



MASTERARBEIT | MASTER'S THESIS

Titel | Title

Täuschend echt? Deepfakes im Diskurs von Forschung, jugendlicher Perspektive und
Unterrichtspraxis

verfasst von | submitted by

Markus Gruber BEd

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Education (MEd)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt |
Degree programme code as it appears on the
student record sheet:

UA 199 506 514 02

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student record
sheet:

Masterstudium Lehramt Sek (AB) Unterrichtsfach Deutsch
Unterrichtsfach Digitale Grundbildung und Informatik

Betreut von | Supervisor:

Univ.-Prof. Dipl.-Ing. Dr. Renate Motschnig

Abstract

Die vorliegende Arbeit beschäftigt sich mit Deepfakes und deren gesellschaftlichen Auswirkungen. Es handelt sich dabei um KI-generierte und manipulierte Medien. Im Wesentlichen wird das Thema aus drei verschiedenen Blickwinkeln betrachtet. Es wird wissenschaftliche Literatur recherchiert, es werden jugendliche Schüler*innen der Oberstufe einer allgemeinbildenden höheren Schule zu Deepfakes befragt und es wird ein Unterrichtskonzept einer Doppelstunde zu KI-generierten und manipulierten Medien geplant, mehrmals durchgeführt und anhand von Design-based Research nach jedem Durchgang überarbeitet.

Erwartet wird, dass Jugendliche sehr viel Zeit auf sozialen Medien verbringen und dadurch auch häufig mit Deepfakes konfrontiert werden, jedoch deren Gefahren und deren Möglichkeiten häufig nicht kennen. Es wird auch erwartet, dass die Oberstufenschüler*innen es häufig auch erkennen, wenn ein Bild oder Video KI-generiert oder manipuliert ist. Bezüglich der Unterrichtsstunden wird erwartet, dass das am Ende der Einheiten geplante Educational Live Action Role Playing (EduLARP) den Schüler*innen eine abwechslungsreiche, aber auch lehrreiche und ansprechende Lernerfahrung bietet.

Mit dieser Arbeit wird ein Beitrag zur Medienpädagogik geleistet und die geplanten Unterrichtsstunden weisen einen Lehrplanbezug in der 5. Klasse der AHS im Abschnitt Informatik, Mensch und Gesellschaft auf. Die Kenntnisse der Auswirkungen der Informatik auf die Einzelnen und die Gesellschaft und das Abwägen von Vor- und Nachteilen werden im Lehrplan genannt und diese Unterrichtsstunden fördern diese Bereiche bezüglich künstlicher Intelligenz. Außerdem Sie können sich inspirieren lassen von den geplanten Unterrichtsstunden und können erkennen, dass auch in diesem technischen Unterrichtsfach ein EduLARP möglich und hilfreich ist, um den Schüler*innen abwechslungsreiche Lernerfahrungen zu bieten.

Inhaltsverzeichnis

1	EINLEITUNG.....	6
1.1	Relevanz des Themas inklusive eigener Motivation	7
1.2	Zielsetzung und Forschungsfragen	10
2	THEORETISCHER HINTERGRUND	11
2.1	Begriffsklärungen	12
2.2	Von der Bildbearbeitung zu Deepfakes.....	14
2.3	Funktionsweise – Deep Learning	15
2.4	Medienkompetenz	16
2.5	Spielerisches Lernen in der Schule – EduLARP	18
3	FORSCHUNGSSTAND ZU GEFAHREN UND CHANCEN	19
3.1	Probleme und Risiken.....	19
3.2	Chancen und Potentiale.....	24
4	METHODIK	25
4.1	Beschreibung des Fragebogens.....	25
4.2	Datenanalyse.....	26
4.3	Design-Based Research	28
4.4	Zielgruppe	29
5	PERSPEKTIVEN DER JUGENDLICHEN.....	30
5.1	Auswertung der Fragebögen.....	30
5.2	Vergleich mit aktuellem Forschungsstand	36

