Abwehr von Scraping und KI-Training

Stand: 30.03.2025
Autor: Marco Urban

Im Oktober 2023 haben wir einen Leitfaden „Über den Umgang mit KI-Bildgeneratoren“ veröffentlicht. Es ging dabei eigentlich weniger um den Umgang mit KI-Generatoren, sondern vielmehr darum, wie man dem Scraping, also dem Auslesen von Daten für das Erstellen der Kataloge, die Grundlage für das Training von generativer KI sind, am besten begegnen kann. Nach wie vor werden wir als Urheber*innen weder gefragt, ob wir das möchten, noch werden wir für diese Nutzung unserer Werke honoriert. Nach wie vor kommt es also zum massenhaften und organisierten Verstoß gegen das Urheberrecht durch riesige IT-Konzerne. Unsere Bilder werden dabei dazu benutzt, um Maschinen zu trainieren, die uns am Ende überflüssig machen sollen.

Hier ist ein fälliges Update der Handlungsoptionen für Fotograf*innen. Das Zitat von Ranga Yogeshwar bleibt aktuell.

Der Wissenschaftsjournalist und Autor wird in der Augsburger Allgemeinen vom 17. Mai 2023 wie folgt zitiert:

»[...] wir erleben im Moment den größten Diebstahl in der Menschheitsgeschichte. Die reichsten Unternehmen der Welt wie Microsoft, Apple, Google, Meta oder Amazon bemächtigen sich der Summe des menschlichen Wissens. Also aller Texte, Kunstwerke, Fotografien und so weiter, die in digital verwertbarer Form existieren, um dieses Weltwissen dann in eigentumsrechtlich geschützten Produkten einzumauern. Es gibt dabei keine klare Offenlegung, mit welchen Lerndaten sie die KI trainieren. [...] Das Urheberrecht wird missachtet — und zwar bewusst. Inzwischen kann per KI eine Massenproduktion von Plagiaten stattfinden, wobei ganze Berufsstände vor ihrem existenziellen Aus stehen.«

Was können und müssen wir angesichts dieser rasanten Veränderung eigentlich tun? Was sind die Handlungsoptionen für Fotograf*innen, sich gegen unerlaubtes Training zu wehren?
Wohlgemerkt muss uns das nicht hindern, selbst generative KI und natürlich KI-Werkzeuge zu nutzen, sei es aus Neugier, zur Arbeitserleichterung oder weil wir uns neue Geschäftsfelder erschließen möchten. Aber im Gegensatz zu den KI-Trainern zahlen wir Lizenzgebühren für die Software, die wir nutzen.

Ich habe die relevanten Punkte zusammengefasst, wobei sich sowohl die Fragestellungen als auch die Antworten sehr schnell ändern können:

1. Werden meine Bilder bereits für KI-Training genutzt?
2. Wie können wir einer Verwendung für das KI-Training widersprechen?
3. Wie können wir Dritten die Nutzung von Werken als KI-Trainingsmaterial untersagen?
4. Wie können wir die Erstellung von Trainingsmodellen sabotieren?
5. Was tut FREELENS in dieser Frage?
6. Fazit

1. Have I been trained?

LAION ist ein in Deutschland ansässiger Verein, der Links und Metadaten in riesigen Katalogen sammelt und als Open Source jedermann zur Verfügung stellt. Der Katalog LAION 5B enthält weit über 5 Billionen Werke. Es ist davon auszugehen, dass nahezu alle KI-Generatoren wie z.B. **Stable Diffusion**, DALL-E3 von OPEN AI oder **Imagen** mit den Bildern und Metadaten aus diesem Katalog trainiert wurden.

Die Website **https://haveibeentrained.com** wird von **https://spawning.ai** betrieben und bietet die Möglichkeit, die Kataloge **Laion-5B** und **Laion-400M** zu durchsuchen.
Es gibt hier auch die Möglichkeit, Bilder für das Training zu sperren.
Allerdings ändert das nichts am Inhalt der mit LAION erstellten Modelle. Ist das KI-Modell erst einmal erstellt, können einzelne Trainingsbilder nicht mehr ausgenommen werden. Bilder in möglicherweise größerer Zahl bei »**Have I been Trained**« einzeln zu sperren, hat also nur für Modelle einen Effekt, die zukünftig erstellt werden.

Darüber hinaus gibt es natürlich viele Crawler, die Bilder und Metadaten aus dem Netz sammeln und deren Datenbanken nicht Open Source sind. Hier ist es zurzeit nicht möglich, zu überprüfen, ob eigene Werke in den Katalogen enthalten sind. Der AI Act der EU sieht zwar künftig eine gewisse Verpflichtung zur Transparenz bezüglich der Trainingsinhalte vor, die KI-Industrie übt jedoch massiven Einfluss aus, die Anforderungen dafür zu verwässern.

