



Leitfaden: Verwendung generativer KI-Systeme bei Maturitäts- und Projektarbeiten an Zürcher Mittelschulen

1 Reglementarischer Rahmen

Der Leitfaden ist gestützt auf den aktuellen reglementarischen Rahmen.¹

2 Generative KI-Systeme

Generative KI-Systeme sind digitale Werkzeuge, deren Technologien auf maschinellem Lernen beruhen. Sie erzeugen auf der Basis bestehender Daten Inhalte in verschiedenen Medienformaten («Output») oder bearbeiten eingegebene Daten (disziplinspezifisch auch «Prompts» genannt) weiter. Der erzeugte Output ist aufgrund der systemimmanenten und dynamischen Eigenschaften von KI nicht oder nur eingeschränkt reproduzierbar.

3 Ausgangslage

Generative Systeme, welche auf künstlicher Intelligenz (KI) basieren, gehören zur neuen Unterrichts- und Berufsrealität an Zürcher Schulen. Im Sinne ihres Bildungs- und Erziehungsauftrages integrieren sie KI-Anwendungen verantwortungsbewusst, konstruktiv und kritisch in ihr Lehren und Lernen.

KI-Tools können sowohl organisatorische Prozesse als auch Lehr- und Lernprozesse vereinfachen, ergänzen und weiterentwickeln. Damit bieten sie vielfältige Chancen, um in der Unterrichtspraxis den Fokus stärker auf kreative, problemlösungsorientierte und soziale Aktivitäten zu lenken und dadurch die Lehr- und Lernqualität positiv zu verändern. Der gezielte und angeleitete Einsatz von KI im Unterricht fördert zudem *digital literacy*, unterstützt das Erlernen von *Umgang mit Wissen* und *computational thinking*. Der praktische Umgang kann ein allgemeines Verständnis darüber fördern, wie KI funktioniert, er trainiert die Fähigkeit, mit KI-Robotern zusammen zu arbeiten, um sie zur Bewältigung von Aufgaben einzusetzen, und er gibt die Gelegenheit zu üben, zwischen menschlichen Kompetenzen und den Fähigkeiten der KI unterscheiden zu können.

KI-Systeme bringen Herausforderungen mit sich. Es stellt sich unter anderem für die Schulen die Frage, welche Voraussetzungen gegeben sein müssen, damit gerechte, gültige und verlässliche summative Leistungsnachweise generiert werden, welche auch tatsächlich diejenigen Kompetenzen überprüfen und bewerten, welche die Schülerinnen und Schüler im Laufe ihrer Schulzeit erwerben sollen.

4 Gegenstand und Zweck

Dieser Leitfaden gilt für summative Leistungsnachweise im Rahmen von Maturitäts- und Projektarbeiten an Zürcher Mittelschulen. Ein Verbot generativer KI-Systeme bei der Erstellung der Arbeiten ist weder praktikabel noch zielführend. Der Leitfaden mit Empfehlungscharakter soll Schulleitungen und Lehrpersonen Orientierung zu Fragen geben, welche in diesem Zusammenhang aufkommen. Er wurde in Zusammenarbeit mit dem Digital Learning Center SekII (DSC) erarbeitet. Stand heute gibt es allerdings noch keine allgemeingültigen, eindeutigen oder abschliessenden Antworten, da die Entwicklungsdynamik der verfügbaren Systeme hoch ist und eine noch fehlende generelle Routine mit KI-Systemen im Kontext von Leistungsnachweisen besteht. Ein längerfristiger Reflexions- und Transformationsprozess im Hinblick auf die Unterrichts- und Prüfungspraxis mit einhergehender Überarbeitung von Leitfaden und konkreten Handreichungen wird deshalb selbstredend nötig sein.