6	DEEPFAKES IM UNTERRICHT	37
6.1	Planung.....	38
6.2	Rahmenbedingungen der Unterrichtsstunden.....	39
6.3	Durchgang 1.....	40
6.4	Durchgang 2.....	47
6.5	Durchgang 3.....	52
7	DISKUSSION.....	56
8	FAZIT UND AUSBLICK	57
9	LITERATURVERZEICHNIS	59
10	HILFSMITTEL UND ETHISCHE ASPEKTE.....	61
11	ANHANG.....	62

Abbildungsverzeichnis

Abbildung 2 Design-Based Research Zyklen (Gess et al. 2014: S.12)	29
Abbildung 3 Befragte Schüler*innen (m/w/d).....	31
Abbildung 4 Diagramm Tägliche Zeit auf sozialen Medien.....	32
Abbildung 5 Anteil an Deepfakes in sozialen Medien	36
Abbildung 6 Charakterkarte Influencer*in	43
Abbildung 7 Charakterkarte Besorgte*r Bürger*in.....	43
Abbildung 8 Charakterkarte Unwissende*r Jugendliche*r.....	44
Abbildung 9 Charakterkarte Geschäftsführer*in	44
Abbildung 10 Charakterkarte Bürgermeister*in	45
Abbildung 11 Frankfurt Dreieck (Brinda, T. et al. 2019: 27)	53

Tabellenverzeichnis

Tabelle 1 Kategorienbildung Beispiel.....	27
Tabelle 2 Unterrichtsmatrix erster Durchgang.....	41
Tabelle 3 Änderung erster Durchgang	46
Tabelle 4 Unterrichtsmatrix zweiter Durchgang.....	51
Tabelle 5 Veränderungen dritter Durchgang.....	54

1 EINLEITUNG

Künstliche Intelligenz hat in den letzten Jahren für sehr viele Schlagzeilen gesorgt. Innerhalb von kurzer Zeit kommunizieren sehr viele Menschen weltweit mit Chatbots wie ChatGPT. Schlagartig gibt es neue Möglichkeiten, vergleichbar mit der Einführung des Internets oder des Buchdrucks. Allerdings gibt es auch eine Kehrseite, da damit auch Risiken verbunden sind. Diese Arbeit setzt sich mit einer ganz bestimmten Form der künstlichen Intelligenz auseinander, Deepfakes werden genauer unter die Lupe genommen. KI generierte oder manipulierte Medien findet man mittlerweile sehr häufig und das sorgt leider auch für große Probleme.

Die Forschung zu Deepfakes ist von großer Bedeutung, da diese Medien sowohl enormes kreatives Potenzial als auch erhebliche Risiken mit sich bringen. An dieser Stelle kann erwähnt werden, dass auch das Internet Gefahren mit sich bringt und dennoch ist es seit langer Zeit eine unverzichtbare Technologie. Derzeit sind viele Menschen noch sehr skeptisch gegenüber künstlicher Intelligenz eingestellt, aber viele erkennen bereits die neuen Möglichkeiten. Einerseits ermöglicht sie Fortschritte in Bereichen wie Filmproduktion, Bildung und Barrierefreiheit, etwa durch realistische Simulationen oder personalisierte Inhalte. Andererseits birgt sie die Gefahr von Desinformation, Identitätsmissbrauch und Manipulation der öffentlichen Meinung, insbesondere in sozialen Medien und politischen Kontexten. Viele Jugendliche verbringen sehr viel Zeit auf Social Media Plattformen. Aus diesem Grund wird diese Altersgruppe in dieser Arbeit auch beforscht. Durch gezielte Forschung können Methoden entwickelt werden, um KI-generierte und manipulierte Medien zuverlässig zu erkennen, ihre Verbreitung einzudämmen und ethische sowie rechtliche Rahmenbedingungen zu schaffen.

