيراً في تبسيط العلاقة بين الحفظ وحق المؤلف". وبشأن القول إن المحتوى المحفوظ لا بد من ترميزه في معلمات النموذج لأنه "لا مكان آخر يمكن أن يكون فيه"، انظر(ي) أيضاً أ فيدر كوبر وجيمس غريملمان (المشار إليهما في هذا المصنف في الصفحة 25).

[21] The New York Times Company v. Microsoft Corp، رقم الدعوى 1:23-cv-11195، مذكرة تأييد طلب المدعى عليهم رفض الدعوى (المحكمة الجزئية للمنطقة الجنوبية من نيويورك، 26 فبراير 2024)؛ الرأي والأمر القضائي (المحكمة الجزئية للمنطقة الجنوبية من نيويورك، 4 أبريل 2025لقاضي شتاين) (برفض الجزء ذي الصلة من طلبات رفض دعاوى الانتهاك المباشر والمساهمة في الانتهاك). للاطلاع على الطريقة التي عرضت بها شركة OpenAI هذه القضية للجمهور، انظر(ي): OpenAI، "Reporting the Facts About the New York Times' Lawsuit"، متاح على الرابط openai.com/new-york-times.

[22] في دعوى انتهاك حقوق المؤلف المتعلقة بشركة OpenAI، رقم الدعوى 1:25-md-03143-SHS-OTW (المحكمة الجزئية للمنطقة الجنوبية من نيويورك)، الرأي والأمر القضائي (القاضي شتاين، 5 يناير 2026)، المؤيّدان للرأي والأمر القضائي (القاضية وانغ، قاضية الصلح، 2 ديسمبر 2025) (برفض طلب شركة OpenAI إعادة النظر)، وكذلك أمر الكشف عن الأدلة الصادر في نوفمبر 2025. وأمرت المحكمة بتقديم عينة عشوائية ومنزوعة الهوية من 20 مليون سجل من سجلات تفاعل مستخدمي ChatGPT، مستمدة من المجتمع المحفوظ، على أساس أن الأدلة الواسعة المتعلقة بمخرجات النموذج الموجَّهة إلى المستهلكين ذات صلة بتحليل الاستخدام المنصف، ولا سيما بمسألة ما إذا كانت مخرجات النموذج تؤدي وظيفة بديل سوقي لمصنفات المدّعين.

أما المجموعة الثانية من الأدلة فهي الأدبيات الأوسع المتعلقة بهجمات استخراج البيانات، التي توثق الظروف التي يمكن فيها استرجاع أمثلة تدريب فردية من نموذج مدرَّب من خلال مطالبات مستهدفة.[23] وتشمل الأمثلة المسترجعة مقاطع نصية طويلة جداً، وصوراً يمكن التعرف عليها، ومعلومات تحدد الهوية الشخصية. وقد استُشهد بهذا المصنف كدليل على أن قابلية الاستخراج قد تكون خاصية هيكلية في النماذج التوليدية الكبيرة لا مجرد نتيجة لتلاعب المستخدم. وتزعم بعض شركات الذكاء الاصطناعي أن هذا التوجيه يشكّل تلاعباً غير منصف أو مخالفة لشروط الخدمة.[24] غير أن التقييمات من هذا النوع أصبحت أكثر تعقيداً خلال العام الماضي: فالنظم المنشورة تضع بصورة متزايدة مرشِّحات وضوابط حماية أخرى بين النموذج والمستخدم لتقليل قابلية استخراج المحتوى المحفوظ؛ وقد اضطرت الدراسات الحديثة إلى اللجوء إلى تقنيات خصومية مع بعض النظم، في حين أتاحت نظم أخرى الحصول على محتوى محمي من دونها.[25]

وقد يكون من المفيد استخدام تشبيه لفهم الموقف. وهو تشبيه يتعلق ببرامج الحاسوب[26]. وحين تُجمع شفرة المصدر لبرنامج ما (غالبًا ما يكون مكتوبًا بلغة تشبه اللغة الطبيعية) في صيغة قابلة للتنفيذ، تختلف النتيجة عن شفرة المصدر ولا يستطيع الإنسان قراءتها (بسهولة). غير أن التجميع لا يخلو من الأصل: فالبرنامج القابل للتنفيذ يحتفظ بمعلومات برنامج المصدر، ويمكن، بالهندسة العكسية، إعادة بناء الأصل إلى حد كبير. وفي هذا الصدد، تظل الشفرة القابلة للتنفيذ مشابهة لبرنامج المصدر. ويمكن القول إن الخطوات التي تنتج نموذجاً مدرَّباً من ذخيرة لغوية تدريبية هي، إذا اتُّبع هذا التشبيه، من نوع مماثل. وفي الحالتين، يُحوَّل مصدر مفهوم (برنامج في شفرة مصدر؛ ذخيرة لغوية من المصنفات) إلى كائن عددي وظيفي لم يعد يشبهه لكنه يظل يحمل معلومات مستمدة منه. غير أن لهذا التشبيه حدوداً: فالتجميع عملية حتمية تنقل برنامج المصدر الكامل إلى برنامج قابل للتنفيذ، في حين أن ترميز بيانات التدريب في نموذج انتقائي وجزئياً احتمالي، بحيث ينطبق التشبيه، في أقصى الأحوال، على المجموعة الفرعية من المصنفات التي ثبت حفظها؛ وتذهب دراسات حديثة إلى أن هذا الفارق بين قابلية الاسترجاع الحتمية وقابلية الاسترجاع الاحتمالية فحسب هو تحديداً ما يرجَّح أن يتوقف عليه التوصيف القانوني للنموذج.[27]

وتشير الأدلة التقنية التي تم استعراضها هنا إلى أن النموذج المدرّب يحتفظ بالمعلومات المستمدة من المصنّفات الموجودة في بيانات التدريب الخاصة به على الرغم من أن التحويل أمر جوهري وأن النفاذ إلى أي مصنف بعينه قد يكون أمراً جزئياً واحتمالياً. أما كيفية ترجمة هذه الملاحظة التقنية إلى توصيف قانوني للنموذج، وما إذا كان النموذج نفسه يُعدّ نسخة من المصنَّفات التي دُرِّب عليها أو مصنَّفاً محوّراً/ مشتقاً عنها، بموجب قانون حق المؤلف، فهي مسألة لا تزال قيد البحث والتحليل أمام المحاكم وفي الأدبيات العلمية المتخصصة، كما توضح هذه الدراسة لاحقاً.

ويمكن استخلاص ملاحظتين في هذه المرحلة. أولاً، يُستخدم الحفظ كحجة في كل من مرحلة التدريب، باعتباره دليلاً على أن تدريب النموذج تضمن ترميز نُسخ، وفي سياق مرحلة المخرجات، باعتباره دليلاً على قدرة النموذج المنشور على توليد مصنَّفات محمية في صورة تحل محل المصنَّفات الأصلية. ثانياً، تتعلق العروض التي يستند إليها المدعون بمصنَّفات بعينها في نماذج محددة. وهي تميل إلى تأكيد أن الحفظ قد يكون كبيراً فيما يتعلق بمصنَّفات محددة بعينها يكثر التعرض لها، لكنها لا تثبت أن هذه الخاصية تتوافر على نحو موحّد في جميع النماذج، أو أن معلمات النموذج لا تعدو، بوجه عام، أن تكون نسخاً مخزّنة. أما التوصيف القانوني لهذه المسائل، فيتناوله الجزء الثالث.[28]

---

[23] ومن بين المصنفات السابقة في هذا المجال: أنطونيا كاراموليغكو وآخرون، " Copyright Violations and Large Language Models"، arXiv:2310.13771 (2023) (التي أثبتت، عبر مجموعة متنوعة من نماذج اللغة، أنه يمكن استعادة أمثلة فردية من بيانات التدريب من خلال هجمات استخراج البيانات).

[24] انظر(ي) إلى نيكولاس كارليني وآخرين، "Extracting Training Data from Large Language Models"، الندوة الثلاثون لأمن يوزينيكس (2021) 2633؛ وميلاد نصر وآخرين، " Scalable Extraction of Training Data from Aligned, Production Language Models"، المؤتمر الدولي لتمثيلات التعلّم (ICLR) لعام 2025 (امتداداً لعمل نصر وآخرين، (2023) arXiv:2311.17035).

[25] أحمد أحمد وآخرون، "Extracting Books from Production Language Models"، arXiv:2601.02671 (2026) (توصلت الدراسة إلى أن استخدام المطالبات التنافسية كان ضروريًّا لاستخراج نصوص الكتب المُخزَّنة في ذاكرة بعض الأنظمة المنشورة، في حين لم يكن ذلك ضروريًّا في أنظمة أخرى).

[26] انظر(ي) جوهر المسألة الحاشية 10 أعلاه، الجزء الثاني، جيم (حيث يُعقد تشبيه بين تدريب نموذج لغوي وتجميع برنامج حاسوب في شكل قابل للتنفيذ، ويُلاحظ أن "المصنف الأصلي المحمي بحق المؤلف" يُترجَم في الحالتين إلى "شكل آخر لا يفهمه البشر مباشرة"، لكن محتوى الأصل "[ليس] غائباً عن تلك النسخة التي تفهمها الآلة").

[27] م أ ليملي وأ ف كوبر، ""النسخ" الاحتمالية في نماذج الذكاء الاصطناعي التوليدي"، Berkeley Technology Law Journal (2026)، متاحة على الرابط: https://ssrn.com/abstract=7067878 (يرى المؤلِّفان، انطلاقاً من أدبيات الاستخراج التجريبي التي نوقشت في هذا الجزء، وتفيد بأنه لا ينبغي اعتبار النموذج يحتوي على "نسخة" من مصنف معين إلا إذا كان من الممكن استخراج المصنف بطريقة مباشرة، وبصورة حتمية في أوضح الحالات؛ وأنه عندما ينجح الاستخراج احتمالياً فقط، بعد محاولات عديدة، فإن ما يتم تخزينه ليس نسخة من المصنَّف بل مجموعة من "المكونات"؛ وأن المسألة، بالتالي، مسألة استدلالية، ويجب الإجابة عنها لكل مصنف على حدة ولكل نموذج على حدة).

[28] بشأن الإجراءات المجمعة، انظر(ي) قضية In re OpenAI, Inc. Copyright Infringement Litigation، رقم 1:25--md-03143-SHS-OTW (المحكمة الجزئية الأمريكية للمنطقة الجنوبية من نيويورك)، التي ضُمّت إليها الدعاوى التي رفعتها صحيفة نيويورك تايمز ومجموعات من المؤلفين ضد

وأخيرًا، تجدر الإشارة أيضًا إلى أن النموذج قد يتغير بعد اكتمال التدريب من خلال الضبط الدقيق، أو التدريب على المواءمة، أو محاولة إجراء "إلغاء" تعلم الآلة الموجه إلى محتوى محدد. وإذا توقف النموذج عن إعادة إنتاج مصنَّف معين بعد إجراء مثل هذا التدخل، فلا يثبت ذلك أنه لم يسبق له ترميز ذلك المصنَّف. وتشير الأدبيات التقنية إلى أن إزالة مصنَّفات محددة من نموذج مدرَّب على نحو موثوق لا تزال أمراً صعباً، وأن هذه التدخلات قد تؤدي إلى تدهور أداء النموذج بطرق غير متوقعة.[29]

3. علاقة المدخلات والمخرجات

بعد نشر النموذج وإتاحته للاستخدام، يعمل من خلال تلقي مدخلات من المستخدم أو الموجِّهات، وإنتاج مخرجات استجابةً لها. وقد تكون هذه المخرجات نصوصاً أو شيفرات برمجية أو صوراً أو مواد صوتية أو مقاطع فيديو أو أشكالاً أخرى من المحتوى. وتشبه العلاقة بين المدخلات والمخرجات، من بعض الوجوه، العلاقة بين الاستعلام والإجابة في محرك بحث، إلا أن هذا التشبيه لا ينطبق إلا جزئياً. فالنموذج لا يرجع إلى مجموعة بيانات مفهرسة ليستخرج منها مقاطع، بل يولّد محتوى جديداً من خلال أخذ عينات من التوزيع الاحتمالي الذي استوعبه أثناء التدريب. وفي بعض الحالات، يرتبط المحتوى الذي يولّده ارتباطاً وثيقاً بمصنَّفات محددة واردة في بيانات التدريب؛ وفي حالات أخرى، يكون المخرج تركيباً جديداً لا يشبه إلى حد كبير أي مثال تدريبي منفرد.

وتكتسب ثلاث ملاحظات أهمية خاصة في هذا السياق. أولاً، قد يكون المخرج استنساخاً حرفياً أو شبه حرفي لمصنَّف وارد في بيانات التدريب، وفي هذه الحالة يشكّل المخرج نسخة من المصنَّف الأصلي.[30] وقد أثبتت ذلك دراسات تناولت إعادة إنتاج النماذج اللغوية مقاطع من كتب محمية بحق المؤلف، وكذلك إعادة إنتاج نماذج توليد الصور صوراً فوتوغرافية وأعمالاً فنية محمية.[31]. أما إذا ما كان هذا الاستنساخ يُعدّ تعدياً على الحقوق، فيتوقف على عدد من المسائل الإضافية، منها مدى تطبيق الاستثناءات والتقييدات، ودور مرشّحات المخرجات التي يطبّقها مطوّر النموذج، وكيفية إسناد المسؤولية بين مطوّر النموذج والمستخدم الذي أدخل الموجِّه. يتم التطرق إلى هذه الأسئلة بإيجاز في جزء لاحق من هذه الدراسة[32].

ثانياً، قد يكون المخرج مشابهاً بدرجة جوهرية لمصنَّف وارد في بيانات التدريب، من دون أن يكون استنساخاً حرفياً له. وقد يشكل ذلك انتهاكاً، على سبيل المثال، بموجب مبدأ التشابه الجوهري، وهو معيار معروف في عدد من القوانين الوطنية بشأن حق المؤلف.

ثالثاً، قد يكون المخرج جديداً تماماً، بمعنى تعذّر تحديد أي مصنَّف بعينه من بيانات التدريب بوصفه مصدراً له. إذ تستطيع النماذج التوليدية إنتاج محتوى لا يطابق أي مثال منفرد من أمثلة التدريب؛ بل إن هذا هو تحديداً ما تفعله في معظم الحالات. وفي هذه الحالات، تضعف الصلة بين المخرج وأي مصنَّف محميّ محدد. أما ما إذا كان المخرج يظل "مستنداً إلى" بيانات التدريب بأي معنى يعتدّ به قانون حق المؤلف، فهي مسألة عُرض على المحاكم في عدة ولايات قضائية، لكن لم يتبلور بشأنها توافق في الآراء حتى الآن.

وعليه، يتناول الجزء الأول أوجه استخدام المصنَّفات المحمية، والمسائل المتعلقة بالاحتفاظ بها أو استنساخها، في مراحل متعددة من تطوير نموذج للذكاء الاصطناعي التوليدي وتشغيله: بدءاً بمجموعة البيانات التي تُجمَع في البداية، مروراً بعمليات التخزين والمعالجة التي تُطبَّق على هذه المجموعة أثناء التدريب، ثم النموذج المدرَّب ذاته، وانتهاءً بالمخرجات التي ينتجها بعد النشر. وقد تنطوي كل مرحلة من هذه المراحل على حقوق استئثارية رئيسية يقرّها التشريع الوطني لحق المؤلف والمعاهدات الدولية. أما مسألة ما إذا كان كل استنساخ قد يشكّل تعدياً ظاهرياً على الحقوق (مع مراعاة خضوعه لأي استثناء أو تقييد)، ومسألة الآليات القائمة أو التي يمكن استحداثها لترخيص الاستخدامات المعنية، فسيجري تناولهما في الأجزاء التالية.

شركتي OpenAI وMicrosoft. ويعكس توصيف النموذج المدرَّب بأنه آلية للتخزين والاسترجاع، وتوصيف المعلمات بأنها شكل مضغوط من المصنَّفات المُدخلة إلى النموذج، ما أورده المدّعون في مرافعاتهم؛ أما التوصيف المقابل للمعلمات بأنها تعميمات إحصائية ترمّز أنماط اللغة، فيعكس موقف المدّعَى عليهم. وللاطلاع على الأدبيات التجريبية التي تحذّر من كلا التوصيفين، انظر(ي) الدراسة التي أعدّها كوبر وآخرون، والمشار إليها أعلاه في هذا الجزء.

[29] أ فيدر كوبر وآخرون، " Machine Unlearning Doesn't Do What You Think: Lessons for Generative AI Policy, Research, and Practice"، المؤتمر التاسع والثلاثون. أنظمة معالجة المعلومات العصبية () (NeurIPS2025).

[30] إذا كان مخرج نموذج لغوي كبير (LLM) يشبه مصنَّفاً سابقاً محمياً لم يستخدمه النموذج في التدريب ولم يتح له النفاذ إليه بأي طريقة أخرى، فإن القول بإنشاء "نسخة" يكون أقل وضوحاً بكثير.

[31] وثّقت دراسات أكاديمية ومذكرات مقدمة في دعاوى قضائية حالات إعادة إنتاج النماذج اللغوية نصوصاً حرفياً؛ كما وُثّقت حالات إعادة إنتاج نماذج توليد الصور صوراً فوتوغرافية وأعمالاً فنية يمكن التعرّف عليها، وذلك في سياق نماذج الانتشار. ولا يرتبط عدد الأمثلة الموثقة ارتباطاً مباشراً بالتوصيف القانوني لأي مخرج بعينه، إذ يتوقف هذا التوصيف على القانون الوطني الواجب التطبيق ووقائع القضية.

[32] للاطلاع على العلاقة بين الموجِّه المُدخل والنموذج والمخرج، والمسألة التي لا تزال مفتوحة بشأن كيفية توزيع المسؤولية بين المطوّر والمستخدم، انظر(ي) الجزء الثالث أدناه.

الجزء الثاني: التفويض والترخيص والتعويض

حدّد الجزء الأول من هذه الدراسة، على الصعيد التقني، المواضع التي قد تنطوي فيها عملية تطوير نموذج للذكاء الاصطناعي التوليدي وتشغيله على مساس بالحقوق الاستئثارية الرئيسية التي يقرّها التشريع الوطني لحق المؤلف والمعاهدات الدولية المتعلقة به. وينتقل هذا الجزء من بحث تلك الحقوق إلى بحث الآليات التي يمكن من خلالها ممارستها أو التنازل عنها أو الاستعاضة عنها بآليات أخرى في سياق الذكاء الاصطناعي. ويُفضي التساؤل عما إذا كانت الاستخدامات المعنيّة مشروعة دون تصريح، وكيفية إضفاء المشروعية على عمليات المطوّرين في الحالات التي لا يجوز فيها إجراء تلك الاستخدامات دون تصريح، وكيفية تعويض أصحاب الحقوق في إطار هذه العملية، إلى مجموعة واحدة من المبادئ: وهي قانون ترخيص حق المؤلف، سواء كان الترخيص فردياً وجماعياً، طوعياً وقانونياً، إلى جانب الاستثناءات والتقييدات التي تحدّد، من الأساس، الحالات التي يلزم فيها الحصول على ترخيص.

ويتكون هذا الجزء على النحو التالي. يتناول القسم ألف القنوات الرئيسية التي يحصل من خلالها مطوّرو الذكاء الاصطناعي حالياً على المواد المؤهلة للحماية بموجب حق المؤلف التي يدرّبون نماذجهم عليها، والوضع القانوني لكل قناة منها. أما القسم باء، فيتناول أبرز مسائل السياسات التي تشكّل ممارسات الترخيص في الوقت الراهن، وهي: الحقوق التي تتأثر بعملية التدريب، والترخيص الطوعي، والترخيص الجماعي بأشكاله الرئيسية الثلاثة، ووضع استثناء تنقيب النصوص والبيانات في الاتحاد الأوروبي، والنقاش العلمي بشأن مدى انطباقه على تدريب نماذج الذكاء الاصطناعي، ومخططات المكافآت. وتُذكر البلدان على سبيل المثال لتوضيح النُهج المتبعة في مناطق مختلفة من العالم؛ لا بوصفها موضوعاً للتحليل؛ كما أن هذا الاستعراض غير شامل.

القسم ألف: الخيارات المتاحة حالياً للمطورين للحصول على مواد مؤهلة للحماية بموجب حق المؤلف لأغراض التدريب

1. مسارات الاستحواذ الخمسة

في الواقع، يحصل مطورو الذكاء الاصطناعي على بيانات التدريب الخاصة بهم بإحدى الطرق الخمس التالية، أو بمزيج منها. أحدها هو استخدام مجموعات البيانات المفتوحة أو المُنسقة التي أنشأتها جهات خارجية، والتي تُتاح أحياناً مجاناً وأحياناً مقابل رسوم. ومن الأمثلة البارزة على ذلك Common Crawl، وهو مستودع لصفحات الويب يُحدّث باستمرار، وتتولى مؤسسة غير ربحية جمع صفحاته آلياً عن طريق الزحف على الويب وإتاحتها للاستخدام. ويشكّل هذا المستودع، بصورة مباشرة أو غير مباشرة، أساساً لمجموعات البيانات النصية المستخدمة في تدريب معظم النماذج اللغوية الكبرى. ويؤدي LAION-5B وأسلافه الدور نفسه بالنسبة إلى نماذج توليد الصور.[33] وهذه المجموعات من البيانات هي نفسها نتاج عمليات استرجاع آلي واسعة النطاق، وقد تستمر فيها الإشكالات المتعلقة بحقوق المؤلف المرتبطة بعملية الاسترجاع الأصلية. وسيتوقف الوضع القانوني على الأساس الذي تم بموجبه الحصول على مواد المصدر، وعلى القانون المعمول به للاستخدامات اللاحقة لمجموعة البيانات. وقد بدأ ظهور نظير في القطاع العام لمجموعات البيانات المُنسقة هذه: ففي جمهورية كوريا، تم تحديث نظام ترخيص البيانات العامة (Gonggong Nuri) ليشمل فئة ترخيص تنص صراحةً على جواز استخدام مجموعات البيانات العامة في تدريب نماذج الذكاء الاصطناعي.[34]

ومن المصادر الأخرى ما يُعرف باسم "مكتبات الظل"، وهي مستودعات كبيرة غير مصرّح بها تضم كتباً ومقالات وموسيقى وغيرها من المصنَّفات، ويُقال أو يُدَّعى أن عدداً من النماذج واسعة الاستخدام قد استعان بها، إما مباشرةً أو من خلال مجموعات بيانات تتضمن محتوياتها. وتُجسّد النتائج القضائية المتعلقة بالقرصنة في الولايات المتحدة، التي يتناولها الجزء الثالث، القسم ألف، أدناه، ما ينطوي عليه الاعتماد على مثل هذه المستودعات من مخاطر قانونية.

ولغرض المقارنة، ميّزت بعض المحاكم بين الحصول على المصنَّفات بطرق غير مشروعة وبين المسألة المنفصلة المتمثلة فيما إذا كان استخدام هذه المصنَّفات في التدريب مشمولاً باستثناء أو قيد قانوني. ففي أحد الأحكام الأمريكية التي ستُناقش أدناه، قُضي بأن تنزيل نسخ مقرصنة والاحتفاظ بها لا يندرجان ضمن الاستخدام المنصف، في حين اعتُبر تدريب نموذج على نسخ من المصنَّفات ذاتها جرى الحصول عليها بطرق مشروعة استخداماً تحويلياً ومنصفاً. ومن هذا المنطلق، فإن مشروعية الحصول على المصنَّفات ومشروعية استخدامها في التدريب مسألتان مستقلتان من الناحية التحليلية. وتكتسب هذه التفرقة أهميتها لأن بيانات التدريب قد تأتي، ضمن النموذج الواحد، من مصادر مشروعة وأخرى غير مشروعة، مما يزيد تعقيد مسائل الترخيص والإنفاذ. ولذلك، قد

---

[33] يرد وصف مستودع Common Crawl على الرابط: https://commoncrawl.org. وتعد مجموعة البيانات LAION-5B (المستخدمة في بناء نموذج ستيبل دفيوجن ونماذج أخرى لتوليد الصور) وسابقتها LAION-400M مجموعتا بيانات صور ونصوص متاحة للجمهور، جمّعتهما الشبكة المفتوحة للذكاء الاصطناعي واسع النطاق. لا تحتوي مجموعة البيانات نفسها على الصور الأصلية، بل تحتوي على أزواج من عناوين URL والأوصاف النصية؛ ويتم استرداد الصور في مرحلة التدريب من مواقعها الأصلية. وللاطلاع على التوصيف القانوني لعمليات تجميع مجموعات البيانات باعتبارها تنقيباً عن النصوص والبيانات، انظر(ي) القضية Kneschke ضد LAION e.V، المحكمة الإقليمية في هامبورغ، 27 سبتمبر 2024 (227/23 O 310)، والاستئناف في القضية ذاتها أمام المحكمة الإقليمية العليا في هامبورغ، 10 ديسمبر 2025 (104/24 U 5).

[34] ويُقصد هنا نظام Gonggong Nuri (الترخيص الكوري للبيانات الحكومية المفتوحة)، الذي تتيح المؤسسات العامة بموجبه إعادة استخدام البيانات العامة وفق فئات موحّدة من التراخيص؛ ويضيف الإطار المحدَّث فئةً تجيز صراحةً استخدام البيانات في تدريب نماذج الذكاء الاصطناعي.

يحتاج المطوّر إلى ضمان مشروعية الحصول على مجموعة البيانات، وكذلك وجود أساس قانوني يجيز عملية التدريب ذاتها، مع مراعاة القانون أو القوانين الوطنية واجبة التطبيق.

أما النهج الثالث، فيتمثل في الترخيص الفردي والجماعي. إذ يمكن للمطوّر شراء حق النفاذ إلى مجموعة بيانات منسّقة أو التفاوض للحصول عليها من ناشر، أو مؤسسة إخبارية، أو هيئة بث، أو مكتبة للصور، أو مشغّل قاعدة بيانات أكاديمية، أو أصحاب حقوق مؤسسيين. وقد أصبح الترخيص ممارسة مألوفة خلال السنوات الثلاث الماضية، إذ أبرم كبار مطوّري الذكاء الاصطناعي اتفاقات مع مؤسسات إخبارية كبرى ومكتبات للصور في أوروبا والأمريكتين وآسيا.[35] وتتمثل السمة الهيكلية للترخيص في أنه يتيح للمطوّر الحصول على سلسلة موثوقة من سندات الملكية المتعلقة بمجموعة البيانات المرخَّص بها، وذلك على الأقل في حدود الحقوق التي يملكها المرخِّص نفسه. وتُناقَش حدود هذه السلسلة في القسم باء.2 أدناه. وفي بعض الحالات، يتضمن العقد تسليم المحتوى.

أما المسار الرابع، الذي يُعدّ من المسارات الرئيسية من حيث حجم المواد التي يتيح الحصول عليها، فيتمثل في جمع المواد مباشرةً من الويب المفتوح، من خلال الزحف إلى الويب واستخلاص محتواه آلياً، على أن يتولى المطوّر نفسه هذه العمليات، لا جهة خارجية. وترد تفاصيل الإجراءات التقنية ذات الصلة في القسم ألف.2 من الجزء الأول من هذه الدراسة. ويُعدّ حجم عمليات الاسترجاع الآلي المعنية كبيراً: إذ تشير بعض التقديرات إلى أن الوكلاء الآليين باتوا مسؤولين عن غالبية حركة الزيارات على الويب عالمياً.[36] وغالباً ما تُجرى عمليات الزحف دون الحصول مسبقاً على موافقة أصحاب الحقوق في المواد المسترجعة، استناداً إلى إحدى ثلاث نظريات قانونية: أن المواد تقع في الملك العام أو أُتيحت للجمهور بموجب ترخيص صحيح بشروط ميسّرة؛ أو أن تكون العملية مشمولة باستثناء أو تقييد ينص عليه قانون حق المؤلف الواجب التطبيق (وهو، في الاتحاد الأوروبي، استثناء تنقيب النصوص والبيانات المنصوص عليه في المادة 4 من التوجيه بشأن بحق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة؛ وفي الولايات المتحدة وبموجب المادة 35-5، في جمهورية كوريا، من قانون حق المؤلف، مبدأ الاستخدام المنصف العام؛ وفي سنغافورة واليابان، أحكام تشريعية محددة)؛ أو أن العملية لا تمس حق النسخ على الإطلاق لأن النُسخ الذي يتم إنشاؤها مؤقتة أو عابرة، أو أن هذه النُسخ، في بعض الولايات القضائية، تندرج ضمن استثناء صريح لأعمال الاستنساخ المؤقتة (كما هو الحال بموجب المادة 5(1) من توجيه مجتمع المعلومات في الاتحاد الأوروبي). وهذه النظريات ليست متنافية، وقد يستند المطوّر إلى نظريات مختلفة فيما يتعلق بأجزاء مختلفة من مجموعة البيانات.

أما المسار الخامس، فيتمثل في البيانات الاصطناعية، وهي البيانات التي يولّدها نموذج قائم للذكاء الاصطناعي وتُستخدم في تدريب نموذج آخر أو ضبطه الدقيق. ويزداد استخدام البيانات الاصطناعية بوصفها أداةً في تطوير النماذج، وغالباً ما تُستخدم في المراحل المتأخرة من هذه العملية، ولا سيّما في الضبط الدقيق الخاضع للإشراف ونمذجة المكافآت.[37] ومن منظور حق المؤلف، تثير البيانات الاصطناعية مسألة مختلفة. فإذا كان النموذج الذي ولّد هذه البيانات قد دُرِّب على مصنَّفات محمية، فهل تعيد مخرجاته الاصطناعية إنتاج التعبيرات المحمية أيضًا أو تكيفها؟ ولم تُعرض هذه المسألة على أي محكمة حتى الآن على نحو يفضي إلى حكم ملزم بشأنها.

2. الترتيبات التعاونية

وأخيراً، إلى جانب هذه المسارات الخمسة العامة، تُلاحظ في بعض الولايات القضائية أشكال من التعاون المنظّم بين مطوّري الذكاء الاصطناعي وأصحاب الحقوق من المؤسسات أو الجهات المملوكة للدولة، ممن ترتبط حقوقهم على نحو خاص بتطبيقات قطاعية معينة. وتُعد الصين أحد الأمثلة على ذلك. فقد شهد تطوير النماذج اللغوية الكبيرة المتخصصة في مجالات بعينها تعاوناً بين أحد مطوّري نماذج الذكاء الاصطناعي ذات الأغراض العامة ومحطات إذاعية وتلفزيونية في هونغ كونغ (للحصول على بيانات صوتية باللغة الكانتونية)، وبين المطوّرين وناشرين إلكترونيين يمتلكون مجموعات كبيرة من المصنفات التي ألّفها كتّاب متعاقدون معهم (لأغراض الروايات الرقمية)، وبين المطوّرين ومستودعات صور ضخمة تضم مليارات الصور التي جرى الحصول عليها بموجب نقل حقوق أو تراخيص استئثارية (لتوفير بيانات التدريب البصري)، فضلاً عن التعاون بين المطوّرين ومشغّلي قواعد

---

[35] فيما يتعلق بصفقات الترخيص الثنائية المبرمة بين مطوري الذكاء الاصطناعي والناشرين، انظر (ي) (ضمن صفقات كثيرة) الاتفاقات التي أبرمتها OpenAI مع وكالة أسوشيتد برس (Associated Press)، و News Corp، و Le Monde، و Prisa Media، و Axel Springer، و Vox Media، و The Atlantic، و Hearst؛ والاتفاقات التي أبرمتها Google مع Reddit ووكالة أسوشيتد برس؛ والاتفاقات التي أبرمتها Anthropic مع ناشرين مختلفين. وقد كُشف عن هذه الصفقات بمستويات متفاوتة من التفصيل، في حين تظل شروطها المالية سرية لأسباب تجارية.

