Chantal Lutz | Benjamin Domenig | Anja Flükiger

Datenschutzkonformer Einsatz von ChatGPT an Schulen

Die Integration von Künstlicher Intelligenz in Bildungseinrichtungen, insbesondere die Verwendung von OpenAI's ChatGPT, steht vor rechtlichen und datenschutztechnischen Herausforderungen. Während die kantonale Schul- und Datenschutzgesetzgebung den Rahmen für den Einsatz solcher Technologien im Unterricht absteckt, erfordert die Praxis eine genaue Betrachtung der Datenbearbeitungsprozesse. Dieser Beitrag untersucht, wie ChatGPT datenschutzkonform in den Unterricht integriert werden kann und welche Einsatzmöglichkeiten aufgrund von Kontrolldefiziten problematisch sind. Besonderes Augenmerk liegt auf der Nutzung über eine API-Schnittstelle, die es Schulen ermöglicht, die Weitergabe sensibler Daten zu steuern und die Privatsphäre der Schülerinnen und Schüler und der Lernenden zu wahren. Indem die Schulen klare Richtlinien für den Einsatz von KI-Technologien erlassen, kann der verantwortungsvolle Umgang mit ChatGPT im Bildungsbereich gewährleistet werden.

Introduire l'intelligence artificielle dans les établissements d'enseignement, en particulier avec l'utilisation de ChatGPT d'OpenAI, présente des défis du point de vue juridique ainsi qu'en matière de protection des données. Alors que les législations cantonales relatives à l'enseignement et à la protection des données définissent le cadre de l'utilisation de telles technologies dans l'enseignement, la pratique requiert un examen minutieux des processus de traitement des données. Cet article examine comment ChatGPT peut être intégré à l'enseignement en conformité avec la protection des données et quelles utilisations restent problématiques en raison d'un manque de contrôle. Une attention particulière est accordée à l'utilisation de cette technologie via une interface API, qui permet aux écoles de gérer la transmission de données sensibles et de préserver la vie privée des élèves et des apprenants. En adoptant des directives claires sur l'utilisation des technologies d'intelligence artificielle, les écoles peuvent garantir une utilisation responsable de ChatGPT dans le domaine de l'éducation.

I. Einleitung

Die Autorenschaft hat ChatGPT die folgende Einstiegsfrage gestellt:

«ChatGPT, machst du dir Sorgen, dass du nicht datenschutzkonform bist?»

Die Antwort lautete in gekürzter Form wie folgt:

«Es liegt in der Verantwortung meiner Entwickler und Betreiber sicherzustellen, dass meine Implementierung und Nutzung datenschutzkonform sind. Als KI-Modell habe ich keinen direkten Zugriff auf persönliche Daten oder Informationen über Benutzer. Meine Funktionsweise besteht darin, Eingaben zu analysieren und generische Textantworten zu generieren, basierend auf dem Musterlernen aus grossen Datensätzen.»[1]

An dieser Stelle sind die Aussagen von ChatGPT zu relativieren. ChatGPT kann tatsächlich mit Personendaten der Benutzenden in Berührung geraten, und zwar über die Chatverläufe und selbstredend über die Accountinformationen.[2] Sobald man einen ChatGPT-Account mit einer E-Mail-Adresse anlegt – anders kann der Dienst auf der Plattform von OpenAI[3] nicht genutzt werden – erhält man die Möglichkeit, eine Datenschutzeinstellung vorzunehmen: Das

Chantal Lutz, MLaw, Rechtsanwältin, Bern.

Benjamin Domenig, Dr. iur., M.A. HSG in Law and Economics, Rechtsanwalt, Bern.

Anja Flükiger, MLaw, Bern.

1 Die Frage wurde am 3. August 2023 unter ‹https://chat.openai.com/› gestellt.

2 ‹https://openai.com/policies/privacy-policy› (abgerufen am 5. November 2023).

3 ChatGPT ist ein Dienst des Unternehmens OpenAI OpCo, LLC mit Sitz in San Francisco, USA, wobei die OpenAI Ireland Limited mit Sitz in Dublin, Irland, als EU-Vertretung im Sinne von Art. 27 DSGVO fungiert (‹https://openai.com/policies/privacy-policy›). OpenAI hat gemäss eigenen Angaben zum Ziel, die KI-Systeme der ganzen Menschheit zugänglich zu machen, damit alle davon profitieren können und den technischen Fortschritt noch mehr fördern (‹https://openai.com/blog/planning-for-agi-and-beyond› [abgerufen am 4. August 2023]).