2. Der Verwendung für das KI-Training widersprechen

2.1 Gesetzliche Grundlage

Dürfen urheberrechtlich geschützte Werke denn überhaupt zum Training einer KI benutzt werden? Als rechtliche Grundlage könnte §44b des Urheberrechtsgesetzes dienen, mit dem 2021 die DSM-Richtlinie der EU umgesetzt wurde. §44b erlaubt “die Vervielfältigung von rechtmäßig zugänglichen Werken” für sogenanntes “Text and Data Mining”, also für die “automatisierte Analyse … um … Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen”. Dies jedoch nur dann, wenn der Rechtsinhaber/die Rechtsinhaberin dem nicht widersprochen hat.
Die vielbeachtete aktuelle Studie “Urheberrecht und Training generativer KI-Modelle” von Dornis und Stober zieht zwar in Zweifel, dass es sich beim KI-Training überhaupt um Text and Data Mining handelt. Doch solange sich diese Erkenntnis nicht auch vor Gericht durchgesetzt hat, muss die im Moment noch herrschende Auslegung des §44b zur Kenntnis genommen werden. Wer verhindern will, dass seine oder ihre Fotos für KI-Training verwendet werden können, muss also selbst aktiv werden und — wie in diesem Paragrafen gefordert — einen Nutzungsvorbehalt (ein Opt-Out) in maschinenlesbarer Form aussprechen.
Was genau mit “maschinenlesbar” gemeint ist, ist weiterhin unklar, noch gibt es keinen verbindlichen technischen Standard dafür. Das Landgericht Hamburg jedenfalls fand in seinem jüngsten Urteil im Fall Kneschke gegen LAION, dass auch ein Opt-Out in natürlicher Sprache ausreichend sein kann.

Wie kann nun ein Opt-Out erklärt werden?

2.2 Kennzeichnen von Fotografien/Bildern, um Scraping/Einlesen von Werken für KI-Trainings-Datenbanken zu verbieten (Asset-basiert)

2.2.1 IPTC und PLUS

IPTC- und PLUS-Standard wurden vor kurzem um ein Feld mit der Bezeichnung “Data Mining” erweitert. Damit kann gekennzeichnet werden, ob und in welchem Umfang die Verwendung eines Fotos als KI-Trainingsmaterial erlaubt ist.
Dies ist eine gute Maßnahme, um der gesetzlichen Forderung nach Widerspruch in maschinenlesbarer Form für jede einzelne Datei nachzukommen. Leider kann man aktuell mit keiner gängigen Software dieses Feld ausfüllen.
Es gibt aber Abhilfe: Peter Hytrek, Mitglied in der AG Fortschritt des Fotorates und Inhaber der DINAX GmbH, hat den DINAX IPTC-Tagger entwickeln lassen, mit dem man dieses und weitere für das Urheberrecht relevante Felder per Stapelverarbeitung sehr leicht ausfüllen. Aktuell werden die Auswahlfelder noch mit den entsprechenden Begriffen ausgefüllt und nicht wie bei IPTC mit Links. Das dürfte aber rechtlich trotzdem wirksam sein. Ob die Crawler auf solche Hinweise reagieren, ist allerdings noch völlig offen.
Den Tagger gibt es hier zum kostenlosen Download: https://dinax.com/iptc/
Mehr Informationen zu IPTC und der korrekten Beschriftung von Bildern findet Ihr auf der FREELENS Website: https://freelens.com/services/wissen-a-z/iptc/thema:routine-und-chancen

2.2.2 Verifiable Credentials — Liccium und der ISCC-Standard (Soft-Binding)

Die Firma Liccium verspricht einen weitergehenden Schutz der Metadaten und damit auch ein maschinenlesbares Opt-Out gegen KI-Training. Auch C2PA- und IPTC-Daten werden unterstützt.
Dabei verwendet Liccium das TDM-AI Protokoll und den internationalen ESCC-ISO Standard.
Mit der Liccium App erzeugen Nutzer*nnen einzigartige ISCC-Codes aus den Bilddateien, bestehend aus Zahlen und Buchstaben (Fingerprint oder Hash-Wert). Diese Codes bilden zusammen mit den Metadaten und Lizenzinformationen die Verifiable Credentials. Sie werden in öffentlich einsehbaren Datenbanken gespeichert. Da die Bilder selbst nicht gespeichert werden (müssen) ist der benötigte Speicherplatz überschaubar.
Mit einem aufgefundenen (oder auch gescrapten) Bild kann nun umgekehrt der ISCC-Code, die Metadaten und die Urheberinformationen gefunden werden.
Es entsteht also eine Art Werkverzeichnis mit dem Dritten — Nutzer, Medien, Plattformen oder eben auch Datensammler — die Authentizität und Integrität digitaler Medieninhalte überprüfen können und auch Rechte, Lizenzen und andere Metadaten einsehen können.
Da die Metadaten an einem zentralen Ort gespeichert werden und nicht (nur) am Werk selbst, wie es bei IPTC der Fall ist, können Sie auch nicht entfernt oder gelöscht werden.
Auch bei leicht veränderten Bildern soll die Zuordnung möglich sein.
Liccium führt diese „Verifiable Credentials“ derzeit in der gesamten Kultur- und Kreativwirtschaft ein. Berufsverbände, Verwertungsgesellschaften oder anderen Vertrauensdienste können verifizierbare Berechtigungsnachweise zur Nutzung dieser Technologie bei Liccium beantragen, um die digitalen Identitäten Ihrer Mitglieder zu verwalten.
https://docs.creatorcredentials.com

Der ISCC-ISO Standard ist Open Source, er wird auch von anderen Anbietern genutzt werden.