[36] انظر (ي) Imperva، تقرير Bad Bot Report لعام 2025 (الطبعة الثانية عشرة، 2025)، الذي أفاد بأن حركة الزيارات الآلية شكّلت 51 في المائة من إجمالي حركة الزيارات على الإنترنت في عام 2024، متجاوزةً حركة الزيارات البشرية للمرة الأولى منذ عشر سنوات، وأن ما يُسمّى "الروبوتات الآلية الضارة" شكّل 37 في المائة من إجمالي الزيارات.

[37] البيانات الاصطناعية هي بيانات يولّدها نموذج قائم للذكاء الاصطناعي وتُستخدم في تدريب نموذج آخر أو ضبطه الدقيق. وتثير هذه البيانات تساؤلات متميزة تتعلق بحق المؤلف عندما يكون النموذج الذي ولّدها قد دُرّب بدوره على مصنَّفات محمية. وللاطلاع على الاستخدام المتزايد للبيانات الاصطناعية في المراحل المتأخرة من تطوير النماذج، انظر (ي): ر. ليو وآخرون، " Best Practices and Lessons Learned on Synthetic Data"، (2024) arXiv:2404.07503، الذي قُدّم في المؤتمر الأول لنمذجة اللغة.

البيانات الأكاديمية لبناء نماذج متخصصة بصورة مشتركة. ويُقال إن هذا النمط واسع الانتشار، مع أن الاتفاقات ذاتها لا تُصاغ غالباً في شكل تراخيص لحق المؤلف، كما أن نطاقها الدقيق لا يُفصح عنه دائماً. وكما سيُناقش في القسم باء.2 أدناه، يتمثل القيد الهيكلي لهذا النهج في أن الناشر أو هيئة البث أو المستودع لا يملك عادةً سوى حقوق معينة منحها له المؤلف الأصلي، فضلاً عن أن مدى شمول هذه الحقوق للتدريب على نماذج الذكاء الاصطناعي يظل غير محسوم.

القسم باء: العناصر الرئيسية للسياسات

1. الحقوق المتأثرة بتدريب نماذج الذكاء الاصطناعي التوليدي

(أ) حق النسخ

تنطوي العمليات التقنية الموصوفة في الجزء الأول، للوهلة الأولى، على عدد من الحقوق الاستئثارية التي تقرّها التشريعات الوطنية لحق المؤلف ومعاهدات حق المؤلف الدولية. ويحتل حق النسخ مكانة محورية في هذا التحليل.[38] فإنشاء مجموعة بيانات للتدريب ينطوي على إنشاء نسخ؛ كما أن تخزين هذه المجموعة ومعالجتها خلال مختلف مراحل التدريب ينطويان على إنشاء نسخ إضافية، بعضها دائم وبعضها عابر؛ ووفقاً للتحليل التقني الوارد في القسم جيم.2 من الجزء الأول، يمكن أيضاً تحليل بعض جوانب النموذج المدرَّب ذاته على أنها نسخ من المصنَّفات التي دُرِّب النموذج عليها. ويبدو، للوهلة الأولى، أن كل عملية من هذه العمليات تدخل في نطاق حق النسخ المعترف به بموجب المادة 9(1) من اتفاقية برن، والمادة 1(4) من معاهدة الويبو بشأن حق المؤلف، والمادة 7 من معاهدة الويبو بشأن الأداء والتسجيل الصوتي ومن معاهدة بيجين، والأحكام المقابلة في التشريعات الوطنية.

وتتطلب معالجة النسخ المؤقتة والعابرة النظر فيها على نحوٍ مستقل، لأن ثمة أسلوبين قانونيين مختلفين يفضيان إلى نتائج عملية متشابهة؛ ففي بعض الولايات القضائية لا تُعامل هذه النسخ على أنها داخلة في نطاق حق النسخ، بينما تُستوعب في ولايات أخرى بموجب استثناء صريح. وتقدّم الصين المثال الرئيسي على الأسلوب الأول: إذ كان يُنظر باستمرار إلى النسخ المؤقت على أنه يقع خارج نطاق الحق. وفي الممارسة القضائية الصينية، تتمثل عمليات النسخ المؤقتة المشمولة في الاستثناء في النسخ العابرة التي تُنشأ في ذاكرة النفاذ العشوائي أو في ذاكرة التخزين المؤقت وتكون ضرورية جداً لعملية معالجة البيانات، وليس ثمة موقف مستقر بشأن المدة التي تبقى فيها النسخة موجودة مع استمرار اعتبارها مؤقتة. وتعرّف المادة 10(1)(5) من قانون حق المؤلف النسخ بأنه الحق في إنتاج نسخة واحدة أو أكثر من مصنف عن طريق وسائل منها الطباعة أو التصوير الضوئي أو النقش أو التسجيل أو التسجيل على أشرطة الفيديو أو النسخ أو التصوير الفوتوغرافي أو التحويل إلى صورة رقمية،[39] وقد صوّت وفد الصين، من بين وفود أخرى، ضد البيان المتفق عليه في المؤتمر الدبلوماسي الذي كان من شأنه أن يُدرج التخزين الرقمي المؤقت بوضوح ضمن نطاق الحق.[40] ولم تتضمن لائحة حماية حق التبليغ للجمهور على شبكات المعلومات (2006) حكماً يعترف بالنسخ المؤقت بوصفه من أعمال النسخ، وقد طبّقت محاكم الصين هذا الفهم في قضايا تتعلق بتحويل الترميز والعمليات الرقمية العابرة المماثلة. ومن هذا المنظور، قد لا يشكّل فعل التدريب ذاته مساساً بحق النسخ بموجب القانون الصيني، بقدر ما يقتصر على نفاذ عابر إلى المصنفات الموجودة في الذاكرة أو ذاكرة التخزين المؤقت، والتي تُمحى عند انتهاء مرحلة التدريب. أما إنشاء قاعدة بيانات تضم المصنفات الأساسية، فلا شك في أنه قد يمسّ هذا الحق. وفي ولايات قضائية أخرى، يتحقق الناتج العملي نفسه عن طريق استثناء صريح: ففي الاتحاد الأوروبي، تتناول المادة 5(1) من التوجيه الأوروبي بشأن حق المؤلف والحقوق المجاورة في مجتمع المعلومات (توجيه مجتمع المعلومات) مسألة النسخ العابرة والعرضية بموجب استثناء محدد بدقة، بدلاً من استبعادها من نطاق الحق منذ البداية. وتُعد النسخ المعنية متطابقة في جوهرها في النظامين؛ ويكمن الفرق في الأسلوب القانوني المتّبع.[41] أما في معظم الولايات القضائية الأخرى التي شملها هذا الاستعراض، فيُنظر عموماً إلى فعل التدريب على أنه يندرج ضمن حق النسخ، وتتمحور المسألة التحليلية حول توافر استثناء أو تقييد.

---

[38] انظر (ي) الجزء الأول، القسمين ألف-1 وألف-2.

[39] قانون حق المؤلف لجمهورية الصين الشعبية، المادة 10(1)(5).

[40] انظر (ي) وثيقة الويبو CRNR/DC/101، محاضر الجلسات الموجزة، الجلسة العامة، الفقرات من 497 إلى 502. ومع ذلك، اعتمد المؤتمر الدبلوماسي البيان المتفق عليه بشأن المادة 1(4) من معاهدة الويبو بشأن حق المؤلف؛ وصوّتت الصين ضده.

[41] المادة 5(1) من التوجيه EC/29/2001 الصادر عن البرلمان الأوروبي والمجلس الأوروبي في 22 مايو 2001 بشأن توحيد جوانب معيّنة من حق المؤلف والحقوق المجاورة في مجتمع المعلومات، OJ L 167/10 [ويُشار إليه في كثير من الأحيان بما يلي "توجيه مجتمع المعلومات"]. ويشمل الاستثناء أفعال النسخ المؤقتة التي تكون عابرة أو عرضية وتشكل جزءاً لا يتجزأ وأساسياً من عملية تكنولوجية، ويكون غرضها الوحيد تمكين الإرسال في إحدى الشبكات بين أطراف خارجية عن طريق وسيط، أو تمكين الاستخدام القانوني، على أن تكون أفعال النسخ هذه ليس لها أهمية اقتصادية مستقلة.

(ب) حق التكييف والمصنفات المشتقة

وبالإضافة إلى حق النسخ، ثمة ثلاثة حقوق أخرى ذات صلة أو قد تكون ذات صلة؛ فحق التكييف، المعترف به في المادة 12 من اتفاقية برن وفي الأحكام الوطنية المقابلة، يُستند إليه في مرحلة المخرجات، إذ يُعد توليد هذه المخرجات بمثابة إجراء تحويل أو تعديل على المصنفات الأصلية. فعلى سبيل المثال، يدرج القانون البرازيلي صراحةً حق التكييف ضمن أشكال الاستخدام التي تتطلب إذناً مسبقاً وصريحاً،[42] وقد اتخذت أمانة حق المؤلف والملكية الفكرية في وزارة الثقافة موقفاً مفاده أن توليد المخرجات يمسّ حق التكييف إلى جانب حق النسخ. ويتوافق موقف القانون الكندي والفرنسي والألماني عموماً مع هذا التوجه إلى حدٍ كبير.

(ج) التبليغ للجمهور والإتاحة

يُفعَّل حق الإتاحة للجمهور، الوارد في المادة 8 من معاهدة الويبو بشأن حق المؤلف وفي المادتين 10 و14 من معاهدة الويبو بشأن الأداء والتسجيل الصوتي، بشكل أساسي في مرحلة المخرجات، عندما يعيد النموذج المنشور المحتوى إلى المستخدم، إذا كان هذا المحتوى يتضمن نسخة من تعبير محمي. وفي البرازيل، التي ليست طرفاً متعاقداً في معاهدة الويبو بشأن حق المؤلف أو معاهدة الويبو بشأن الأداء والتسجيل الصوتي، قضت محكمة العدل العليا (Superior Tribunal de Justiça) في عام 2017، في القضية REsp 1.559.264-RJ، بأن بثّ الموسيقى عبر الإنترنت يمسّ حق التبليغ للجمهور، وهو حق يشمل ما قد تصنّفه ولايات قضائية أخرى ضمن حق الإتاحة.[43] ويتسق ذلك مع "الحل الشامل" الذي اعتُمد في المؤتمر الدبلوماسي لعام 1996، الذي يمكن بموجبه تنفيذ حق الإتاحة عن طريق حق التبليغ للجمهور أو حق التوزيع أو الجمع بينهما. ولم يُختبر بعد أمام المحاكم البرازيلية تطبيق استدلال مماثل على مخرجات نموذج ذكاء اصطناعي منشور.[44] ويتمتع فنانو الأداء ومنتجو التسجيلات الصوتية بحقوق إتاحة موازية بموجب المادتين 10 و14 من معاهدة الويبو بشأن الأداء والتسجيل الصوتي المشار إليها أعلاه، ما يعني أن نطاق هذا التحليل يتجاوز مجرد حقوق المؤلفين.

(د) الحقوق المعنوية

قد تكون الحقوق المعنوية محل اعتبار أيضاً؛ ويُستند في أغلب الأحيان إلى حق السلامة. ففي البرازيل، يُنظر إلى تجزئة المصنفات في أثناء عملية التدريب على أنها تدخل محتمل في الحق في السلامة، كما يُنظر إلى عدم ذكر أسماء مؤلفي المواد المستخدمة على أنه تدخل محتمل في حق الإسناد.[45] ويعترف قانون حق المؤلف الكندي بكل من الحق في السلامة وحق الإسناد بوصفهما من الحقوق المعنوية للمؤلف.[46] وفي عدد من الولايات التي تطبق حق *المؤلف*، طُرحت في التعليقات الأكاديمية مسألة إذا ما كان استخدام مصنف في ذخيرة بيانات التدريب من دون إسناد، وعلى نحو قد يولّد مخرجات تُعدّ نسخاً مشوّهة أو مزدرية للأصل، يشكل مساساً بالحقوق المعنوية، وهي مسألة لم تجد لها إجابة حاسمة حتى الآن.

وأخيراً، في الاتحاد الأوروبي وفي بعض الولايات القضائية التي اعتمدت حماية مماثلة، قد يكون الحق الخاص المستقل المقرَّر لقواعد البيانات بموجب التوجيه EC/9/96 محلَّ تطبيق إذا أُنشئت ذخيرة بيانات التدريب عن طريق استخراج أجزاء جوهرية من قاعدة بيانات محمية وإعادة استخدامها.[47] ويُعد هذا الحق مستقلاً عن حق المؤلف ويخضع لمعاييره الخاصة المتعلقة بالاستثمار الجوهري و"الجزء الجوهري" (فيما يتعلق "بحق الاستخراج"). ولا يزال نطاق انطباقه على إنشاء ذخائر بيانات التدريب الذكاء

---

[42] قانون حق المؤلف البرازيلي (القانون رقم 1998/9,610)، المادة 29، البندان 3 و4 (التكييف والتوزيع الموسيقي والترجمة وغيرها من أشكال التحويل). وقد عبرت وزارة الثقافة عن موقفها من بين أمور أخرى، في مذكرة تقنية وُجّهت إلى الهيئة الوطنية لحماية البيانات (Autoridade Nacional de Proteção de Dados) في أكتوبر 2024 بشأن تدريب نماذج أحد المطورين الكبار، استناداً إلى المادة 29(3) (التحويل لإنتاج مصنفات مشتقة)؛ انظر(ي) التقرير على الرابط https://nucleo.jor.br/curtas/2024/10/10/ministerio-cultura-prejuizos-direitos-ia-meta/، تاريخ النفاذ: 7 يوليو 2026.

[43] محكمة العدل العليا، (2017) REsp 1.559.264-RJ (التي قضت بأن بثّ الموسيقى عبر الإنترنت يُعد شكلاً من أشكال الأداء العام ويمسّ حقوق التبليغ للجمهور، وأن للمكتب المركزي للتحصيل والتوزيع (ECAD) الحق في التحصيل على هذا الأساس).

[44] قانون حق المؤلف البرازيلي (القانون رقم 1998/9,610)، المادة 5(5) (تعريف "التبليغ للجمهور") والمادة 29(7) (التوزيع بأي نظام يتيح للمستخدم اختيار المصنف).

[45] قانون حق المؤلف البرازيلي (القانون رقم 1998/9,610)، المادة 24، البنود من 1 إلى 4. ويتمتع فنانو الأداء أيضاً بحقوق معنوية تتعلق بسلامة أدائهم ونسبته إليهم بموجب المادة 92.

[46] قانون حق المؤلف، R.S.C. 1985, c. C-42، المادتان 1.14 و1.28.

[47] التوجيه رقم 96/9/EC للبرلمان الأوروبي وللمجلس الأوروبي المؤرخ 11 مارس 1996 بشأن الحماية القانونية لقواعد البيانات، OJ L 77/20، المادة 7 (الحق الخاص). وانظر(ي) أيضاً محكمة العدل التابعة للاتحاد الأوروبي، القضية *British Horseracing Board* ضد *William Hill*, ECLI:EU:C:2004:695، C-203/02 بشأن شروط معيار الاستثمار الجوهري.

الاصطناعي مسألة غير محسومة، بما في ذلك في سياق قواعد بيانات المعلومات القانونية ومجمّعات المجلات العلمية ومكتبات الصور الكبيرة.

ومن ثم، قد تمتد الحقوق التي قد يمسّها تدريب نموذج ذكاء اصطناعي توليدي إلى ما يتجاوز حق النسخ. وعليه فإن مدى كفاية الترخيص الثنائي أو الجماعي لتغطية الاستخدامات ذات الصلة يتوقف على الحقوق التي يشملها؛ فالترخيص المقتصر على حق النسخ قد لا يعالج المطالبات القائمة على التكييف أو الإتاحة للجمهور أو الحقوق المعنوية أو الحق الخاص بقواعد البيانات حيثما انطبق. وثمة نقطة أخرى يتناولها الجزء الثالث في سياق الإنفاذ العابر للحدود، هي أن حزمة الحقوق والنطاق الدقيق لكل حق يختلفان من ولاية قضائية إلى أخرى، إذ قد يُعرض أحد ترتيبات الترخيص المصُمّم لتخليص الحقوق في إحدى الولايات القضائية، المطوّر للمساءلة في ولاية أخرى.[48]

(هـ) *RMI* وتدابير مناهضة التحايل

بموجب المادتين 11 و12 من معاهدة الويبو بشأن حق المؤلف، والمادتين 18 و19 من معاهدة الويبو بشأن الأداء والتسجيل الصوتي، والمادتين 15 و16 من معاهدة بيجين، يحق لأصحاب الحقوق الحصول على حماية قانونية كافية وسبل انتصاف قانونية فعالة ضد التحايل على التدابير التقنية الفعالة التي يستخدمونها في سياق ممارسة حقوقهم بموجب تلك الصكوك واتفاقية برن، وكذلك الحق في سبل انتصاف قانونية ملائمة وفعالة ضد أي شخص يرتكب عن علم، أو في سياق سبل الانتصاف المدنية لديه أسباب معقولة تدعوه إلى العلم، أي من الأفعال التالية، إذا كان من شأنها أن تحرض أو تمكّن أو تُسهل أو تخفي انتهاك أي حق مشمول بتلك المعاهدات أو باتفاقية برن:

"1" حذف أو تغيير، دون إذن، أية معلومات إلكترونية ضرورية لإدارة الحقوق؛

"2" توزيع المصنفات أو نُسخ المصنفات، أو التسجيلات الصوتية أو العروض، أو استيرادها بغرض التوزيع، أو بثها أو إيصالها إلى الجمهور، دون إذن، و(بموجب الصكوك ذات الصلة) مع العلم بأن معلومات إدارة الحقوق الإلكترونية قد حُذفت أو غُيرت دون إذن.

على حد علمنا، لم تُرفع حتى وقت كتابة هذه الدراسة أي دعوى قضائية ناجحة ضد الشركات فيما يتعلق بتدريب نماذج الذكاء الاصطناعي.

2. الترخيص الطوعي (الجماعي والفردي)

(أ) الترخيص الفردي والجماعي لاستخدامات الذكاء الاصطناعي

يشمل الترخيص الطوعي التفاوض على تصريح بين أحد مطوّري الذكاء الاصطناعي وصاحب حقوق أو ممثل لعدة أصحاب حقوق، بما في ذلك منظمة الإدارة الجماعية أو أي وسيط آخر، من دون أي تفويض قانوني. وهذا هو الأصل في جميع الولايات القضائية التي شملتها الدراسة.

وتؤثر سمتان هيكليتان على الشكل الذي يتخذه الترخيص الطوعي في الممارسة. تتمثل السمة الأولى في عدم التماثل بين العدد الصغير من أصحاب الحقوق المؤسسيين الكبار القادرين على التفاوض مع المطورين بشروط مجدية تجارياً، والعدد الكبير جداً من المؤلفين الأفراد وصغار أصحاب الحقوق الذين كثيراً ما يكون هذا التفاوض غير عملي بالنسبة إليهم بسبب تكاليف المعاملات ونقص الخبرة أو عوامل أخرى. وقد أُبرمت الصفقات الثنائية الكبرى خلال السنوات الثلاث الماضية بين مطوري الذكاء الاصطناعي وكبار الناشرين والهيئات الإخبارية وهيئات البث ومكتبات الصور؛ ولا توجد بنية مماثلة للمؤلفين الأفراد أو المصورين أو الرسامين أو الموسيقيين، الذين كثيراً ما يعتمدون على الصفقات التي يبرمها ناشروهم وغيرهم من المرخَّص لهم للحصول على أي نسبة من القيمة المتولدة.

وقد أصبح حجم بعض هذه الترتيبات جوهرياً الآن. فقد أُفيد بأن اتفاقاً واحداً مدته خمس سنوات بين ناشر إخباري كبير ومطوّر للذكاء الاصطناعي تجاوزت قيمته 250 مليون دولار أمريكي،[49] وأصبح ترخيص المحتوى للاستخدام الحاسوبي مصدراً كبيراً

---

[48] انظر(ي) *Toronto Star Newspapers Ltd* ضد *OpenAI Inc.*, 2025 ONSC 6217 (القاضية كيميل جيه.، 7 نوفمبر 2025) (حيث رُفض دفع OpenAI المتعلق بالاختصاص القضائي استناداً إلى وجود صلة حقيقية وجوهرية بأونتاريو).

[49] ذكرت صحيفة وول ستريت جورنال أن قيمة الاتفاق الذي مدته خمس سنوات بين News Corp وOpenAI تبلغ أكثر من 250 مليون دولار أمريكي. ونظراً لأن الشروط المالية لمثل هذه الاتفاقات عادةً ما تكون سرية، فإن هذا الرقم ورد في الصحافة ولم يجري تأكيده علناً من الطرفين. طالع "OpenAI, WSJ Owner News Corp Strike Content Deal Valued at Over $250 Million" (وول ستريت جورنال، 22 مايو 2024) <https://www.wsj.com/business/media/openai-news-corp-strike-deal-23f186ba> تاريخ النفاذ: 3 يوليو

للإيرادات لمكتبات الصور الكبرى: فقد أبلغت شركة Shutterstock, Inc.في نتائجها للسنة الكاملة 2025 عن إيرادات من البيانات والتوزيع والخدمات بلغت 203.3 مليون دولار أمريكي، أي ما يعادل نحو 20.5 في المائة من إجمالي الإيرادات البالغ 989.9 مليون دولار أمريكي؛ وأبلغت Getty Images على حدة عن إيرادات للسنة الكاملة 2025 بلغت 981.3 مليون دولار أمريكي. وقد أعلنت الشركتان اندماجاً بين طرفين متكافئين في 7 يناير 2025؛ وأعلنت Getty Images في 30 يونيو 2026 أنها ستنهي اتفاق الاندماج ، بعد أن اشترطت هيئة المنافسة والأسواق في المملكة المتحدة الموافقة على الاندماج بالتخارج من أعمال Shutterstock في مجال المحتوى التحريري. وأُفيد بأن الاتفاق أُنهي في 7 يوليو 2026.[50] ويتجلى الطابع الانتقائي لهذا السوق في الصحافة الفرنسية: ففي مارس 2024، أبرمت صحيفة لو موند اتفاقاً مع OpenAI وصفته بأنه الأول بين وسيلة إعلام فرنسية ومطوّر رئيسي للذكاء الاصطناعي التوليدي.[51] وفي وقت لاحق من العام نفسه، سعت رابطة الصحافة الإخبارية العامة (APIG) إلى إجراء مناقشات جماعية مع OpenAI، وإن كانت التقارير المتاحة تشير إلى أن تلك المناقشات لم تفضِ إلى ترتيب ترخيص أوسع نطاقاً في ذلك الوقت.[52] وقد بدأت يظهر رد فعل على عدم التماثل هذا في شكل وسطاء يجمعون حقوق صغار الناشرين: إذ أفادت إحدى الجهات المزوّدة للمعلومات التجارية بأن عدة آلاف من الناشرين المنضوين ضمن شبكتها، التي تضم عدداً أكبر بكثير من الناشرين، قد اختاروا منح تراخيص لاستخدام محتواهم في الذكاء الاصطناعي،[53] ويفيد أحد مطوّري محركات الإجابات التي تعتمد على المحتوى المرخّص بأنه جمع مجموعة مماثلة من المنشورات المشاركة.[54]

أما السمة الثانية فهي مسألة سلسلة الملكية. ففي كثير من الولايات القضائية، يكتنف الغموض القانوني مسألة إذا ما كانت الحقوق التي يمنحها المؤلف الأصلي للناشر أو هيئة البث أو المستودع بموجب عقد ترخيص أو تنازل أُبرم قبل ظهور الذكاء الاصطناعي التوليدي تشمل استخدام المصنف في تدريب نماذج الذكاء الاصطناعي، إذالم يتضمن العقد نصاً صريحاً بشأن هذا الاستخدام. ولا يقتصر هذا الغموض على الذكاء الاصطناعي؛ فثمة مسائل مماثلة تنشأ في إطار الاتفاقات بين المؤلفين وناشري الموسيقى، وبين فناني الأداء ومنتجي التسجيلات الصوتية، وبين منتجي المصنفات السمعية البصرية والمواهب المتعاقد معها لإنتاج عمل ما، وذلك في الحالات التي تظهر فيها وسيلة استغلال لم تكن متوقعة وقت إبرام العقد؛ أما ما يميّز سياق الذكاء الاصطناعي فهو النطاق الذي تُثار فيه هذه المسألة إلى جانب صعوبة إعادة التفاوض عملياً على الاتفاقات. ففي الصين، على سبيل المثال، يمنح معظم المؤلفين ناشريهم الحق الاستئثاري في النشر المطبوع، وبصورة متزايدة في النشر الرقمي. وتشمل هذه الحقوق حق نسخ المصنف في شكل ورقي، وتوزيع النسخ الورقية، ونسخ النسخ الرقمية وتوزيعها. ولا يُفسَّر أي من هذه الأنواع من الاستخدام بطبيعته على أنه يشمل استخدام المصنف في تدريب نموذج ذكاء اصطناعي. ومن ثم فإن موقف مزودي الذخائر ومطوري الذكاء الاصطناعي الذين يحصلون على تراخيص حق المؤلف من الناشرين هو أن مشروعية الأنشطة المرخَّص بها لا تزال غير مؤكدة فيما يتعلق بالمؤلفين الأصليين للمصنفات. وحتى لو حصل المطوّر على ترخيص من صاحب الحقوق في المؤسسة، يظل خطر مقاضاته من قِبل

2026؛ و"Sam Altman's OpenAI signs content agreement with News Corp" (رويترز، 22 مايو 2024) <https://finance.yahoo.com/news/sam-altmans-openai-signs-content-201808782.html> تاريخ النفاذ: 3 يوليو 2026.

[50] Shutterstock، "Shutterstock Reports Full Year 2025 and Fourth Quarter Financial Results"17 فبراير 2026)<https://investor.shutterstock.com/node/14551/pdf> تاريخ النفاذ: 3 يوليو 2026؛ Getty Images Holdings, Inc, "Getty Images Reports Fourth Quarter and Full Year 2025 Results"16 مارس 2026) <https://investors.gettyimages.com/node/15726/pdf> تاريخ النفاذ: 3 يوليو 2026؛ Getty Images Holdings, Inc و Shutterstock, Inc, "Getty Images and Shutterstock to Merge, Creating a Premier Visual Content Company"7 يناير 2025) <https://www.sec.gov/Archives/edgar/data/1898496/000110465925001732/tm252186d1_ex99-1.htm> تاريخ النفاذ: 3 يوليو 2026. وبشأن إنهاء الاندماج، طالع هيئة المنافسة والأسواق، "Getty abandons Shutterstock merger plans" (يوليو 2026) <https://www.gov.uk/government/news/getty-abandons-shutterstock-merger-plans> تاريخ النفاذ: 31 يوليو 2026.

[51] وصفت لو موند الاتفاق، الذي أُبرم في مارس 2024، بأنه الأول بين وسيلة إعلام فرنسية ومطوّر رئيسي لنماذج الذكاء الاصطناعي التوليدي. طالع "Le Monde signs artificial intelligence partnership agreement with OpenAI" (لو موند، 13 مارس 2024) <https://www.lemonde.fr/en/about-us/article/2024/03/13/le-monde-signs-artificial-intelligence-partnership-agreement-with-open-ai_6615418_115.html> تاريخ النفاذ: 3 يوليو 2026.

[52] الجهة المعنية هي رابطة الصحافة الإخبارية العامة (APIG) التي تقدمت بطلب رسمي مع نقابة ناشري الصحافة الدورية (SEPM)، لفتح مفاوضات مع OpenAI في يونيو 2024؛ وبحسب ما ورد في التقارير رد المطوّر ردّ بأنه لا يرى حاجة إلى إبرام اتفاقات ترخيص جديدة مع الناشرين الفرنسيين في تلك المرحلة. طالع "Intelligence artificielle : un accord de partenariat entre Le Monde et OpenAI" (لو موند، 13 مارس 2024) <https://www.lemonde.fr/le-monde-et-vous/article/2024/03/13/intelligence-artificielle-un-accord-de-partenariat-entre-le-monde-et-openai_6221836_6065879.html> تاريخ النفاذ: 3 يوليو 2026.

[53] أفادت شركة Factiva التابعة لداو جونز بأن نحو 5,000 ناشر، من أصل شبكة تضم نحو 30,000 ناشر، اختاروا بحلول نهاية عام 2025 ترخيص محتواهم للاستخدام في نظم الذكاء الاصطناعي التوليدي. انظر(ي) داو جونز، " Factiva Sees Growing Publisher Participation in AI Licensing" (داو جونز، 2025) <https://www.dowjones.com> تاريخ النفاذ: 3 يوليو 2026.

[54] تصف شركة ProRata.ai شبكتها، بعبارتها هي، بأنها تضم "أكثر من 1,000 جهة نشر ومنشئ محتوى".

المؤلف الأصلي قائماً. وبحسب الولاية القضائية، لا تقتصر المساءلة على المرخِّص؛ فقد يكون المرخِّص مسؤولاً عن الإخلال بالعقد أو بالقانون، وقد يظل المطوّر، بوصفه الطرف الذي يستفيد في نهاية المطاف من الاستخدام، عرضة بدوره لدعاوى التعدي من المؤلفين الأصليين رغم وجود الترخيص. وفي قطاع المصنفات السمعية البصرية، تشير التقارير بأن اتفاقات الترخيص النموذجية تعالج هذا الخطر تعاقدياً: إذ يقرّ المرخِّص ويضمن أنه يحوز جميع الحقوق اللازمة لاستخدام التدريب. وفيما يتعلق بالمقالات الصحفية التي كتبها الصحفيون والمحررون قبل أن يبدأ سريان التعديل ذي الصلة على قانون حق المؤلف الصيني في 1 يونيو 2021، فقد كان حق المؤلف فيها يعود عموماً إلى الصحفي أو المحرر، لا إلى وكالة الأنباء. وهذا يعني أنه حتى الترخيص الصادر عن وكالة الأنباء لم يكن ليجيز استخدام المقال في تدريب الذكاء الاصطناعي. وقد تعيق قيود من هذا النوع بحد ذاتها ظهور سوق للترخيص الطوعي، لأن الترخيص الصادر عن وسيط تكتنف حقوقه ذاتها شكوك يضيف حلقة أخرى من التعدي المحتمل إلى سلسلة الملكية. كما أن لهذه المسألة بُعداً يتعلق بالتوزيع، فحيثما يتم الترخيص عبر وسطاء من ناشرين وهيئات بث ومستودعات، فإن التفاوت في القوة التفاوضية بين المؤلفين الأفراد أو فناني الأداء من جهة، وتلك المؤسسات من جهة أخرى، قد يحول دون وصول عوائد الترخيص إلى الأفراد الذين تُستخدم مصنفاتهم، ويكتسب هذا الاعتبار أهمية خاصة عند المفاضلة بين نماذج الترخيص التي يناقشها القسم باء.