Löschen der Chatverläufe und damit die Unterbindung von deren Verwendung für das Training der Sprachmodelle von OpenAI. Wird diese Einstellung aktiv vorgenommen, werden die Chatverläufe innert 30 Tagen gelöscht.[4]

Die Autorenschaft hat ChatGPT anschliessend die Kontrollfrage «*Bist du datenschutzkonform?*» gestellt, um zu testen, ob ChatGPT unter leicht anderem Prompting[5] anders reagiert. Die gekürzte Antwort war die folgende: «*Ich bin nur ein Programm, das auf den Servern von OpenAI läuft und keine persönlichen Daten speichert oder besitzt. Als künstliche Intelligenz folge ich den Datenschutzbestimmungen und den Nutzungsbedingungen von OpenAI.*»[6]

II. Ausgangslage und Fragestellung

Der vorliegende Diskussionsbeitrag befasst sich mit der Frage, wie ChatGPT an Schulen und insbesondere im Unterricht datenschutzkonform eingesetzt werden kann. Bei Schulen handelt es sich um unselbständige öffentlich-rechtliche Anstalten ohne Rechtspersönlichkeit.[7] Bei den Volksschulen agieren in der Regel die Gemeinwesen als Trägerorganisation, bei den Mittel- und Berufsschulen der Kanton.[8] Die Schülerinnen und Schüler und die Lernenden befinden sich in Bezug auf die Schule in einem sog. Sonderstatusverhältnis, das regelmässig eine Einschränkungen der Grundrechte bedeutet. Dabei sind die Anforderungen an Normstufe und Normdichte dort weniger streng, wo Grundrechtsbeschränkungen in Frage stehen, welche sich in voraussehbarer Weise aus dem Zweck des Sonderstatusverhältnisses ergeben. Ausserdem dürfen die Einschränkungen nicht schwer wiegen. Sind diese Voraussetzungen erfüllt, dürfen Rechte und Pflichten der Schülerinnen und Schüler und der Lernenden, auch in einer Verordnung oder einem Anstaltsreglement enthalten sein. Dies gilt insbesondere für organisatorische Regelungen oder für solche, welche die Einzelheiten des Benutzungsverhältnisses regeln. Dabei sind der Anstaltszweck, die Verhältnismässigkeit und die Rechtsgleichheit zu wahren. Wesentliche Rechte und Pflichten sind hingegen stets in einem Gesetz im formellen Sinn zu regeln und die weiteren Aspekte von Art. 36 BV sind auch bei geringeren Anforderungen an die Normstufe und -dichte zu beachten.[9]

In Bezug auf den Einsatz von ChatGPT[10] an Schulen kommen als mögliche Grundrechtseingriffe eine Ungleichbehandlung durch dem Sprachmodell zugrunde liegende «bias»[11] (Art. 8 BV), Willkür aufgrund des Phänomen des «Halluzinierens»[12] (Art. 9 BV) und eine Verletzung der Privatsphäre bzw. der Missbrauch von persönlichen Daten der Schülerinnen und Schüler und der Lernenden (Art. 13 BV) in Frage. Gerade in Sonderstatusverhältnissen kann die Grundrechtssensibilität erhöht sein,[13] womit dem Einsatz von neuen und die nötige Transparenz vermissen lassenden Technologien wie ChatGPT[14] Aufmerksamkeit geschenkt werden sollte. Datenschutzfragen sind dabei von besonderer Relevanz. Der vorliegende Beitrag legt den Fokus somit auf den Schutz von Personendaten und die Vermeidung ihres Missbrauchs bei der Nutzung von ChatGPT im schulischen Umfeld.

III. Funktionsweise von ChatGPT

1. Trainingsdaten

ChatGPT ist ein von OpenAI entwickelter Chatbot, der auf künstlicher Intelligenz (KI), im englischen «Artificial Intelligence» (AI) genannt, basiert. Eine KI zeichnet sich dadurch aus, dass sie autonom arbeiten kann und sich die entsprechenden Verarbeitungsprozesse selbstredend weiterentwickeln.[15] Gemäss Art. 3 des Vorschlages für eine Verordnung des europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (KI-Verordnung-E) handelt es sich bei einem KI-System um «*… eine Software, die mit einer oder mehreren (…) Techniken und Konzepte entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren.*»[16]

4 ‹https://help.openai.com/en/articles/7730893-data-controls-faq› (abgerufen am 5. November 2023).

5 Definition gemäss Cambridge Dictionary: «*the act of trying to make someone say something*» (‹https://dictionary.cambridge.org/de/worterbuch/englisch/prompting› [abgerufen am 5. November 2023]). Gemäss dem Gabler Wirtschaftslexikon bedeutet dieser Begriff bei generativer KI einen Input des Benutzers, zu dem das System einen Output erzeugt (‹https://wirtschaftslexikon.gabler.de/definition/prompt-125087› [abgerufen am 5. November 2023]).