2.2.3 Content Credentials — Content Authenticity Initiative (Hard Binding)

Die Content Credentials der **Content Authenticity Initiative** CAI nach dem CP2A-Standard haben einen ähnlichen Ansatz, allerdings geht es hier primär darum, die Authentizität der Fotografien nachzuweisen. Die dafür nötigen Daten sind nur an der Datei gespeichert, sog. Soft-Binding. Aktuell gibt es hier keine Möglichkeit ein Opt-Out zu erklären, technisch wäre das natürlich problemlos möglich.
Mehr zu den Content Credentials und zum Nachweis der Authentizität findet Ihr in der Information zu IPTC-Daten auf der FREELENS Website: https://freelens.com/services/wissen-a-z/iptc/thema:routine-und-chancen

2.2.4 Adobe Content Authenticity - Adobe

Adobe, der Hersteller von Photoshop und Lightroom, verbindet Soft- und Hard Binding. Bei dem neuen Produkt Adobe Content Authenticity werden Urheberdaten, Social Media Profile und auch ein Nutzungsvorbehalt in die Bilddatei eingebettet aber auch in der öffentlichen Adobe Content Credentials-Cloud gespeichert werden. Dafür werden Fotograf*innen Ihre Bilder auf einen Server, Adobes Creative Cloud, hochladen müssen, der die Informationen in die Bilder einbettet und auch in die Cloud-Datenbank übernimmt. Dann kann man sie — mit den eingebetteten Daten - wieder herunterladen. Aktuell kann man das mit bis zu 50 Dateien gleichzeitig machen, unterstützt werden JPG- und PNG-Format.
Adobe Content Authenticity ist, Stand Januar 2025, eine Betaversion. Man kann sich, wenn man es nutzen möchte, dafür auf einer Warteliste eintragen. Aktuell ist das noch kostenlos, aber sicher wird dafür nach erfolgreicher Testphase eine Gebühr erhoben werden.
Vermutlich wird der Vorgang zumindest mit Cloud-Versionen der Adobe Software wie Lightroom einfacher gestaltet sein, denn dann sind die Bilder ja bereits hochgeladen.
Die Glaubwürdigkeit wird dadurch hergestellt, dass es durch die verwendete Hardware und das Adobe Konto immer möglich ist, den Urheber (bzw. seine Hardware) zu identifizieren. Vermutlich binet man sich damit aber in irgendeiner Form an Adobe.

Erste Informationen findet Ihr hier: https://contentauthenticity.adobe.com
Und hier: https://helpx.adobe.com/de/creative-cloud/help/cai/adobe-content-authenticity.html

2.3 Kennzeichnen von Webseiten, um Scraping/Einlesen von Werken für KI-Trainings-Datenbanken zu verbieten. (Location- oder Domain-basiert)

Die oben beschriebenen Nutzungsvorbehalte am Werk selbst, also an euren Fotos, werden zwar langfristig wirkungsvoller sein. Um gegen die Verwendung eigener Fotos für KI-Training gerichtlich vorgehen zu können, ist es aber wichtig, möglichst frühzeitig auch einen entsprechenden Widerspruch auf der eigenen Website zu platzieren.

2.3.1 Impressum — Der juristische Hinweis

Als erstes Gericht in Europa hat das Landgericht Hamburg im Prozess von Robert Kneschke gegen den Datensammler LAION e.V. die Meinung vertreten, dass auch ein Text in natürlicher Sprache auf einer Website als maschinenlesbarer Vorbehalt zu werten ist. Noch steht eine höchstrichterliche Entscheidung dazu aus. Aber bis dahin kann es jedenfalls nicht schaden, z.B. im Impressum der eigenen Website einen Vorbehalt gemäß §44b (3) UrhG auszusprechen, um die Nutzung der eigenen Fotos für KI-Training zu untersagen.
Mehr Informationen zum Impressum findet Ihr auf der FREELENS Website:
https://freelens.com/services/wissen-a-z/impressum

2.3.2 robots.txt-Datei, ai.txt-Datei, TDMRep — Die Verbotsschilder

2.3.2.1 Robots.txt

Für die Crawler lesbar ist ein solcher Hinweis dann, wenn er in der robots.txt-Datei steht. Robots.txt ist eine Textdatei, die in den Backend-Code einer Website eingefügt werden kann, um Web-Crawlern mitzuteilen, was sie durchsuchen dürfen und was nicht. Wer seine Website nicht von Google durchsuchen lassen möchte, schreibt den entsprechenden Befehl in die robots.txt-Datei der Website.