بيد أن الممارسات في الصين تبيّن أنه حتى في إطار هذه القيود الهيكلية، يجري الترخيص الطوعي على نطاق واسع في سياقات قطاعية وفي مجالات رأسية متخصصة. وتتعاون محطات الإذاعة والتلفزيون في هونغ كونغ مع مطوّري الذكاء الاصطناعي لتوفير بيانات صوتية عالية الجودة باللغة الكانتونية. وتتوفر الروايات الرقمية من خلال التعاون بين المطورين وناشرين إلكترونيين يتعاقدون مع أعداد كبيرة من الكتّاب. كما تتوفر الصور من خلال التعاون بين المطورين ومستودعات الصور الكبرى، التي تحوز تلك الصور بموجب تنازلات أو تراخيص استئثارية من المصورين. وتتاح الأوراق العلمية من خلال التعاون بين المطورين ومشغّلي قواعد البيانات الأكاديمية، الذين يصبحون في بعض الحالات شركاء في تطوير النموذج المتخصص الناتج في مجال بعينه. كما ينتج بعض مزودي ذخيرة البيانات موادتدريب خاصة بهم، فيتعاقدون مع مصورين وعارضين ومعلّمي مدارس لنقل المصنفات الناتجة أو ترخيصها لاستخدامها في مجموعات بيانات التدريب.

وتُظهر ولايات قضائية أخرى أنماطاً مماثلة، وإن اختلفت الترتيبات المؤسسية. ففي البرازيل، يتبع النظام القانوني للترخيص الطوعي القواعد العامة لقانون حق المؤلف. ويعني اشتراط الحصول على تصريح مسبق وصريح بموجب المادة 29، والتفسير المقيد لمعاملات حق المؤلف بموجب المواد 4 و31 و49، أنه حتى لو كانت الحقوق المالية قد أُحيلت إحالة كاملة قبل ظهور الذكاء الاصطناعي التوليدي، فإن استخدام المصنف في تدريب الذكاء الاصطناعي سيظل محل نزاع، إذ لم يكن من الممكن أن ينص العقد على وجه التحديد على استخدام تكنولوجيا لم تكن موجودة وقت إبرامه. وتوجد مبادئ مماثلة تقوم على التفسير المقيد للنصوص في النظم القانونية التي تتبع نهج *حق المؤلف*، كما هو الحال في فرنسا وإيطاليا وغيرهما.[55]

وحيثما ينص القانون على حق صريح للمؤلف فيما يتعلق بالعروض المسرحية أو عروض الأداء أو غيرها من المصنفات المحددة، قد يكون مشهد الترخيص الطوعي أكثر وضوحاً لأن الاستحقاقات الأساسية تكون محددة بجلاء أكبر. أما في حالة المصنفات التي يشارك في إبداعها عدد كبير من المساهمين، مثل المصنفات السمعية والبصرية والتسجيلات الصوتية- فقد تزداد الأمور تعقيداً؛ إذ قد يؤدي الغموض في النصوص القانونية أو في العقود المبرمة بين المنتجين والمساهمين الأفراد إلى عرقلة الترخيص الطوعي، أو توجيه عوائده بعيداً عن هؤلاء المساهمين. ويُعد القانون السنغالي، الذي يعترف صراحةً بأن العروض المسرحية مصنفات محمية،[56] مثالاً على ذلك، على الرغم من أن التطبيق العملي لهذا الاعتراف على تدريب الذكاء الاصطناعي لم يُختبر بعد. في كينيا، يعترف قانون حق المؤلف (الفصل 130) بمجموعة قياسية من الحقوق الاستئثارية وينص على الإدارة الجماعية لحق المؤلف من خلال منظمات الإدارة الجماعية، ولكن البنية التحتية المؤسسية لم تُصمم للاستخدام الحاسوبي، كما أن الترخيص الطوعي لتدريب الذكاء الاصطناعي لم يصبح بعد ممارسة موثّقة. وفي كندا، يُعدُّ الترخيص الطوعي الآلية الافتراضية، ويُعد التفاوض على العقود النشاط المحوري. وقد نشأت الدعوى القضائية الرئيسية الحالية، شركة Toronto Star Newspapers Ltd ضد شركة OpenAI Inc.، في سياق نزاع حول الاستخدام غير المصرح به وليس في سياق الترخيص بحد ذاته، ويتم مناقشتها في الجزء الثالث من هذه الدراسة.

ومن السمات المتكررة للترخيص الطوعي، الفردي أو الجماعي، تكييف شروط الترخيص مع نمط استهلاك البيانات في النظام الذي تُطلب البيانات من أجله، ويختلف هذا النمط بشكل كبير عبر فئات أنظمة الذكاء الاصطناعي التوليدي المستخدمة حالياً. عندما تُستخدم البيانات مرة واحدة فقط، في التدريب الأولي أو الضبط الدقيق لنموذج يتم بعد ذلك تثبيت أوزانه، فإن الهيكل الشائع هو

---

[55] فيما يتعلق قاعدة الملكية السابقة لعام 2021 في قانون حق المؤلف الصيني للمقالات الصحفية التي كتبها الصحفيون والمحررون، يُرجى مطالعة قانون حق المؤلف في جمهورية الصين الشعبية (بصيغته المعدلة في عام 2020، الذي أصبح نافذاً اعتباراً من 1 يونيو 2021)، المادة 18 (المصنفات المُعدَّة بالتعاقد).

[56] قانون حق المؤلف السنغالي (القانون رقم 2008-09 بتاريخ 25 يناير 2008)، المادة 2. وتوجد أحكام مماثلة في اتفاق بانغي المتعلق بالمنظمة الأفريقية للملكية الفكرية (القانون المنقح لعام 2015)، الملحق السابع.

رسم يُدفع مرة واحدة أو على أساس حصة من الإيرادات مقابل ذخيرة لغوية محددة. وفي الحالات التي يحتاج فيها النظام إلى النفاذ المستمر إلى المواد المرخصّة عند توليد مخرجات، مثل التوليد المعزَّز بالاسترجاع والأنظمة التي تجري تحليلات مستمرة للنصوص والبيانات على مكتبة ما، لا يكون النفاذ حدثاً منفرداً، بل علاقة مستمرة. وتعكس نماذج التعويض التي تطورت في الممارسة العملية ذلك، بدءاً من الرسوم المستمرة وصولاً إلى الرسوم المفروضة على كل عملية زحف أو كل عملية استدلال. وبصرف النظر عن السعر، فإن النقاط الرئيسية التي يتعين على الأطراف الاتفاق عليها هي نطاق الاستخدام المسموح به، ومدة الترخيص، ونطاقه الجغرافي، وما يحدث في حالة إنهائه، وما إذا كان يمكن التنازل عن النموذج لجهات خارجية. ويثير سياق الذكاء الاصطناعي تساؤلاً حول ما إذا كان إنهاء الترخيص يعني أن على المطور التوقف عن استخدام نموذج تم تدريبه بالفعل على المواد المرخصة، أو إعادة تدريبه، أو إتلاف النسخ المخزّنة، حيث لا يتناسب أي من ذلك بشكل واضح مع سبل الانتصاف التي وُضعت لعلاقات الترخيص السابقة. ويُجسِّد الإنهاء هذه الصعوبة: ففي الترخيص التقليدي، يُعيد الإنهاء السيطرة لصاحب الحقوق على الاستخدامات المستقبلية؛ لكن صاحب الحقوق الذي يُنهي ترخيصاً لتدريب الذكاء الاصطناعي لا يمكنه، من الناحية العملية، إلزام المطوّر على التوقف عن استخدام ما تم تثبيته بالفعل في النموذج، ولا سيما عندما تكون المجموعة المعنية صغيرة مقارنةً بالذخيرة اللغوية؛ ولا بد أن يعمل سبيل الانتصاف، إن أمكن، من خلال إعادة التدريب أو محاولات إلغاء التعلم الآلي، والتي تمت الإشارة إلى حدودها في الجزء الأول. وحتى في الحالات التي تثبت فيها جدوى مثل هذه التدخلات، فإنها تستتبع تكاليف تشغيلية وتكاليف معالجة، قد يتعين على الأطراف تخصيصها في الاتفاقية ذاتها، مما يضيف صعوبة إضافية إلى التفاوض بشأن مثل هذه التراخيص. وعندما يتم الإفصاح عن أوزان النماذج هذه أو بيعها للمستخدمين في المراحل اللاحقة، تبرز مسألة قابلية النقل: هل يترتب على أوزان النماذج الالتزام التعاقدي المرتبط بترخيص ما؟ وثمة مسألةٌ ذات صلة ولكنها مستقلة، وهي ما إذا كان من الممكن احتساب المكافأة في المراحل اللاحقة انطلاقاً من النموذج ذاته، على سبيل المثال بالرجوع إلى مساهمة المواد المرخصة في الأوزان. ولم تُحسم هذه المسائل بعد في العقود المنشورة، وهي ذات صلة بالاختيار المحتمل بين نماذج المكافآت التي تمت مناقشتها في القسم باء 6 أدناه.

وفيما يتعلق بالقطاع عالي القيمة، اتخذ الترخيص الطوعي في قطاعي الموسيقى والوسائط السمعية البصرية في الغالب شكل شراكات فردية تربط الترخيص بشروط تجارية تتجاوز بكثير مجرد استخدام المصنفات المحمية. وخلال عام 2025، تحوّلت كبرى شركات التسجيلات الموسيقية من التقاضي إلى الترخيص مع كبار مطوّري أدوات الموسيقى القائمة على الذكاء الاصطناعي: فقد سوّى اثنان من هؤلاء المطوّرين دعاوى التعدي المرفوعة ضدهما واتفقا على بناء منصات جديدة مرخَّصة، حيث تضمنت إحدى التسويتين استحواذ المطور على خدمة لاكتشاف الحفلات الموسيقية من شركة التسجيلات، كما اتفقت إحدى خدمات البث بشكل منفصل مع شركات التسجيل الثلاث الكبرى واثنين آخرين من أصحاب الحقوق لتطوير منتجات موسيقية قائمة على الذكاء الاصطناعي على أساس الاشتراك الطوعي.[57] وقد سبقت هاتين التسويتينِ بياناتُ مبادئ جماعية في القطاع: فقد حددت تحالفات تضم هيئات صناعية ومنظمات فنية وشركات تكنولوجية مبادئ تشترط الحصول على إذن وترخيص لاستخدام المستودع المحمي في تطوير الذكاء الاصطناعي، وتدعو إلى الشفافية في ممارسات التدريب.[58] وفي قطاع المحتوى السمعي البصري، أُعلن في ديسمبر 2025 عن ترتيب ترخيص يقضي بإتاحة استوديو لأكثر من مئتي شخصية من شخصياته على منصة لتوليد مقاطع الفيديو، وأن يستحوذ على حصة ملكية بقيمة مليار دولار أمريكي، وأُفيد بأن هذا الترتيب أُلغي في مارس 2026، عندما سحبت الشركة المطوّرة المنصة.[59] تتعلق هذه الترتيبات جزئياً بوضع المخرجات، وهو أمر يخرج عن نطاق هذه الدراسة؛ وإنما أُشير إليها هنا على سبيل التوضيح فحسب، بوصفها أمثلة على هياكل الترخيص المخصَّصة وعالية القيمة التي بدأت تظهر حالياً، ولما تتسم به من هشاشة.

---

[57] سوّت كلٌّ من Universal Music Group و Warner Music Group دعواها ضد شركة Udio في أواخر عام 2025، وأبرمتا ترتيبات ترخيص بشأن منصات جديدة؛ كما سوّت Warner Music Group مع شركة Suno، ترتيباً شمل استحواذ Suno على خدمة اكتشاف الحفلات الموسيقية Songkick. كما أبرمت شركة تطوير أخرى، وهي KLAY Vision، اتفاقيات مع كبرى شركات التسجيلات الموسيقية وشركات النشر التابعة لها. أعلنت Spotify عن هذا التعاون في مجال البث، مع Sony Music Group و Universal Music Group و Warner Music Group و Merlin و Believe، في أكتوبر 2025.

[58] وتُصيغ حملة Human Artistry Campaign، التي تشكّلت في عام 2023 بواسطة تحالف يضم أكثر من 150 منظمة، من بينها هيئات كبرى في صناعة الموسيقى ومجموعات تمثّل الفنانين، سبعة مبادئ أساسية <https://www.humanartistrycampaign.com> جرى النفاذ بتاريخ 25 أغسطس 2026؛ مبادئ الإبداع الموسيقي باستخدام الذكاء الاصطناعي، التي طرحتها شركة رولاند ومجموعة يونيفرسال الموسيقية وأيدتها أكثر من خمسين منظمة <https://aiformusic.info> جرى النفاذ بتاريخ 25 أغسطس 2026. تتم مناقشة هاتين المبادرتين ووضعهما في سياق إطار عمل أوسع للتصميم المسؤول في بحث دي بيرارديتيس وآخرين، بعنوان " RAIM: Operationalising Trustworthy AI in Music Generation"، المشار إليه في الجزء الأول، الحاشية رقم 17، من هذه الدراسة.

[59] كان الترتيب المبرم بين شركة والت ديزني وشركة OpenAI، الذي أُعلن عنه في 11 ديسمبر 2025، يتعلق بأكثر من 200 شخصية واستثماراً مخططاً في حصص الملكية بقيمة مليار دولار أمريكي؛ وقيل إنه تم إلغاؤه عقب سحب شركة OpenAI منتجها Sora المخصص لتوليد مقاطع الفيديو في 24 مارس 2026، دون أن يتم إتمام أي استثمار.

(ب) النماذج الرئيسية للترخيص الجماعي (الطوعي، والموسع، والإلزامي)

والشكل الأكثر شيوعاً للترخيص الجماعي هو الإدارة الجماعية الطوعية. تُعهد إلى منظمة الإدارة الجماعية، بموجب عقد، إدارة حقوق معينة نيابةً عن أعضائها؛ وتتفاوض بشأن عقود الترخيص وتبرمها في نطاق صلاحياتها؛ ويتم توزيع العائدات على أعضائها. تضم معظم الولايات القضائية التي تتمتع بنظام متطور لحق المؤلف منظمات إدارة جماعية تطوعية، ولكن الحقوق والمصنفات التي تديرها تختلف اختلافاً كبيراً. وتعد جمعيات الأداء العلني للموسيقى المثال الأكثر شيوعاً على المستوى العالمي. في الصين، توجد خمس منظمات إدارة جماعية تعمل الآن أو في مرحلة بدء العمل الكامل: الجمعية الصينية لحق المؤلف في مجال الموسيقى، وجمعية حق المؤلف الصوتية والمرئية في الصين، والجمعية الصينية للمصنفات المكتوبة، وجمعية حقوق الصور في الصين، والجمعية الصينية لحق المؤلف للأفلام.[60] وقد أعربت منظمات الإدارة الجماعية علناً عن استعدادها لمنح تراخيص لتدريب الذكاء الاصطناعي، ولكن اعتباراً من تاريخ إعداد هذه الدراسة، لم يتم إبرام أي اتفاقية ترخيص جماعي كبرى لتدريب الذكاء الاصطناعي في الصين.[61] والسبب الهيكلي، في السياق الصيني، هو أن ولاية منظمة الإدارة الجماعية تقتصر على حقوق أعضائها، ولا يزال المطوّر يواجه مخاطر متبقية فيما يتعلق بالمنصفات التي لا ينتمي مؤلفوها الأساسيون إلى المنظمة.

ويمكن ملاحظة سمات مماثلة في جمهورية كوريا، حيث توجد جمعية حق المؤلف والنشر للموسيقى الكورية والهيئات المماثلة. يبدو أن الطبيعة المحدودة للمستودع الجماعي ليست سوى جزء من التفسير لفشل إبرام التراخيص. في ظل الأنظمة التنظيمية الحالية، قد تكون لدى المطورين حوافز محدودة للسعي للحصول على التراخيص طواعيةً. يبدو أن ثمة عاملين آخرين يفسران سبب استمرار إبرام الصفقات في الولايات المتحدة رغم ذلك. أولاً، تتيح بنية السوق لعدد قليل من كبار أصحاب الحقوق ومزوِّدي المحتوى التفاوض مباشرة مع المطوِّرين. ثانياً، يشكِّل خطر التعرض للتقاضي، بما في ذلك التعويضات القانونية، حافزاً للتوصل إلى اتفاق. وتُعدّ اتفاقات الترخيص المرتبطة بالتسويات، المبيَّنة في الفرع باء 2، أمثلة على ذلك. ولذلك فإن تفسير عدم اكتساب الترخيص زخماً حتى الآن في الصين وجمهورية كوريا يجب أن يوازن بين هذه العوامل السوقية والقانونية وبين قيود التفويض الجماعي. في تاريخ إجراء هذه الدراسة، كانت وزارة الثقافة والرياضة والسياحة واللجنة الكورية لحق المؤلف في جمهورية كوريا تتشاركان في هدف تحسين الإطار التنظيمي.

ومن الناحية المؤسسية، تمتلك الولايات المتحدة شبكة واسعة من هيئات الإدارة الجماعية القطاعية. تتولى منظمات حقوق الأداء (الجمعيّة الأمريكيّة للملحّنين والمؤلّفين والناشرين وشركة البث الموسيقي وشركة سيساك وجلوبال ميوزيك رايتس وAllTrack) إدارة حقوق الأداء العلني للمصنفات الموسيقية؛ في حين تتولى منظمة Mechanical Licensing Collective، التي أُنشئت بموجب قانون تحديث الموسيقى لعام 2018، وإدارة التراخيص الميكانيكية الشاملة للاستخدامات الرقمية للمصنفات الموسيقية؛[62] وتدير منظمة SoundExchange الإتاوات القانونية المتعلقة بالأداء الرقمي للتسجيلات الصوتية بموجب المادة 114 من قانون حق المؤلف؛ وتدير شركة Clearance Center, Inc. (CCC) حقوق النسخ في المواد النصية المنشورة؛ وتدير جمعية حقوق الفنانين (ARS) حقوق النسخ المعينة في الفنون البصرية. تعمل كل منظمة في إطار قطاعها وبشروط يتم التفاوض عليها مع الجهات التي تمثلها. استجابةً لظهور تدريب الذكاء الاصطناعي كأحد أوجه استخدام المصنفات المحمية، قامت شركة CCC بتوسيع نطاق ترخيصها السنوي للشركات ليشمل بعض استخدامات الذكاء الاصطناعي، وأعلنت عن خطط لإصدار عدة تراخيص إضافية مخصصة للذكاء الاصطناعي.[63]

وبعيداً عن الصين والولايات المتحدة، شهدت السنتان الماضيتان اتخاذ منظمات الإدارة الجماعية والوسطاء الآخرين العديد من الخطوات الملموسة لإدراج تدريب الذكاء الاصطناعي ضمن صلاحيات الترخيص لديهم، ولم يُعلن عن العديد منها إلا في أواخر عام 2025. في السويد، أطلق مكتب الموسيقى الدولي للملحنين السويديين ترتيب ترخيص جماعي في سبتمبر 2025، بمشاركة شركة Songfox المتخصصة في الموسيقى القائمة على الذكاء الاصطناعي، ومُزوّد خدمات الإسناد Sureel، الذي تتيح تكنولوجيته تتبع المخرجات المولَّدة بالذكاء الاصطناعي وإسنادَها إلى المصنفات التي أثّرت فيها.[64] وفي النرويج، أبرمت جمعية كوبينور اتفاقية في

---

[60] ويُشار إلى المنظمات الثلاث الأولى عادةً بالاختصارات MCSC وAVCA وCWWCS؛ أما جمعية حقوق الصور في الصين والجمعية الصينية لحق المؤلف للأفلام فهما المنظمتان اللتان لم تدخلا بعد مرحلة العمل الكامل.

[61] وقد أعربت كبرى منظمات الإدارة الجماعية الصينية علناً عن استعدادها لمنح تراخيص لتدريب الذكاء الاصطناعي، لكن لم يتم إبرام أي اتفاقية ترخيص ذات شأن حتى الآن.

[62] قانون أورين ج هاتش-بوب غودلات لتحديث الموسيقى، المنشور القانون العام رقم 115-264، 132 القانون الأساسي 3676 (2018)، المُدرج في الباب 17 من مدونة الولايات المتحدة، المواد 115 وما يليها. تم تعيين Mechanical Licensing Collective بموجب المادة 115(د)(3) من الباب 17 من مدونة الولايات المتحدة.

[63] انظر(ي) www.copyright.com/ai.

[64] الجمعية السويدية لحقوق الأداء () "STIM launches world's first collective AI music licence" (STIM9 سبتمبر 2025)، وفقاً لتقرير صادر عن الاتحاد الدولي لجمعيات المؤلفين والملحنين، https://www.cisac.org/Newsroom/society-news/stim-launches-worlds-first-collective-ai-music-licence. يُغطي ترخيص الاشتراك الطوعي كلاً من تدريب النماذج والاستخدام اللاحق لمخرجات الذكاء

ديسمبر 2025، تسمح بموجبها للمكتبة الوطنية النرويجية باستخدام محتوى الصحف لتدريب النماذج باللغة النرويجية، بتمويل من الدولة النرويجية.[65] في الدانمرك، أبرمت منظمة الإدارة الجماعية للمنشورات الصحفية الدانمركية شراكة ترخيص قائمة على الاشتراك الطوعي مع شركة ProRata AI لاستخدام محتوى أعضائها في محرك إجابات قائم على الاسترجاع، وذلك أيضاً في ديسمبر 2025، في حين تعمل منظمة دانمركية أخرى، وهي Tekst&Node، على تطوير نموذج ترخيص ضمن إطار نظام الترخيص الجماعي الموسع الدانمركي.[66] وفي هولندا، تعاونت جمعية ناشري الأخبار الخاصة، بالتعاون مع وكالة الأنباء ANP، مع منظمات عامة لتطوير نموذج لغوي محلي، هو GPT-NL، منذ يوليو 2025.[67] وفي إسبانيا، انضم المركز الإسباني لحقوق النسخ إلى مبادرة الترخيص البسيط للغاية لتطوير نموذج ترخيص قائم على الذكاء الاصطناعي للسوق الإسبانية.[68] وفي المملكة المتحدة، تعمل وكالة ترخيص حق المؤلف، بالتعاون مع خدمات ترخيص الناشرين وجمعية التراخيص وجمع الإتاوات للمؤلفين، على تطوير ترخيص جماعي مخصص للذكاء الاصطناعي التوليدي، يشمل التدريب، والضبط الدقيق، والتوليد المعزّز بالاسترجاع.[69] وفي أماكن أخرى، أدرجت منظمات حقوق النسخ حقوق إعادة استخدام الذكاء الاصطناعي في تراخيص حق المؤلف الرقمية الحالية الخاصة بها أو تطوّر حلولاً مماثلة: ففي اليابان، أضافت الجمعية الأكاديمية اليابانية هذه الحقوق إلى ترخيص حق المؤلف الرقمي الخاص

الاصطناعي؛ ويجمع التعويض بين دفعة مقدمة وحصة من إيرادات المطوّر، ويتتبع مُزوّد خدمات الإسناد Sureel المخرجات المولّدة بالذكاء الاصطناعي ويسندها إلى المصنفات التي أثّرت فيها.

[65] "Kopinor signs historic agreement on newspaper content for AI in Norway" (19 ديسمبر 2025)، وفقاً لتقرير صادر عن الاتحاد الدولي للمنظمات المعنية بحقوق النسخ، https://ifrro.org/page/article-detail/ifrro-member-kopinor-signs-historic-agreement-on-newspaper-content-for-ai-in-norway/. وقد أبرمت جمعية كوبينور هذه الاتفاقية نيابة عن رابطة شركات الإعلام النرويجية، وتمنح الاتفاقية المكتبة الوطنية النرويجية حق النفاذ إلى محتوى الصحف التي مضى على نشرها أكثر من عام، وذلك لتدريب النماذج باللغة النرويجية، بتمويل من الدولة النرويجية بمبلغ 45 مليون كرونة نرويجية سنوياً. وهو يختلف عن اتفاقية Bookshelf السابقة المتعلقة بنشر الكتب رقمياً.

[66] شركة ProRata AI ومنظمة الإدارة الجماعية لمنشورات الصحافة الدانمركية (DPCMO)، "ProRata Partners with Danish Publishers Group DPCMO" (10 ديسمبر 2025)، https://www.prnewswire.com/news-releases/prorata-partners-with-danish-publishers-group-dpcmo-to-launch-the-first-decentralized-sovereign-ai-answer-engine-302637182.html. أبرمت منظمة الإدارة الجماعية لمنشورات الصحافة الدانمركية، التي تُصرّح بأنها تمثل 99 في المائة من قطاع الأخبار الدانمركي، شراكة ترخيص جماعي قائمة على الاشتراك الطوعي، تُمكّن من استخدام محتوى أعضائها في محرك الإجابات القائم على الاسترجاع لدى لشركة ProRata على أساس تقاسم الإيرادات. وتطور منظمة دانمركية أخرى (Tekst&Node) نموذجاً لترخيص الذكاء الاصطناعي ضمن إطار الترخيص الجماعي الموسع الدانمركي، يسمح لأصحاب الحقوق بترخيص استخدامات معينة للمصنفات الأدبية في تنقيب النصوص والبيانات وفيما يتعلق بالذكاء الاصطناعي، شريطة إتاحة خيار الانسحاب والموافقة من جانب وزارة الثقافة. انظر(ي): تيريز فوغيد، "[DK] First extended collective licence approval for text and data mining and the use of works by AI" (ملاحظات IRIS القانونية الصادرة عن المرصد الأوروبي للوسائط السمعية البصرية، 2026) <HTTPS://merlin.obs.coe.int/article/10505> جرى النفاذ بتاريخ 3 يوليو 2026.

[67] NDP Nieuwsmedia وTNO، "Large dataset news organizations for Dutch AI language model GPT-NL" (17 يوليو 2025)، https://www.ndpnieuwsmedia.nl/2025/07/17/large-dataset-news-organizations-for-dutch-ai-language-model-gpt-nl/. يعمل أعضاء NDP Nieuwsmedia، وهي جمعية ناشري الأخبار الخاصة الهولندية، بالتعاون مع وكالة الأنباء ANP، على إتاحة أجزاء من أرشيفاتهم لتدريب GPT-NL، وهو نموذج لغوي غير ربحي باللغة الهولندية طورته مؤسسات TNO وNFI وSURF؛ على أن يتلقى الناشرون المساهمون مكافأة عند إطلاق النموذج.

[68] المركز الإسباني لحقوق النسخ، بيان صحفي بشأن اتفاقية تعاونه مع RSL Collective (يناير 2026)، https://www.cedro.org/sala-de-prensa/noticias. انضم المركز الإسباني لحقوق النسخ، وهو المنظمة الإسبانية المعنية بحقوق النسخ، إلى RSL Collective بصفتها شريكاً له في إسبانيا، وتشارك في اللجنة الفنية التابعة لـ RSL، حيث تعمل على تكييف معيار الترخيص القابل للقراءة آلياً مع السوق الإسبانية، حتى يتسنى لأعضائها تحديد الشروط التي تُستخدم بموجبها مصنفاتهم في تدريب الذكاء الاصطناعي. وقد نشر المركز الإسباني لحقوق النسخ بشكل منفصل توصيات بشأن التنازل عن الحقوق المتعلقة بتدريب الذكاء الاصطناعي في 16 ديسمبر 2025.

[69] وكالة ترخيص حق المؤلف، "CLA announces development of Generative AI Training Licence" (2025)، https://cla.co.uk/development-of-cla-generative-ai-licence/؛ خدمات ترخيص الناشرين "PLS and ALCS agree to development of pioneering CLA generative AI licence" (أبريل 2025)، https://www.pls.org.uk/news-events-policy/news/pls-and-alcs-agree-to-development-of-pioneering-cla-generative-ai-licence/. تعمل وكالة ترخيص حق المؤلف، بالتعاون مع جمعية ترخيص الناشرين وجمعية التراخيص وجمع الإتاوات للمؤلفين، على تطوير ترخيص جماعي تطوعي لاستخدام النصوص في الذكاء الاصطناعي التوليدي، يشمل التدريب، والضبط الدقيق، والتوليد المعزّز بالاسترجاع؛ ويشارك أصحاب الحقوق من خلال تسجيل مصنفاتهم.

بها في أبريل 2025؛[70]وتعمل وكالة حق المؤلف في أستراليا على تطوير حلول مماثلة؛[71]وفي ألمانيا، وسعت مؤسسة VG Wort نطاق ترخيص حق المؤلف الرقمي الخاص بها ليشمل إعادة استخدام الذكاء الاصطناعي، في الوقت الذي اقترحت جمعية التحصيل جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي (GEMA) على نحوٍ منفصل نموذجاً مخصصاً لم يكن قد إبرام أي اتفاق حتى تاريخ إعداد هذه الدراسة.[72] لا تزال هذه التطورات متفاوتة، وتكتسي، في حالات عدة، طابعاً تجريبياً، لكنها تشير إلى أن مسار الترخيص الجماعي يخضع للاختبار الفعلي وليس مجرد موضوع نقاش.

ويُعد الترخيص الجماعي الموسع نموذجاً للترخيص نشأ في الولايات القضائية في البلدان الشمالية في ستينيات القرن العشرين، واعتُمد في عدة بلدان أخرى. ويعالج هذا النموذج مشكلة المخاطر المتبقية عن طريق مدّ أثر ترخيص منظمة الإدارة الجماعية ليشمل أصحاب الحقوق غير الأعضاء ضمن مستودع محدَّد، ويتيح حق الانسحاب. وقد ذُكر هذا النموذج في مناقشات السياسات الأوروبية بشأن تدريب الذكاء الاصطناعي باعتباره وسيلة لتخليص الحقوق في المجموعات النصية والمرئية الكبيرة التي يحتاج المطوِّر إلى استخدامها ولا يكون الترخيص الفردي لها عملياً. وقد أشارت إليه صراحةً دراسة أجراها البرلمان الأوروبي عام 2025 حول الذكاء الاصطناعي التوليدي وحق المؤلف.[73] في الولايات المتحدة، حدّد مكتب حق المؤلف، في الجزء الثالث من تقريره عن حق المؤلف والذكاء الاصطناعي، الترخيص الجماعي الموسَّع كخيار وسيط مستهدف ممكن لتدريب الذكاء الاصطناعي، يختلف عن الترخيص الإلزامي.[74] وينعكس هذا الاقتراح ذاته، مع بعض الاختلافات، في الأدبيات الأكاديمية المتعلقة بالمكافآت: إذ يقترح مصنف الأستاذ سينفتليين حول الذكاء الاصطناعي التوليدي ومكافأة المؤلفين، نظاماً تقوم بموجبه منظمات الإدارة الجماعية بتحصيل مكافآت مقابل استخدام خدمات الذكاء الاصطناعي التوليدي التي تحلّ مخرجاتها محل مصنفات مؤلفين يمكن تحديد هويتهم أو تتنافس معها، وتقوم بتوزيع العائدات على أعضائها.[75]

بالإضافة إلى الأمثلة المذكورة سابقاً، ثمة مثالان آخران يوضحان إمكانية تكييف الإدارة الجماعية الطوعية الحالية لأغراض تدريب الذكاء الاصطناعي، من خلال العمل ضمن الهياكل المؤسسية القائمة بدل انتظار تشريعات جديدة. وفي البرازيل، يتركز نظام الأداء العام للموسيقى في مكتب واحد وهو المكتب المركزي للتحصيل والتوزيع، حيث يحظى بدعم قانوني وقضائي لتجميع وتحصيل وتوزيع إتاوات الأداء العلني عن المصنفات الموسيقية والتسجيلات الصوتية، سواء فيما يخص حقوق المؤلفين أو الحقوق

---

[70] الجمعية الأكاديمية اليابانية لتخليص حق المؤلف (JAC) وشركة RightsDirect Japan، "JAC and RightsDirect Japan Announce the Availability of AI Re-Use Rights for Digital Copyright License" (9 أبريل 2025)، https://www.businesswire.com/news/home/20250409032666/en/. أضافت الجمعية الأكاديمية اليابانية حقوق إعادة استخدام الذكاء الاصطناعي إلى ترخيص حق المؤلف الرقمي الخاص بها، وهو ترخيص للاستخدام الداخلي على مستوى المؤسسة للشركات التي يقع مقرها الرئيسي في اليابان، ويستند إلى مستودع مصنفات مقدمة من الناشرين والجمعيات الأكاديمية في جميع أنحاء العالم.