¹ Reglement der EDK über die Anerkennung von gymnasialen Maturitätsausweisen (MAR) 2018 Verordnung über die Anerkennung von gymnasialen Maturitätszeugnissen (Maturitäts-Anerkennungsverordnung, MAV) 2018; Reglemente für die Maturitätsprüfungen an den Gymnasien und des schweizerisch-italienischen Liceo artistico (Kunstgymnasium) des Kantons Zürich 1998; Promotionsreglemente des Kantons Zürich, ; Schulische Reglemente zur Nutzung der IT-Infrastruktur; Schulische Leitfäden zur Maturitätsarbeit



5 Projekt- und Maturitätsarbeiten als Leitungsnachweise

Projekt- und Maturitätsarbeiten sind Leistungsnachweise, die -im Unterschied zu Prüfungen- über einen längeren Zeitraum hinweg von Schülerinnen und Schülern weitgehend selbständig und unbeaufsichtigt erbracht werden. Die Schülerinnen und Schüler werden von der betreuenden Lehrperson im Entstehungsprozess der Arbeit begleitet. Die Einzelheiten zur Einreichung, Präsentation und Bewertung der Maturitätsarbeit sind in den Richtlinien der Schulen geregelt. Die Note für die Maturitätsarbeit wird gemäss eidgenössischen und kantonalen Vorgaben gesetzt und ist eine Maturitätsnote.

Es wird den Schulen empfohlen:

- die schulischen Leitfäden resp. Richtlinien zu überprüfen und zu überarbeiten hinsichtlich Rollen, Verantwortung bzw. Verbindlichkeiten der Beteiligten.
- die Summe der beiden Einzelnoten «Prozess» und «Präsentation» gesamthaft stärker zu gewichten als die Note des «Produkts».
- die «Präsentation» in ein «bewertetes Kolloquium» umzudefinieren und dabei den möglichen Risiken bei der Beurteilung -wie üblich bei mündlichen Prüfungen- zu begegnen.
- die Ansprüche an die sprachliche Qualität des Produkts zu überprüfen und gegebenenfalls zu erhöhen.

6 Verwendung generativer KI-Systeme bei Maturitätsarbeiten

Die Generierung von Texten oder gestalterischen Werken mit Hilfe von KI kann als ein Prozess der «Ko-Kreation» von Mensch und Maschine aufgefasst werden, wobei der Mensch die Verantwortung für den Text und das Werk behält. Die Schülerinnen und Schüler sind somit sowohl für die generierten Inhalte (den «Faktencheck») als auch für die Einhaltung von urheberrechtlichen Bestimmungen sowie die wissenschaftliche Integrität ihres Leistungsnachweises verantwortlich. Die erlaubten Hilfsmittel werden von der betreuenden Lehrperson festgelegt und den Schülerinnen und Schülern frühzeitig mündlich und schriftlich bekannt gegeben. Die Verwendung unerlaubter Hilfsmittel gilt als Unredlichkeit und kann (Disziplinar-)Massnahmen nach sich ziehen. Die Arbeit kann von der Schule auch zum Zweck der Feststellung von Plagiaten in anonymisierter Form erfasst werden.

Im Sinne der Eigenleistung bzw. wissenschaftlichen Integrität muss die Verwendung von KI bestmöglich transparent gemacht werden. Es muss für Dritte erkennbar sein, welche Teile in welchem Ausmass von einer KI generiert wurden. Grundsätzlich gilt daher eine Deklarationspflicht für alle KI-Systeme, welche die Qualität einer Arbeit auf inhaltlicher Ebene beeinflussen. Dies ist immer dann gegeben, wenn KI-Systeme über die Funktion als Gedankenanstoss (Inspirationsquelle), Sprachunterstützung (Formulierungshilfe, Rechtschreibhilfe) bzw. Sprachergänzung hinaus verwendet werden.

Eigenleistung, wissenschaftliche Integrität und der kreative und kritisch-reflektierte Einsatz von Hilfsmitteln für die Erstellung des Produktes kann während des Prozesses und bei der Präsentation mit anschliessendem Fachgespräch besonders gut beurteilt werden.