Diese Form des Nutzungsvorbehalts wird — im Gegensatz zu einem Hinweis im Impressum — von den meisten Datensammlern akzeptiert. Auch der deutsche IT-Verband Bitkom spricht sich dafür aus. Allerdings müssen hier alle Crawler namentlich aufgeführt werden, was in der Praxis schwierig ist, weil viele von ihnen gar nicht bekannt sind oder ihre Namen auch jederzeit ändern können.
Die New York Times, CNN und Australiens ABC **blockieren auf diese Weise** den Zugriff auf ihre Inhalte für OpenAIs GPTBot und andere Webcrawler.
Eine Anleitung findet Ihr auf der FREELENS Website unter Wissen A-Z:
https://freelens.com/services/wissen-a-z/ki-webcrawler-ind-bots/thema:ki
Technisch gesehen können die Befehle in der robots.txt allerdings ignoriert werden. Sie sind lediglich ein Hinweis an die Web-Crawler/Bots. Böswillige Bots werden sie ignorieren.

2.3.2.2 ai.txt-Datei

Einen ähnlichen Ansatz hat das Projekt **spawning.ai** entwickelt: Ihr könnt eine ai.txt Datei generieren lassen, um die Verwendung von Website-Inhalten für das Training von KI-Modellen zuzulassen oder zu verbieten. Die Generierung dauert keine Minute und die Datei muss dann nur noch ins Hauptverzeichnis der Website geladen werden.
Die Idee ist sehr gut, aber es ist derzeit nicht sicher, welche Crawler diese Datei tatsächlich auswerten. Und wie bei robots.txt: es ist nur ein Hinweis an die Bots.
Vor Gericht aber sollte diese Form des Nutzungsvorbehalts hilfreich sein.
Das Tool findet Ihr hier.

2.3.2.3 TDM Reservation Protocol (TDMRep)

Das TDMRep-Protokoll verfolgt einen ähnlichen Ansatz wie die vorgenannten, aber hier soll nicht nur ausgedrückt werden, ob Schutzrechte für bestimmte Webinhalte vorbehalten sind oder nicht, sondern auch, wie die Rechteinhaber kontaktiert werden können und welche Lizenzen gegebenenfalls verfügbar sind.
Das kann auf drei Arten passieren:
1. Analog zur robots.txt kann eine spezielle Datei unter dem Namen /.well-known/tdmrep.json auf dem Webserver angelegt werden, die alle Informationen enthält.
2. Die Informationen können in den HTTP-Header der Serverantwort eingebaut werden.
3. Es können entsprechende Meta-Tags im Kopf von HTML-Seiten verwendet werden.
Für eine schnelle und einfache Umsetzung empfiehlt sich die erste Variante. Um den Nutzungsvorbehalt auszudrücken, wird einfach über einen der drei Mechanismen der Wert von »tdm-reservation« auf 1 gesetzt. Zusätzlich kann über den Eintrag »tdm-policy« eine maschinenlesbare Lizenzierungsrichtlinie (Policy) verlinkt werden. Diese nutzt das Format der Open Digital Rights Language (ODRL) und enthält Informationen zu Rechteinhabern sowie Details über verfügbare TDM-Lizenzen. TDM-Akteuren wird es so erleichtert, mit Rechteinhabern von Inhalten in Kontakt zu treten und TDM-Lizenzen zu erwerben. Da das Format maschinenlesbar ist, könnte dies auch automatisiert erfolgen. Damit bietet sich hier eine sehr interessante Lösung, die sogar über den einfachen Nutzungsvorbehalt hinaus geht. Inwiefern diese jedoch schon von Crawlern und KI-Industrie beachtet wird, ist nicht klar. Angesichts des sehr geringen Aufwands empfiehlt sich die Umsetzung für Rechteinhaber*innen auf jeden Fall. Je weiter sich die Nutzung verbreitet, desto stärker dürfte der Druck auf die KI-industrie werden TDMRep zu beachten.
https://w3c.github.io/tdm-reservation-protocol/spec/

2.3.3 Cloudflare und Kudurru — Die Türsteher

2.3.3.1 Cloudflare

Der Serveranbieter Cloudflare bietet kostenlos an, Bots abzuweisen. Das geht also einen Schritt weiter als robots.txt. Dabei werden Anfragen an Websites erst einmal auf den Cloudflare Server umgeleitet und dort mit einer aktuellen Liste von Bots abgeglichen. Unseriöse Bots und Datensammler werden abgewiesen, erwünschte Anfragen, wie die von Suchmaschinenbots, auf die Website zurückgeleitet. Cloudflare kann als großer Serveranbieter schnell feststellen, welcher Crawler welchen Zweck verfolgen und so seine Listen immer auf aktuellem Stand halten.
https://blog.cloudflare.com/de-de/cloudflare-ai-audit-control-ai-content-crawlers/
Eine Anleitung findet Ihr auf der FREELENS-Website: https://freelens.com/services/wissen-a-z/ki-bots-blockieren