[71] وكالة حق المؤلف، "AI and Copyright in Australia" والأسئلة والأجوبة المصاحبة لها، https://www.copyright.com.au/membership/ai-and-copyright-in-australia/. تُفيد وكالة حق المؤلف، وهي عضو في المجموعة المرجعية المعنية بحق المؤلف والذكاء الاصطناعي التابعة للحكومة الأسترالية، بأنها تطوّر حلول ترخيص لاستخدام المواد المحمية فيما يتعلق بالذكاء الاصطناعي التوليدي في أستراليا.

[72] فيما يتعلق بمؤسسة VG Wort، تم توسيع نطاق ترخيص حق المؤلف الرقمي الذي تديره RightsDirect (إحدى الشركات التابعة لشركة CCC) ليشمل إعادة الاستخدام الداخلي لأغراض الذكاء الاصطناعي؛ انظر(ي) https://www.rightsdirect.com/license/vg-wort-digital-license-2/. اقترحت جمعية التحصيل "جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي (GEMA)" نموذجاً للترخيص بشأن الذكاء الاصطناعي التوليدي في عام 2024؛ ولم يكن قد أُبرم أي اتفاق في تاريخ إعداد هذه الدراسة، كما رفعت جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي (GEMA) في الوقت نفسه دعاوى انتهاك ضد مطوّري الذكاء الاصطناعي (انظر(ي) *جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي (GEMA) ضد شركة OpenAI*، *المشار إليه أعلاه*). انظر(ي) جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي (GEMA)، " AI and music: GEMA calls for music creators to share in AI providers' revenues and presents first licensing model" (26 سبتمبر 2024) <HTTPS://www.cisac.org/Newsroom/society-news/ai-and-music-gema-calls-music-creators-share-ai-providers-revenues-and>؛ تم النفاذ بتاريخ 3 يوليو 2026؛ GEMA، "Suno AI and Open AI: GEMA sues for fair compensation" (20 يناير 2025) <https://www.gema.de/en/news/ai-and-music/ai-lawsuit> تم النفاذ بتاريخ 3 يوليو 2026.

[73] البرلمان الأوروبي، المديرية العامة للسياسات الداخلية، دراسة بعنوان "Generative AI and Copyright: Training, Creation, Regulation"، IUST_STU(2025)774095_EN (2025)، وهي دراسة توصي بالترخيص القانوني باعتباره الإطار الأمثل لتدريب الذكاء الاصطناعي في الاتحاد الأوروبي.

[74] مكتب حق المؤلف بالولايات المتحدة، *حق المؤلف والذكاء الاصطناعي، الجزء الثالث: تدريب الذكاء الاصطناعي التوليدي* (2025).

[75] م. ر. ف. سينفتليين، "Generative AI and Author Remuneration"، 54 IIC 1535 (2023).

المجاورة.[76] وفي إطار قانون حق المؤلف القائم، ومن دون تغيير تشريعي، سيكون من الممكن قانوناً ممارسة حق محدد يتعلق باستخدام المصنفات والتسجيلات الصوتية في تدريب نماذج الذكاء الاصطناعي، وفق نموذج مماثل للنموذج المعمول به حالياً بالنسبة إلى حق الأداء العلني للموسيقى. ويقتضي ذلك أن ينضم أصحاب الحقوق إلى منظمة إدارة جماعية وأن تنص هذه المنظمة في نظامها الأساسي على تمثيل صاحب الحقوق فيما يتعلق بحقوق المؤلف الأخرى، لا فيما يتعلق بحقوق الأداء العلني للموسيقى وحدها؛ كما يقتضي إذناً مسبقاً من وزارة الثقافة، وهي الجهة الاتحادية الحكومية المسؤولة عن الإشراف على منظمات الإدارة الجماعية. وتتيح المادة 97 من قانون حق المؤلف البرازيلي لمنظمات الإدارة الجماعية توسيع نطاق عملها ليشمل الترخيص بالحقوق التي ينطوي عليها التدريب وتنفيذ النموذج (حق النسخ، والتبليغ للجمهور أو الإتاحة للجمهور، وحق التكييف في مرحلة المخرجات، بحسب الموقف الذي اتخذته وزارة الثقافة)، شريطة الحصول على التصريحات اللازمة. بيد أن شرط التحصيل المُوحَّد لا ينطبق إلا على حقوق الأداء العام للموسيقى؛ لذا، فحتى لو قام المكتب المركزي للتحصيل والتوزيع أو الشركات التابعة له بتركيز تمثيل حق تدريب الذكاء الاصطناعي داخل الهيئة، فإن ذلك لن يحول دون تشكيل كيانات جماعية أخرى للغرض نفسه.

وفي كينيا، تنظم لوائح حق المؤلف (الإدارة الجماعية) لعام 2020 تسجيل منظمات الإدارة الجماعية وتشغيلها والإشراف عليها.[77] وبالرغم من عمومية الإطار وكونه لم يُوضع خصيصاً ليتناسب مع الاستخدامات الحاسوبية، فإنه وفر أساساً مؤسسياً يمكن البناء عليه لوضع نظام ترخيص جماعي طوعي لتدريب الذكاء الاصطناعي. ويتضمن مشروع قانون حق المؤلف والحقوق المجاورة لعام 2026، الذي نشره مجلس حق المؤلف الكيني في مارس 2026 للتعليق العام، على أن تنتهي مهلة التعليقات في نهاية ذلك الشهر، قسماً (القسم 34) يجيز تحليل النصوص والبيانات لأغراض البحث غير التجاري، لكنه لا يورد تعريفاً قانونياً واضحاً لعبارة "غير تجاري"، ولا يستحدث آلية اعتراض على الاستخدام لأصحاب الحقوق، ولا يتضمن أحكاماً محددة بشأن استخدام المصنفات المحمية بحق المؤلف في تدريب الذكاء الاصطناعي ولا بشأن قابلية المصنفات المولَّدة بالذكاء الاصطناعي للحماية بحق المؤلف. ويُعد الوضع في كينيا مثالاً على نمط أوسع في النظم القانونية التي يجري فيها تكييف إطار حق المؤلف على مراحل؛ فالبنية التحتية القائمة لمنظمات الإدارة الجماعية توفر وسيلة ممكنة للترخيص الجماعي لتدريب الذكاء الاصطناعي، لكن هذا التكييف يقتضي خطوات تشريعية أو تعاقدية إيجابية لم تُتخذ بعد.

3. الترخيص المجتمعي والترخيص القائم على الإشراف للبيانات اللغوية الأولية

تعمل آليات الترخيص الموصوفة أعلاه على بيانات تدريب تكون مثبَّتة ومنشورة بالفعل عند اطلاع المطور عليها. وتنشأ فئة متمايزة من بيانات التدريب على نحو مغاير؛ إذ لا يمكن بناء نماذج الكلام واللغة لمعظم اللغات الأفريقية انطلاقاً من ذخائر نصية ضخمة المستمدة من الويب، نظراً لندرة هذه النصوص. وعوضاً عن ذلك، تُجمع البيانات بوصفها بيانات أولية، مثل تسجيلات صوتية وتفريغات نصية تُجمع من المتحدثين عبر العمل الميداني وجلسات التسجيل المجتمعية والشراكات مع هيئات البث. وينشأ حق المؤلف عند لحظة تثبيت البيانات وليس عند لحظة جمعها آلياً عبر الزحف على الويب. فالتسجيل محمي بوصفه تسجيلاً صوتياً، والنص المفرغ محمي بوصفه مصنفاً أدبياً، وقد يتمتع المتحدث بحقوق فناني الأداء.[78] ومن ثم تبرز مسائل الإذن والتعويض لحظة الجمع، قبل أن تُطرح أي مسألة تتعلق بتنقيب النصوص والبيانات.

وتستدعي هذه الفئة من البيانات أيضاً طبقة إضافية من الملكية الفكرية؛ فتسجيلات التراث الشفهي والحكايات الشعبية والأمثال، تُعد في آن واحد مصنفات محمية بحق المؤلف تؤول إلى من ثبَّتها، وتعابير ثقافية تملك فيها المجتمعات المحلية حقوقاً جماعية بموجب تشريعات فريدة من نوعها. وفي كينيا، يعترف قانون حماية المعارف التقليدية وأشكال التعبير الثقافي، رقم 33 لعام 2016، الذي يُفعِّل المواد 11 و40 و69(1)(ج) من دستور كينيا لعام 2010، بالمجتمعات المحلية بصفتها أصحاب حقوق في تعابيرها الثقافية، ويشترط الموافقة المسبقة المستنيرة وإبرام اتفاقات مع المستخدمين المصرح لهم باستغلالها، وينص على تقديم تعويضات تؤول إلى المجتمعات؛ وتُلزم المادة 11(3)(أ) من الدستور البرلمان بكفالة حصول المجتمعات على الإتاوات مقابل استخدام ثقافاتها وتراثها الثقافي.[79] وحين تدخل بيانات لغوية من هذا النوع في ذخيرة بيانات تدريب، يعمل تسلسل حق المؤلف

---

[76] فيما يتعلق بهيكل نظام الإدارة الجماعية البرازيلي وتركيز تحصيل إتاوات الأداء العلني في المكتب المركزي للتحصيل والتوزيع، طالع قانون حق المؤلف البرازيلي (القانون رقم 1998/9,610)، المواد 97-100، والمرسوم رقم 2018/9,574. وقد أكدت المحكمة الاتحادية العليا دستورية هذا النموذج في (2003) ADI 2.054 وفي الحكم المشترك بشأن ADI 5062 و(2016) ADI 5065.

[77] لوائح حق المؤلف (الإدارة الجماعية) لعام 2020 (كينيا). وترد البنية التحتية للترخيص في القسمين 33 و34 والجزء السابع من قانون حق المؤلف (الفصل 130).

[78] قانون حق المؤلف (الفصل 130)، قوانين كينيا، القسم 2 (الذي يعرّف، ضمن جملة أمور، "التسجيل الصوتي" و"المصنف الأدبي") والأحكام المتعلقة بحقوق فناني الأداء. وبشأن حقوق فناني الأداء على الصعيد الدولي، انظر(ي) المادتين 10 و14 من معاهدة الويبو بشأن الأداء والتسجيل الصوتي، اللتين ترد مناقشتهما في القسم باء.1 أعلاه.

[79] قانون حماية المعارف التقليدية والتعابير الثقافية، رقم 33 لعام 2016 (كينيا)؛ دستور كينيا لعام 2010، المواد 11 و40(5) و69(1)(ج).

والحقوق المستحقة للمجتمع المحلي بالتوازي، لذا لا يكفي الحصول على ترخيص من الجهة التي ثبتت التسجيل لمعالجة الوضع القانوني لذلك المجتمع.

وقد نشأت ممارسة ترخيص مميزة استجابةً لذلك. وتُدخل التراخيص المفتوحة القياسية، ولا سيما التخصيصات للملك العام، هذه البيانات في مسارات التدريب من دون أن تمنح المجتمع المساهم سلطة حوكمة أو عائداً اقتصادياً. أما التراخيص التي تركز على المجتمع فتُعدل هذا الوضع مع الحفاظ على سمة الانفتاح. ويطبق ترخيص البيانات المفتوحة نوليتي أوبودو (Nwulite Obodo) شروط استخدام متدرجة تتباين بحسب السياق الجغرافي والإنمائي للمستخدم، وتتضمن التزامات بتقاسم المنافع؛ وقد طُبق على مجموعات بيانات كلام أفريقية، منها مجموعة بيانات الكلام دونام دولوو (DhoNam Dholuo) التي وُضعت مع جامعة ماسينو.[80] تسمح رخصة إيستو (Esethu)، التي تم تطويرها لمجموعة بيانات الكلام المستمدة من صحيفة فوكوزينزيل بلغة الإيكوسا، بالاستخدام لأغراض البحث وبالاستخدام التجاري من جانب الكيانات الأفريقية من دون مقابل، مع اشتراط رسوم ترخيص من سائر المستخدمين التجاريين، على أن تُعاد العائدات إلى الاستثمار في إنشاء مزيد من البيانات اللغوية.[81] ويجسد ترخيص كايتياكيتانغا (Kaitiakitanga) الصادر عن Te Hiku Media، الذي ينظم البيانات المتعلقة باللغة الماورية، النهج نفسه في منطقة أخرى.[82] وتشكل هذه الأدوات آلية ثالثة، إلى جانب الترخيص الفردي والترخيص الجماعي، يجري من خلالها تنظيم الإذن والتعويض عن استخدام البيانات اللغوية الأولية في تدريب الذكاء الاصطناعي، وتوضح كذلك كيف يمكن لإطار قانوني وطني خاص بالأصول الثقافية المملوكة جماعياً أن يوفر الأساس الذي تستند إليه الممارسة التعاقدية. وفي كينيا، تتفاعل هذه الطبقة أيضاً مع استثناء تحليل النصوص والبيانات المقترح في مشروع قانون حق المؤلف والحقوق المجاورة لعام 2026، الذي ترد مناقشته في القسم الفرعي السابق.

4. الاستثناءات القانونية لتنقيب النصوص والبيانات والأحكام العامة

ويتضمن قانون حق المؤلف في الاتحاد الأوروبي استثناءات صريحة لتنقيب النصوص والبيانات ذات صلة بالعمليات المستخدمة في تدريب الذكاء الاصطناعي، وإن ظل تطبيقها ومدى كفايتها في هذا السياق محل خلاف. وتضع المادتان 3 و4 من التوجيه (الاتحاد الأوروبي) 790/2019 بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة (توجيه CDSM) استثناءين إلزاميين لتنقيب النصوص والبيانات. وتتعلق المادة 3 باستثناء لأغراض البحث العلمي لصالح مؤسسات البحث ومؤسسات التراث الثقافي، من دون إتاحة خيار الاعتراض على الاستخدام. أما المادة 4 فتُعد استثناءً عاماً ينطبق على أي مستفيد، مع جواز اعتراض صاحب الحقوق على الاستخدام شريطة أن يكون "بطريقة مناسبة، مثل الوسائل القابلة للقراءة آلياً في حالة المحتوى المتاح للجمهور على الإنترنت".[83] ويُدرج قانون الذكاء الاصطناعي هذا الإطار عن طريق الإحالة، فيُلزم مقدمي نماذج الذكاء الاصطناعي ذات الأغراض العامة بوضع سياسة للامتثال لقانون حق المؤلف للاتحاد، وبتحديد التحفظات على الحقوق المعبَّر عنها عملاً بالمادة 4(3) من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة، وبنشر ملخص مفصل بقدر كاف للمحتوى المستخدم في التدريب.[84]

---

[80] ترخيص البيانات المفتوحة نوليت أوبودو (Nwulite Obodo Open)، الإصدار 1.0 (مختبر قانون علوم البيانات، جامعة بريتوريا، 2025) <HTTPS://datasciencelawlab.africa/nwulite-obodo-open-data-license/> تاريخ النفاذ 31 يوليو 2026؛ I.C. أوكوري وم. أومينو " Addressing Inequitable Openness in Licences for Sharing African Data and Datasets Through the Nwulite Obodo Open Data Licence"، مجلة "Law, Technology and Humans"(2025) 7(3) ص 94. تُتاح مجموعة بيانات الكلام دونام دولوو (DhoNam Dholuo)، التابعة لمركز الذكاء الاصطناعي الأفريقي بجامعة ماسينو، بموجب تلك الرخصة: https://mozilladatacollective.com/datasets/cmjepxo6t08nmmk07iauvua6v تاريخ النفاذ 31 يوليو 2026.

[81] شركة ليلابا إيه آي (Lelapa AI)، وشركة واي وذ ووردز (Way With Words)، والمجموعة البحثية علم البيانات من أجل الأثر الاجتماعي (Data Science for Social Impact) (جامعة بريتوريا)، " The Esethu Framework: Reimagining Sustainable Dataset Governance and Curation for Low-Resource Languages"، وقائع الاجتماع السنوي الثالث والستين لجمعية اللغويات الحاسوبية (2025) <https://aclanthology.org/2025.acl-long.1487/> تاريخ النفاذ: 31 يوليو 2026.

[82] Te Hiku Media، ترخيص كايتياكيتانغا (Kaitiakitanga) <https://tehiku.nz/te-hiku-tech/te-hiku-dev-korero/25141/data-sovereignty-and-the-kaitiakitanga-license>، تاريخ النفاذ: 31 يوليو 2026.

[83] التوجيه رقم ((EU)790/2019 الصادر عن البرلمان الأوروبي وعن المجلس المعقود في 17 أبريل 2019، بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة ، الجريدة الرسمية L 130/92 ("توجيه CDSM")، المادة 3 (استثناء إلزامي لتنقيب النصوص والبيانات لأغراض البحث العلمي تستفيد منه منظمات البحث ومؤسسات التراث الثقافي) والمادة 4 (استثناء عام يجوز فيه الاعتراض على الاستخدام لتنقيب النصوص والبيانات في المصنفات التي يمكن النفاذ إليها بصورة مشروعة).

[84] اللائحة رقم ((EU)1689/2024 الصادرة عن البرلمان الأوروبي والمجلس المعقود في 13 يونيو 2024، التي تحدد قواعد موحدة بشأن الذكاء الاصطناعي (قانون الذكاء الاصطناعي)، الجريدة الرسمية L، 2024/1689، 12.7.2024، المادة 53(1)(ج) و(د)، والبندان 105 و106.

وقد تناول حكمان ألمانيان تطبيق تلك الاستثناءات على تدريب الذكاء الاصطناعي. في قضية Kneschke ضد LAION e.V، قضت المحكمة الإقليمية في هامبورغ في سبتمبر 2024 بأن استخدام صورة فوتوغرافية في إنشاء مجموعة البيانات LAION-5B، المستخدمة في تدريب نماذج توليد الصور، يندرج ضمن الاستثناء الوارد في المادة 3 من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة، الخاص بتنقيب النصوص والبيانات لأغراض البحث العلمي، بصيغته المعتمدة في القانون الألماني. واعتبرت المحكمة LAION منظمة بحث علمي بالمعنى المقصود في الحكم، ورأت أن التحليل الحاسوبي المطروح يشكّل تنقيب نصوص وبيانات لأغراض علمية. وأشارت المحكمة أيضاً في معرض تعليق عرضي، إلى أن استثناء تنقيب النصوص والبيانات العام الوارد في المادة 4 لن ينطبق في أي حال، وذلك لأن شرط التحفظ على الحقوق، الوارد ضمن شروط استخدام المنصة المصدر، يُعدّ في رأي المحكمة- قابلاً للقراءة آلياً، في ضوء القدرات التي تتمتع بها النماذج اللغوية الكبيرة في الوقت الحالي.[85] وفي الاستئناف، في قضية Kneschke ضد LAION e.V، أيدت المحكمة الإقليمية العليا في هامبورغ النتيجة لكنها قدمت معالجة أكثر تفصيلاً لشرط القابلية للقراءة آلياً بموجب المادة 4(3).[86] ورأت محكمة الاستئناف أن القابلية للقراءة آلياً لا تقتضي فقط أن يكون التحفظ قابلاً لأن تلتقطه آلة، بل أيضاً أن يكون قابلاً للتفسير بواسطة عملية آلية بحيث لا تتم معالجة المحتوى المشمول بالتحفظ. وباختصار، يجب أن يكون التحفظ قابلاً للقراءة آلياً وللتنفيذ آلياً في آن واحد. وأشارت المحكمة كذلك أن القابلية للقراءة آلياً ليست خاصية ثابتة، بل يجب تقييمها في ضوء القدرات التكنولوجية السائدة وقت حدوث الاستخدام محل النزاع؛ وبحسب وقائع هذه القضية، مستوى التطور التقني الذي كان سائداً في النصف الثاني من عام 2021. ولا تحسم حيثيات محكمة الاستئناف مسألة ما يُعدّ قابلاً للقراءة آلياً في عام 2025 وما بعده، حين تصبح النماذج القادرة على استخلاص المعاني من النصوص اللغوية الطبيعية غير المهيكلة متاحة على نطاق واسع.[87]

وقبل ذلك بنحو شهر، في قضية GEMA ضد OpenAI، كانت المحكمة الإقليمية في ميونيخ الأولى قد خلصت إلى نتيجة مختلفة بشأن مسألة ذات صلة لكنها متمايزة.[88] وكانت جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي (GEMA) قد رفعت دعوى ضد OpenAI زاعمة أن ChatGPT حفظ كلمات تسع أغانٍ ألمانية مشهورة، وأنه يمكن استحثاثه على نسخها. واعتبرت محكمة ميونخ أن إدخال كلمات الأغاني في ذخيرة بيانات التدريب وحفظها اللاحق داخل النموذج يشكلان فعلين من أفعال النسخ بالمعنى المقصود في التنفيذ الألماني لتوجيه مجتمع المعلومات؛ ورفضت تطبيق استثناء تنقيب النصوص والبيانات العام الوارد في المادة 44ب من قانون حق المؤلف الألماني، استناداً إلى أن النماذج التوليدية تحفظ كلمات الأغاني حفظاً دائماً وأن ذلك يتجاوز تحليل المعلومات النصية والقائمة على الأنماط الذي وضع الاستثناء للسماح به؛ واعتبرت أن المخرجات التي يعيد النموذج إنتاجها دليلاً على أن كلمات الأغاني مخزنة داخل النموذج على نحوٍ يضاهي تخزينها في مكتبة رقمية. وأصدرت المحكمة أمراً زجرياً، وألزمت بتقديم المعلومات، وقضت بالتعويض عن الأضرار.[89] القرار قيد الاستئناف؛ وقد اعتبرته بعض الأوساط أول حكم أوروبي رئيسي بشأن تدريب الذكاء الاصطناعي لصالح أصحاب الحقوق؛ بينما انتقده آخرون بدعوى أنه اختزل سلسلة من العمليات المتمايزة من الناحية التكنولوجية في فعل نسخ واحد، وأنه أساء تطبيق استثناء تنقيب النصوص والبيانات بخلطه بين الأدلة على الحفظ في مرحلة المخرجات والواقع التقني لعملية التدريب.[90]

---

[85] *Kneschke* ضد *LAION e.*ضد، المحكمة الإقليمية في هامبورغ، 27 سبتمبر 2024 (227/23 O 310) (قضت بأن بناء مجموعة بيانات LAION-5B لاستخدامها في تدريب نماذج توليد الصور يندرج في نطاق التنفيذ الألماني لاستثناء المادة 3 من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة المتعلق بتنقيب النصوص والبيانات لأغراض البحث العلمي (المادة 60d من قانون حق المؤلف الألماني؛ وأشارت أيضاً، في معرض تعليق عرضي، إلى أن استثناء تنقيب النصوص والبيانات العام الوارد في المادة 4 ما كان ليُطبَّق، على أساس أن التحفظ على الحقوق الوارد في شروط المنصة المصدر كان قابلاً للقراءة آلياً بالمعنى ذي الصلة). في مرحلة الاستئناف، طالع *Kneschke* ضد *LAION e.*ضد، محكمة الاستئناف في هامبورغ، 10 ديسمبر 2025 (104/24 U 5)، والتي سيتم مناقشتها أدناه.

[86] *Kneschke* ضد *LAION e.*ضد، محكمة الاستئناف في هامبورغ، 10 ديسمبر 2025 (104/24 U 5).

[87] ذكرت محكمة الاستئناف في هامبورغ أن القابلية للقراءة آلياً تقتضي أن يكون التحفظ "قابلاً لأن تفسره آلة على نحو ... لا يُعالَج معه، في إطار عملية آلية، المحتوى المشمول بالتحفظ [على الحقوق]" (ترجمة المؤلفين). انظر(ي) أيضاً: ب. كيلر، "LAION Round 2: Machine-Readable but Still Not Actionable - The Lack of Progress on TDM Opt-Outs"، (Kluwer Copyright Blog 17 و18 ديسمبر 2025)، الجزء الأول.

[88] GEMA ضد OpenAI، المحكمة الإقليمية في ميونيخ الأولى، 11 نوفمبر 2025 (42 O 14139/24)، حكم نهائي (Endurteil). المدعية هي جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي GEMA؛ أما المدعى عليهما فهما شركتا OpenAI OpCo LLC و OpenAI Ireland Limited. أصدرت المحكمة أمراً زجرياً وبتقديم المعلومات وبدفع تعويضات.

[89] المرجع السابق نفسه. واعتبرت المحكمة أن إدخال كلمات الأغاني في ذخيرة بيانات التدريب وحفظها اللاحق داخل النموذج يشكلان فعلين من أفعال النسخ بالمعنى المقصود في المادة 16 من قانون حق المؤلف الألماني (UrhG)، التي تنفذ المادة 2 من التوجيه 2001/29/EC.

[90] للاطلاع على تعليق نقدي بشأن هذه النقطة، انظر(ي): ج. فروزيو "Copyright in Formaldehyde: How GEMA v OpenAI Freezes Doctrine and Chills AI"، (Kluwer Copyright Blog 10 و11 ديسمبر 2025)، الجزءان 1 و2، حيث يرى الكاتب أن محكمة ميونخ تخلط بين الأدلة على الحفظ في مرحلة المخرجات والواقع التقني للتدريب، وأن الرأي الأرجح هو اعتبار النسخ التقنية الوسيطة المشروعة مندرجة ضمن المادة 5(1) من توجيه مجتمع المعلومات أو استثناءات تنقيب النصوص والبيانات.

وفضلاً عن السوابق القضائية، ثمة خلاف أكاديمي بشأن إذا ما كان استثناء المادة 4 من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة ينطبق أصلاً على تدريب الذكاء الاصطناعي. ويسود في الأدبيات القانونية رأي مفاده أن تدريب الذكاء الاصطناعي شكل من أشكال تنقيب البيانات تشمله المادتان 3 و4 من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة. وكثيراً ما يستند هذا الرأي إلى الإشارة إلى هاتين المادتين في قانون الذكاء الاصطناعي. وفي المقابل، طرح عدد من الباحثين الأوروبيين من بينهم البروفيسور تيم دورنيس وجهة نظر مغايرة؛ إذ يجادل دورنيس، في ورقة بحثية نُشرت في المجلة الأوروبية للملكية الفكرية وفي رأي قانوني أكثر شمولاً أعدّه بالاشتراك مع عالم الحاسوب شتوبر، بأن الفحص الدقيق للأسس التكنولوجية لتدريب نماذج الذكاء الاصطناعي التوليدي يكشف أن العمليات المنجزة تتجاوز نطاق ما صُمم استثناء تنقيب النصوص والبيانات لتغطيته. وتسير حجته بوجه عام على النحو الآتي: يتيح استثناء تنقيب النصوص والبيانات استخلاص المعلومات الدلالية والواقعية والإحصائية من المصنفات؛ وفي المقابل، يستخلص تدريب نماذج الذكاء الاصطناعي التوليدي السمات التركيبية والتعبيرية ويشفّرها أيضاً، على نحو يُستفاد منه لاحقاً في إنتاج مخرجات تحل محل المصنفات الأصلية؛ ومن ثم، فإن ما ينشأ عن ذلك من مساس بالمصالح المشروعة لأصحاب الحقوق يتجاوز نطاق ما وُضع الاستثناء للسماح به.[91]

وقد طرحت الأستاذة إليونورا روزاتي حجة مماثلة، ولكن من منطلق مختلف؛ إذ ترى أن تنقيب النصوص والبيانات قد يكون جزءاً من عمليات تدريب الذكاء الاصطناعي لكنه ليس مرادفاً لتدريب الذكاء الاصطناعي كما أنه لا يشكل كل ما يتضمنه تدريب الذكاء الاصطناعي؛ وأن حصر النقاش الأوروبي في المادة 4 من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة يُعد طرحاً ضيقاً للغاية، بالنظر إلى تعدد الأفعال المقيَّدة بحق المؤلف التي قد ينطوي عليها تدريب الذكاء الاصطناعي؛ وأنه، في أي حال، قد يؤدي اختبار الخطوات الثلاث الوارد في المادة 7(2) من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة والمادة 5(5) من توجيه مجتمع المعلومات إلى استبعاد تدريب الذكاء الاصطناعي من الاستثناء، بحجة أن الاستخدام يتعارض مع الاستغلال العادي للمصنفات ويضر ضرراً غير مبرر بالمصالح المشروعة لأصحاب الحقوق. وإن صح ذلك، فسيظل الترخيص جزءاً لازماً من التحليل.[92] وقد أشار البروفيسور ماتياس لايستنر بشكل منفصل إلى أن آلية الاعتراض على الاستخدام الواردة في المادة 4(3) من المرجح أن تصبح التحدي الرئيسي في قانون حق المؤلف في الاتحاد الأوروبي، وأنها غير قابلة للتطبيق عملياً في الوقت الحالي، وهو موقف تم تبنيه في النقاش السياسي الأوروبي.[93]

[94]ويجادل البروفيسور سينفتليبن، بدلاً من ذلك، بأن المادة 4 لا تنطبق فقط في سياق تدريب الذكاء الاصطناعي، بل إنها قد تكون ذات صلة مباشرة بتحديد ما يشكل الاستغلال العادي لمصنف ما في سياق الذكاء الاصطناعي، وهو ما سيكون ذا صلة بتطبيق اختبار الخطوات الثلاث.

وقد تعاملت المفوضية الأوروبية، في تفسيرها لقانون الذكاء الاصطناعي، مع المادة 4 من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة بوصفها منطبقة على تدريب الذكاء الاصطناعي، بمعنى أن نهجها يركز على إعمال آلية الاعتراض على الاستخدام لا على استبعاد تدريب الذكاء الاصطناعي من نطاق الاستثناء. ويرد تناول البنية التحتية لآلية الاعتراض على الاستخدام في الجزء الثالث من هذه الدراسة، في سياق الشفافية والإنفاذ.[95]

---

[91] ت. و. دورنيس، "The Training of Generative AI Is Not Text and Data Mining"، 47 المجلة الأوروبية للملكية الفكرية 65 (2025). انظر(ي) أيضاً ت. و. دورنيس وس. شتوبر، *Urheberrecht und Training generativer KI-Modelle - technologische und juristische Grundlagen*, Initiative Urheberrecht (نوموس، سبتمبر 2024).