7 Eigenständigkeitserklärung

Mit der Einreichung der Maturitätsarbeit ist eine persönlich unterzeichnete Erklärung abzugeben. Dies ist auch für Projektarbeiten zu empfehlen, falls sie als Leistungsnachweise gelten. Die Eigenständigkeitserklärung bestätigt, dass sämtliche zitierte Quellen und andere verwendete Hilfsmittel korrekt gekennzeichnet sind und der Leistungsnachweis nur mit den deklarierten Hilfsmitteln erstellt wurde. Durch diese prinzipienbasierte und technologieneutrale Gestaltung der Eigenständigkeitserklärung ist die Verwendung generativer KI-Systeme eigentlich miteingefasst. Ein ergänzender spezifischer Passus zum Umgang mit KI-Systemen kann den Rahmen noch verdeutlichen.

8 Datenschutz und Verfügbarkeit generativer KI-Systeme

Bei der Nutzung von auf künstlicher Intelligenz basierenden Diensten wie ChatGPT müssen neben den pädagogischen und ethischen Aspekten auch die rechtlichen und die Informationssicherheit betreffenden Aspekte berücksichtigt werden. Obwohl die Nutzung aus pädagogischer Sicht sinnvoll und



Digital Learning Hub Sek II, Riesbachstrasse 11, 8090 Zürich
vernetzt@dlh.zh.ch

nützlich sein kann, bestehen auch Risiken, die jedem bei der Nutzung bewusst sein sollten. Die nachfolgenden ersten Informationen und Empfehlungen aus Sicht der Informationssicherheit dienen als Hilfestellung, um einen angemessenen Umgang mit diesen Technologien zu ermöglichen.

Aktuell fehlt im Kanton Zürich aber auch europaweit die rechtliche Grundlage, worauf sich die Nutzung von ChatGPT stützen könnte. Rechtsexpertinnen und -experten des Kantons Zürich führen eine Abklärung zum Thema durch. Die Ergebnisse laufender Rechtsgutachten können die Gültigkeit der nachstehenden Informationen beeinflussen. Sobald neue Erkenntnisse vorhanden sind, werde diese u.a. über den DLH kommuniziert.

8.1 Rechtliche Rahmenbedingungen kennen

Dienste wie ChatGPT erfordern immer die Eingabe von Informationen (Daten); diese werden an den Anbieter und ggf. an Dritte übermittelt. Im Fall von Chat GPT von OpenAI werden Name, Vorname, E-Mailadresse und Telefonnummer bei der Registrierung verlangt. Diese Informationen werden in den USA gehostet. Der Anbieter und ggf. Dritte können mit den erhaltenen Daten z.B. Analysen durchführen und allgemeine Nutzerstatistiken an Dritte weitergeben. Die datenschutzrechtliche Beurteilung hängt ausserdem von den Daten ab, die eingegeben werden. Dies gilt unabhängig von den bei der Anmeldung übermittelten Informationen. Es muss also sichergestellt werden, dass keine Personendaten oder besondere Personendaten verarbeitet werden.

- Aus Datenschutzperspektive ist aktuell eine Nutzung von Diensten wie ChatGPT kritisch, wenn sie auf einem persönlichen Nutzeraccount basiert, d.h. wenn beim Login persönliche Informationen zur Eingabe nötig waren. Die Integration von ChatGPT in der Suchmaschine Bing von Microsoft wird in den nächsten Wochen/Monaten schrittweise implementiert. Die Nutzung mit Schulaccounts ist (Stand heute, 13.6.2023) noch nicht möglich.
- Die technischen Gegebenheiten sowie die zu erwartenden Konsequenzen in Sachen Datenschutz werden derzeit geprüft.
- Wenn die Nutzung nicht über eine datenschutzfreundliche Lösung erfolgt (siehe "Alternative"), kann sie deshalb aktuell nicht im Rahmen des Unterrichts verlangt werden.
- Die Risiken, welche mit der Nutzung von Diensten wie ChatGPT über einen persönlichen Nutzeraccount verbunden sind, tragen die Nutzerinnen und Nutzer.