2.3.3.2 Kudurru

Eine ähnliche Methode wendet Kudurru an. Es überwacht Scraping-Verhalten populärer KI-Datensammler und koordiniert sich im Netzwerk, um Bots schnell zu identifizieren. Wird ein Bot identifiziert, wird seine Identität an alle geschützten Kudurru-Websites übermittelt. Alle Kudurru-Websites blockieren dann gemeinsam den Bot für das Herunterladen von Inhalten von ihrem jeweiligen Host. Wenn der Scraper seine Arbeit beendet hat, informiert Kudurru das Netz, und der Datenverkehr kann wie gewohnt fortgesetzt werden.
Dabei werden die Bots nicht nur zurückgewiesen, sondern man kann auch alternative Bilder anstelle der von den Bots angeforderten Bilder ausgeben lassen. Das kann dazu führen, dass KI-Trainingsmodelle falsche Assoziationen entwickeln, und in der Folge Bilder produzieren, die nicht zu den Prompts der KI-Anwender passen.
Kudurru blockiert keine Suchmaschinen-Crawler oder -Bots, wie z. B. Google Bot. Diese Bots sind von Google genau definiert (so dass man sie absichtlich nicht blockieren kann) und werden von Kudurru ignoriert. Kudurru hat keinen Einfluss auf das SEO-Ranking oder die Auffindbarkeit der Website.

Das Kudurru-Netzwerk besteht aus mehr als tausend aktive Websites, auf denen Millionen von Medien gehostet werden.
Es gibt ein erstes Plug-In für WordPress-Websites und es sollen weiter für andere Plattformen entwickelt werden. Nach der Beta-Phase soll der Quellcode als Open Source Software zur Verfügung gestellt werden.
Die Macher von Kudurru sind, wie bei ai.txt, Nightshade und Glaze (siehe 4.2) Spawing.ai, Allerdings scheint es seit Herbst 2023 keine Fortschritte zu geben. Die Software ist immer noch Beta.
https://kudurru.ai

Für alle genannten Maßnahmen gilt aber: Der Schutz kann natürlich nur für Fotografien auf der eigenen Website gelten oder wirken, nicht aber, wenn die gleichen Fotos an anderer Stelle veröffentlicht werden.

3. Kund*innen die Nutzung von Werken als KI-Trainingsmaterial untersagen

3.1 Untersagen der Verwendung von Fotografien als KI-Trainingsmaterial durch Kund*innen.

Für unsere Kund*innen ist es genau so schwierig wie für uns, sicherzustellen, dass Fotografien, die wir in ihrem Auftrag erstellt haben, nicht als KI-Trainingsmaterial verwendet werden. Selbst Fotografien von Druckerzeugnissen werden genutzt. Davon abgesehen interessiert unsere Kund*innen das Thema häufig gar nicht.

Fotograf*innen sollten jedoch durch AGB oder Verträge verbieten, dass die Kund*innen selbst die von uns erstellten Fotografien als Trainingsmaterial verwenden, zum Beispiel, um mit Hilfe von Porträtfotos KI-generierte Porträtbilder der gleichen Personen zu erstellen (oder sich diese besondere Nutzung zumindest angemessen vergüten lassen).
Solche Porträts generiert beispielsweise **Remini** oder **Generated Photos**.
Andererseits mag es aber fraglich sein, ob wir Kund*innen gänzlich verbieten sollten, Fotografien mit KI-Tools zu bearbeiten. Die App **https://piktid.com** anonymisiert beispielsweise Gesichter, indem sie diese durch KI-generierte ersetzt. Vermutlich lassen sich so Probleme mit Persönlichkeitsrechten umgehen, was bei Fotoaufträgen im öffentlichen Raum hilfreich sein mag.

3.2 Posten von Bildern auf Social Media Plattformen vermeiden

Posten Fotograf*innen eigene Fotografien auf Social-Media-Plattformen, werden umfangreiche Nutzungsrechte an die Plattformen übertragen. Die Formulierungen über den genauen Umfang in den Online-Geschäftsbedingungen sind oft unklar.
In der Regel wird den Plattformen das Recht eingeräumt, sie auf jede erdenkliche Weise zu nutzen. Dies wird sicherlich auch die Integration in KI-Daten-Trainingseinheiten beinhalten.
X (Twitter) hat seine **Nutzungsbedingungen dahingehend angepasst**, so dass Posts für KI-Training benutzt werden dürfen.
Bei Meta (Facebook und Instagram) kann man der Nutzung von Inhalten als KI-Trainingsmaterial widersprechen. Eine Anleitung findet Ihr auf der FREELENS Website:
https://freelens.com/services/wissen-a-z/ki-meta-widerspruch/thema:ki