[92] إ. روزاتي، "Is text and data mining synonymous with AI training?"، 19 مجلة قانون الملكية الفكرية وممارسته 851 (2024)؛ إ. روزاتي، " The German LAION decision: A problematic understanding of the scope of the TDM exceptions and thetransition from TDM to AI training"، IPKat (7 أكتوبر 2024)؛ إ. روزاتي، " Copyright Exceptions and Fair Use Defences for AI Training Done for 'Research' and 'Learning', or the Inescapable Licensing Horizon"، ورقة عمل على "SSRN" رقم 5331405 (2025).

[93] م. لايستنر، "TDM und KI-Training in der Europäischen Union"، "GRUR" 2024، 1665 (حيث يرى أن آلية الاعتراض على الاستخدام يرجح أن تصبح التحدي الرئيسي في قانون حق المؤلف في الاتحاد الأوروبي وأنها غير قابلة للتطبيق عملياً في الوقت الحالي).

[94] Senftleben, M. Text and Data Mining, Generative AI, and the Copyright Three-Step Test. IIC 57, 67, at 79 (2026). https://doi.org/10.1007/s40319-026-01680-2

[95] قانون الذكاء الاصطناعي، المادة 53(1)(ج) (الالتزام بوضع سياسة للامتثال لقانون الاتحاد الأوروبي بشأن حق المؤلف، وتحديد التحفظات على الحقوق المُعبَّر عنها وفقاً للمادة 4(3) من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة واحترامها، بما في ذلك باستخدام أحدث التقنيات).

نظرت الهيئة التشريعية الكورية في استثناء تنقيب البيانات والنصوص، لكنها لم تقرها؛ وقد عُرضت مشاريع قوانين بهذا الشأن على الجمعية الوطنية منذ عام 2021.[96] وفي حالة عدم وجود حكم قانوني، يتعين تقييم مسألة التدريب في ضوء بند الاستخدام المنصف العام الوارد في المادة 35-5 من قانون حق المؤلف، وهو حكم صيغ على غرار المادة 107 من قانون حق المؤلف بالولايات المتحدة، وأُدرج عقب إبرام اتفاق التجارة الحرة بين جمهورية كوريا والولايات المتحدة. ويتابع النقاش المحلي عن كثب الاعتبارات التي شملها الاستقصاء في هذا القسم. يشير مؤيدو أحد الاستثناءات القانونية إلى اليقين الذي ستوفره الصياغة القانونية الصريحة؛ أما المعارضون لاستثناء قانوني فيردون بأن الصياغة القانونية تخلق أوجه غموض تفسيرية بحد ذاتها، كما يتضح من الدعوى القضائية الألمانية المتعلقة بقابلية القراءة الآلية التي نوقشت أعلاه، وأن توضيح المحكمة العليا الكورية في عام 2024 للمعايير التي تحكم المادة 35-5 يمنح البند العام محتويات قابلة للتطبيق، وأن استثناءً واسع النطاق قد يعود بفائدة غير متناسبة على المطورين الأجانب، في حين يحرم أصحاب الحقوق في المحتوى باللغة الكورية من القيمة.[97]

وفيما يتعلق بالتدريب، تم إضفاء محتويات عملية على البند العام من خلال التوجيهات الإدارية. وفي فبراير 2026، أصدرت وزارة الثقافة والرياضة والسياحة ولجنة حق المؤلف الكورية بشكل مشترك دليل قابلية تطبيق مبدأ الاستخدام المنصف على تدريب نماذج الذكاء الاصطناعي التوليدي (المشار إليه في الجزء الأول من هذه الدراسة)، والذي يطبّق العوامل الأربعة الواردة في المادة 35-5 على مراحل مسار التدريب الموضحة في الجزء الأول من هذه الدراسة. ويوضح الدليل أن الطريقة التي تم بها الحصول على المادة التدريبية تعتبر عاملاً ذا صلة، بحيث إن استخراج البيانات غير المصرح به، وتجاوز تدابير الحماية التكنولوجية، والتحايل على حواجز الدفع، وتجاهل التحفظات المُعبَّر عنها من خلال بروتوكول استبعاد "الروبوتات" تُثقل كفة عدم الاعتراف بالاستخدام المنصف، وقد تم تقديم أمثلة توضيحية لكل من جانبي هذا الحد الفاصل.[98]

5. أنظمة المكافآت

(أ) أنواع أنظمة المكافآت

قد تثار مسألة المكافآت، حيثما يُسمح باستخدام المصنفات المحمية في تدريب الذكاء الاصطناعي، سواء بموجب استثناء أو ترخيص أو كليهما. وهناك عدة نماذج قيد التطبيق، تستند إلى الخبرة الأوسع نطاقاً لقانون حق المؤلف في التكيّف مع أنماط الاستغلال الجديدة.

ويتمثل أول نموذج في نظام رسوم الترخيص، التي يتم التفاوض عليها فردياً أو جماعياً، ويدفعها المطور عند مرحلة الذخيرة اللغوية أو التدريب. وتكمن قوة هذا النموذج في أن صاحب الحقوق يتلقى تعويضاً عن استخدام المصنف في أغراض التدريب، دون أي شروط إضافية؛ أما نقطة الضعف في حالة الترخيص الفردي فهي أن هذا النموذج يتجاهل تمثيل الشريحة الطويلة من أصحاب الحقوق الأفراد الذين لا يكون التفاوض المباشر معهم مجدياً.

ويتمثل النموذج الثاني في نظام قائم على الضريبة، حيث تُفرض رسوم على مطوري الذكاء الاصطناعي أو على فئات معينة من مستخدمي خدمات الذكاء الاصطناعي التوليدي، وتُوزَّع العائدات على أصحاب الحقوق من خلال منظمة إدارة جماعية أو هيئة مماثلة. وتوصي دراسة البرلمان الأوروبي المشار إليها أعلاه في القسم باء 3 بشكل من أشكال الترخيص القانوني لتدريب الذكاء الاصطناعي في الاتحاد الأوروبي، والذي سيتضمن ميزات مخطط قائم على الضرائب مقترنة بإدارة جماعية إلزامية. وقد قُدمت مقترحات مماثلة في ولايات قضائية خارج الاتحاد الأوروبي، غالباً على غرار ضرائب النسخ الخاص المعمول بها في العديد من البلدان. وقد أثيرت تساؤلات بشأن توافق هذه المخططات مع اختبار الخطوات الثلاث المنصوص عليه في المادة 9(2) من اتفاقية برن، والمادة 10 من معاهدة الويبو بشأن حق المؤلف، والمادة 13 من اتفاق تريبس، وكذلك، بقدر ما تشمل هذه الأنظمة حقوقاً أخرى غير حق النسخ، مع الشروط المحددة الواردة في المادة 11*(ثانياً)*(2) والمادة 13(1) من اتفاقية برن.

أما النموذج الثالث، وهو الأحدث والأكثر تميزاً من الناحية المفاهيمية، فيتمثل في المكافآت على مستوى المخرجات. وبناءً على هذا النهج، لا ينصب تركيز المكافآت على استخدام التدريب ذاته، بل إنتاج مخرجات حسب النموذج المنشور، والتي تحل محل مصنفات مؤلفين يمكن تحديد هويتهم أو تتنافس معها. ويربط النموذج المكافآت مباشرة بإمكانية استبدال مصنف صاحب الحقوق في السوق ذات الصلة. إلا أن تنفيذ ذلك يبدو أنه يتطلب آليات تقنية لتحديد المخرجات البديلة وتتبعها حتى الوصول إلى المصنفات التي يمكن تحديدها في بيانات التدريب.

---

[96] بدءاً من مشروع القانون الشامل لتعديل قانون حق المؤلف الذي عُرض على الجمعية الوطنية في 15 يناير 2021، والذي تضمن استثناءً لتنقيب نصوص وبيانات؛ وقد سقط مشروع القانون مع انتهاء ولاية الجمعية الوطنية الحادية والعشرين، ولا تزال المقترحات اللاحقة قيد الدراسة.

[97] المحكمة العليا الكورية، القرار 2021Da272001 (2024)، الذي يحدد بشكل قاطع المعايير المنصوص عليها في المادة 35-5 من قانون حق المؤلف.

[98] دليل الاستخدام المنصف في كوريا، الملاحظة رقم 10 أعلاه.

وهناك مسألتان أخريان تثيران الاهتمام. عندما يكون عدد المصنفات المحددة كبيراً جداً، قد تكون المكافأة لكل مصنف ضئيلة لدرجة أنها تصبح رمزية. كما أن أساس الحساب يحتاج إلى توضيح: ففي معظم المقترحات، لن تعتمد المكافأة على إثبات أن مخرجاً معيناً يعيد إنتاج مصنف محدد أو يكيِّفه، بل على إبلاغ المطور بالطرق التي استُخدمت المصنفات من خلالها في النظام، وهو نهج يفترض مسبقاً آليات الشفافية التي نوقشت في القسم جيم من هذا الجزء.[99] علاوة على ذلك، قد يتأثر المؤلفون بظاهرة الاستبدال، سواء تم استخدام مصنفاتهم في التدريب أم لا. وكما ذُكر أعلاه، في ظل وجود تقنيات التصفية الحالية، فمن غير المرجح في الواقع أن تشكِّل المخرجات انتهاكاً لأي مصنف معين محمي بحق المؤلف، ومن ثم فإن الضرر الذي يتم النظر فيه هو ضرر نظامي وليس انتهاكاً نموذجياً لحق المؤلف. ولا يعني نسخ يتم في مرحلة التدريب، أو يظل موجوداً في النموذج، بالضرورة أن المخرجات النهائية تشكِّل انتهاكاً: فتحليل الانتهاك لأي مُخرج معيّن يعتمد على مدى تشابه ذلك المُخرج مع مصنف محمي.

ويتزايد توثيق حجم المخاوف المتعلقة بالاستبدال في قطاع الموسيقى: في أبريل 2026، أفادت شركة ديزر أن المقاطع المولّدة بالكامل بواسطة الذكاء الاصطناعي شكّلت ما يقارب 44 في المائة من التحميلات اليومية على منصتها، أي حوالي 75,000 مقطع يومي، في حين أفادت شركة سبوتيفاي بإزالة أكثر من 75 مليون مقطع عشوائي خلال الاثني عشر شهراً الماضية،[100] وتوقعت دراسة صناعية أن نسبة كبيرة من إيرادات المبدعين الموسيقيين قد يتم استبدالها بمحتوى مولّد بالذكاء الاصطناعي بحلول عام 2028.[101] ورداً على ذلك، جادل البروفيسور سينفتليين بأن اختبار الاستبدال يمكن أن يكون الأداة المفاهيمية للتمييز بين الاستخدام المشروع للذكاء الاصطناعي لدعم الإبداع البشري، وهو لا يستلزم أي مكافأة، وبين استخدام الذكاء الاصطناعي كبديل للإبداع البشري، وهو ما يستلزم دفع المكافأة. ويمكن القول إن نهجًا مشابهًا يشكل الأساس لنسخة مجلس الشيوخ من مشروع القانون البرازيلي رقم 2023/2338، الذي يربط بين آلية الانسحاب بحق في المكافأة يمكن لصاحب الحقوق ممارسته بشكل فردياً أو جماعياً.[102]

(ب) طرق وأسس المكافآت

في البرازيل، يمكن ملاحظة تأثير الاستبدال الذي تحدثه الإبداعات الاصطناعية على المصنفات البشرية في قطاعات ملموسة، بما في ذلك تصميم الصوت المحيط، حيث سعى المستخدمون إلى تشغيل موسيقى مولّدة بالكامل بأنظمة الذكاء الاصطناعي بهدف التهرب من دفع رسوم الأداء العام. وستسفر القضايا المعلقة ضد المكتب المركزي للتحصيل والتوزيع (ECAD) بشأن هذه المسألة، في الوقت المناسب، عن توجيهات قضائية حول ما إذا كان الأداء العلني للموسيقى المولدة بالذكاء الاصطناعي في المباني التجارية يستلزم أي التزام بالمكافآت. ويوضح هذا المثال التفاعل بين المكافآت المخصصة لتدريب الذكاء الاصطناعي والمكافآت المرتبطة بمخرجات الذكاء الاصطناعي.

وقد طُرح سؤالان بشأن نظام المكافآت المحتمل خلال المناقشات الجارية. السؤال الأول هو ما إذا كان ينبغي ربط المكافأة بالتدريب، أم بتأثير الاستبدال، أم بكليهما. تسعى دراسة البرلمان الأوروبي والعديد من المقترحات الأكاديمية إلى الجمع بين الأمرين؛ أما مشروع القانون البرازيلي رقم 2023/2338، بصيغته الحالية، فيُفهم على أنه يربط التعويض بعملية الاستيعاب لأغراض التدريب: فالمادة 65 تنص على التعويض عن استخدام المحتويات المحمية في تنقيب البيانات وفي تدريب وتطوير أنظمة الذكاء الاصطناعي، بحيث ينشئ مشروع القانون آلية تعويض مرتبطة بعملية التدريب ولا تمتد إلى استغلال المخرجات.[103] . والسؤال

---

[99] في مجال الموسيقى، تم تحديد سلامة البيانات الوصفية على مستوى المصنف ومستوى أصحاب الحقوق المرفقة بمواد التدريب باعتبارها الشرط المسبق للتصميم، سواء فيما يتعلق بالإسناد أو أي تقاسم للإيرادات المتأتية من المُخرج المولّد: دي بيراردينيس وآخرون، " RAIM: Operationalising Trustworthy AI in Music Generation"، مقتبس في الجزء الأول، الحاشية 17، من هذه الدراسة.

[100] ديزر، " ("AI-generated tracks now represent 44% of all new uploaded music" 19 أبريل 2026) <https://newsroom-deezer.com/2026/04/ai-generated-tracks-represent-44-of-new-uploaded-music/> تم النفاذ إليه في 3 يوليو 2026؛ Spotify, "Spotify Strengthens AI Protections for Artists, Songwriters, and Producers" (25 سبتمبر 2025) <https://newsroom.spotify.com/2025-09-25/spotify-strengthens-ai-protections/> تم النفاذ إليه في 3 يوليو 2026.

[101] الاتحاد الدولي لجمعيات المؤلفين والملحنين (CISAC)، دراسة أجرتها شركة PMP Strategy (ديسمبر 2024)، تشير إلى أن حوالي 24 في المائة من إيرادات المبدعين الموسيقيين قد تكون معرضة للخطر بحلول عام 2028 نتيجة للذكاء الاصطناعي التوليدي، مع ارتفاع هذه النسبة بشكل خاص فيما يتعلق بالإيرادات الرقمية.

[102] مشروع قانون مجلس الشيوخ رقم 2023/2338، بصيغته التي أقرها مجلس الشيوخ الاتحادي البرازيلي في 10 ديسمبر 2024، والذي أحيل إلى مجلس النواب. وحتى تاريخ إعداد هذه الدراسة، لا يزال مشروع القانون معروضاً على لجنة خاصة تابعة لمجلس النواب؛ ولم تصبح الصيغة التي قد تُعرض في نهاية المطاف على الرئيس للموافقة عليها نهائية بعد.

[103] هذا هو نص المادة 65 من مشروع القانون رقم 2023/2338، بصيغته التي وافق عليها مجلس الشيوخ الاتحادي، والتي تم اعتمادها في هذه الدراسة. كما فُسر مشروع القانون على أنه يجمع بين أساس الاستيعاب ومراعاة استبدال المخرجات، لا سيما لأن المعايير المتعلقة بالتعويض تشير إلى الآثار السوقية. وبناءً على الصيغة المعتمدة في النص، فإن التعويض يتعلق بتنقيب البيانات والتدريب والتطوير.

الثاني هو ما إذا كان نظام المكافآت يفترض وجود حقوق استئثارية معترف بها تم ترخيصها أو التنازل عنها، أم أنه ينبغي النظر إليها بدلاً من ذلك على أنها مطالبة مستقلة بالحصول على مكافآت منصفة، ربما على غرار مكافأة النسخ الخاص الموجودة في العديد من الولايات القضائية.[104] وكما أشير أعلاه، قد تشمل التشريعات البرازيلية الحالية مرحلة المخرجات من خلال حق النسخ أو حق التكييف أو حق التبليغ للجمهور، حتى في حالة عدم إقرار مشروع القانون رقم 2023/2338، وبناءً على ذلك، سيكون من الضروري الحصول على تصريح مسبق وصريح من أصحاب الحقوق.[105] وبخلاف ذلك، يُعدُّ النسخ الخاص شكلاً من أشكال التعويض، وليس ترخيصاً، لأن أعمال النسخ الخاصة تُعتبر بطبيعتها غير قابلة للترخيص لأسباب متنوعة، منها خصوصية المستخدمين.

في البرازيل، هناك ملاحظة أخرى تتعلق بتصميم نظام المكافآت: يمكن ملاحظة تأثير الاستبدال الذي تحدثه الإبداعات الاصطناعية على المصنفات البشرية في قطاعات ملموسة، بما في ذلك تصميم الصوت المحيط، حيث سعى المستخدمون إلى تشغيل موسيقى مولّدة بالكامل بأنظمة الذكاء الاصطناعي بهدف التهرب من رسوم الأداء العام. وستسفر القضايا المعلقة ضد المكتب المركزي للتحصيل والتوزيع (ECAD) بشأن هذه المسألة، في الوقت المناسب، عن توجيهات قضائية حول ما إذا كان الأداء العلني للموسيقى المولدة بالذكاء الاصطناعي في المباني التجارية يستلزم أي التزام بمكافأة أعضاء المكتب المركزي للتحصيل والتوزيع (ECAD). ويوضح هذا المثال التفاعل بين المكافآت المخصصة لتدريب الذكاء الاصطناعي والمكافآت المرتبطة بمخرجات الذكاء الاصطناعي؛ فقد تؤثر الخيارات المتخذة في أحد المجالين على تصميم المجال الآخر.[106]

وأخيراً، في الولايات القضائية التي لا يزال الإطار التنظيمي فيها قيد التشكيل، تظل مسألة المكافآت مطروحة في الوقت الحالي. إن دليل عام 2026 بشأن قابلية انطباق مبدأ الاستخدام المنصف على تدريب نماذج الذكاء الاصطناعي التوليدي، الذي أصدرته بشكل مشترك وزارة الثقافة والرياضة والسياحة ولجنة حق المؤلف الكورية، والمشار إليه في الجزء الأول، يتناول بشكل أساسي مسألة عمليات التدريب التي يمكن إجراؤها دون ترخيص؛ ويترك المسألة المقابلة المتعلقة بالمكافآت لتتم معالجتها في إطار عملية الإصلاح الأوسع نطاقاً الجارية حالياً.[107] إلا أن جزءاً كبيراً من النقاش الداخلي ينطلق من فرضية أن الاستخدام التجاري لنماذج الذكاء الاصطناعي ينبغي أن يترتب عليه تعويض للمبدعين، وبالتالي فإن السؤال الذي لا يزال مطروحاً هو آلية المكافآت. تتشابه الأوضاع في كينيا والبرازيل والصين، كلٌّ على طريقتها الخاصة: ففي كل منها، تجري العملية التشريعية، وتظل الخيارات السياسية مفتوحة، والنموذج الذي سيظهر لم يتحدد بعد. ويعود الجزء الثالث من هذه الدراسة، الذي يتناول مسألة الإنفاذ، إلى الآثار العملية التي تنجم عن حالة التغير القانوني هذه بالنسبة إلى أصحاب الحقوق الذي يسعى إلى إثبات مطالباته.

الجزء الثالث: الإنفاذ

غالباً ما تنطوي العمليات التقنية الموصوفة في الجزء الأول على حقوق استئثارية معترف بها في تشريعات حقوق المؤلف الوطنية وفي المعاهدات الدولية المتعلقة بحقوق المؤلف؛ وتوفر ترتيبات الترخيص التي تم استعراضها في الجزء الثاني آليات يمكن من خلالها تخليص تلك الحقوق في حالة عدم انطباق أي استثناء أو قيد. ويتناول الجزء الثالث ما يحدث في حالة عدم وقوع أي من الأمرين: عندما يرى صاحب حقوق أن مصفنه قد استُخدم في تدريب نموذج ذكاء اصطناعي توليدي من دون تصريح ومن دون أن ينطبق عليه أي استثناء، ويسعى إلى الدفاع عن حقوقه من خلال المحاكم أو الإجراءات الإدارية.

وحتى وقت قريب، لم يكن هناك سوى القليل جداً من السوابق القضائية التي يمكن الرجوع إليها. لم يعد الأمر كذلك. فحتى 22 أغسطس 2026، تم رفع 137 دعوى ضد مطوري الذكاء الاصطناعي في الولايات المتحدة تتعلق باستخدام أنظمة الذكاء الاصطناعي التوليدية لمصنفات محمية، مع تسجيل ما يقارب 30 دعوى أخرى في ولايات قضائية أخرى.[108] وأصدرت المحاكم في

---

[104] ولا يشمل نطاق هذه الدراسة بحثاً مستفيضاً في الأنظمة الفريدة المحتملة، كما لا يتناول تحديد أوجه الاستخدام التي يمكن أن تغطيها تلك الأنظمة.

[105] هذا هو نص المادة 65 من مشروع القانون رقم 2023/2338، بصيغته التي وافق عليها مجلس الشيوخ الاتحادي، والتي تم اعتمادها في هذه الدراسة. كما فُسر مشروع القانون على أنه يجمع بين أساس الاستيعاب ومراعاة استبدال المخرجات، لا سيما لأن المعايير المتعلقة بالتعويض تشير إلى الآثار السوقية. وبناءً على الصيغة المعتمدة في النص، فإن التعويض يتعلق بتنقيب البيانات والتدريب والتطوير.

[106] مشروع القانون رقم 2023/2338، المادة 62 وما يليها (بالصيغة التي أُحيلت إلى مجلس النواب). ويُستثني السر التجاري من التزام الشفافية، وهو موضوع نقاش جوهري في إطار العملية التشريعية.

[107] وزارة الثقافة والرياضة والسياحة ولجنة حق المؤلف الكورية، *دليل حول قابلية انطباق مبدأ الاستخدام المنصف على تدريب نماذج الذكاء الاصطناعي التوليدي* (فبراير 2026)؛ انظر(ي) الجزء الأول، الحاشية 10، من هذه الدراسة.

[108] وفقاً لمتتبع الدعاوى القضائية الذي يديره البروفيسور إدوارد لي على موقع chatgptiseatingtheworld.com، أظهرت خريطة الولايات المتحدة 156 دعوى قضائية حتى 22 أغسطس 2026، مع تسجيل ما يقرب من 30 دعوى أخرى في ولايات قضائية أخرى. هذا الرقم قابل للتغير، وهو معلن بتاريخ إعداد هذه الدراسة. انظر(ي) أيضاً خريطة الدعاوى القضائية التي تُديرها رابطة حق المؤلف.

البرازيل وكندا والصين وألمانيا والمملكة المتحدة والولايات المتحدة وأماكن أخرى أحكاماً تتناول جوانب متعددة من التحليل: حول ما إذا كان النموذج المدرَّب هو ذاته المصنفات التي تم تدريبه عليها أو يحتوي على نسخ منها؛ وحول ما إذا كانت الاستثناءات في القانون الأوروبي تشمل تدريب الذكاء الاصطناعي؛ وحول ما إذا كانت المحاكم الوطنية لديها ولاية قضائية بشأن عمليات التدريب التي تُجرى في الخارج؛ وحول ما إذا كانت المخرجات التي تحل محل المصنفات الموجودة في المجموعة التدريبية تنطوي على حق صاحب الحقوق في النسخ أو حق الإتاحة للجمهور؛ وحول ما إذا كانت المخرجات ذاتها قابلة لحماية حق المؤلف (وهو سؤال يقع خارج نطاق هذه الدراسة، ويُشار إليه هنا فقط كسياق).

يتمحور هذا الجزء حول ثلاثة أسئلة رئيسية. يستخلص القسم ألف الدروس المستفادة من مجموعة السوابق القضائية في صيغتها الحالية، مرتبةً حسب الموضوع الذي تتناوله كل قضية، وليس حسب الولاية القضائية. يتناول القسم باء المسألة الهيكلية المتعلقة بإنفاذ القانون ضد مطوري الذكاء الاصطناعي، بما في ذلك السمات عبر الحدود التي تعقّد عملية الإنفاذ. بينما يحدد القسم جيم التحديات العملية الرئيسية التي تواجه أصحاب الحقوق الساعين إلى إنفاذ حقوقهم، بما في ذلك الشفافية، وانتشار المعايير التقنية المتعلقة بالانسحاب، والصعوبة التشغيلية في جمع الأدلة.

القسم ألف: الدروس المستفادة من السوابق القضائية الحالية

تعدُّ السوابق القضائية التي تراكمت على مدار الثمانية عشر شهراً الماضية سوابق مجزأة، وتختلف من حيث الولاية القضائية، ولم تستقر بعد من ناحية المبدأ في أي نظام قانوني واحد. ولم تتناول بعد جميع المسائل. ومع ذلك، فإنها تسلط الضوء على الأقل على أربع قضايا متميزة، بدأت المحاكم في مختلف الولايات القضائية تتفق بشأن كل منها، أو تختلف في بعض الحالات. سيتم تناول هذه القضايا بالترتيب أدناه.

1. مسألة النموذج كنسخة

المسألة الأولى تكمن فيما إذا كان نموذج الذكاء الاصطناعي المُدرّب، من منظور قانون حق المؤلف، يحتوي في حد ذاته على نُسخ من المصنفات التي تم تدريبه عليها. وهذا يختلف عن المسألة السابقة المتعلقة بما إذا كانت النسخ تُنتَج في أثناء إنشاء الذخيرة اللغوية والتدريب. وقد نوقشت الطبيعة التقنية للمسألة المتعلقة بالحفظ في الجزء الأول من هذه الدراسة. هناك ثلاثة قرارات صدرت خلال الثمانية عشر شهراً الماضية تقدم إجابات مختلفة بشكل ملحوظ.

في قضية جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي (GEMA) ضد شركة OpenAI، التي تمت مناقشتها في الجزء الثاني، القسم باء 5 أعلاه، قضت محكمة ميونخ الإقليمية في نوفمبر 2025 بأن تدريب نموذج لغوي على كلمات الأغاني، وما تبع ذلك من حفظ تلك الكلمات داخل النموذج، يمثلان أعمال النسخ بالمعنى المقصود في القانون الألماني الذي ينفذ التوجيه الأوروبي بشأن حق المؤلف، ورفضت المحكمة تطبيق الاستثناء المنصوص عليه في المادة 4 من استثناء قانون حق المؤلف والحقوق المجاورة في السوق الرقمية (CDSM).[109]

وفي مسألة النموذج كنسخة ذات الصلة التي تتضمن تكنولوجيا مختلفة، توصلت محكمة العدل العليا لإنجلترا وويلز إلى استنتاج مختلف قبل بضعة أيام فقط. في قضية Getty Images (US) Inc. ضد Stability AI Ltd.، قضت القاضية جوانا سميث، الحائزة على وسام الإمبراطورية البريطانية (DBE)، في 4 نوفمبر 2025 بأن ترجيحات نموذج ستيبل دفيوجن، وهو نموذج نشر كامن يُستخدم في توليد الصور، لا تشكّل نسخاً متعددة للصور الفوتوغرافية التي تم تدريب النموذج عليها.[110] وتخلّى المدعي عن مطالباته المتعلقة بالانتهاك الأساسي وحقوق قاعدة البيانات عند اختتام المحاكمة، معترفاً بأن الأدلة لم تثبت أن التدريب قد جرى في المملكة المتحدة؛ وبالتالي، تركز النزاع على الانتهاك الثانوي، الأمر الذي استلزم من المحكمة أن تقرر ما إذا كان النموذج يمثّل مادة تُعد في حد ذاتها نسخة متعددية. وقد خلصت المحكمة إلى أن الدليل التقني المقدم خلال المحاكمة أثبت أن ترجيحات نموذج النشر لا تخزّن ببيانات التدريب، وأنه في ظل عدم وجود دليل على الحفظ بالمعنى ذي الصلة، فإن النموذج لا يُعد نسخة متعددية.[111] وأضافت المحكمة، في ملاحظة جانبية، أنه لو كانت قد توصلت إلى استنتاج مخالف، لكانت شركة Stability AI قد

---

[109] *قضية جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي ضد شركة OpenAI*، محكمة ميونخ الأولى، 11 نوفمبر 2025 (O 42 14139/24)، الحكم النهائي. كما تمت مناقشة هذا القرار في الجزء الثاني، القسم باء 5.

[110] *Getty Images (US) Inc ضد Stability AI Ltd.*، [2025] EWHC 2863 (Ch) (القاضية سميث، 4 نوفمبر 2025). ويتناول القرار ادعاءات حق المؤلف والعلامة التجارية معاً؛ وتقتصر المناقشة الحالية على تحليل حق المؤلف. وقد مُنح الإذن باستئناف النتيجة بشأن التعدي الثانوي في ديسمبر 2025 ([2025] EWHC 3343 (Ch))؛ وكان الاستئناف لا يزال قيد النظر في تاريخ هذه الدراسة.

[111] بشأن بنية نماذج النشر، طالع الجزء الأول، القسم باء-1، والمراجع المذكورة فيه.

أصبحت مسؤولة بموجب إطار الانتهاك الثانوي؛ لأن المطور كان يعلم أو كان لديه سبب يدعوه للاعتقاد بأن مجموعة التدريب تحتوي على مصنفات تم استخراجها من شبكة الإنترنت دون موافقة.[112]

وثمة قرار ثالث، هو حكم محكمة الإنترنت بهانغتشو في القضية المعروفة بقضية "هانغتشو ألترامان"، يتناول وقائع مشابهة لكنها متميزة. فقد كان المدعى عليه في تلك القضية منصة تستضيف نماذج ذكاء اصطناعي مفتوحة المصدر يتاح للمستخدمين النفاذ إليها؛ وكان بعض المستخدمين قد درّبوا تلك النماذج على المانغا اليابانية ألترامان المشمولة بحق المؤلف، وحمّلوا النماذج المدرَّبة ليستخدمها مستخدمون آخرون. وقد رفع مالك حق المؤلف دعوى على المنصة، وقضت المحكمة بمسؤولية المنصة عن التعدي بالمشاركة. وفي معرض حديثها، أعربت المحكمة عن رأي مفاده أنه يمكن تحليل فعل التدريب ذاته على أنه استخدام منصف بموجب القانون الصيني، استناداً إلى أن الغرض من استخدام المصنفات المحمية بحق المؤلف في تدريب نموذج ذكاء اصطناعي توليدي هو تعلّم السمات اللغوية والخصائص الأسلوبية وغيرها من العناصر وتحليلها، لا نسخ التعبير الإبداعي الأصلي. ولم تكن المحكمة بحاجة إلى البتّ في هذه النقطة، ولم يكن هذا التعليل مؤثراً في النتيجة، غير أن ما ورد في معرض حديثها أشار إلى احتمال اعتبار التدريب مشمولاً بمبدأ الاستخدام المنصف موجب القانون الصيني.[113]

والقرارات الثلاثة ليست متسقة تماماً لأسباب تقنية بحتة، لأنها تتعلق بأنواع مختلفة من النماذج (نموذج لغوي قائم على بنية المحوِّل في ميونخ، ونموذج نشر كامن في لندن، ومنصة تستضيف نماذج لتوليد الصور ضبطها المستخدمون في هانغتشو) فضلاً عن اختلاف الأساليب الإجرائية. فقد واجهت محكمتا ميونخ ولندن بصعوبة مع أدبيات علوم الحاسوب بشأن الحفظ في النماذج التوليدية الكبيرة. ويشمل ذلك سلسلة المصنفات التي أعدّها الأساتذة كوبر وغريملمان ولملي وآخرون بشأن استخراج المحتوى المحفوظ من النماذج مفتوحة الأوزان.[114] وقد فسّرت محكمة ميونخ المصنفات نفسها على أنها تثبت أن التدريب يرقى إلى النسخ على نطاق واسع، بينما رأت محكمة لندن أنها تدعم الاستنتاج القائل إن أوزان نماذج الانتشار لا تُعدّ، من حيث المبدأ، نسخاً متعدية لحقوق المؤلف. وتوضح المقارنة أن الأدلة التجريبية المتاحة بشأن الحفظ لا تحسم بمفردها نتيجة التحليل القانوني حسماً كاملاً. أما مسألة إذا ما كان ينبغي تحليل قدرة نموذج مثبتة على نسخ تسلسلات من بيانات تدريبه على أنها دليل على أن النموذج يُعد نسخة من مجموعة التدريب، أو على أنها عيب سلوكي ينبغي التخفيف منه بمرشِّحات المخرجات وتقنيات إلغاء التعلّم، فهو مسألة توصيف قانوني، وقد خلصت بشأنها محاكم مختلفة إلى استنتاجات متباينة، استناداً إلى سجلات أدلة استندت إلى أدبيات علمية أساسية متشابهة، وإن اختلفت في طبيعتها وفي الأدلة التي عُرضت فعلياً على كل محكمة.[115] وكما ورد في الجزء الأول، فإن مسألة ما إذا كان النموذج يحتوي على نسخة من مصنفات محمية أو موضوعات حقوق مجاورة يجب البت فيها على أساس كل نموذج على حدة.