6 ‹https://chat.openai.com/›.

7 VGer Zürich vom 5. Oktober 2021, VB.2021.00689, Regeste.

8 Am Beispiel des Kantons Bern: Art. 34 Abs. 1 Volksschulgesetz (VSG; BSG 432.210), Art. 33 Abs. 1 Mittelschulgesetz (MiSG; BSG 433.12) und Art. 16 Abs. 1 des Gesetzes über die Berufsbildung, die Weiterbildung und die Berufsberatung (BerG; BSG 435.11).

9 S. A. Bernet, Der Lehrplan – Rechtsnatur und Bedeutung, Zürich 2021, 44 f.

10 Chat Generative pre-trained Transformer (ChatGPT) ist ein sog. Large Language Model, das einer enormen Menge an frei verfügbaren Daten aus dem Internet trainiert wurde. Dabei wurde «Reinforcement Learning from Human Feedback (RLHF)» verwendet, eine Trainingsmethode, die menschlichen Input/Bewertungen verwendet und das Modell für korrekte Antworten belohnt (‹https://openai.com/blog/chatgpt› [abgerufen am 5. November 2023]).

11 Der Ausdruck «bias» bedeutet zu Deutsch Vorurteil, Voreingenommenheit oder Verzerrung. In Zusammenhang mit ChatGPT bedeutet dies, dass durch die beigebrachten Trainingsinformationen bereits sprachliche Annahmen beispielsweise über das Geschlecht oder die Herkunft gemacht werden (M. Kurpicz-Briki, Cultural Differences in Bias? Origin and Gender Bias in Pre-Trained German and French Word Embeddings, Biel 2020, 1).

12 Eine Halluzination bedeutet, dass von der KI falsche Informationen erzeugt werden, die von externen Faktoren abweichen oder von der kontextuellen Logik der vorliegenden Fragestellung (‹www.computerweekly.com/de/definition/KI-Halluzination#:~:text=Was%20sind%20KI%2DHalluzinationen%3F,kontextueller%20Logik%20oder%20beidem%20sein.› [abgerufen am 6. November 2023]).

13 Bernet (Fn. 9), 47.

14 Gemäss dem Foundation Model Transparency Index der Stanford University für Large Language Models erreicht das Sprachmodell ChatGPT-4 von OpenAI lediglich ein Rating von 48% und landet damit auf dem 3. Platz. Erstrebenswert seien jedoch Werte von mind. 80%, die aktuell keines der verglichenen Sprachmodelle erreicht (‹https://hai.stanford.edu/news/introducing-foundation-model-transparency-index› [abgerufen am 5. November 2023]).

15 P. Müller-Peltzer, Künstliche Intelligenz und Datenschutzrecht: Ein Blick auf die neue KI-Verordnung, DATENSCHUTZ-BERATER, 2022, 230.

16 ‹https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF› (abgerufen am 4. August 2023).

ChatGPT wurde so entwickelt, dass der Chatbot die Fragen und Anweisungen der Nutzer verstehen und beantworten kann. Dies geschieht, indem er eine riesige Menge an vorhandenem Text «liest» und lernt, wie Wörter dazu neigen, im Zusammenhang mit anderen Wörtern zu erscheinen. Anschliessend verwendet er das Gelernte, um das am besten passende Wort vorherzusagen, das als Antwort auf eine Frage angezeigt werden könnte sowie jedes nachfolgende Wort.[17] Es handelt sich dabei um eine Art neuronales Netz, die Algorithmen und Strukturen gleichen der Funktionsweise des menschlichen Gehirns.[18]

Explizit bedient sich ChatGPT an Daten, die im Internet öffentlich zugänglich sind, also nicht dafür bezahlt werden muss oder sich nicht im Darknet befinden. Ausserdem bedient er sich an Informationen, welche OpenAI von Dritten lizenziert und an Informationen, die die Nutzenden oder menschliche Trainer zur Verfügung stellen.[19] Gesamthaft wurde ChatGPT-3 mit rund 45 Terabyte Text trainiert.[20]

Für das Training von ChatGPT werden personenbezogene Daten verwendet. Gemäss OpenAI werden diese Daten nur als Trainingsinformationen gebraucht, damit ChatGPT menschliche Sprache lernt, diese versteht und darauf reagiert.[21]

Die Datenschutzrichtlinien von OpenAI liefern in Bezug auf die Verwendung von Trainingsdaten kaum Informationen. Sie verweisen jedoch auf einen Blogpost von OpenAI, der das folgend Statement enthält: «*We use training information only to help our models learn about language and how to understand and respond to it. We do not and will not use any personal information in training information to build profiles about people, to contact them, to advertise to them, to try to sell them anything, or to sell the information itself.*»[22].