3.3 Bildagenturen erstellen KI-Modelle

Getty Images startete Ende September 2023 **seine eigene KI-Anwendung**, mit der Kund*innen selbst KI-Bilder generieren können. Der Trainings-Datensatz besteht aus Bildern aus dem Getty-Archiv. Versprochen wird eine Kompensation der Urheber*innen, sichere Verwendung bei unbegrenzter Haftungsfreistellung.
Wer also Fotos über Getty Images distribuiert, sollte klären, ob diese auch für das KI-Training verwendet werden und wie das honoriert wird.
Dass Adobe Stock die Fotos im Archivbestand für KI-Training nutzt, ist bekannt. Mit diesem Modell wird auch die Funktion »Generative Füllung« möglich gemacht.
Die Agentur laif hat sich **klar gegen KI-generierte Bilder** ausgesprochen.
Von den großen, internationalen Bildagenturen ist jedoch ziemlich sicher zu erwarten, dass sie ihre Archive gegen entsprechendes Entgelt anderen für das KI-Training zur Verfügung stellen werden oder dies wie Getty gleich selbst machen.

3.4 Adobe widersprechen

Auch Adobe möchte sehr gerne alle Daten, die seine Kunden erstellen, für das Training seiner eigenen KI nutzen. Diese wird zum Beispiel bei den KI-Tools der Adobe-Software genutzt, aber auch bei dem Bildgenerator Adobe Firefly.
Man kann (und soll) dem widersprechen.
Eine Anleitung findet Ihr auf der FREELENS Website: https://freelens.com/services/wissen-a-z/ki-adobe-widerspruch/thema:ki

4. Die Erstellung von Trainingsmodellen sabotieren

4.1 Verwenden von Wasserzeichen

Getty Images verwendet für Preview-Ansichten grundsätzlich ein deutlich sichtbares Wasserzeichen mit dem Getty Images Logo. Offenbar wurden die Bilder von Getty Images in großer Menge eingelesen und als KI-Trainingsmaterial verwendet.
Die KI hat den offenbar sehr hohen Anteil an Getty-Fotos derart interpretiert, dass Fotos von Fussballspielen mit hoher Wahrscheinlichkeit ein Logo von Getty Images beinhalten. Getty Images konnte zum einen somit die Verwendung ihrer Fotografien leicht nachweisen, zum anderen waren die Bilder dadurch auch unbrauchbar. Das funktioniert natürlich in erster Linie bei einer sehr großen Anzahl Bilder oder wenn Prompts zu Fotos führen, die alle ein gleiches Wasserzeichen haben.
Grundsätzlich ist es aber von Vorteil, wenn schon auf dem Foto erkennbar ist, wer der Urheber ist und dass es urheberrechtlich geschützt ist. Aber: es gibt bereits KI-Anwendungen, mit denen man Wasserzeichen leicht entfernen lassen kann. Das geschieht nicht beim oder für das Training, aber Nutzer, die Bilder nicht rechtmäßig erworben haben, können diesen Schutz mittlerweile leicht entfernen.

Die folgenden Maßnahmen sind eher eine Nerd-Spielerei. Aber um zu zeigen, was alles möglich ist, führen wir sie hier auf:

4.2 Nightshade und Glaze — vergiftete Köder für die Scraper

Eine ziemlich robuste Maßnahme gegen unerlaubte Nutzung von Fotografien für das KI-Training ist es, diese zu „vergiften". Das bieten Nightshade und Glaze von Spawing.ai an, einem Label des Departments of Computer Science an der University of Chicago.
Damit können Urheber*innen für Menschen unsichtbare Änderungen an den Pixeln ihrer Werke vornehmen, bevor sie diese online stellen. Wird bildgenerative KI mit solchen Bildern trainiert, kann es passieren, dass die Ergebnisse auf chaotische und unvorhersehbare Weise unbrauchbar werden. (heise online - https://www.heise.de/news/Gift-fuer-Trainingsdaten-Neues-Tool-soll-Bilder-vor-KI-Bildgeneratoren-schuetzen-9343354.html)

Eine Warnung ist aber angebracht: letztlich ist das eine technische Spielerei und diese Veränderung der Bilder ist durchaus sichtbar, je nach Motiv sogar auffällig. Für Portfolios oder andere Präsentationen ist das also nicht geeignet.
Echte Aktivisten im Widerstand gegen unerlaubtes KI-Training könnten aber extra Webseiten einrichten, auf denen mit Nightshade oder Glaze bearbeitete Bilder als Köder ausgelegt werden.