وقد أُثيرت مسألة النموذج بوصفه نسخة في صورة أخرى متميزة من الناحية الإجرائية في الولايات المتحدة، حيث ذهب بعض المدّعين إلى أن عملية التدريب لم تُنتج نسخاً فحسب، بل إن النموذج المدرَّب ذاته يُعد مصنفاً مشتقاً متعدِّ. ففي قضية Kadrey ضد Meta Platforms, Inc.، دفع المؤلفون بأن نماذج LLaMA لا يمكنها العمل بدون المعلومات التعبيرية المستخرجة من كتبهم، وبأن هذه المعلومات محتفظ بها داخل أوزان النموذج، ومن ثم فإن النماذج تُعد مصنفات مشتقة غير مصرح بها، وهو توصيف من شأنه أن يُرتّب المسؤولية على النموذج ذاته من دون الحاجة إلى إثبات وجود تطابق بين أي مخرَج بعينه والمصنفات التي دُرِّب عليها النموذج. ورفضت المحكمة قبول تلك النظرية في مرحلة النظر في المرافعات. وفي نهاية المطاف، كسبت الجهة

---

[112] للاطلاع على تعليق، انظر(ي) إ ديركلاي، " Getty Images v Stability AI: A landmark judgment reinforcing the need for the UK government to amend its copyright laws"، Kluwer Copyright Blog (20 نوفمبر 2025)، متاح على الإنترنت: https://legalblogs.wolterskluwer.com/copyright-blog/getty-images-v-stability-ai-a-landmark-judgment-reinforcing-the-need-for-the-uk-government-to-amend-its-copyright-laws/.

[113] محكمة الإنترنت بهانغتشو، Zhe 0192 Min Chu No. 1587 (2024) (الابتدائية)؛ ومحكمة هانغتشو الوسطى الشعبية بمقاطعة تشجيانغ، Zhe 01 Min Zhong No. 10332 (2024) (الدرجة الثانية).

[114] أ فيدر كوبر وج غريملمان، "The Files are in the Computer: Copyright, Memorization, and Generative AI"، ورقة عمل SSRN رقم 4803118 (2024). بشأن البحوث التجريبية حول الحفظ في النماذج اللغوية، انظر(ي) أيضاً دراسة كوبر وآخرين المذكورة في الجزء الأول، الحاشية 18، من هذه الدراسة. وانظر(ي) أيضاً م أ ليملي وأ ف كوبر، " Probabilistic 'Copies' in Generative AI Models"، المذكورة في الجزء الأول، الحاشية 25، من هذه الدراسة.

[115] للاطلاع على تعليق نقدي على تعليل جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي (GEMA) في هذه المرحلة، انظر(ي) ج فروسيو، " Copyright in Formaldehyde: How GEMA v OpenAI Freezes Doctrine and Chills AI"، Kluwer Copyright Blog (10 و11 ديسمبر 2025)، الجزآن 1 و2.

المدّعى عليها الدعوى على أساس مبدأ الاستخدام المنصف، الذي سيُناقش تالياً في القسم 4، بعد أن دفعت بأن الأوزان مجرد معلمات إحصائية مستمدة من بيانات التدريب وليست تعبيراً مخزَّناً.[116]

وطُرحت نظرية مماثلة، فيما يتعلق بنموذج نشر كامن، في قضية Andersen ضد Stability AI Ltd.، حيث وصف المدّعون النموذج بأنه وسيلة بديلة لتخزين نسخ من صور التدريب؛ وقد واجه هذا التوصيف صعوبات إثباتية في غياب مخرجات تطابق مصنفاً محمياً بدرجة وثيقة، وهي صعوبة تضاهي ما خلصت إليه المحكمة الإنكليزية في قضية Getty بأن أوزان نموذج النشر لا تخزّن بيانات التدريب. وتشير التجربة الأمريكية إلى أن إثبات أن المصنفات المحمية قد جرى إدخالها في التدريب لا يكفي بحد ذاته لاعتبار النموذج المدرَّب مشتق من تلك المصنفات.[117] وحتى وقت كتابة هذه الدراسة، لم يصدر أي قرار استئنافي في أي من المسائل المذكورة أعلاه.

2. استثناء تنقيب النصوص والبيانات

تتمثل المسألة الثانية، التي تنشأ في الاتحاد الأوروبي وفي الولايات القضائية التي تأثرت قوانين حق المؤلف فيها بالإطار الأوروبي (ومنها بلدان المنطقة الاقتصادية الأوروبية)، في إذا ما كانت استثناءات تنقيب النصوص والبيانات الواردة في المادتين 3 و4 من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة تمتد لتشمل تدريب نماذج الذكاء الاصطناعي التوليدي. ويُتناول خلاف المبادئ الموضوعي في الجزء الثاني، القسم باء-5، من هذه الدراسة. وتُلخَّص هنا السوابق القضائية الألمانية بشأن هذه المسألة استكمالاً للطرح، مع تسليط الضوء على ما تقدمه من دروس بشأن إنفاذ مطالبات أصحاب الحقوق.

وقد نوقش في الجزء الثاني، القسم باء-5، أعلاه الحكمان الألمانيان الصادران في قضية Kneschke ضد LAION e.ضد بشأن تطبيق استثناء تنقيب النصوص والبيانات على تدريب الذكاء الاصطناعي (في الدرجة الأولى، محكمة هامبورغ الإقليمية في 27 سبتمبر 2024، وفي الاستئناف، محكمة هامبورغ الإقليمية العليا في 10 ديسمبر 2025).

ويمكن استخلاص درسين فيما يتعلق بالإنفاذ؛ أولهما أنه في الولايات القضائية التي ينطبق فيها استثناء تنقيب النصوص والبيانات على تدريب الذكاء الاصطناعي، يتأثر موقف صاحب الحقوق في الإنفاذ بآلية الاعتراض على الاستخدام. فحيثما أُعرب عن اعتراض صحيح على نحوٍ يتيح لعملية آلية رصده والتصرف بناءً عليه، جاز لصاحب الحقوق إنفاذ حقوقه عبر إجراءات التقاضي المعتادة. أما في حال عدم الاعتراض، أو حيث لا يستوفي التحفظ المتطلبات المنطبقة، فقد ينطبق استثناء تنقيب النصوص والبيانات، مع مراعاة شروطه الأخرى، ما يؤدي إلى تقييد دعوى تعدٍّ قائمة على حق النسخ. وفي هذا الصدد، تتوقف قدرة صاحب الحقوق على التحفظ بشأن حقوقه جزئياً على الشكل التقني لإشارة الاعتراض.

أما الدرس الثاني فهو أن معيار قابلية القراءة آلياً ليس ثابتاً. فقد قضت محكمة الاستئناف في هامبورغ صراحةً بأن قابلية القراءة آلياً يجب أن تُقدَّر في ضوء القدرات التكنولوجية السائدة وقت الاستخدام المتنازع عليه. وفي سياق وقائع القضية، كان ذلك يعني الرجوع إلى مستوى التقدم التقني في النصف الثاني من عام 2021، قبل الانتشار الواسع للنماذج اللغوية الكبيرة القادرة على استخلاص المعنى من النصوص غير المهيكلة بلغة طبيعية. ويدل الحكم على أن خيار الاعتراض على الاستخدام الذي لم يكن قابلاً للقراءة آلياً في عام 2021 قد يصبح قابلاً للقراءة آلياً الآن، وبالمثل، فإن خياراً قد لا يكون قابلاً للقراءة آلياً في عام 2025 قد يصبح كذلك في تاريخ لاحق. وتتمثل النتيجة العملية لذلك بالنسبة إلى أصحاب الحقوق في أنهم قد يلزمهم مراجعة خيارات الاعتراض على استخدام مصنفاتهم، بل وإعادة صياغة هذه الاعتراضات في كثير من الحالات، مع تطور التكنولوجيا؛ أما بالنسبة لمطوري الذكاء الاصطناعي، فتعني هذه النتيجة أن مسألة ما إذا كان إجراءٌ ما مشروعاً وقت تنفيذه لا تُحسم بناءً على أي تغيرات لاحقة في القدرات التقنية ذات الصلة.[118]

---

[116] للاطلاع على بيان لموقف المطورين خارج سياق التقاضي، انظر(ي) غوغل، " A Policy Agenda for Responsible Progress in Artificial Intelligence" (ورقة بيضاء بشأن تنظيم الذكاء الاصطناعي)، متاحة على الرابط https://blog.google/company-news/outreach-and-initiatives/public-policy/white-paper-ai-regulation/ تاريخ النفاذ: 7 يوليو 2026.

[117] بشأن نظرية المصنف المشتق في الولايات المتحدة، انظر(ي) Kadrey ضد Meta Platforms, Inc.، رقم. 3:23-cv-03417-VC (محكمة المقاطعة الشمالية في كاليفورنيا، القاضي تشابريا، 20 نوفمبر 2023) (التي رفضت في مرحلة النظر في المرافعات الادعاء بأن نماذج LLaMA تُعد مصنفات مشتقة متعدية بذاتها، وسمحت في الوقت نفسه بالمضي قماً في دعوى التعدي المباشر القائمة على النسخ لأغراض التدريب)؛ وقضية Andersen ضد Stability AI Ltd.، رقم. 3:23-cv-00201-WHO (محكمة المقاطعة الشمالية في كاليفورنيا، القاضي أوريك)، والأمرين الصادرين في 30 أكتوبر 2023 و12 أغسطس 2024 (اللذان تناولا، من بين أمور أخرى، الادعاء بأن النموذج يخزّن نسخاً مضغوطة من صور التدريب). وتُناقش النتيجة التي انتهت إليها المحكمة في قضية Kadrey بشأن الاستخدام المنصف في القسم ألف-4 أدناه.

[118] ب كيلر، "LAION Round 2: Machine-Readable but Still Not Actionable - The Lack of Progress on TDM Opt-Outs"، (Kluwer Copyright Blog17 و18 ديسمبر 2025)، الجزآن 1 و2.

3. الاختصاص القضائي على التدريب خارج الإقليم

تتعلق المسألة الثالثة، التي تتزايد أهميتها العملية، باختصاص المحاكم الوطنية على عمليات التدريب التي تُجرى على بنية تحتية تقع في بلد آخر. وقد صدر الحكم الكندي الرئيسي بشأن هذه المسألة، في قضية Toronto Star Newspapers Ltd ضد OpenAI Inc.، في 7 نوفمبر 2025 عن القاضية كيميل من المحكمة العليا لمقاطعة أونتاريو.[119] رفع تحالف من كبرى المؤسسات الإعلامية الكندية دعوى على OpenAI، بدعوى أن المدّعى عليها استخدمت محتواه لتدريب ChatGPT بما ينتهك قانون حق المؤلف الكندي وتدابير الحماية التكنولوجية التي تحمي ذلك المحتوى وشروط الاستخدام الخاصة بالمواقع الإلكترونية المعنية. ودفعت OpenAI بعدم الاختصاص طالبةً ردّ الدعوى، محتجّةً بأن أفعال التدريب أُجريت على خوادم تقع في الولايات المتحدة، ومن ثم فهي خارج النطاق الإقليمي لقانون حق المؤلف الكندي. ورفضت المحكمة هذه الحجة، إذ قضت بأن ثمة صلة حقيقية وجوهرية بين الإجراءات ومقاطعة أونتاريو، وبأن مسألة إذا ما كانت الأفعال المزعومة تمسّ قانون حق المؤلف الكندي هي مسألة تتعلق بالموضوع وليس بتحليل الاختصاص. ومضت الدعوى إلى مرحلة الكشف في أوائل عام 2026؛ وقد طعن المدّعى عليهم في حكم الاختصاص، وكان الطعن لا يزال قيد النظر عند تاريخ إعداد هذه الدراسة.

ولم يتحدد بعد نطاق اختبار الصلة الحقيقية والجوهرية في سياق تدريب الذكاء الاصطناعي. فقد أشار التعليق الأكاديمي على حكم Toronto Star أن هذا الاختبار وُضع في القانون الكندي في سياق حق التبليغ للجمهور، مثل قضيتي SOCAN ضد CAIP وSOCAN ضد ESA، وأن تطبيقه على حق النسخ ليس بالضرورة أمراً واضحاً.[120] وقد جرى العرف على تحديد موقع حق النسخ، لأغراض الاختصاص القضائي، في المكان الذي يحدث فيه التثبيت أو النسخ. أما العلاقة بين النطاق الإقليمي لحق النسخ ومبدأ المعاملة الوطنية المنصوص عليه في المادة 5 من اتفاقية برن فتتناولها الفقرة باء-2 أدناه.

وتبيّن تطورات أخرى في كندا تنوّع الاستجابات الإجرائية المتاحة. ففي قضية CanLII ضد Caseway AI، كانت CanLII قد رفعت دعوى في عام 2024 على شركة Caseway AI الناشئة التي يوجد مقرها في فانكوفر بسبب التنزيل الضخم لقاعدة بياناتها التي تضم القرارات القضائية بغرض تدريب مساعد ذكاء اصطناعي للمعلومات القانونية. ودفعت CanLII بأنه، وإن كانت القرارات الأساسية في الملك العام، فإن البنية المنظَّمة لقاعدة البيانات والبيانات الوصفية وطريقة الاختيار في قاعدة بياناتها تشكل مجموعة محمية بحق المؤلف. وقد توصل الطرفان إلى تسوية في مارس 2026 بشروط سرية؛ ونظراً لهذه التسوية، لم يصدر أي حكم قضائي بشأن موضوع تلك المسائل. وثمة دعاوى جماعية قيد النظر أيضاً؛ ففي كولومبيا البريطانية، رفع المؤلف ج ب ماكينون دعاوى جماعية على Meta وNvidia وAnthropic وغيرها، زاعماً أن تلك الشركات استخدمت نسخاً غير مشروعة من مصنفات (مستمدة من مجموعة بيانات Books3) لتدريب نماذجها؛ وفي كيبيك، تقدمت آن روبيلارد بطلب للحصول على التصريح برفع دعاوى جماعية على Meta وOpenAI، ساعيةً إلى الحصول على تعويضات قانونية تصل إلى 20,000 دولار كندي عن كل مصنف، وهو الحد الأقصى المنصوص عليه في القانون فيما يخص التعدي التجاري. وقد ثبت أن القيد الهيكلي ذاته كان حاسماً في الدعوى الإنكليزية لقضية Getty، المشار إليها في القسم ألف-1 أعلاه، حيث جرى التخلي عن الدعاوى الأساسية بالتعدي على حقوق المؤلف بعدما لم تثبت الأدلة أن التدريب قد جرى في المملكة المتحدة.[121]

وفي الهند، أيدت محكمة دلهي العليا اختصاصها القضائي على التدريب الذي أُجري على خوادم موجودة في الولايات المتحدة في قضية «ANI Media Pvt. Ltd. ضد OpenAI OPCO LLC»، وقررت، استنادًا إلى الأدلة الظاهرة، أن تخزين المصنفات في الخارج لم يكن سوى "الخطوة النهائية في سلسلة الأحداث" التي تبدأ بالوصول إليها من الهند، وأن اعتراض المطور "لا يمكن الدفاع عنه لأنه من شأنه أن يؤدي إلى التحايل على قانون حقوق النشر الهندي"؛[122] وترد مناقشة مضمون القرار في القسم ألف-4

---

[119] *Toronto Star Newspapers Ltd ضد OpenAI Inc.*، 2025 ONSC 6217 (القاضية كيميل، 7 نوفمبر 2025). ويشمل المدّعون Toronto Star وGlobe and Mail وPostmedia وهيئة الإذاعة الكندية. وقد استند الدفع بعدم الاختصاص إلى القول بأن أفعال التدريب المزعومة أُجريت على خوادم تقع في الولايات المتحدة، ومن ثم فهي خارج النطاق الإقليمي لقانون حق المؤلف الكندي.

[120] Society of Composers، Authors and Music Publishers of Canada ضد Canadian Association of Internet Providers، SOCAN") 2004 SCC 45 ضد CAIP")؛ وانظر(ي) أيضاً Society of Composers، Authors and Music Publishers of Canada ضد Entertainment Software Association، 2022 SCC 30. وقد طُرحت في التعليق الأكاديمي على حكم Toronto Star الملاحظة المتعلقة بالمبادئ التي تفيد بأن الاختبار وُضع في سياق حق التبليغ للجمهور، وقد لا يمكن نقله بسهولة إلى حق النسخ.

[121] *انتهت قضية CanLII ضد Caseway AI* بتسوية سرية في مارس 2026. أما الدعاوى الجماعية التي رفعها ج ب ماكينون في كولومبيا البريطانية ضد Meta وNvidia وAnthropic وDatabricks، والتي رفعتها آن روبيلارد في كيبيك ضد Meta وOpenAI، فهي في مرحلة التصديق حتى تاريخ هذه الدراسة. وللاطلاع على الحد الأقصى المنصوص عليه في القانون فيما يخص التعدي التجاري، انظر(ي) قانون حق المؤلف، RSC 1985, c C-42، القسم 38.1(1)(أ).

[122] للاطلاع على المقتطفات الواردة في النص، انظر(ي)(ي) إلى الفقرة 51 من القضية "ANI Media Pvt. Ltd. v. OpenAI OPCO LLC, High Court of Delhi, CS(COMM) 1028/2024 (Bansal J., judgment of 24 July 2026), 2026 SCC OnLine Del 5291". صدر الحكم بناءً على طلب المدعي إصدار أمر قضائي مؤقت؛ ولا تزال الدعوى معلقة، وفي تاريخ إعداد هذه الدراسة، كان الاستئناف معروضًا أمام هيئة قضائية في المحكمة العليا.

أدناه. وفي البرازيل، أُثيرت مسألة التدريب خارج الإقليم في الدعوى التي رفعتها صحيفة Folha de São Paulo ضد OpenAI في أغسطس 2025، والتي تزعم فيها الصحيفة أن الشركة المدّعى عليها استخدمت محتواها الصحفي في تدريب ChatGPT وفي تقديم الخدمات للمستخدمين البرازيليين دون تصريح.[123] وقد جرت الإجراءات في إطار السرية القضائية؛ وتوصل الطرفان إلى تسوية القضية في عام 2026، ولم يصدر أي حكم في الموضوع. وفي الأرجنتين، رفعت Editorial Perfil، وهي دار ناشر إخبارية كبرى، دعوى على OpenAI وMicrosoft في أغسطس 2026 أمام المحاكم الاتحادية المدنية والتجارية، زاعمةً الاستخدام غير المصرح به لمحتواها الصحفي في تدريب خدمات الذكاء الاصطناعي التوليدي وتشغيلها.[124] وبصفة أعم، طُرحت حجة أخرى لم تُختبر بعد أمام المحاكم البرازيلية في سياق الذكاء الاصطناعي، قياساً على السوابق القضائية للمحكمة الاتحادية العليا بشأن المنصات الرقمية الأجنبية العاملة في البرازيل.[125] ووفقاً لهذه الحجة، فإن تقديم خدمات الذكاء الاصطناعي في البرازيل يُخضع التدريب الأساسي لقانون حق المؤلف البرازيلي لأن توليد المخرجات اللاحقة على يد مستخدمين برازيليين يشكل استمراراً لفعل التدريب؛ ومن هذا المنظور، إذا جرى التدريب على نحو غير مشروع، فإن تقديم الخدمة التي تعتمد عليه يخضع لقانون حق المؤلف البرازيلي بموجب مبدأ المعاملة الوطنية في المادة 5 من اتفاقية برن.[126] وتُعد هذه الحجة في جوهرها، صيغة من المبدأ القائل إن إحدى المخالفات التي تبدأ في الخارج لكن تُحدث آثارها في نطاق الولاية القضائية يمكن إثباته في تلك الولاية، وذلك بتطبيقه على سياق تكنولوجي يحدث فيه التثبيت المعني قبل وصوله إلى المستخدم.

4. مسألة الاستبدال

تتعلق المسألة الرابعة بالتوصيف القانوني للمخرجات التي تحل محل المواد المحمية التي دُرِّب عليها النموذج أو تنافسها. وتكتسب هذه المسألة أهميتها لأنها تحدد ما إذا كان الضرر الاقتصادي الذي يدّعي أصحاب الحقوق تكبّده يمكن معالجته في إطار الفئات القانونية القائمة لحق المؤلف، التي تتعلق بأفعال النسخ والتبليغ للجمهور، أم أنه يتطلب آلية مستقلة، نوقشت في الجزء الثاني، للتعويض عن آثار الاستبدال التي لا يترتب على أي مخرجات منفردة، في حد ذاتها، أساس قانوني يجيز المطالبة بالتعويض عنها. وترد الأسس التقنية في الجزء الأول، القسم جيم-3.

وفي الولايات المتحدة، ناقش عدد من المحاكم هذا الموضوع بالفعل. ففي قضية The New York Times Company ضد Microsoft Corp.، التي جرى ضمها الآن إلى قضية *In re* OpenAI, Inc. Copyright Infringement Litigation، دفع المدّعى عليه في وقت مبكر من الإجراءات بردّ دعاوى التعدي المباشر على حق المؤلف، على أساس أن المدّعي استخدم ما وصفه المدّعى عليه بأنه تعليمات مُحكمة الصياغة ومتلاعَب بها لافتعال الاجترار الذي يشتكي منه. ورفضت محكمة المقاطعة الجنوبية في نيويورك ردّ تلك الدعاوى على هذا الأساس في أبريل 2025، ومضت القضية إلى مرحلة الكشف. وقد هيمن على مرحلة الكشف نزاع بشأن تقديم سجلات الاستخدام الموجهة للمستهلكين؛ وفي سلسلة من الأوامر القضائية انتهت بتأييد القاضي شتاين في 5 يناير 2026 أمراً سابقاً لقاضية الصلح وانغ، أمرت المحكمة بتقديم عيّنة عشوائية ومجهولة الهوية من 20 مليون سجل لتفاعلات مستخدمي ChatGPT.[127] ويحوّل الأمر محور الإثبات من مسألة إذا ما كان يمكن استحثاث النموذج على نسخ المحتوى عبر تعليمات استفزازية، إلى مسألة مدى تكرار حدوث ذلك أثناء الاستخدام الاعتيادي من جانب المستهلكين. وعدّت المحكمة الأدلة

---

[123] Folha de São Paulo ضد OpenAI، القضية رقم 96.2025.8.26.0100-1107237 (TJSP). رُفعت القضية في أغسطس 2025 وتسير في إطار السرية القضائية؛ ولم يصدر أي حل علني حتى تاريخ هذه الدراسة.

[124] Editorial Perfil S.A. v OpenAI, L.L.C., Microsoft Corporation and Microsoft de Argentina S.R.L. (المحاكم الاتحادية المدنية والتجارية، الأرجنتين، رُفعت في 24 أغسطس 2026). ولم تتوافر معلومات علنية تُذكر عن الدعوى حتى وقت كتابة هذه الدراسة حول الشكوى، ولكن يمكن مطالعة Editorial Perfil, "Perfil demanda a OpenAI y Microsoft por el uso de contenidos periodísticos sin autorización" (24 أغسطس 2026) <https://www.perfil.com/noticias/judiciales/perfil-demanda-a-openai-y-microsoft-por-el-uso-de-contenidos-periodisticos-sin-autorizacion.phtml> تاريخ النفاذ: 29 سبتمبر 2026. وانظر(ي) أيضاً EFE, "El diario argentino Perfil demanda a OpenAI y Microsoft por uso indebido de sus contenidos" (25 أغسطس 2026) <https://www.swissinfo.ch/spa/el-diario-argentino-perfil-demanda-a-openai-y-microsoft-por-uso-indebido-de-sus-contenidos/91953813> تاريخ النفاذ: 29 سبتمبر 2026.

[125] المحكمة الاتحادية العليا (STF)، RE 1037396 وRE 1057258 (مضمومتان). وقد قضت المحكمة الاتحادية العليا بأن المنصات قد تُسأل عن المحتوى الذي ينشره المستخدمون إذا لم تتخذ إجراء بعد تلقيها إخطاراً بالأمر، وبأن المنصات التي تقدم خدمات في البرازيل يجب أن تمتثل للتشريعات البرازيلية، وبأن هذه الشركات يجب أن يكون لها ممثل قانوني في البلد.

[126] اتفاقية برن لحماية المصنفات الأدبية والفنية، المادة 5(1) (المعاملة الوطنية)؛ وانظر(ي) أيضاً المادة 5(2) (استقلال الحماية).

[127] *In re OpenAI, Inc. Copyright Infringement Litigation*، رقم. 1:25-md-03143-SHS-OTW (محكمة المقاطعة الجنوبية في نيويورك). تجمع إجراءات التقاضي متعدد المقاطعات عدداً من الدعاوى المرفوعة ضد OpenAI من قِبل مؤسسات إخبارية ومؤلفين ومدّعين آخرين. بشأن الأمر بالكشف عن 20 مليون سجل من سجلات ChatGPT، انظر(ي) المراجع الواردة في الجزء الأول من هذه الدراسة، الحاشية 20.

الواسعة على مخرجات النموذج في ظروف الاستخدام العام غير القائمة على نهج غير استفزازي، ذات صلة بتحليل الاستخدام المنصف، لا سيما فيما يتعلق بمدى عمل مخرجات النموذج كبدائل تجارية لمصنفات المدعين.

ثم، توصلت محكمتان اتحاديتان، في حكمين موجزين صادرين عن الدائرة الشمالية في كاليفورنيا في يونيو 2025، بفارق يومين بينهما، في قضيتي Bartz ضد Anthropic PBC وKadrey ضد Meta Platforms, Inc، إلى النتيجة نفسها بشأن المسألة الرئيسية، ولكن استناداً إلى تعليلين مختلفين اختلافاً جوهرياً. في قضية Bartz، قضى القاضي ألسوب بأن تدريب نموذج لغوي كبير على كتب جرى الحصول عليها بصورة قانونية يُعد استخداماً تحويلياً بدرجة كبيرة، ويشكّل استخداماً منصفاً بموجب القسم 107 من قانون حق المؤلف؛ وقضى في الوقت نفسه بأن تنزيل نسخ مقرصنة من تلك الكتب من موقعي LibGen وPiLiMi والاحتفاظ بها لا يشكّلان استخداماً منصفاً، واقتصر اعتماد الدعوى الجماعية في 17 يوليو 2025 على مطالبات القرصنة. وتوصل الطرفان لاحقاً إلى تسوية بقيمة 1.5 مليار دولار أمريكي، حصلت على الموافقة النهائية في يوليو 128.2026 في قضية Kadrey، أصدر القاضي تشابريا بدوره حكماً موجزاً لصالح المدعى عليه بشأن مسألة الاستخدام المنصف،[129] لكن التعليل أولى أهمية أكبر للعامل الرابع من عوامل الاستخدام المنصف (الضرر بالسوق)، وأهمية أقل للعامل الأول (الاستخدام التحويلي)، محاكياً منهجية المحكمة العليا في قضية Andy Warhol Foundation ضد Goldsmith. جرى تناول هذه القضية في التعليقات اللاحقة على أنها تتمحور أساساً حول إخفاق المدّعي في تقديم أدلة على الإحلال السوقي، بما يوحي بأن سجلاً إثباتياً مختلفاً قد يؤدي إلى نتيجة مغايرة.[130] لا يُقرّر الحكمان مبدأً عاماً مفاده أن تدريب الذكاء الاصطناعي يُشكّل استخداماً منصفاً أو لا يُشكّل استخداماً منصفاً.