Somit werden junge Menschen besonders häufig mit Deepfakes konfrontiert, weil sie einen großen Teil ihrer Freizeit in digitalen Räumen wie sozialen Netzwerken, Video-Plattformen und Messaging-Apps verbringen, in denen solche Inhalte schnell verbreitet werden. Zudem sind sie oft früh mit neuen Technologien vertraut und Teil von Online-Communities, in denen Trends wie KI-generierte Bilder oder Videos rasch aufgegriffen und geteilt werden. Gleichzeitig muss aber gesagt werden, dass es vielen noch an Erfahrung im kritischen Umgang mit Medieninhalten fehlt, wodurch es für sie schwieriger sein kann, manipulierte Inhalte sofort zu erkennen. Aus diesem Grund

werden für diese Arbeit Jugendliche aus der Sekundarstufe 2 zu Deepfakes befragt und es werden außerdem Unterrichtsstunden zu diesem Thema gehalten und analysiert.

Das nachfolgende Kapitel soll die aktuelle Relevanz des Themas aufzeigen und die eigene Motivation zur Forschung in diesem Bereich erläutern.

1.1 Relevanz des Themas inklusive eigener Motivation

Die rasante Entwicklung künstlicher Intelligenz hat in den letzten Jahren eine neue Qualität der Medienmanipulation hervorgebracht. Manche Medien wurden schon länger manipuliert und bearbeitet, jedoch braucht man dafür jetzt nicht zwangsweise Fachkenntnisse und Programme wie PhotoShop, sondern sogenannte Deepfakes können ganz einfach mithilfe von künstlicher Intelligenz erstellt werden. Dabei handelt es sich um täuschend echte, mithilfe von KI erzeugte oder veränderte Audio-, Bild- oder Videoinhalte (Kleemann, 2023). Diese Technologie bietet zwar kreative und auch nützliche Anwendungsmöglichkeiten, birgt jedoch zugleich erhebliche Risiken für Individuen, Institutionen und die Gesellschaft insgesamt. Die App Grok steht beispielsweise häufig unter Kritik, da auch erotische Inhalte damit erstellt wurden. Viele Zeitungsartikel berichteten in den letzten Monaten darüber. Aus diesem Grund ist es besonders wichtig, intensiv zu diesem Thema zu forschen und Deepfakes auch im schulischen Unterricht umfassend zu behandeln.

Ein zentraler Grund für die Forschung im Bereich Deepfakes liegt in der wachsenden Gefahr von Betrug und Manipulation. Moderne KI-Modelle sind inzwischen in der Lage, Stimmen und Gesichter so realistisch zu imitieren, dass selbst aufmerksame Menschen Schwierigkeiten haben, echte von gefälschten Inhalten zu unterscheiden. Davon wird häufig in den Medien berichtet und deshalb ist die Forschung zu den Gefahren von künstlicher Intelligenz von besonders großer Bedeutung.

Kriminelle Menschen nutzen diese Möglichkeiten gezielt aus, etwa durch gefälschte Anrufe oder Videos, um Vertrauen zu erschleichen und Menschen zu finanziellen Handlungen zu bewegen. Der Enkeltrick kann an dieser Stelle genannt werden. Solche Anrufe gab es auch zuvor bereits, die einige Personen dazu brachten, hohe Geldsummen an fremde Menschen zu überweisen. Durch das Nachahmen von

bekannten Stimmen werden diese Anrufe nun viel authentischer und noch schwieriger, zu erkennen (Mößle et al., o.D.,).

Die Weiterentwicklung solcher Technologien schreitet sehr schnell voran, weshalb auch die Forschung Schritt halten muss. Nur durch kontinuierliche wissenschaftliche Arbeit können wirksame Methoden zur Erkennung und Abwehr solcher Fälschungen entwickelt werden. Ohne diese Forschung würde die Gesellschaft den technischen Fortschritten der Betrüger*innen schutzlos gegenüberstehen. Besonders gute Deepfakes können sonst vermutlich nur noch durch forensische Prüfung oder durch genaue Recherche der Herkunft eines Mediums identifiziert werden. Das wäre sehr bedenklich und ist ein wesentlicher Motivationsgrund zur Auseinandersetzung mit diesem Thema.