4.2.1 Glaze

Viele KI-Modelle werden auch dazu verwendet, den Stil einzelner Künstler zu imitieren oder zu kopieren. Anwender können Kunstwerke von menschlichen Künstlern nehmen, um eine „Feinabstimmung" oder LoRA an Modellen wie Stable Diffusion vornehmen und so ein Modell erhalten, das in der Lage ist, beliebige Bilder im „Stil" des Zielkünstlers zu produzieren, wenn es mit dessen Namen als Eingabeaufforderung aufgerufen wird.
Glaze soll die Nachahmung von Stilen durch KI verhindern. Mithilfe von KI-Algorithmen werden an menschlichen Kunstwerken minimalen Änderungen berechnet. Für das menschliche Auge erscheinen sie unverändert, aber für KI-Modelle wirken sie wie ein dramatisch anderer Kunststil. So könnte beispielsweise ein mit Glaze verändertes Kohleporträt im Stil des Realismus für menschliche Augen unverändert erscheinen, während ein KI-Modell die Glaze-Version als modernen abstrakten Stil à la Jackson Pollock interpretieren. Wenn also jemand das Modell auffordert, den Zeichner zu imitieren, erhält er etwas ganz anderes als das, was er erwartet hat.

Nach Aussagen des Glaze Teams kann man die Effekte von Glaze nicht einfach durch einen Screenshot, Zuschneiden des Bildes, Kompression oder ähnliche Veränderungen austricksen. Bei der Technologie handelt es sich nicht um ein Wasserzeichen oder eine versteckte Nachricht (Steganografie), vielmehr beschreiben die Macher Glaze als eine weitere Dimension des Kunstwerks.
Leider, so Entwickler Prof. Ben Zhao, ist Glaze keine dauerhafte Lösung gegen KI-Nachahmung. Systeme wie Glaze stehen vor der Herausforderung, zukunftssicher zu sein. Es ist immer möglich, dass Techniken, die wir heute verwenden, von einem zukünftigen Algorithmus überwunden werden. Daher wird Glaze nicht als Allheilmittel verstanden, sondern ein notwendiger erster Schritt auf dem Weg zu künstlerischen Schutzinstrumenten, die KI-Nachahmungen widerstehen, bis längerfristige (rechtliche, regulatorische) Bemühungen greifen.
Weitere Informationen unter https://glaze.cs.uchicago.edu/webglaze.html

4.2.2. Nightshade

Nightshade verwendet eine ähnliche Technik: Es „vergiftet" Bilder durch KI Berechnung, so dass Modelle unvorhersehbare Verhaltensweisen erlernen, die von den erwarteten Normen abweichen, z. B. könnte eine Eingabeaufforderung, die nach einem Bild einer im Raum fliegenden Kuh fragt, stattdessen ein Bild einer im Raum schwebenden Handtasche erhalten.
Je mehr derart manipulierte Bilder in einem KI-Trainingssatz sind, desto dramatischer die Auswirkungen. In einem ausführlichen Forschungs-Paper legen die Wissenschaftler dar, dass bereits wenige Bilder ausreichen, um eine KI falsch zu trainieren.
Während Glaze als defensives Werkzeug bezeichnet wird, das einzelne Künstler verwenden können, um sich gegen Stilimitationsangriffe zu schützen, soll Nightshade ein offensives Werkzeug sein, das Künstler als Gruppe verwenden können, um Modelle zu stören, die ihre Bilder ohne Zustimmung nutzen und so alle Künstler vor diesen Modellen schützen. Während Glaze bei jedem Werk verwendet werden sollte, dass Urheber*innen online stellen, ist Nightshade eine optionale Funktion, die verwendet werden kann, um Scraper abzuschrecken, die Opt-Out Hinweise und Nutzungsvorbehalte ignorieren.

Weitere Informationen unter https://nightshade.cs.uchicago.edu/whatis.html

Nightshade soll künftig in „Glaze" implementiert werden. Die Verbindung soll dann besonders effektiv sein.
Die Auswirkungen von Glace und Nightshade wurden an Stable Diffusion Modellen getestet. Hier haben bereits 300 manipulierte Bilder gereicht, um die KI-Trainings zu verwirren. (Connect Living - https://www.connect-living.de/news/nightshade-ki-tool-bilder-schutz-kuenstler-urheberrecht-3206362.html)
Mehr dazu in diesem Artikel der MIT Technology Review:
https://www.technologyreview.com/2023/10/23/1082189/data-poisoning-artists-fight-generative-ai/

4.3 Nepenthes — KI-Crawler in die Teergrube locken

Aaron B. entwickelte „Nepenthes" (benannt nach einer fleischfressenden Pflanze) als Werkzeug, das darauf abzielt, KI-Crawler mit irrelevanten Daten zu beschäftigen und dadurch die Qualität ihrer gesammelten Trainingsdaten zu untergraben.
Die Funktionsweise von Nepenthes ist einfach, aber effektiv: Es generiert automatisch große Mengen an nutzlosen oder irreführenden Links, die speziell darauf ausgelegt sind, KI-Crawler sinnfrei zu beschäftigen. Der Crawler lädt eine URL herunter, und wenn er Links zu anderen URLs sieht, lädt er auch diese herunter. Nepenthes generiert zufällige Links, die immer zu ihm selbst zurückführen und führt den Crawler so quasi im Kreis herum.