في قضية Thomson Reuters Enterprise Centre GmbH ضد *Ross Intelligence Inc.*، أصدرت محكمة الاستئناف الأمريكية للدائرة الثالثة أول قرار استئنافي أمريكي بشأن الاستخدام المنصف وتدريب الذكاء الاصطناعي، وذلك بعد تأييدها حكماً جزئياً في الموضوع صدر بطريق الحكم الموجز، في إطار استئناف تمهيدي. كان المدعى عليه قد استخدم الموجزات الخاصة بالمدعي، وهي ملخصات تحريرية للمسائل القانونية أُعدّت لمنصة البحث القانوني Westlaw، لتدريب أداة للبحث القانوني قائمة على الذكاء الاصطناعي وغير توليدية. فيما يتعلق بالعامل الأول، رأت المحكمة أن الاستخدام لم يُحدث، في أفضل الأحوال، سوى قدر ضئيل للغاية من التحوّل في طبيعة الاستخدام، إذ كان يخدم، إلى حد كبير، الغرض نفسه المتمثل في البحث القانوني الذي يخدمه استخدام المدعي للموجزات. وقد أدّى هذا الاستنتاج، مقترناً بالطابع التجاري للاستخدام، إلى ترجيح العامل الأول ضد اعتبار الاستخدام استخداماً منصفاً. ومَيَّزت القضايا المتعلقة بالنسخ الوسيط على أساس أن النسخ فيها كان ضرورياً للوصول إلى العناصر الوظيفية للبرمجيات غير المشمولة بالحماية، في حين لم توجد ضرورة مماثلة تبرر نسخ الموجزات. وقد رجح العامل الثاني لصالح المدعى عليه لكنه، كما أشارت المحكمة، نادراً ما يكون له وزن كبير؛ أما العامل الثالث فرجح ضد الاستخدام المنصف لأن ملخصات تمهيدية كاملة أُخذت دون مسوّغ تحويلي. فيما يتعلق بالعامل الرابع، وجدت المحكمة ضرراً على ثلاثة أصعدة: بقيمة الموجزات باعتبارها عاملاً جاذباً للاشتراك في منصة المدعي؛ وبالسوق الأصلي لمنصات البحث القانوني، الذي كان المدعى عليه ينافس فيه مباشرةً؛ وفي ختام تحليلها للاستخدام المنصف، بسوق مشتقة محتملة لترخيص الموجزات باعتبارها بيانات لتدريب الذكاء الاصطناعي، والتي وصفتها المحكمة بأنها سريعة التطور، مشيرةً إلى أن المدعي استخدم موجزاته لتدريب منتجاته الخاصة القائمة على الذكاء الاصطناعي، وأن قراره عدم منح تراخيص للآخرين لا ينفي وجود السوق، وأن النسخ غير المصرح به استولى على فرصة المدعي لدخول هذا السوق. ومَيَّزت المحكمة النماذج التوليدية محل النظر في قضيتي Bartz و In re OpenAI، ولذلك قد يختلف تقييم الطبيعة التحويلية للاستخدام في القضايا المتعلقة بالنماذج التوليدية. ومع ذلك، يؤكد القرار، على مستوى محاكم الاستئناف، أن وجود سوق ناشئة لترخيص المحتوى الذي يطوّره المبدعون عادةً أو يمنحون تراخيص للآخرين لتطويره يُعد اعتباراً يمكن أخذه في الحسبان بموجب العامل الرابع.[131] لم تُنهِ موجة التسويات في قطاع الموسيقى الدعاوى

---

[128] قضية *Bartz ضد Anthropic PBC*، رقم cv-05417-3:24 (محكمة المقاطعة الشمالية في كاليفورنيا، القاضي ألسوب، 23 يونيو 2025) (منح حكماً موجزاً جزئياً لصالح المدعى عليه على أساس الاستخدام المنصف فيما يتعلق بالتدريب على المصنفات المكتسبة بصورة مشروعة، ورفض منح حكم موجز للمدعى عليه فيما يتعلق بتنزيل والاحتفاظ بنسخ مقرصنة من المصنفات المأخوذة من موقعي LibGen وPiLiMi). وأُقرت الدعوى الجماعية فيما يتعلق بمطالبات القرصنة فقط. وحصلت تسوية بقيمة 1.5 مليار دولار أمريكي على مواقة مبدئية في 25 سبتمبر 2025؛ وعُقدت جلسة الاستماع النهائية بشأن الإنصاف في 14 مايو 2026 أمام القاضية مارتينيز-أولغين، التي حجزت الدعوى للمداولة والدراسة. وكانت المواقة النهائية قيد الانتظار حتى تاريخ إعداد هذه الدراسة. وحصلت التسوية على المواقة النهائية من محكمة المقاطعة في يوليو 2026، وتَمّ إغلاق القضية.

[129] *قضية Kadrey. ضد Meta Platforms, Inc*، رقم cv-03417-VC-3:23 (محكمة المقاطعة الشمالية في كاليفورنيا، القاضي تشابريا، 25 يونيو 2025).

[130] حول التباين بين تعليلَي القاضيَيْن، يُراجع (من بين التعليقات العديدة) التحليل الوارد في منشور مكتب Wiggin and Dana الصادر في 18 سبتمبر 2025، والمتاح على الموقع wiggin.com.

[131] *قضية Thomson Reuters Enterprise Centre GmbH ضد Ross Intelligence Inc.*، رقم cv-613-1:20 (محكمة مقاطعة ديلاوير، القاضي بيباس، 11 فبراير 2025) (منح حكماً موجزاً جزئياً للمدعي بشأن الانتهاك المباشر ورفض دفع الاستخدام المنصف). وأيّد حكم محكمة المقاطعة في استئناف فرعي: رقم 25-2153 (الدائرة الثالثة، جرت المرافعة في 11 يونيو 2026، وصدر الحكم في 29 سبتمبر 2026) (الرأي الصادر عن القاضية مونتغومري-ريفز). وفي حاشية، أشارت محكمة الاستئناف إلى بيان المصلحة المُقدم من وزارة العدل الأمريكية في قضية In re OpenAI, Inc.

القضائية فيه. في يناير 2026، رفعت مجموعة من ناشري الموسيقى دعوى أخرى ضد المطوّر في إطار النزاع القضائي المتعلق بقضية Bartz، مدعيةً استخدام أكثر من عشرين ألف مصنف موسيقي، ومطالبةً بتعويضات قانونية تتجاوز 3 مليارات دولار أمريكي، كما واصلت جمعيات الإدارة الجماعية في ألمانيا والدانمرك ملاحقة مطوّر آخر لأنظمة توليد الموسيقى في دعوى منفصلة.[132]

وخارج الولايات المتحدة، تناولت المحاكم أيضًا هذا الموضوع. وقد طرحت قضية Folha de São Paulo ضد OpenAI الموصوفة أعلاه هذه المسألة إلى أن توصل الأطراف إلى تسوية بشأنها ومن الأمثلة التي لا تزال قيد النظر حتى وقت كتابة هذا التقرير القضية المنظورة أمام محكمة الإنترنت في بيجين، والتي رفع فيها أربعة رسامين دعوى ضد Xiaohongshu، مدعين أن المنصة استخدمت مصنفاتهم الأصيلة لتدريب نموذج الذكاء الاصطناعي لتوليد الصور Trik دون تصريح، وأن النموذج قادر على إنشاء صور متطابقة جوهرياً مع مصنفات المدعين.[133]

وفي يناير 2025 في جمهورية كوريا، رفعت هيئات البث التلفزيوني الأرضي الثلاث الكبرى (KBS وMBC وSBS) دعوى قضائية ضد Naver أمام محكمة مقاطعة سول المركزية، مدعية استخدام محتواها الإخباري دون تصريح لتدريب مجموعة النماذج اللغوية HyperCLOVA التابعة لشركة Naver. وتُعد هذه أول دعوى قضائية كورية تُطرح فيها مسألة التدريب بشكل مباشر، ونظراً للأصل المشترك بين المادة 35-5 من قانون حق المؤلف الكوري والقسم 107 من قانون حق المؤلف الأمريكي، فإن السوابق القضائية الأمريكية المتعلقة بالاستخدام المنصف والمستعرضة أعلاه تُشكل جزءاً من الدفوع المقدمة أمام المحكمة الكورية.[134] وفي البرازيل، أقام مؤدي صوتي دعوى قضائية ضد مركز تجاري، مدعياً استخدام صوته دون تصريح في الإعلانات؛ ودفع المركز التجاري في المقابل بأن الصوت المعني تم توليده بواسطة الذكاء الاصطناعي، وقد قُبل هذا الدفع أمام محكمة الدرجة الأولى. غير أنه في عام 2024، ألغت محكمة استئناف ساو باولو الحكم، وقضت بأن "استخدام الذكاء الاصطناعي في حد ذاته لا ينفي (بل على العكس، يفاقم) خطر الاستخدام غير المشروع لحقوق الغير"، بالنظر إلى أن هذه الأدوات يُجرى تدريبها باستخدام محتوى مستخرج من شبكة الإنترنت عن طريق تقنية استخلاص البيانات دون تصريح من أصحاب الحقوق. كما رأت المحكمة ضرورة جمع المزيد من الأدلة، وأعادت القضية إلى محكمة الدرجة الأولى؛ ولم يُصدر فيها حكم نهائي في الموضوع حتى تاريخ إعداد هذه الدراسة.[135]

أما في الهند، على النقيض من ذلك، فقد حظيت مسألة التدريب بالفعل بأكثر معاملة قضائية شاملة وقائمة على الحجج المنطقية حتى الآن خارج الولايات المتحدة وألمانيا. وفي قضية ANI Media Pvt. Ltd ضد OpenAI OPCO LLC، زعمت وكالة أنباء أن

Copyright Infringement Litigation، رقم 1:25-md-3143 (محكمة المقاطعة الجنوبية في نيويورك، 1 سبتمبر 2026)، والذي دفع بأن تدريب نموذج لغوي توليدي كبير يُعدّ استخداماً تحويلياً، وميّزت المحكمة السياق التوليدي عن القضية المعروضة أمامها.

[132] رفعت كلٌ من Universal Music Publishing Group وConcord وABKCO دعوى قضائية ثانية ضد Anthropic أمام محكمة الولايات المتحدة للمقاطعة الشمالية في كاليفورنيا في 28 يناير 2026، مدعية انتهاك حقوق المؤلف لأكثر من 20,000 مصنف. وتُعدّ هذه الدعوى مستقلة عن الدعوى الأصيلة التي أقامها الناشرون في عام 2023 بشأن كلمات الأغاني ضد Anthropic، والتي رُفعت في البداية أمام محكمة المقاطعة الوسطى في تينيسي قبل تحويلها لاحقاً إلى محكمة المقاطعة الشمالية في كاليفورنيا. يُراجع Music Business Worldwide، مقال " UMG, Concord and ABKCO sue Anthropic for $3bn in what could be single-largest non-class action copyright case" (28 يناير 2026) <https://www.musicbusinessworldwide.com/umg-concord-and-abkco-sue-anthropic-for-3bn-in-what-could-be-single-largest-non-class-act/>تم النفاذ إليه في 3 يوليو 2026؛ وحول دعوى عام 2023، يُراجع Rolling Stone، مقال "Music Publishers Sue Amazon-backed AI Platform for Copyright Infringement" (18 أكتوبر 2023) <https://www.rollingstone.com/music/music-news/umg-concord-abkco-sue-anthropic-ai-company-1234856761/>تم النفاذ إليه في 3 يوليو 2026؛ وMusic Business Worldwide، مقال "Universal, Concord and ABKCO ask court for injunction to stop Anthropic AI from using their lyrics" (16 نوفمبر 2023) <https://www.musicbusinessworldwide.com/universal-concord-and-abkco-ask-court-for-injunction-to-stop-anthropic-ai-from-using-their-lyrics/>، تم النفاذ إليه في 3 يوليو 2026. واستمرت GEMA (ألمانيا) وKoda (الدانمرك)، من بين جهات أخرى، في مقاضاة Suno على الرغم من التسويات التي توصلت إليها شركات التسجيلات مع ذلك المطوّر ومع Udio في أواخر عام 2025.

[133] محكمة الإنترنت في بيجين (رُفعت في نوفمبر 2023؛ وقيد النظر حتى تاريخ إعداد هذه الدراسة)؛ والمدعون الأربعة هم رسامون يدعون أن Xiaohongshu استخدمت مصنفاتهم الأصيلة لتدريب نموذج الذكاء الاصطناعي لتوليد الصور Trik دون تصريح.

[134] قضية KBS، وMBC، وSBS ضد Naver Corp.، محكمة مقاطعة سول المركزية (رُفعت في 13 يناير 2025). وقد أُقيمت المطالبات بموجب قانون حق المؤلف وقانون منع المنافسة غير المشروعة، ويطالب المدعون بالتعويض عن الأضرار وبإصدار أمر قضائي بمنع أي استخدام مستقبلي. ورفعت هيئات البث ذاتها دعوى مماثلة ضد OpenAI أمام المحكمة نفسها في فبراير 2026.

[135] محكمة العدل في ساو باولو، الاستئناف المدني رقم 41.2023.8.26.0100-1119021 (الدائرة السادسة للقانون الخاص، القاضي المقرر كوستا نيتو، حكم صادر في 31 أكتوبر 2024)، عُرض في منتدى قضاة الملكية الفكرية التابع للويبو لعام 2025؛ يُراجع <https://www.wipo.int/wipolex/en/judgments/details/3217> تم النفاذ في 19 سبتمبر 2026.

مصنفاتها قد استُخدمت دون إذن، سواء لتدريب نماذج اللغة الخاصة بالمدعى عليه أو في إنتاج المخرجات. كما أن محكمة دلهي العليا، في حكمها بشأن طلب الوكالة إصدار أمر قضائي مؤقت، وبعد أن أيدت اختصاصها الإقليمي على عمليات التدريب التي أجريت على خوادم موجودة في الولايات المتحدة (انظر(ي) القسم أ-3 أعلاه)، قد قضت بأن عملية التدريب تنطوي على حق الاستنساخ، إذ أن الفقرة الفرعية "1" من المادة 14(أ) من قانون حق المؤلف لعام 1957 توسع نطاق هذا الحق ليشمل تخزين المصنف "في أي وسيط باستخدام وسائل إلكترونية"، كما أن حذف النسخ بعد انتهاء التدريب لا يلغي عملية التخزين. غير أن المحكمة رأت، مبدئياً، أن عملية التخزين تندرج ضمن أحكام الاستخدام العادل الواردة في المادة 52(1)(أ)، التي تشمل الاستخدامات لأغراض "الاستخدام الخاص أو الشخصي، بما في ذلك البحث". وقد فُسر مصطلح "البحث" بطريقة حديثة على النحو التالي: "يمكن اعتبار عملية تدريب النماذج اللغوية الكبيرة التي يقوم عليها برنامج ChatGPT شكلاً من أشكال البحث الرامي إلى توليد معرفة جديدة"، و"لا يمكن حصر أنشطة البحث الإضافية في الأعمال التي يؤديها البشر وحدهم"، لأن "البحث، في نهاية المطاف، يتم بناءً على طلب البشر ولصالحهم". كما أن الطابع التجاري للمدعى عليه لم يجعل هذا الاستخدام خارج نطاق أحكام القانون: فقد وصفت المحكمة هذا الاستخدام بأنه "خاص تماماً"، إذ بقيت نسخ التدريب في بيئة مغلقة يمكن للنماذج فقط الوصول إليها دون الجمهور، ولا تتضمن المادة 52(1)(أ) أي استثناء للاستخدامات التجارية.

وفيما يتعلق بمعيار الإنصاف، رأت المحكمة أن استخدام المصنفات اقتصر على تدريب النماذج، وأن ملف القضية لم يتضمن أي دليل على حدوث استبدال في السوق أو خسائر في الإيرادات بما يثبت الادعاءات، وأن الفوائد العامة لهذه التكنولوجيا أُخذت بعين الاعتبار؛ كما أشارت إلى أن المدعي كان بإمكانه حجب عناكب البحث الخاصة بالمدعى عليه، لكنه لم يفعل ذلك. وقد رُفض الأمر القضائي المؤقت استنادًا إلى سبب إضافي مفاده أن أي خسارة، في حال ثبوتها خلال المحاكمة، ستكون قابلة للتعويض عن طريق دفع تعويضات، إذ أشارت المحكمة إلى أن اشتراط الحصول على تراخيص من كل مصدر إخباري مماثل من شأنه أن يجعل تطوير النماذج اللغوية الكبيرة غير مجدي اقتصاديًا. وهذا القرار هو قرار مؤقت: فالدعوى ستُعرض على المحكمة، وفي تاريخ إعداد هذه الدراسة، كان الاستئناف معروضاً أمام هيئة قضائية للنظر فيه. ومع ذلك، فإن أهميته من الناحية المقارنة كبيرة. ففي حين طرحت المحاكم الألمانية سؤالاً بشأن ما إذا كان الاستثناء الخاص باستخراج النصوص والبيانات لأغراض محددة يشمل تدريب الذكاء الاصطناعي، وطرحت المحاكم الأمريكية سؤالاً بشأن ما إذا كان التدريب يندرج ضمن الاستخدام العادل، فإن محكمة دلهي العليا توصلت إلى نفس السؤال من خلال بند التعامل العادل الذي ينص على "الاستخدام الخاص أو الشخصي، بما في ذلك البحث"، وأجابت عليه، في هذه المرحلة، لصالح المطور، استناداً إلى الحجج التالية (عدم وجود ضرر سوقي مثبت، وتوافر خيار الانسحاب غير المستخدم، والفوائد العامة للتكنولوجيا)، وهو ما يوازي العوامل التي تم تقييمها في الأحكام الأمريكية التي تم استعراضها أعلاه.

القسم باء: إنفاذ الحقوق ضد مطوري الذكاء الاصطناعي

تُوضح السوابق القضائية المستعرضة في القسم ألف التطورات الموضوعية التي شهدتها الثمانية عشر شهراً الماضية. وينتقل القسم باء إلى المسألة الهيكلية المتعلقة بكيفية إعمال أصحاب الحقوق لمطالباتهم ضد مطوري الذكاء الاصطناعي، في ظل الخصائص العابرة للحدود لهذه التقنية وأوجه عدم التماثل بين الأطراف المشاركة عادةً في مثل هذه النزاعات.

1. الخصائص الهيكلية لبيئة الإنفاذ

تتكرر ثلاث خصائص هيكلية في هذا الصدد. تتمثل الخاصية الأولى في استمرار حدة المشكلة العابرة للحدود وتفاقمها، على النحو الموضح في القسم ألف - 3 أعلاه. فعادةً ما يُجرى تدريب نموذج الذكاء الاصطناعي التوليدي عبر بنية تحتية تقع في ولاية قضائية تختلف عن الولاية القضائية لصاحب الحقوق؛ ثم يُنشر النموذج ويُستخدم في العديد من الولايات القضائية في وقت واحد؛ في حين تُنتَج المخرجات في الولاية القضائية الخاصة بكل مستخدم. وقد تخضع كل مرحلة من هذه المراحل لقانون بلد مختلف. ولم تُسفر السوابق القضائية حتى الآن عن إجابة موحدة بشأن كيفية تفاعل هذه الأنظمة القانونية الوطنية المختلفة، وقد يواجه المطورون التزامات متضاربة عبر الولايات القضائية التي يُقدمون فيها خدماتهم.

وتتمثل الخاصية الثانية في دور مبدأ المعاملة الوطنية بموجب المادة 5(1) من اتفاقية برن. وبموجب هذا المبدأ، يتمتع المؤلفون في بلدان الاتحاد - غير بلد منشأ المصنف - بالحقوق التي تمنحها قوانين تلك البلدان لمواطنيها، بالإضافة إلى الحقوق التي تمنحها الاتفاقية بشكل خاص. ويتمثل هذا المبدأ في عدم التمييز، ولا يحدد في حد ذاته النطاق الإقليمي للقانون الوطني.

وتتمثل الخاصية الثالثة في استخدام الآليات الإجرائية لضم الدعاوى والمطالبات، ويُعدّ التقاضي متعدد المقاطعات في الولايات المتحدة المثال الرئيسي على ذلك. وقد أدى توحيد القضايا المرفوعة ضد شركة OpenAI من قِبل المؤسسات الإخبارية والمؤلفين وغيرهم من المدعين في محكمة المقاطعة في المنطقة الجنوبية من نيويورك ضمن قضية تقاضٍ واحدة متعددة المقاطعات إلى تركيز الإجراءات التي كانت ستتشتت لولا ذلك عبر دعاوى فردية. وبالمثل، قامت الدعوى الجماعية في قضية Bartz بتجميع المطالبات المتعلقة بالنسخ المقرصنة لنحو 482,000 كتاب، وأسفرت عن تسوية ذات حجم وأهمية بالغين، كما ورد أعلاه. أما في الولايات القضائية التي تغيب فيها آليات الدعاوى الجماعية (أو"المطالبات الجماعية") أو التقاضي متعدد المقاطعات أو تكون أقل تطوراً، فقد تكون تكلفة الإنفاذ الفردي باهظة للغاية، لا سيما بالنسبة للمؤلفين الأفراد وصغار أصحاب الحقوق. وتُعد الدعاوى الجماعية في كيبيك الموصوفة في القسم ألف - 3 محاولة لتطبيق الآلية ذاتها في بيئة إجرائية مختلفة؛ كما تُقدم الدعاوى

الجماعية في كولومبيا البريطانية صورة مماثلة. وفي الولايات القضائية التي تفتقر إلى مثل هذه الآليات، تتأثر القدرة العملية لأصحاب الحقوق الأفراد على إنفاذ حقوقهم ضد المطورين ذوي الموارد الضخمة بشكل كبير.

2. أوجه عدم التماثل بين الأطراف

وإلى جانب هذه الخصائص، تُحدد السوابق القضائية عدداً من أوجه عدم التماثل التي تُعقد عملية الإنفاذ من الناحية العملية. يتمثل أول وجه لعدم التماثل في "عدم التماثل المعلوماتي": حيث يمتلك المطورون السجلات التقنية (بيانات مجموعات البيانات، وسجلات الزحف، وإعدادات التدريب) التي يمكن من خلالها -من حيث المبدأ- إثبات مكونات ذخيرة بيانات التدريب، حتى لو كانت تلك السجلات غير مكتملة؛ في حين لا يتوفر لأصحاب الحقوق نفاذ مماثل، إلا في الحالات التي يؤدي فيها الاستخراج أو العروض التجريبية العملية الأخرى إلى تقديم أدلة غير مباشرة. وهذا هو موضوع مناقشة الشفافية في القسم جيم-1 أدناه. ويتمثل وجه عدم التماثل الثاني في "عدم تماثل الموارد": حيث يمتلك المطورون، لا سيما مطورو النماذج المتقدمة الكبرى، موارد تتجاوز بكثير ما يمتلكه معظم أصحاب الحقوق. إذ تُعد تكلفة إقامة دعوى إنفاذ مستمرة، إلى جانب عمل الخبراء الفني والاستشاري المتوقع الآن في مثل هذه القضايا، تكلفة كبيرة. ويتمثل وجه عدم التماثل الثالث في عدم اليقين الفقهي والقانوني: فعندما يكون القانون غير مستقر، قد تتأثر السوابق القضائية المنشورة بآثار الانتقاء، لأن الأطراف يُميلون إلى تسوية بعض النزاعات والتقاضي في أخرى، في حين تظل قدرة أصحاب الحقوق على رفع دعاوى نموذجية اختبارية مقيدة بتكاليف التقاضي. ومن ثم، ينبغي تقييم السوابق القضائية الناتجة عن ذلك مع أخذ آثار الانتقاء هذه في الاعتبار.

القسم جيم: التحديات في إنفاذ الحقوق

تواجه أصحاب الحقوق الساعين إلى إنفاذ حقوقهم ضد مطوري الذكاء الاصطناعي ثلاثة تحديات رئيسية في ظل الوضع الحالي للقانون والتكنولوجيا. ونتناول كلاً منها على التوالي.

1. الشفافية

يتمثل التحدي الأول في محدودية الشفافية المتعلقة بعملية التدريب. فقد امتنع مطورو الذكاء الاصطناعي، إلا في حالات نادرة جداً، عن الكشف عن مكونات ذخائر بيانات التدريب الخاصة بهم؛ وحتى في الحالات التي قُدمت فيها عمليات كشف موجزة، كانت بأسلوب عام لا يُمكّن صاحب حقوق من تحديد ما إذا كان مصنفه المعين قد أُستخدم أم لا. ويتعزز هذا الغموض بالادعاء بوجود أسرار تجارية: ففي أوائل عام 2026، على سبيل المثال، قدم مطور نظام لتوليد الموسيقى طلباً لإبقاء حجم مجموعة بيانات التدريب الخاصة به طي الكتمان والسرية في دعوى قضائية أقامتها شركات موسيقى كبرى، متعللاً بالحفاظ على التنافسية وتجنب الضرر التنافسي.[136] وأحد الأسباب هو أن مكونات ذخيرة بيانات التدريب قد تُشكل أصلاً تجارياً، وقد يؤثر الكشف الشامل عن محتوياته على الموقف التنافسي للمطور. والنتيجة هي عدم تماثل في المعلومات يُمكن أن يحد من القدرة العملية لأصحاب الحقوق على تحديد الانتهاكات المحتملة ومباشرة إجراءات الإنفاذ.

وقد عالج الاتحاد الأوروبي هذه المسألة من خلال اشتراط قانوني صريح بالكشف. حيث تقتضي المادة 53(1)(د) من قانون الذكاء الاصطناعي من مزودي نماذج الذكاء الاصطناعي عامة الأغراض إعداد وإتاحة "ملخص تفصيلي بدرجة كافية" للجمهور بشأن المحتوى المستخدم لتدريب النموذج، وذلك وفقاً للنموذج المقترح من مكتب الذكاء الاصطناعي التابع للاتحاد الأوروبي.[137] ويتوقف التنفيذ جزئياً على مستوى التفصيل الذي يُعدّ كافياً لأغراض هذا النص. علاوة على ذلك، يتقيد هذا النص بالاعتراف الصريح بضرورة احترام الأسرار التجارية والمعلومات التجارية السرية، ويُعد تنفيذ هذا الشرط حالياً موضوعاً لجهود السياسات والتقنيات القياسية الجارية.[138] ويُشكل هذا التوتر بين التزام الكشف واستثناء الأسرار التجارية موضوعاً لنقاش واسع. ومن المرجح أن يعتمد النطاق التشغيلي النهائي لهذا الالتزام على كل من تطبيق النموذج المنشور من قِبل مكتب الذكاء الاصطناعي في يوليو 2025 والتفسير الذي تضفيه السلطات القضائية والوطنية الأخرى على هذا الاستثناء.

---

[136] يُراجع Music Business Worldwide، مقال " Suno moves to keep size of its AI training data sealed in UMG and Sony's copyright case, citing competitive harm" (2026) <https://www.musicbusinessworldwide.com/suno-moves-to-keep-size-of-its-ai-training-data-sealed-in-umg-and-sonys-copyright-case-citing-competitive-harm/> ، تم النفاذ إليه في 7 يوليو 2026.

[137] اللائحة (الاتحاد الأوروبي) 1689/2024، المادة 53(1)(د)، المشار إليها في الجزء الثاني، الحاشية 82، من هذه الدراسة.

[138] قانون الذكاء الاصطناعي، الحيثية 107، التي تبيّن بوضوح أنه ينبغي صوغ الالتزام بنشر ملخص على نحو يحترم الأسرار التجارية والمعلومات التجارية السرية ولا يستلزم الكشف عنها.

واتبعت ولايات قضائية أخرى نهجاً في معالجة الشفافية من خلال أدوات تشريعية مختلفة؛ وفي الصين، يُطبق نظام شامل من التزامات وضع العلامات على المحتوى المُنتج بواسطة الذكاء الاصطناعي[139]ولا يتجه النظام الصيني لوضع العلامات في المقام الأول إلى حماية حق المؤلف بل إلى منع الغش والخداع والالتباس لدى الجمهور؛ ومع ذلك فقد كان الأثر العملي لأصحاب الحقوق إتاحة قدر من الاستدلال الرجعي، إذ إن وضع العلامات على المحتوى المولَّد بالذكاء الاصطناعي يتيح تحديد مخرجات بعينها بأنها مولَّدة بالذكاء الاصطناعي والتحقق انطلاقاً من تلك المخرجات ما إذا كان من المرجح استخدام مصنفات محمية في تدريب النموذج.

وفي جمهورية كوريا، يشترط القانون الإطاري بشأن تطوير الذكاء الاصطناعي وإرساء أساس للثقة، النافذ منذ 22 يناير 2026، إخطار مستخدمي خدمات الذكاء الاصطناعي التوليدي مسبقاً ووضع علامات على المحتوى الذي ينتجه الذكاء الاصطناعي.[140] ولا تهدف التزامات وضع العلامات، على غرار تلك المعمول بها في الصين، إلى حماية حق المؤلف في المقام الأول؛ لكنها قد تيسّر النوع نفسه من الاستدلال الرجعي انطلاقاً من المخرجات. تُعالج مسائل حق المؤلف بموجب قانون حق المؤلف، وليس بموجب التشريع الإطاري، وهو توزيع للاختصاصات بين التشريعين يعكس العلاقة القائمة في الاتحاد الأوروبي بين قانون الذكاء الاصطناعي وتشريعات حق المؤلف.

وفي البرازيل، يتجسد الرد التشريعي في مشروع القانون رقم 2023/2338، الذي يلزم، بصيغته الحالية، مطوري الذكاء الاصطناعي بالكشف عن استخدام المواد المحمية في تطوير نظمهم عن طريق نشر ملخص؛ ويخضع الالتزام لتحفظ مفاده وجوب مراعاة الأسرار التجارية والصناعية.141 سيعتمد الأثر العملي لالتزام الكشف على مدى اتساع نطاق تفسير استثناء الأسرار التجارية. وتنطبق الملاحظة نفسها على الإطار الأوروبي مع إجراء ما يلزم من تعديلات.

وثمة ملاحظتان تتعلقان بالشفافية باعتبارها مسألة تتعلق بالإنفاذ. الأولى هي أن الشفافية قد تيسّر الإنفاذ، لكنها ليست كافية بحد ذاتها لتحقيق ذلك. فنشر ملخّص، مهما بلغت درجة تفصيله، لا ينشئ في حد ذاته سبباً قانونياً لإقامة دعوى؛ إذ يظل على صاحب الحقوق إثبات أن مصنَّفه قد استُخدم، وأن هذا الاستخدام مسّ حقوقه الاستئثارية، وأنه لا يندرج ضمن أي استثناء. ومن ثم، فإن التزام الشفافية يحدّ من التفاوت في المعلومات، لكنه لا يزيله تماماً. أما الملاحظة الثانية فتتعلق بكيفية تأثير تصميم التزام الشفافية في تصميم آلية الاعتراض على الاستخدام. فإذا كان صاحب حقوق قد أعلن اعتراضه على استخدام مصنَّفه، ثم تبيّن من ملخّص الشفافية أن مصنَّفه قد أُدرج ضمن مجموعة مواد التدريب، فقد يتوافر له أساس أوضح للمطالبة بإنفاذ حقوقه، مع مراعاة القانون الواجب التطبيق وسائر شروط الاستثناء المعني؛ ويتمثل التساؤل العملي في مدى تحديد الملخّص لهوية المصنَّفات على نحو يكفي للتعرّف عليها.

2. الهجمات في استخراج البيانات، والتصفية، وتعدد معايير الاعتراض على الاستخدام

يتعلق التحدي الثاني بالتقنيات التي لجأ إليها أصحاب الحقوق والمطوّرون والوسطاء لمعالجة الخصائص التقنية لعمليتي التدريب والاستدلال التي تؤثر في الإنفاذ. ويمكن التمييز بين أربع فئات مستقلة من هذه التقنيات.

تتمثل الفئة الأولى في هجمات الاستخراج، التي طُوّرت في إطار الأدبيات المتعلقة بأمن المعلومات، ويُستند إليها الآن في التقاضي. وهي تقنيات تهدف إلى حثّ نموذج مدرَّب على استنساخ محتوى من بيانات التدريب الخاصة به، من خلال صياغة موجِّهات مدروسة بعناية. وقد استُشهد بمجموعة المصنّفات في هذا المجال، التي يستعرضها الجزء الأول من هذه الدراسة، في أدبيات أمن المعلومات بوصفها دليلاً على أن قابلية استخراج المحتوى قد تكون خاصية بنيوية للنماذج التوليدية الكبيرة، لا مجرد نتيجة لاستراتيجية معيّنة لصياغة الموجِّهات.[142] ومن منظور أصحاب الحقوق، تُعدّ هجمات الاستخراج وسيلةً لجمع الأدلة: إذ قد توفّر أدلة على إدراج مصنَّفات معينة ضمن بيانات التدريب أو إمكانية استرجاعها من النموذج. أما من منظور المطوّرين، فتمثل الهجمات نفسها ثغرةً أمنية ينبغي الحد منها من خلال إزالة التكرار من بيانات التدريب، وترشيح المخرجات، وتقنيات إلغاء التعلّم.