Per Default werden die Chatverläufe von Privatpersonenaccounts für das Training von ChatGPT verwendet, solange die Benutzenden dies in ihrem OpenAI-Account nicht aktiv unterbinden. Per Default unterbunden ist diese Datenverwendung hingegen im Rahmen der Nutzung der API[23]-Version oder der Enterprise-Version von ChatGPT.[24]

2. Umgang mit persönlichen Daten

ChatGPT nutzt zum Trainieren auch persönliche Daten, welche aus den verwendeten Quellen hervorgehen. Gemäss ChatGPT werden diese nur zum Trainieren verwendet, nicht aber um diese dann primär für die Beantwortung von Anfragen bereitzuhalten. Wenn dennoch persönliche Informationen über Personen als Antwort erscheinen, kann durch ein auf der OpenAI Webseite auffindbares Formular[25] Widerspruch erhoben werden. Ausserdem haben Einzelpersonen das Recht, auf ihre personenbezogenen Daten, die in den Schulungsinformationen enthalten sein können, zuzugreifen, sie zu korrigieren, einzuschränken, zu löschen oder zu übertragen. Diese Rechte können ausgeübt werden, indem sich die betroffene Person mit ihrem Anliegen an die E-Mail-Adresse dsar@openai.com wendet.[26]

Um zu testen, wie ChatGPT persönliche Informationen herausgibt, wurde der Name einer der Autorinnen plus deren Wohnort in die ChatGPT-Suche eingegeben und gefragt, wer diese Person sei. ChatGPT hat die Autorin nicht gekannt und geantwortet: «*Als KI-Sprachmodell habe ich keinen Zugriff auf aktuelle Daten oder Informationen über spezifische Personen, die nach meinem Wissensstichtag im September 2021 aufgetreten sind.*»

Bezüglich der Aufbewahrung der Informationen äusserte sich OpenAI eher vage: «*Wir bewahren diese Informationen nur so lange auf, wie wir sie benötigen, um den beabsichtigten Zweck zu erfüllen. Wie lange wir diese Informationen aufbewahren, hängt von Faktoren wie ihrer Menge, Art und Sensibilität, dem Risiko eines Schadens durch unbefugte Nutzung oder Weitergabe ab, ob die Informationen noch notwendig oder nützlich sind, um unsere Modelle zu trainieren oder zu aktualisieren, und von etwaigen gesetzlichen Anforderungen.*»[27]

Doch werden mit dieser Handhabung die einschlägigen Datenschutzbestimmungen eingehalten?

17 ‹https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed› (abgerufen am 3. August 2023).

18 S. Klaus, KI trifft Datenschutz – Risiken und Lösungsansätze, in: A. Epiney, S. Rovelli (Hg.), Forum Europarecht Band Nr. 42 – Künstliche Intelligenz und Datenschutz, 2021, 82.

19 ‹https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed› (abgerufen am 3. August 2023).

20 J. Rudolph/S. Tan/S. Tan, ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?, Journal of Applied Learning & Teaching 2023, 344.

21 Wenn ChatGPT gefragt wird, wie er trainiert wurde, folgt die Antwort: «*Die Daten, die für das Training verwendet wurden, umfassen eine breite Palette von Quellen, wie Bücher, Artikel, Webseiten, Foren, Gespräche und vieles mehr. Der Prozess des Trainings erfolgte in mehreren Schritten*:
1. *Datensammlung: Eine Vielzahl von Texten wurde aus dem Internet und anderen Quellen gesammelt, um eine breite Wissensbasis zu erstellen.*
2. *Datenvorverarbeitung: Die gesammelten Textdaten wurden bereinigt und vorverarbeitet, um unnötige Informationen zu entfernen und eine konsistente Darstellung zu gewährleisten.*
3. *Training: Mit den vorverarbeiteten Daten wurde ein maschinelles Lernmodell, wie GPT-3.5, trainiert. Während des Trainings lernte das Modell, Muster und Zusammenhänge zwischen den Wörtern und Sätzen zu erkennen, um in der Lage zu sein, sinnvolle und kontextbezogene Antworten zu generieren.*
4. *Validierung: Das Modell wurde mit validierten Daten überprüft, um seine Leistung und Genauigkeit zu verbessern.*
5. *Iteratives Training: Der obige Prozess wurde mehrmals iteriert, um die Leistung und Qualität des Modells kontinuierlich zu verbessern.*»
(‹https://chat.openai.com/›).

22 ‹https://openai.com/policies/privacy-policy›.

23 Application Programming Interface, was auf Deutsch Programmierschnittstelle bedeutet (vgl. auch ‹www.computerweekly.com/de/definition/Programmierschnittstelle-API› [abgerufen am 6. November 2023]).