Der Webseite 404 Media erklärte Aaron B., Nepenthes sei „ein unendliches Labyrinth, in dem ein Minotaurus gefangen ist, wobei der Crawler der Minotaurus ist, der nicht entkommen kann" und weiter „Nehmen wir an, Sie haben Rechnerleistung und Bandbreite zur Verfügung und wollen einfach nur sehen, wie diese KI-Modelle verbrennen. Nepenthes hat, was Sie brauchen ... Kurz gesagt, lassen Sie sie so viel Mist aufsaugen, wie sie Platz auf der Festplatte haben, und ersticken Sie daran."

Das hört sich nicht unbedingt nach einer praxisnahen Anwendung an, aber es zeigt, dass es auch kreative Möglichkeiten zur Abwehr von KI-Crawlern gibt.
Der Entwickler begründet sein Handeln so: „„Es ist auch eine Art Kunstwerk, in dem ich meiner schieren, unverfälschten Wut darüber, wie die Dinge laufen, freien Lauf lasse. Ich hatte einfach die Nase voll davon, wie sich das Internet zu einem Panoptikum der Geldabschöpfung entwickelt, wie die Welt als Ganzes in den Faschismus abrutscht und Oligarchen das Sagen haben - und es ist schlimm genug geworden, dass wir uns nicht durch Boykott oder Wahlen herauswinden können, wir müssen anfangen, denen da oben echten Schmerz zuzufügen, damit sich etwas ändert."

Bericht auf Golem.de (24.01.2025):
https://www.golem.de/news/vorgehen-gegen-ki-konzerne-neues-tool-beschaeftigt-crawler-mit-viel-bloedsinn-2501-192710.html

Bericht auf 404 media (23.01.2025):
https://www.404media.co/email/7a39d947-4a4a-42bc-bbcf-3379f112c999/

Die Entwicklerwebsite:
https://zadzmo.org/code/nepenthes/

5. FREELENS, der Fotorat und die Initiative Urheberrecht

Welche Rolle spielt FREELENS bei all den Fragen von Text and Data Mining und der europäischen und deutschen Gesetzgebung?
FREELENS ist ein Stakeholder in diesen Prozessen. Als Berufsverband vertreten wir die Interessen unserer Mitglieder in erster Linie durch Mitglieder des Vorstands, die in den Dachverbänden Deutscher Fotorat, Initiative Urheberrecht sowie in der VG Bild-Kunst aktiv sind. Sie übernehmen dort Aufgaben, recherchieren, wirken an Positionspapieren mit, nehmen an Veranstaltungen teil und sprechen mit Politikern.
An der Ausgestaltung des geplanten Praxisleitfadens für den AI Act sind wir mit dem Fotorat und der Initiative Urheberrecht unmittelbar beteiligt. Der Deutsche Fotorat ist in diesem Prozess übrigens die einzige Stimme für die Interessen der europäischen Fotograf*innen, wir haben an der erfolgreichen Bewerbung dafür mitgewirkt. Die Initiative Urheberrecht ist einer der wichtigsten Vertreter der europäischen Urheber.
Wir versuchen dort unsere Punkte zu platzieren, uns gegenüber stehen Lobbyverbände wie BITKOM und IT-Giganten wie Google oder Amazon. Sie haben quasi unbeschränkte finanzielle und personelle Mittel.
Aber wir werden gehört und die Politiker haben erkannt, dass es wichtig ist, die Rechte von Urhebern und Kreativen zu schützen.

6. Fazit

Die Abweisung von Crawlern durch Vorbehalte an der eigenen Website wie robots.txt und ai.txt sind nur wirksam, wenn Crawler sich daran halten. Sie können aber vor Gericht als wirksamer Opt-Out-Hinweis gewertet werden und somit Scraping überhaupt erst zu einer illegalen Handlung werden lassen.
Scraping tatsächlich verhindern kann man nur mit Lösungen wie Cloudflare, die Crawler blockieren und gar nicht erst bis zur Webseite kommen lassen.
Diese Maßnahmen wirken aber nur bei Bildern auf der entsprechenden Website. Da aber gerade Fotografien überwiegend auf fremden Webseiten publiziert werden, z.B. auf den Seiten der Kunden, können wir nur am Werk selbst ein Opt-Out erklären — und hoffen, dass Kunden diese Informationen nicht löschen und Crawler sie zur Kenntnis nehmen.
Vieles bleibt also auch heute noch vage und unklar. Dennoch solltet ihr jedenfalls schon jetzt zumindest die einfachen Maßnahmen ergreifen, wenn ihr die Verwendung eurer Fotos als Trainingsdaten nicht wünscht oder sie angemessen honoriert haben möchtet:

1. vollständige Urheberangaben und Vorbehalt gegen KI-Training in den IPTC-Daten
2. Vorbehalt gegen KI-Training im Impressum und in den robots.txt.

Diese Aufzählung stellt den Stand der Dinge im März 2024 dar. Es gibt fast täglich neue Entwicklungen zum Thema. Wir werden euch weiterhin auf dem Laufenden halten.

»Photograph the world as it is. Nothing's more interesting than reality.«
Mary Ellen Mark, Magnum-Fotografin