---

[139] أحكام إدارة خدمات التوصية الخوارزمية لخدمات المعلومات على الإنترنت (جمهورية الصين الشعبية)، النافذة اعتباراً من 1 مارس 2022؛ أحكام إدارة التوليف العميق في خدمات المعلومات على الإنترنت (جمهورية الصين الشعبية)، النافذة اعتباراً من 10 يناير 2023؛ التدابير المؤقتة لإدارة خدمات الذكاء الاصطناعي التوليدي (جمهورية الصين الشعبية)، النافذة اعتباراً من 15 أغسطس 2023؛ تدابير وضع العلامات على المحتوى المولَّد والمركَّب بالذكاء الاصطناعي (جمهورية الصين الشعبية)، النافذة اعتباراً من 1 سبتمبر 2025؛ والمعيار الوطني الإلزامي الذي يحمل العنوان نفسه والذي أقرته إدارة الدولة لتنظيم السوق وتوحيد المعايير في عام 2025.

[140] القانون الإطاري بشأن تطوير الذكاء الاصطناعي وإرساء أساس للثقة (جمهورية كوريا)، الذي صدر في يناير 2025 ونفذ منذ 22 يناير 2026.

[141] مشروع القانون رقم 2023/2338 (البرازيل)، المادة 62 (بالصيغة المحالة إلى مجلس النواب)؛ انظر(ي) الجزء الثاني، الحاشية السفلية 100، من هذه الدراسة، للاطّلاع على الوضع التشريعي لمشروع القانون.

[142] للاطلاع على المواد المتعلقة بهجمات الاستخراج، انظر(ي) الحاشية السفلية 18 من الجزء الأول (كوبر، وجوكاسلان، وأحمد، وسايفرت، ودي سا، وليملي، وهو وليانغ) والحاشية السفلية 22 (كارليني وآخرون (2021) ونصر وآخرون (2025)).

وتتمثل المسألة القانونية المطروحة حالياً أمام المحاكم في عدة ولايات قضائية، على النحو الموضح في القسم جيم-2 من الجزء الأول من هذه الدراسة، في ما إذا كان ينبغي أن تؤثر الطريقة التي يُستحثّ بها النموذج على إخراج مُخرج من المخرجات المحفوظة في ذاكرته في التحليل الموضوعي لحق المؤلف.

وتتمثل الفئة الثانية في مجموعة مرشحات المخرجات والمدخلات. فقد عمد المطورون، بوتيرة متزايدة، إلى استخدام مرشحات تمنع النموذج من استنساخ محتوى معين تم تحديده (مرشحات المخرجات) أو تستبعد محتوى معيناً من ذخيرة بيانات التدريب في مرحلة إعداد البيانات (مرشحات المدخلات).[143] ولا يزال دور هذه المرشحات في التحليل الموضوعي غير محسوم. ففي الولايات المتحدة، تناولت الأحكام الصادرة في قضيتي Bartz وKadrey دور المرشّحات بطرائق مختلفة؛ إذ اعتبر القاضي تشابريا ج، على وجه الخصوص، أن مسألة إحلال المخرجات في السوق تتوقف جزئياً على مدى فاعلية المرشّحات في منع النموذج من إنتاج مخرجات يمكن أن تحل محل مصنَّفات أصلية محدّدة. أما في الاتحاد الأوروبي، فتشير التزامات الشفافية المنصوص عليها في قانون الذكاء الاصطناعي إلى السياسات التي يعتمدها المطوّرون فيما يتعلق بحق المؤلف؛ وقد يكون لكيفية تنفيذ هذه الالتزامات أثر في التكييف القانوني لممارسات الترشيح باعتبارها جزءاً من الموقف العام الذي يتّخذه المطوّر من الامتثال للقانون.

تتمثل الفئة الثالثة في مجموعة تقنيات تسميم البيانات، التي يُدخل أصحاب الحقوق أو الجهات الخارجية، بموجبها، بيانات خبيثة أو خطأ أو محرّفة إلى مجموعات البيانات، بهدف الإضرار بأداء أي نموذج يُدرَّب على تلك المجموعات. وفي الصين، تعاملت إدارة الفضاء السيبراني مع تسميم البيانات باعتباره تهديداً أمنياً، وأطلقت إجراءً خاصاً لمكافحته (يُعرف باسم "تشينغلانغ: تصحيح الفوضى في تطبيقات الذكاء الاصطناعي")؛ ويشترط المعيار الوطني الموصى به بشأن المتطلبات الأمنية الأساسية لخدمات الذكاء الاصطناعي التوليدي ترشيح بيانات التدريب لإزالة المعلومات غير القانونية والضارة.[144] ولم يظهر حتى الآن في الولايات القضائية الأخرى التي شملتها الدراسة أي نهج تنظيمي مماثل للتعامل مع تسميم البيانات. كذلك، لم يُقيَّم بعد، على نحو موثوق ومعتمد، مدى الأهمية العملية لتسميم البيانات بوصفه آليةً تتيح لأصحاب الحقوق حماية حقوقهم بأنفسهم؛ ولا تزال المسائل التقنية والقانونية التي تثيرها مفتوحة في جوهرها: ومنها ما إذا كان استخدام تقنيات التسميم ضد مطوّري الذكاء الاصطناعي يشكّل اعتداءً مقرّاً به قانوناً على أي حق يتمتع به المطوّر، أم أنه ممارسة مشروعة للحقوق التي يملكها صاحب الحقوق نفسه.

والنقطة الرابعة تكتسب أهمية خاصة في سياق الإنفاذ في أوروبا. فهي تتمثل في مسألة المعايير الخاصة بالتعبير عن الاعتراض على الاستخدام بصيغة قابلة للقراءة آلياً.[145] فقد قضى حكم محكمة الاستئناف في هامبورغ في قضية Kneschke ضد LAION e.V.، التي نوقشت أعلاه في القسم باء-5 من الجزء الثاني، بأن صحة الاعتراض على الاستخدام بموجب المادة 4(3) من التوجيه بشأن حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة تقتضي أن يكون الاعتراض قابلاً للقراءة آلياً وقابلاً للتنفيذ آلياً في آن واحد. ولم يُحسم بعد كيفية تنفيذ هذا المطلب من الناحية التقنية. ويجري حالياً العمل على وضع مفردات موحدة لتفضيلات الاستخدام المتعلقة بالذكاء الاصطناعي ضمن مجموعة العمل المعنية بتفضيلات الذكاء الاصطناعي التابعة لفرقة العمل المعنية بهندسة الإنترنت، وكان هذا العمل مستمراً حتى منتصف عام 2026؛ وتظل مسوّدات العمل قيد الإعداد، ولا تحدّد سوى عدد محدود من فئات الاستخدام (هي في جوهرها تدريب الذكاء الاصطناعي والاستخدامات المرتبطة بالبحث، مع وجود فئة قيد المناقشة تتعلق بالتوليد المعزّز بالاسترجاع)، ولم يُعتمد أي معيار حتى الآن.[146] ومن المسائل ذات الصلة التي قد يكون من المفيد دراستها هي كيفية توافق الالتزام بالاعتراض على الاستخدام مع حظر الإجراءات الشكلية بشأن وجود الحقوق المحمية بموجب اتفاقية برن أو ممارستها. وقد تعتمد الإجابة على التكوين الفعلي لآلية الاعتراض على الاستخدام وأثرها القانوني.

وقد ظهر بالتوازي عدد من المعايير المتنافسة، ويمكن تقسيمها إلى مجموعتين يسهل الخلط بينهما. المجموعة الأولى تُعبر فقط عن الاحتفاظ بالحقوق، أي الاعتراض على الاستخدام. ويشمل ذلك بروتوكول الاحتفاظ بالحقوق في تنقيب النصوص والبيانات (TDMRep)، وبروتوكول TDM·AI، وكلاهما موجّه تحديداً إلى إتاحة إبداء الاعتراض على الاستخدام بما يتوافق مع المادة 4(3)

---

[143] قد يشمل الترشيح في مرحلة الإدخال استبعاد المصنَّفات التي تشير بياناتها الوصفية إلى الاحتفاظ بالحقوق، وإزالة التكرار من ذخيرة البيانات المستخدمة، للحد من مخاطر الحفظ في الذاكرة، واستبعاد المحتوى الوارد من مصادر معروف أنها بدون ترخيص. أما الترشيح في مرحلة الإخراج، فقد يشمل اكتشاف المخرجات التي تتطابق بدرجة كبيرة مع مصنَّفات محمية محددة وحجبها. ويختلف مدى التركيز على الترشيح في مرحلتي الإدخال والإخراج باختلاف المطوّرين وأنواع النماذج.

[144] إدارة الفضاء السيبراني الصينية، الإجراء الخاص "تشينغلانغ · تصحيح الفوضى في تطبيقات الذكاء الاصطناعي"؛ *تكنولوجيا الأمن السيبراني - المتطلبات الأمنية الأساسية لخدمات الذكاء الاصطناعي التوليدي*، المعيار الوطني الموصى به، الذي يدخل حيز التنفيذ في 1 نوفمبر 2025، المادة 1.2.4 (ترشيح محتوى بيانات التدريب).

[145] فيما يتعلق بحكم الاستئناف في قضية LAION II وتعدد معايير الاعتراض على الاستخدام، انظر(ي) القسم باء-5 من الجزء الثاني، والإشارات الواردة فيه إلى تعليق كيلر.

[146] فيما يتعلق بمجموعة العمل المعنية بتفضيلات الذكاء الاصطناعي التابعة لفرقة العمل المعنية بهندسة الإنترنت، انظر(ي) https://datatracker.ietf.org/wg/aipref/about/. ومن المبادرات الرئيسية الأخرى العاملة في المجال نفسه: بروتوكول الاحتفاظ بالحقوق في تنقيب النصوص والبيانات (TDMRep)، وبروتوكول TDM·AI، وبروتوكول الترخيص الجماعي Really Simple Licensing، وآلية contentsignals.org التي وضعتها Cloudflare، وC2PA، وIPTC PLUS.

من توجيه حق المؤلف والحقوق المجاورة في السوق الرقمية الموحدة؛ وآلية إشارات المحتوى من Cloudflare، وهي امتداد خاص لبروتوكول استبعاد الروبوتات (robots.txt)؛ ومبادرات أوسع نطاقاً مثل C2PA و IPTC PLUS، التي تتيح وظائف الاعتراض على الاستخدام من خلال مفرداتها الخاصة. تُستخدم إشارات من هذا النوع بالفعل في الممارسة العملية لمنع برامج الزحف الخاصة بمحركات البحث الحوارية التي تعمل بالذكاء الاصطناعي: فقد أفادت التقارير أن نسبة كبيرة من المواقع الإخبارية الرائدة تحجب برنامج الزحف الذي تستخدمه إحدى هذه المحركات.[147] أما المجموعة الثانية فتذهب إلى أبعد من ذلك، حيث تتيح لصاحب الحقوق تضمين شروط ترخيص صريحة وقابلة للقراءة آلياً في البنية التحتية نفسها للويب. وبذلك، لا يقتصر السؤال على ما إذا كان يجوز استخدام مصنف، بل يمتد إلى الشروط التي يجوز استخدامه بموجبها، بما في ذلك المقابل المالي والاختيار بين الدفع مقابل كل عملية زحف والدفع مقابل كل عملية استدلال. ويُعد معيار Really Simple Licensing، وهو مبادرة تدعمها دور النشر الإلكترونية وشبكات توزيع المحتوى في الولايات المتحدة، مثالاً بارزاً على ذلك؛ ومن الأمثلة الأخرى نموذج الدفع مقابل كل عملية زحف الذي تعتمده Cloudflare، ويعمل على مستوى طلب النفاذ نفسه، من خلال إرجاع استجابة تفيد بوجوب الدفع إلى برنامج زحف غير مرخّص له. وفي هذه المجموعة الثانية، لا تؤدي الإشارة التي يصادفها الوكيل الآلي وظيفة الاعتراض على الاستخدام فحسب، بل تُعدّ أيضاً عرضاً لمنح ترخيص. وقد يتيح إدخال وكلاء الذكاء الاصطناعي إمكانات أوسع للترخيص الآلي في هذا المجال.

ويزيد من تعقيد أمر الامتثال والفصل في النزاعات اجتماع هذين النوعين من الاحتفاظات ضمن مناقشة موحدة لم تُحسم بعد بشأن المعايير. قد يعمل المزوّد نفسه في كلا النوعين، كما تفعل شركة Cloudflare من خلال آليتي إشارات المحتوى والدفع مقابل الزحف، بحيث لا يتضح دائماً من ظاهر الإشارة إذا ما كانت لا تعدو أن تكون احتفاظاً بالحقوق أم تمثّل عرضاً لمنح ترخيص. وقد أطلقت المفوضية الأوروبية عملية تشاور بشأن بروتوكولات الاحتفاظ بالحقوق في عمليات تنقيب النصوص والبيانات بموجب قانون الذكاء الاصطناعي ومدونة الممارسات الخاصة بالذكاء الاصطناعي للأغراض العامة، بهدف إعداد قائمة بالبروتوكولات المتفق عليها عموماً والقابلة للقراءة آلياً، والتي يتعين على مزودي نماذج الذكاء الاصطناعي للأغراض العامة الالتزام بها للوفاء بالتزاماتهم بموجب قانون الذكاء الاصطناعي.[148]

ويُمثّل تجزؤ بيئة المعايير تحدياً في حد ذاته أمام الإنفاذ. يشير حكم محكمة الاستئناف في هامبورغ إلى أن قابلية قراءة الاعتراض على الاستخدام آلياً ينبغي تقييمها في ضوء القدرات التكنولوجية المتاحة وقت الاستخدام محل النزاع؛ غير أنه في ظل وجود معايير متعددة غير متوافقة، لا تكون الإجابة عن السؤال العملي المتعلق بالمعايير التي يتعين على المطوّر مراعاتها للامتثال للقانون واضحة أو محسومة. ويتمثل أحد المخاطر التي أشار إليها التعليق على الحكم في أن الاعتراضات على الاستخدام القابلة للقراءة آلياً قد تتحول، في غياب معايير مشتركة تخضع لحوكمة مفتوحة، إلى أساس لطبقة جديدة من الوسطاء تُرسّخ نموذج النفاذ إلى الويب المفتوح مقابل الدفع عن كل عملية زحف أو كل نقرة. ويتوقف تحقق هذا الخطر من عدمه على الخيارات التي ستُتخذ خلال السنوات المقبلة في هيئات وضع المعايير وفي مجال صنع السياسات.

وتشكل آلية الاعتراض على الاستخدام جزءًا من بنية تحتية متعددة المستويات لإدارة الحقوق. تتمثل الطبقة الأولى في المعرّفات والبيانات الوصفية. وهي تحتوي على معرّفات معيارية دولية معتمدة للمصنَّفات والمساهمين في إنشائها، فضلاً عن البيانات الوصفية القابلة للقراءة آلياً، والمتعلقة بإدارة الحقوق ومصدر المحتوى، اولتي يمكنها ربط المصنَّف بمعلومات عن وضع حقوقه. ويجري العمل على تجميع الممارسات الجيدة المتعلقة ببيانات حق المؤلف تتسم بقابلية التشغيل البيني والموثوقية وإمكانية قراءتها آلياً بما يلائم بيئة الذكاء الاصطناعي، ومن أبرز هذه الجهود عمل فرقة العمل المعنية بالبنية التحتية لحق المؤلف التي شكّلتها وزارة التعليم والثقافة الفنلندية[149]، انطلاقاً من مبدأ أن التحديد الموثوق لهوية المصنَّفات شرط مسبق، سواء للتعبير عن تفضيلات استخدامها أو لتتبّع أوجه استخدامها بعد وقوعها. وتستند هذه الطبقة إلى أحكام معاهدة الويبو بشأن حق المؤلف ومعاهدة الويبو بشأن الأداء والتسجيل الصوتي ومعاهدة بيجين، المتعلقة بمعلومات إدارة الحقوق المذكورة أعلاه.

والطبقة الثانية هي طبقة التحكم في عمليات النفاذ. وهي تشمل واجهات برمجة التطبيقات وتدابير مكافحة الزحف، والتي يسعى من خلالها أصحاب الحقوق والمنصات إلى ضمان أن يقتصر النفاذ إلى محتواهم على الوكلاء المصرّح لهم فقط، إلى جانب تدابير الحماية التكنولوجية المستخدمة في قطاعي البث والنشر لمنع الاستخراج غير المصرّح به. وتستند هذه الطبقة إلى أحكام معاهدة

---

[147] أظهرت دراسة نُشرت في عام 2026 أن العديد من المواقع الإخبارية الرائدة في الولايات المتحدة والمملكة المتحدة تحظر برنامج الزحف من شركة Perplexity، وPerplexityBot، في ملفات «robots.txt» الخاصة بها. انظر(ي) مقالة " Most Major News Publishers Block AI Training & Retrieval Bots" من) BuzzStream17 ديسمبر 2025، نُشرت في يناير 2026) <https://www.buzzstream.com/blog/publishers-block-ai-study/> تم النفاذ بتاريخ 3 يوليو 2026.

[148] المفوضية الأوروبية، مشاورة حول بروتوكولات الاحتفاظ بالحقوق في سياق تنقيب البيانات والنصوص بموجب قانون الذكاء الاصطناعي ومدونة الممارسات الخاصة بالذكاء الاصطناعي للأغراض العامة (2025)؛ انظر(ي) https://digital-strategy.ec.europa.eu.

[149] نشرت فرقة العمل المعنية بالبنية التحتية لحق المؤلف (CITF)، التي شُكّلت تحت رعاية وزارة التعليم والثقافة الفنلندية، تقريرها الأول بعنوان "Interoperable, trustworthy and machine-readable copyright data in the AI era"، في عام 2025.

الويبو بشأن حق المؤلف ومعاهدة الويبو بشأن الأداء والتسجيل الصوتي ومعاهدة بيجين، المتعلقة باستراتيجيات مكافحة التحايل المذكورة أعلاه.

أما الطبقة الثالثة فهي طبقة الإسناد، ويُقصد بها تقييم مساهمة مصنفات بعينها في مخرج معين بهدف توزيع المكافآت حيثما ينطبق ذلك. وينبغي التمييز بين الإسناد بهذا المعنى والكشف. فتكنولوجيات الكشف تسعى إلى تحديد إذا ما كان ملف ما قد أُنتج بواسطة إنسان أم آلة، بدرجة معينة من الاحتمال؛ لكنها لا تحدد المصنَّفات المصدرية، ولا تكفي في حد ذاتها لإسناد مطالبة بموجب حق المؤلف. أما تكنولوجيات الإسناد، فتسعى إلى تقدير حصة المصنَّفات المحدّدة في مخرج مولَّد، ولذلك تعتمد على قدر من التعاون من جانب المطوّر وعلى شفافيته. وقد بدأت الاستثمارات في هذه القدرات بالفعل؛ إذ استحوذت شركة كبرى لتسجيلات الموسيقى في عام 2026 على أحد مزوّدي تكنولوجيات الإسناد، الذي كان مشاركاً بالفعل في ترتيب الترخيص الجماعي السويدي المشار إليه في القسم الثاني من الجزء الثاني.[150] ترتبط طبقة الإسناد ارتباطاً مباشراً بنماذج المكافآت على مستوى المخرجات المذكورة في الجزء الثاني، القسم باء-6. ولا تستند بعض هذه النماذج إلى وقوع انتهاك عند إنتاج المخرج، بل إلى الأثر الإحلالي لمخرجات نماذج معينة.

وثمة مجموعة مستقلة من أدوات تحديد المصدر ووضع العلامات المائية، إلى جانب مبادرات لوسم المحتوى المؤلف بشرياً، تتناول مسألة أصالة المخرجات ومصدرها بدلاً من استخدام المصنّفات المحمية في التدريب؛ وتتصل هذه الأدوات أساساً بالمسائل المتعلقة بالمُخرَجات التي لا تتناولها هذه الدراسة، ولا يُشار إليها هنا إلا لوضع طبقة الإسناد في سياق البنية التحتية الأوسع نطاقاً.

3. السوابق القضائية في مجملها

تتيح مجموعة السوابق القضائية التي استعرضتها هذه الدراسة استخلاص عدة ملاحظات عامة مشتركة بشأن الإنفاذ الحالي لحق المؤلف في سياق تدريب نماذج الذكاء الاصطناعي التوليدي.

وتتمثل الملاحظة الأولى في أن المحاكم توخت، بوجه عام، قدراً من الحذر، وامتنعت عموماً عن إصدار أحكام يُفترض أن تحسم بصورة شاملة الإطار القانوني الناظم لهذه المسائل. فقد تناولت الأحكام الصادرة حتى الآن تكنولوجيات معينة (الانتشار الكامن في قضية Getty، ونماذج اللغة في قضية جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي، والكتب التي جرى الحصول عليها بطرق مشروعة وغير مشروعة في قضية Bartz)؛ وأوضاعاً إجرائية معينة (التماس إسقاط الدعوى في تقاضي صحيفة نيويورك تايمز، والحكم المستعجل في قضيتي Bartz وKadrey)؛ ومسائل معينة (النموذج باعتباره نسخة في قضيتي جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي وGetty، والاستخدام المنصف في قضيتي Bartz وKadrey، من بين قضايا أخرى). ولم يتمخّض بعد عن هذه الأحكام إرساء إطار شامل في أي نظام قانوني بمفرده. والموقف الألماني، وإن كان يتجلى حالياً في ثلاثة قرارات، لا يتّسم بالاتساق الداخلي في حد ذاته: إذ يقبل التوجه القضائي في قضية LAION بتطبيق استثناء تنقيب النصوص والبيانات على التدريب، في حين يتبنى توجه جمعية الأداء الموسيقي وحقوق الاستنساخ الميكانيكي نهجاً مختلفاً. أما الموقف في الولايات المتحدة فهو متباين عل نحو مماثل، بانتظار المراجعة أمام محاكم الاستئناف، وينتظر في نهاية المطاف صدور رأي قضائي واحد أو أكثر عن المحكمة العليا للولايات المتحدة.

والملاحظة الثانية هي أن مجموعة السوابق القضائية قد أسفرت مع ذلك عن بعض نقاط التوافق. إذ يوجد توافق بشأن الشرط المنصوص عليه في اتفاقية برن والمتعلق بضرورة وجود مؤلف بشري لتوليد الأصالة اللازمة لحماية مخرجات الذكاء الاصطناعي. ويمكن ملاحظة ذلك في الولايات المتحدة وكينيا[151] وألمانيا على مستوى المحاكم الابتدائية، وكذلك في الاتحاد الأوروبي حيث يفترض معيار الأصالة، الذي أكدت عليه محكمة العدل الأوروبية في قضية Mio، وهي قضية تتعلق بحماية حق المؤلف للأثاث، وجود مؤلف بشري[152]. كما يوجد توافق في الآراء بشأن تطبيق حق الاستنساخ من خلال إنشاء مجموعات بيانات تدريبية (في

---

[150] استحوذت Warner Music Group في يونيو 2026، على شركة Sureel، المزوّدة لتكنولوجيا الإسناد. وتُستخدم تكنولوجيا الشركة المزوّدة نفسها في ترتيب الترخيص المتعلق بالجمعية السويدية لحقوق الأداء (STIM) الذي جرى مناقشته في القسم باء-3 الوارد في الجزء الثاني.

[151] قضية Aryeh Movement Limited ضد Cynthia Beldina Akoth Okello، رقم COPTA/E001/2025 (محكمة حق المؤلف في كينيا) (خلصت المحكمة إلى أن المادة 22(3) من قانون حق المؤلف (الفصل 130) تفترض تدخلاً بشرياً في بذل "جهد كافٍ"، ومن ثم لا يُمكن للمصنّف المولَّد ذاتياً بواسطة نظام ذكاء اصطناعي استيفاء الحد الأدنى القانوني ما لم يُسبغ عليه التدخل البشري الكافي طابع الأصالة).

[152] *قضية Mio وآخرون*، رقم C-580/23، المنضمّة إلى *قضية USM* ضد *konektra*،رقم C-795/23 (محكمة العدل التابعة للاتحاد الأوروبي، الحكم الصادر بتاريخ 4 ديسمبر 2025). تعلقت القضيتان بالحماية بموجب حق المؤلف للمصنَّفات الخاصة بالفنون التطبيقية، وتحديداً قطع الأثاث (ومنها طاولات الطعام "Palais Royal" في القضيةرقم C-580/23). وأكدت المحكمة مجدداً أن موضوع الحماية لا يُعدّ مصنفاً محمياً إلا إذا كان أصيلاً، بمعنى أنه يشكّل إبداعاً فكرياً خاصاً بمؤلفه، يعكس خيارات حرة وإبداعية تعبّر عن شخصيته، وأن المصنَّفات الخاصة بالفنون التطبيقية تخضع لمعيار الأصالةنفسه الذي تخضع له سائر المصنَّفات. ولم يتطرق الحكم إلى موضوع الذكاء الاصطناعي؛ وإنما يُستمد اشتراط وجود مؤلف بشري لحماية مخرجات الذكاء الاصطناعي من معيار الأصالة هذا، لا من حكم أصدرته المحكمة بشأن الذكاء

ألمانيا والصين وكندا). وأخيرًا، هناك بعض التوافق بشأن أهمية استبدال المخرجات في التحليل (في الولايات المتحدة، ومن المتوقع أن يكون الأمر كذلك في الاتحاد الأوروبي من خلال تطبيق اختبار الخطوات الثلاث). أما المسائل التي لا تزال السوابق القضائية منقسمة بشأنها، فتشمل ما إذا كانت النماذج المدرّبة تحتوي على نسخ من المصنّفات المستخدمة في التدريب وإعادة التدريب والضبط الدقيق، وما إذا كان استثناء تنقيب النصوص والبيانات يمتد ليشمل تدريب الذكاء الاصطناعي وكيف يتم ذلك، ومدى امتداد الاختصاص القضائي عبر الحدود.

والملاحظة الثالثة هي أن التحديات العملية المحدَّدة في القسمين جيم-1 وجيم-2 أعلاه، وهي الشفافية، وتعدد معايير الاعتراض على الاستخدام، والتفاوت بين المطوّر وصاحب الحقوق من حيث التكاليف والمعلومات المتاحة، من المرجح أن تؤثر في تطور السوابق القضائية في الأشهر والسنوات المقبلة. وحتى في الحالات التي تبت فيها المحاكم في المسائل الموضوعية المتعلقة بقانون حق المؤلف، فإن إمكانية إنفاذ أصحاب الحقوق لمطالباتهم عملياً ستظل رهينة بالآليات التي تتيح لهم التحقق من استخدام مصنَّفاتهم، وإبداء الاعتراض على استخدامها، وتقديم الأدلة في صيغة تمكّن المحاكم من الاستناد إليها واتخاذ ما يلزم بناءً عليها. وفي هذا المجال، يرتبط التصميم التقني ارتباطاً وثيقاً بالتحليل القانوني، وستتأثر ممارسات الإنفاذ بالخيارات المتخذة في كلا المجالين.

الختام

وترصد هذه الدراسة التطور السريع والمتفاوت في مجالين: من ناحية، الولايات القضائية التي تُجري مزيجاً مختلفاً من الاستثناءات وآليات الترخيص وخطط المكافآت؛ ومن ناحية أخرى، مجموعة من السوابق القضائية التي سلّطت الضوء على هذه القضايا دون أن تفضي إلى إرساء إطار مستقر لها حتى الآن. وإن الدول الأعضاء التي تدرس حالياً اتخاذ تدابير تشريعية أو تنظيمية تُقدِم على ذلك في ظل مشهد أكثر تطوراً مما كان عليه حتى قبل اثني عشر شهراً؛ فالمسائل التي أجمعت عليها الخبرات والسوابق القضائية تشكل أساساً جزئياً يمكن الاستناد إليه، في حين تتطلب المسائل التي لا تزال محل خلاف اتّخاذ خيارات سياسات سعت هذه الدراسة إلى عرضها، لا إلى حسمها. وسيستمر اتخاذ هذه الخيارات على المستوى الوطني، وسيجري مناقشتها على المستوى الدولي، إذا ما قررت الدول الأعضاء ذلك، من خلال اللجنة الدائمة المعنية بحق المؤلف والحقوق المجاورة، والإطار الأوسع للويبو.

وتُظهر الدراسة أن التوازن بين مصالح أصحاب الحقوق ومصالح مطوري الذكاء الاصطناعي قد يتم تحقيقه بطرق مختلفة في كل دولة عضو. ويتضح من التحليل خمسة أسئلة أكثر دقة.

الأول هو ما إذا كانت عمليات النسخ التي تتم في سياق تجميع مجموعة نصوص وتدريب نموذج تتطلب الحصول على إذن، أم أنها تندرج ضمن استثناء أو قيد، سواء كان ذلك خاصًا بالذكاء الاصطناعي، أو ضمن أحكام استخراج النصوص والبيانات، أو ضمن بند عام، وتحت أي شروط، بما في ذلك الوصول القانوني، والغرض من التدريب، والطابع التجاري أو غير التجاري للاستخدام.

والثاني هو تصميم أي آلية للتحفظ حيثما يتم استخدامها. ويشمل ذلك شكل آلية الاعتراض على الاستخدام، مع الأخذ في الاعتبار أن التحفظات القابلة للقراءة آليًّا تفترض وجود قدرات قد لا يمتلكها الأفراد من المؤلفين والمؤدين. وعلى العكس من ذلك، فإن نظام التحفظ الذي لا يمكن تنفيذه آليًّا على نطاق واسع لا يوفر للمطورين مسارًا واضحًا للامتثال. كما ينبغي النظر في مدى توافق آلية الاعتراض على الاستخدام مع المادة 5(2) من اتفاقية برن.

والسؤال الثالث هو ما إذا كان ينبغي أن يرتبط التعويض بالاستهلاك، أم باستبدال الإنتاج، أم بكليهما؛ وما إذا كان ينبغي أن يتخذ شكل ترخيص سوقي قائم على حقوق حصرية أم مطالبة قانونية بالحصول على أجر؛ وفي كلتا الحالتين، من خلال أي مؤسسات ينبغي أن يتم دفع هذا التعويض. ويشمل ذلك ما إذا كانت صلاحيات الإدارة الجماعية تشمل الاستخدامات التدريبية، وما إذا كانت هناك حاجة إلى أشكال موسعة أو جديدة من الترخيص الجماعي، وكيف يمكن تصميم سلاسل العقود بحيث تصل القيمة إلى المؤلفين والفنانين الذين تُستخدم أعمالهم.

والسؤال الرابع يتعلق بالشفافية: ما هي إجراءات حفظ السجلات والإفصاح عن المواد التدريبية اللازمة لتمكين ممارسة الحقوق وإدارة المكافآت، وكيف يمكن التوفيق بين هذه الالتزامات وبين الأسرار التجارية المشروعة والمخاوف المتعلقة بالمنافسة.

والسؤال الخامس والأخير هو: هل تظل سبل الانتصاف فعالة ومتناسبة في الحالات التي يكون فيها إزالة الأعمال من نموذج مدرب أمراً غير ممكن تقنياً، وكيف ينبغي ممارسة الاختصاص القضائي عندما يتم التدريب في الخارج بينما تُقدَّم الخدمات الناتجة عنه محليًّا.

الاصطناعي تحديداً. وقد يواجه المؤلفون، فضلاً عن ذلك، عقبة جديدة في إجراءات الإنفاذ، تتمثل في تقديم دليل على أن مصنَّفهم لم يُنشأ بواسطة الذكاء الاصطناعي، أو على الأقل على أنهم أسهموا فيه إسهاماً كافياً لتبرير حمايته بحق المؤلف.

[نهاية الوثيقة]