24 ‹https://openai.com/enterprise-privacy› (abgerufen am 6. November 2023).

25 ‹https://share.hsforms.com/1UPy6xqxZSEqTrGDh4ywo_g4sk30› (abgerufen am 3. August 2023).

26 ‹https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed›, (abgerufen am 7. August 2023).

27 ‹https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed› (abgerufen am 4. August 2023).

IV. Datenschutzkonformer Einsatz von ChatGPT an Schulen

1. Ausgangslage

Gemäss dem Bundesgesetz über den Datenschutz (DSG) dürfen ohne entsprechende Rechtsgrundlage Personendaten erhoben werden, solange die Bearbeitungsgrundsätze eingehalten werden.[28] Schulen profitieren von dieser liberalen Grundhaltung des DSG wiederum nicht, da sie dem Datenschutzrecht ihres Kantons unterstehen. Demnach ist eine Datenbearbeitung stets nur auf der Basis einer gesetzlichen Grundlage bzw. zwecks Erfüllens einer gesetzlichen Aufgabe zulässig, wobei für die Bearbeitung von besonders schützenswerten Personendaten eine formell-gesetzliche Grundlage notwendig ist.[29] Diese Aufgabenumschreibung findet sich in den jeweiligen Schulgesetzen, welche Schulen in der Regel auch zur Bearbeitung von besonders schützenswerten Personendaten ermächtigen.[30] Gemäss einer Vorgabe des Mittelschul- und Berufsbildungsamts des Kantons Bern fallen darunter – auch wenn in Art. 3 Abs. 1 KDSG nicht explizit genannt – praxisgemäss auch Noten und Beurteilungsberichte, Akten betreffend Disziplinarmassnahmen bei Schülerinnen und Schülern sowie unter Umständen Akten von Beschwerdeverfahren.[31]

2. Rechtmässigkeit

ChatGPT kann, wenn pädagogisch sinnvoll eingesetzt, Schulen bei der Vermittlung von Wissen und Fähigkeiten an die Schülerinnen und Schüler und die Lernenden unterstützen. Insbesondere kann ChatGPT diese befähigen, Inhalte kritisch zu hinterfragen. ChatGPT kann für Lehrpersonen eine Entlastung bei der Vorbereitung des Unterrichts bieten, da der Chatbot für die Aufgabenerstellung verwendet werden kann. In diesem Sinne kann sich die Verwendung von ChatGPT tatsächlich als hilfreiches Mittel zum Erreichen der gesetzlich festgelegten Bildungsziele erweisen.[32] Die grundsätzliche Zulässigkeit der Verwendung von KI-gestützten Computeranwendungen im Unterricht lässt sich damit aus der Aufgabenstellung der Schulen ableiten. Die Autorenschaft kommt daher zum Schluss, dass für die Datenbearbeitung im Rahmen des Einsatzes von KI an Schulen eine genügende gesetzliche Grundlage besteht.[33] Im Sinne der herabgesetzten Anforderungen an die Normstufe im Sonderstatusverhältnis wäre es weiter zulässig, die Nutzung von ChatGPT in einer schulinternen Richtlinie zu regeln.

Das vorstehend Ausgeführte gilt für den Fall, dass ChatGPT über die API- oder Enterprise-Version genutzt wird und die Benutzungsdaten (d.h. auch Chatverläufe) nicht zu Trainingszwecken an OpenAI weitergegeben werden. Nicht zulässig wäre die Bekanntgabe von Schülerdaten an OpenAI im Rahmen der Chatverlaufsübermittlung zwecks Training des LLM, weil hierdurch eine Zweckentfremdung der Personendaten und ein rechtlicher Kontrollverlust der Schule über die Schülerinnendaten erfolgt.[34] Eine solche Drittbekanntgabe ist weder gesetzlich vorgesehen noch kann sie über eine Einwilligung der Schülerinnen und Schüler oder der Lernenden im Einzelfall sinnvoll legitimiert und praxistauglich umgesetzt werden. Auch liegt eine solche Datenweitergabe nicht im Interesse der Schülerinnen und Schüler und der Lernenden.[35]

