KI-Trainingsdaten enthalten private Informationen

Originalartikel

Backup

<html> <div class=„seitenkopfmedia columns twelve m-twelve ts-picturewrapper“><noscript></noscript> <p class=„textabsatz columns twelve m-ten m-offset-one l-eight l-offset-two“>Trainingsdaten sind der Rohstoff für KI-Systeme. Sie bestehen aus riesigen Mengen an Bildern und Texten aus dem Netz. Eine BR-Recherche zeigt nun: Darunter sind viele privaten Daten - ein Problem für den Datenschutz. Von Katharina Brunner und Elisa Harlan, BR <p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>Das Nacktbild eines Niederländers: In der Bildbeschreibung stehen sein Vor- und Nachname und die Geokoordinaten des Aufnahmeortes. Sucht man mit diesen Informationen im Internet nach dem Mann, ist er schnell gefunden. <p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>Der Niederländer ist kein Einzelfall. Bei der Analyse des weltweit wohl größten Trainingsdatensatzes für KI-Bildgenerierung haben BR-Datenjournalistinnen massenweise Daten gefunden, mit denen sich Personen identifizieren lassen: Gesichter und Namen, Geokoordinaten oder E-Mails, sogar Kontonummern. Der LAION5B-Datensatz, eine Abkürzung für „Large-scale Artificial Intelligence Open Network“, übersetzt: „Groß angelegtes offenes Netzwerk für künstliche Intelligenz“, besteht aus fünf Milliarden Links auf Bilder und ihren Beschreibungen im Internet. Er ist zugleich der einzige vergleichbare Trainingsdatensatz für KI-Modelle, der öffentlich zugänglich ist. <h2 id=„Problematische-Daten-bereits-in-Stichprobe“ class=„meldungsubhead columns twelve m-ten m-offset-one l-eight l-offset-two liveblog–anchor“>Problematische Daten bereits in Stichprobe</h2><p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>KI-Trainingsdatensätze bestehen üblicherweise aus enormen Mengen an Texten und Bildern. Sie liefern den Rohstoff für KI-Systeme, die Texte und Bilder generieren, und derzeit von Millionen Menschen auf der Welt benutzt werden. Sogenannte Bildgeneratoren wie Stable Diffusion, Midjourney oder Dall-e von OpenAI funktionieren alle ähnlich: Menschen beschreiben mit kurzen Texteingaben, sogenannten Prompts, welches Motiv auf dem gewünschten Bild zu sehen sein soll. Die Programme erstellen dann mit Hilfe der Trainingsdaten ein neues Bild, oder so viele man möchte. <p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>Ende Juni 2023 reichte in den USA eine anonyme Gruppe Klage gegen das Unternehmen OpenAI ein, das unter anderem auch ChatGPT betreibt: Sie wirft der Firma unter anderem massenhafte Verletzungen der Privatsphäre vor. <p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>Die BR-Analyse zeigt, dass sich auch im deutschsprachigen Teil der LAION-Datensätze für 20 Millionen Bilder Zusatzinformationen finden, sogenannte Exif-Metadaten. Exif steht für „Exchangeable Image File Format“ und bezeichnet Informationen, die in den Bilddateien gespeichert sind. Das Aufnahmegerät hält zum Beispiel zusätzlich zum Bild den Zeitpunkt der Aufnahme fest, das Modell der Kamera und oft auch den genauen Standort. Solche Positionsangaben können automatisch bei allen Aufnahmen entstehen. <p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>Ein Ergebnis der BR-Analyse: Zu 310.000 Bildverweisen im deutschsprachigen LAION-Teil

konnte das Team von BR Data den exakten Aufnahmeort auslesen.</p><p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>Diese Exif-Daten, in denen die Ortsinformationen gespeichert sind, werden auch als „Restinformationen“ bezeichnet. Sofern es keine begründeten Argumente gibt, sollen solche Daten nach Meinung des Bundesamts für Sicherheit und Informationstechnik (BSI) vernichtet werden, wenn Dateien weitergegeben werden.</p><h2 id=„Datenschuetzer-Kleinfeld-Hochproblematisch“ class=„meldungsubhead columns twelve m-ten m-offset-one l-eight l-offset-two liveblog–anchor“>Datenschützer Kleinfeld: „Hochproblematisch”</h2><p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>Für Eike Kleinfeld, tätig beim Hamburger Datenschutzbeauftragten, ist die massenhafte Verbreitung von Ortsinformationen aus Exif-Metadaten ein Problem, sofern es sich um sensible Informationen handelt, die einen Rückschluss auf natürliche Personen erlauben: „Die absolute Zahl ist natürlich hoch problematisch, wenn man sich vergegenwärtigt, dass da Millionen Bilder mit solchen Informationen liegen.“</p><p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>LAION-Mitgründer Christoph Schuhmann war sich bisher des Problems nicht bewusst: „Auf das Problem werden wir jetzt das erste Mal hingewiesen“, so Schuhmann gegenüber dem BR.</p><h2 id=„LAION-setzt-auf-Transparenz“ class=„meldungsubhead columns twelve m-ten m-offset-one l-eight l-offset-two liveblog–anchor“>LAION setzt auf Transparenz</h2><p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>LAION ist ein Zusammenschluss von Freiwilligen aus Europa und Nordamerika. Schuhmann sagt: „Wir haben LAION aus Begeisterung für diese KI-Technologie gegründet und dem Wunsch, dass das demokratisiert wird und es am Ende nicht nur zwei, drei großen Firmen gibt.“</p><p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>Die Methode der radikalen Transparenz unterscheidet LAION von der Konkurrenz wie Microsoft, Google, Midjourney oder OpenAI. Über deren Trainingsmaterial und Methodik ist wenig bekannt. Die vier Unternehmen ließen Fragen, wie Trainingsdaten zusammengesetzt und verwendet werden, unbeantwortet. Will man die Lieferkette von populären KI-Bildgeneratoren untersuchen, sind LAIONs Datensätze, Modelle und Werkzeuge aktuell die einzige Möglichkeit - für Wissenschaftler und Journalisten.</p><h2 id=„EU-Gesetzgebung-nimmt-Trainingsdaten-in-den-Blick“ class=„meldungsubhead columns twelve m-ten m-offset-one l-eight l-offset-two liveblog–anchor“>EU-Gesetzgebung nimmt Trainingsdaten in den Blick</h2><p class=„textabsatz m-ten m-offset-one l-eight l-offset-two columns twelve“>Mehr Transparenz in Sachen Trainingsdaten sieht die EU in der geplanten Gesetzgebung zur Künstlichen Intelligenz vor, dem sogenannten AI Act. Unklar ist aber, wie genau die künftigen Regelungen aussehen sollen. Derzeit wird noch verhandelt: „Was im AI Act stehen wird, ist in Teilen noch offen. Momentan ist ein wichtiges Diskussionsthema, ob generative KI als Hochrisikogruppe gewertet und reguliert wird“, sagt Sandra Wachter, Professorin am Oxford Internet Institute. Bis Ende des Jahres sollen die EU-Regeln stehen und 2025 in Kraft treten. Bis dahin laufen die Systeme weiter wie bisher.</p> </html>