8.2 Über Risiken und Mehrwert aufklären

Der Einsatz von KI-basierten Systemen ist bereits eine Realität. Um die Nutzerinnen und Nutzer so gut wie möglich zu schützen, ist es wichtig, sie für die folgenden Punkte zu sensibilisieren:

- Anwendungsfälle sollten vor Nutzung klar definiert werden (Bsp. Wird ChatGPT im Rahmen einer Deutschlektion im Klassenverband verwendet oder wird ChatGPT von einzelnen Schülerinnen und Schülern im Rahmen der Maturitätsarbeit verwendet?)
- Da Informationen übertragen werden, sollte das Risiko vorab von Lehrpersonen bewertet werden und die Risikofolgen sollten den Nutzerinnen und Nutzern bewusst sein
- Ein Verständnis für eine bewusste und sachgerechte Nutzung sollte von den Lehrpersonen geschaffen werden
- Eine grundsätzliche Auseinandersetzung über Nutzen, Chancen und Gefahren sollte in den Schulen erfolgen.

8.3 Alternative

Das Digital Service Center SekII prüft derzeit unterschiedliche Lösungen, welche eine Nutzung von ChatGPT und anderer KI gestützter Systeme datenschutzkonform ermöglichen könnten. Diese Lösungen basieren auf der Verwendung von KI-Technologien (z. B. von OpenAI und Microsoft) in einem angepassten technischen Rahmen, um den Schutz personenbezogener Daten zu gewährleisten. Dies hat Auswirkungen auf die konkreten Nutzungsmodalitäten der Tools. Bis jetzt wurden zwei



Digital Learning Hub Sek II, Riesbachstrasse 11, 8090 Zürich
vernetzt@dlh.zh.ch

Lösungen geprüft und als datenschutz-konform evaluiert. Andere Lösungen werden weiterhin evaluiert.

- <https://gptschule.de/privacy>: GPTSchule.de verwendet u.a. ChatGPT von Open AI, aber Open AI erhält keine personenbezogene Daten. Dennoch muss erwähnt werden, dass keine Personendaten in der Konversation mit ChatGPT genannt werden dürfen.
- <https://tools.fobizz.com/>: bietet auch Möglichkeiten für eine datenschutzkonforme Nutzung im Unterricht.

Der nächste Schritt wird sein, eine Bewertung dieser beiden Lösungen aus pädagogischer Sicht vorzunehmen, um festzustellen, ob sie den Bedürfnissen und Erwartungen entsprechen.

9 Bewertung von Leistungsnachweisen mit generativen KI-Systemen

Verwenden Lehrpersonen generative KI-Systeme zur Bewertung von Vertiefungs- und Abschlussarbeiten, müssen sie neben dem Datenschutz auch urheberrechtliche Bestimmungen einhalten. So ist die Eingabe von schöpferischen Arbeiten von Lernenden in ein KI-System nur erlaubt, wenn es die eingegebenen Daten nicht als Trainingsdaten weiterverwendet oder anderweitig nutzt. D.h. für ChatGPT beispielsweise dürfen *keine* schöpferischen Daten von Lernenden eingegeben werden. Zudem dürfen KI-Systeme bei der Bewertung nur als Hilfsmittel eingesetzt werden. Deren Output muss von der Lehrperson und der Expertin/dem Experten gemäss geltendem Bewertungsraster (unabhängig davon, ob KI-Systeme verwendet werden oder nicht) überprüft werden. Eine Anwendung unterschiedlicher Bewertungsraster entlang der Unterscheidung, ob KI-Systeme verwendet wurden oder nicht, ist nicht zulässig.

Der vorliegende Leitfaden lehnt sich eng an die von der Zürcher Hochschule für angewandte Wissenschaften ZHAW am 1. April 2023 veröffentlichte «Richtlinie KI bei Leistungsnachweisen» ([Richtlinie KI bei Leistungsnachweisen \(zhaw.ch\)](#)). Wir bedanken uns erneut bei der ZHAW für diese Möglichkeit.