3. Auslagerung der Datenbearbeitung an OpenAI

Die Datenbearbeitung in der Schule kann zusammen mit einem oder durch einen Dritten wahrgenommen werden, allerdings bleibt die Schule für die Bearbeitung verantwortlich. Wird ChatGPT über API oder eine Enterprise-Subscription genutzt, handelt es sich um eine Auftragsdatenbearbeitung und die Schule hat entsprechende vertragliche Kontrollmechanismen mit OpenAI zu vereinbaren.[36] Zu beachten ist, dass bei der Nutzung der API wie auch der Enterprise-Version Personendaten der Benutzenden an die OpenAI-Server in den USA übermittelt werden.[37] Als Garantie im Sinne von bspw. Art. 14 Abs. 2 lit. a KDSG bietet OpenAI ein Standard Data Processing Addendum (DPA) mit Einbezug der Standardvertragsklauseln der Europäischen Union an.[38] Um den Rahmen dieses Beitrags nicht zu überdehnen, wird die Frage betreffend die Zulässigkeit der Übermittlung von Schülerdaten wie bspw. Accountinformationen oder Chatinhalte basierend auf dem DPA von OpenAI nicht näher beleuchtet. Es müsste im Rahmen eines risikobasierten Ansatzes beleuchtet werden, ob die von OpenAI bereitgestellten vertraglichen Garantien genügend rechtliche Kontrollmöglichkeiten für Schulen bieten und ob zusätzliche technische Massnahmen zu ergreifen sind, um eine Identifizierung der Schülerinnen und Schüler und der Lernenden durch OpenAI zu unterbinden.

---

28 B. Domenig/C. Mitscherlich/C. Lutz, Datenschutzrecht für Schweizer Unternehmen, Stiftungen und Vereine, 2. Aufl., Bern 2022, 19.

29 Am Beispiel des Kantons Bern: Art. 5 und 6 des Datenschutzgesetzes (KDSG; BSG 152.04).

30 Am Beispiel des Mittelschulgesetzes des Kantons Bern: Art. 67 MiSG, wobei das Mittelschul- und Berufsbildungsamt Art und Umfang der an den Mittelschulen erfassten Personendaten bestimmt (Art. 85 Mittelschulverordnung [MiSV; 433.121]).

31 Vorgabe Nr. 900.90.900.4 zum Umgang mit Personendaten des Mittelschul- und Berufsbildungsamts (MBA) (‹www.bkd.be.ch/content/dam/bkd/dokumente/de/ueber-uns/organisation/mba/mba-vorgaben/900.90.900.4-Umgang-mit-Personendaten.pdf› [abgerufen am 5. November 2023]).

32 Am Beispiel des Kantons Bern: vgl. die Wirkungsziele in Art. 1 MiSG, bspw. das Vermitteln einer breiten und vertieften Allgemeinbildung und die Befähigung der Schülerinnen und Schüler zur Übernahme verantwortungsvoller Aufgaben in der Gesellschaft.

33 Vgl. auch R. von Thiessen/S. Volz, Künstliche Intelligenz in der Bildung, Rechtliche Best Practices, Standortförderung AWA, Kanton Zürich, Verein Metropolitanraum Zürich, Innovation Zurich (Hg.), Zürich 2023, 6, (‹www.zh.ch/content/dam/zhweb/bilder-dokumente/themen/wirtschaft-arbeit/wirtschaftsstandort/dokumente/best_practices_ki_bildung_DE.pdf›) [abgerufen am 7. November 2023]).

34 von Thiessen/Volz (Fn. 33), 11.

35 Am Beispiel des Kantons Bern: vgl. die Anforderung an eine Datenbekanntgabe an Private in Art. 11 KDSG.

36 von Thiessen/Volz (Fn. 33), 6.

37 ‹https://openai.com/policies/privacy-policy› (abgerufen am 7. November 2023).

38 ‹https://openai.com/policies/data-processing-addendum› (abgerufen am 7. November 2023).

4. Verhältnismässigkeit

Die Schulen trifft die Pflicht, die mit dem Einsatz von ChatGPT zusammenhängende Datenbearbeitung auf das Notwendige zu beschränken. In diesem Sinne gibt Art. 5 Abs. 3 KDSG für den Kanton Bern vor, dass die Personendaten und die Art des Bearbeitens für die Aufgabenerfüllung geeignet und notwendig sein müssen.

Bei der Registrierung für die Nutzung von ChatGPT[39] muss neben der E-Mail-Adresse der Namen sowie das Geburtsdatum zwecks Bestätigung des Alters angegeben werden. Auch wird man darauf hingewiesen, dass man sich beim Klicken auf «continue» mit den «Terms», also den Nutzungsbedingungen sowie der «Privacy policy», den Datenschutzrichtlinien, einverstanden erklärt. Im Anschluss muss das Profil noch durch eine Handynummer verifiziert werden. Im Lichte dieser Datenerhebung werden die Nutzungsvarianten von ChatGPT im Unterricht nachfolgend anhand ihrer Datenschutzkonformität bewertet, wobei 1 die schlechteste, und 4 die beste Variante darstellt:

1. Die Nutzung des Dienstes durch Schülerinnen und Schüler sowie Lernende mit der eigenen E-Mail-Adresse, der eigenen Handynummer, des Geburtsdatums sowie des eigenen Gerätes, sei dies ein Laptop, ein Handy oder ein Tablet. Die Schule hat hiermit keinerlei rechtliche oder faktische Kontrollmöglichkeit bei der Datenbearbeitung durch OpenAI, weshalb von dieser Nutzungsart im Unterricht abgeraten wird.
2. Die Nutzung des Dienstes über einen Account einer Lehrperson, wobei diese ein explizites Opt-Out für die Nutzung der Chatverläufe zu Trainingszwecken durch OpenAI vornehmen müsste. In diesem Szenario müsste weder die Handynummer noch die E-Mail-Adresse oder das Geburtsdatum der Schülerinnen und Schüler oder der Lernenden angegeben werden. Problematisch in diesem Beispiel ist jedoch, dass die Chats für die Lehrperson zugänglich wäre und diese je nachdem Kenntnis von sensiblen Informationen der Benutzenden erhält. Die Lehrperson müsste sicherstellen, dass konkrete Anweisungen für die Inhalte, die mit ChatGPT geteilt werden, vorliegen und dass die Chatverläufe je nach Informationsgehalt nach dem Unterricht wieder gelöscht werden.[40]
3. Die Nutzung der API- oder Enterprise-Version von ChatGPT. Hierdurch wäre sichergestellt, dass die Chatverläufe per se nicht zum Training des LLM benutzt werden. Insbesondere bei der API-Nutzung ist der Dienst im Sinne des Grundsatzes von «Privacy by Design» so zu konzipieren, dass die Schülerinnen und Schüler und die Lernenden möglichst wenige Personen erfassen müssen (insb. kein Geburtsdatum und keine Handynummer).[41] In diesem Zusammenhang ist auch die von OpenAI festgelegte Altersgrenze für die Nutzung von 13 Jahren zu beachten.[42] Wird ChatGPT über private Geräte von Schülerinnen und Schülern sowie Lernenden verwendet, sollte der Dienst von der Schule in einer Weise angeboten werden, dass eine Anonymisierung der IP-Adresse erfolgt.
4. Nutzung von ChatGPT via Microsoft 365 Copilot[43], sobald der Dienst für Schulen verfügbar ist.[44] Es handelt sich dabei um LLM, die direkt in die Anwendungen von Microsoft eingebunden werden. Der Dienst verwendet die Azure OpenAI services, welche auf den Servern von Microsoft in Europa und nicht auf denjenigen von OpenAI in den USA laufen. Gemäss den Äusserungen von Microsoft werden die bestehenden Datenschutz- und Datensicherheitsrichtlinien im eigenen Azure Tenant übernommen und es wird versprochen, dass Daten innerhalb des Tenants isoliert werden. Die über Microsoft 365 Copilot an die LLM übermittelten Geschäftsdaten werden nicht zu Trainingszwecken verwendet. Hierdurch hat die Schule sowohl rechtliche Kontrollmöglichkeiten über die bestehenden Verträge mit Microsoft als auch faktische Kontrollmöglichkeiten, indem der Datenzugriff auf die Schuldaten via Berechtigungsmanagement in Azure und dem M365-Anwendungen gesteuert werden kann.[45] So kann auch sichergestellt werden, dass Copilot nicht auf besonders schützenswerte Personendaten der Schülerinnen und Schüler und der Lernenden, wie Gesundheitsdaten, Noten und Beurteilungsberichte oder Akten von Disziplinarmassnahmen zugreift.

Erwägt die Schule eine der oben dargestellten Einsatzmöglichkeiten, müsste sie unter Umständen eine sog. Vorabkontrolle bei der zuständigen Datenschutzaufsichtsbehörde vornehmen lassen.[46] Hierfür müsste die Schule je nach Kanton eine sog. Datenschutz-Folgenabschätzung bzw. ein ISDS-Konzept erstellen.[47]

39 ‹https://chat.openai.com/auth/login› (abgerufen am 7. August 2023).
40 von Thiessen/Volz (Fn. 33), 8; ‹www.klicksafe.de/news/chatgpt-in-der-schule-wie-damit-umgehen› (abgerufen am 6. November 2023).
41 von Thiessen/Volz (Fn. 33), 8.
42 ‹www.klicksafe.de/news/chatgpt-in-der-schule-wie-damit-umgehen›, (abgerufen am 6. November 2023).
43 ‹https://adoption.microsoft.com/de-de/copilot/›, (abgerufen am 6. November 2023).
44 Für den öffentlichen Sektor werde Microsoft 365 Copilot im Sommer 2024 erwartet: ‹https://techcommunity.microsoft.com/t5/public-sector-blog/elevating-government-productivity-microsoft-365-copilot-at-the/ba-p/3969398› (abgerufen am 7. November 2023).
45 ‹https://learn.microsoft.com/de-de/microsoft-365-copilot/microsoft-365-copilot-privacy›, (abgerufen am 6. November 2023).
46 Am Beispiel des Kantons Bern: Art. 17a KDSG, wonach beim Einsatz von technischen Mitteln mit besonderen Risiken für die Rechte und Freiheiten der betroffenen Personen eine Stellungnahme der Aufsichtsstelle einzuholen ist. Die Gemeindeträgerwesen der Volksschulen können unter Umständen darauf verzichten, bspw. wenn weniger als 500 Personen von der Datenbearbeitung betroffen sind (vgl. Art. 8 Abs. 1 lit. a Datenschutzverordnung [DSV; BSG 152.040.1]).
47 von Thiessen/Volz (Fn. 33), 9.

Zusammenfassung

Die Nutzung von ChatGPT an Schulen ist nicht per se unzulässig. Die bestehenden Schulgesetze der jeweiligen Kantone können eine genügende gesetzliche Grundlage für den Einsatz von KI im Unterricht bieten. Der Beitrag hat sich mit der Datenbearbeitung durch OpenAI auseinandergesetzt. Er zeigt auf, welche Varianten sich für ein Schulprojekt eignen und bei welcher Nutzungsform aufgrund mangelnder Kontrollmöglichkeit durch die Schule ein Einsatz von ChatGPT im Unterricht unzulässig wäre. Empfohlen wird namentlich ein Einsatz über eine API-Schnittstelle, da hier die Schule mittels der technischen Einbindung der API in eigene Dienstleistungen Einfluss darauf nehmen kann, ob eine Übermittlung von IP-Adressen oder Accountinformationen und damit Identifizierungsmerkmalen der Schülerinnen und Schüler und der Lernenden erfolgt. Mittels klarer Vorgaben für die Nutzung von ChatGPT auf Schulstufe kann zudem verhindert werden, dass Schülerinnen und Schüler sowie Lernende sensible Informationen in den Chatbot eingeben. So oder anders werden Chatverläufe bei der API- und auch Enterprise-Variante nicht für das Training der zugrunde liegenden LLM verwendet. Microsoft 365 Copilot bietet künftig eine sichere Integration der LLM-Nutzung in den eigenen Azure Tenant, womit eine Datenübermittlung an die Server von OpenAI in die USA ausser Frage steht. In dieser letzten Variante erhält die Schule die faktische Kontrollmöglichkeit über die mittels Microsoft 365 Copilot verwendeten Schuldaten, indem Copilot in Übereinstimmung mit den bestehenden Berechtigungsstrukturen eingesetzt wird. Die Nutzung von ChatGPT im schulischen Kontext kann datenschutzkonform umgesetzt werden, vorausgesetzt, die Schulen schaffen die erforderlichen technischen und organisatorischen Voraussetzungen.

Résumé

L'utilisation de ChatGPT dans les écoles n'est pas interdite en soi. Les lois scolaires existant dans les différents cantons peuvent offrir une base légale suffisante pour l'utilisation de l'IA dans l'enseignement. Cet article aborde la thématique du traitement des données par OpenAI. Il montre quelles variantes des modèles proposés sont adaptées à un projet scolaire et quelles utilisations ne seraient pas autorisées en raison du manque de possibilités de contrôle par l'école. L'article recommande notamment une utilisation via une interface API, car l'école peut alors, grâce à l'intégration technique de l'API dans ses propres services, permettre ou empêcher la transmission des adresses IP ou des informations de compte, caractéristiques qui permettraient l'identification des élèves et des apprenants. Des directives claires pour l'utilisation de ChatGPT au niveau scolaire permettent en outre d'éviter que les élèves et les apprenants n'entrent des informations sensibles dans le chatbot. Quoi qu'il en soit, dans les variantes API et Enterprise, les historiques de chat ne sont pas utilisés pour entraîner le LLM (grand modèle de langage) sous-jacent. Microsoft 365 Copilot permettra à l'avenir d'intégrer de manière sûre l'utilisation du LLM dans son propre locataire Azure, ce qui élimine la question de la transmission de données aux serveurs d'OpenAI aux États-Unis. Dans cette dernière variante, l'école peut ainsi contrôler de fait les données scolaires utilisées au moyen de Microsoft 365 Copilot, car Copilot est utilisé en accord avec les structures d'autorisation existantes. L'utilisation de ChatGPT dans le contexte scolaire peut être mise en œuvre conformément à la protection des données, à condition que les écoles créent les conditions techniques et organisationnelles nécessaires.