Darüber hinaus hat das Thema eine große gesellschaftliche und politische Bedeutung. Deepfakes können gezielt eingesetzt werden, um falsche Informationen, sogenannte Fake News, zu verbreiten oder öffentliche Meinungen zu beeinflussen. Davon wird immer wieder in den Medien berichtet. Besonders im Kontext von Wahlen oder politischen Konflikten besteht die Gefahr, dass manipulierte Bilder, Videos oder Aussagen das Vertrauen in demokratische Prozesse untergraben und das sorgt für ein großes Problem. Wenn Menschen nicht mehr sicher sein können, ob eine Aussage tatsächlich von einer bestimmten Person stammt, wird die Grundlage für eine informierte Meinungsbildung geschwächt. Forschung ist hier notwendig, um Mechanismen zu verstehen, wie sich solche Inhalte verbreiten, wie sie wirken und wie man ihnen effektiv begegnen kann. Gleichzeitig gibt es aber auch ethische Aspekte, die bei diesem sehr komplexen Thema genau bedacht werden müssen.

Ein weiterer wichtiger Aspekt ist der Schutz von Individuen. Deepfakes können enorme Schäden anrichten, viele Menschen denken dabei schnell an Rufschädigung, psychischer Belastung und somit vielen sozialen Konsequenzen. Besonders problematisch ist, dass solche Inhalte sich im Internet sehr schnell verbreiten und oft schwer vollständig zu entfernen sind. Jemand schickt ein Medium einer anderen Person, die es dann in einen Gruppenchat weiterleitet und schon verbreitet es sich sehr schnell. Man kann dann auch nur sehr schwer herausfinden, wer von den Leuten den Inhalt tatsächlich glaubt und wer es als Fake enttarnt. Forschung kann dazu beitragen, bessere Schutzmechanismen zu entwickeln und Betroffenen Unterstützung

zu bieten. Gleichzeitig können technische Lösungen entstehen, um Deepfakes leichter erkennen zu können.

Neben der Forschung spielt auch die Bildung eine entscheidende Rolle. Der schulische Unterricht ist ein zentraler Ort, um junge Menschen auf die Herausforderungen der digitalen Welt vorzubereiten. Deepfakes sind ein aktuelles Beispiel dafür, wie wichtig Medienkompetenz geworden ist. Für Schüler*innen ist es relevant zu lernen, Informationen kritisch zu hinterfragen und nicht alles, was sie im Internet sehen oder hören, ungeprüft zu glauben. Es ist sehr positiv hervorzuheben, dass in der Oberstufe der allgemeinbildenden höheren Schulen in Österreich zukünftig mehr Informatik und künstliche Intelligenz unterrichtet wird, da diese Thematiken sehr aktuell und wichtig sind. Somit können die Jugendlichen ein besseres Verständnis dafür entwickeln, wie solche Fälschungen entstehen und auch welche Absichten dahinterstecken.

Darüber hinaus fördert die Auseinandersetzung mit Deepfakes wichtige Kompetenzen wie beispielsweise das kritische Denken. Es können immer glaubwürdigere Fake News im Internet bzw. auf sozialen Netzwerken kursieren, sodass kritisches Denken eine sehr wichtige Kenntnis ist. Lernende können beispielsweise üben, Hinweise auf manipulierte Inhalte zu erkennen oder verschiedene Quellen miteinander zu vergleichen. Ersteres wird in den geplanten Unterrichtsstunden im Rahmen dieser Arbeit geübt. Wer versteht, wie leicht Inhalte manipuliert werden können, wird dann auch eher vorsichtig sein, bevor sie*er etwas weiterverbreitet oder selbst erstellt. Auch das ist ein wesentliches Ziel dieser geplanten Doppelstunde.

Ein weiterer Vorteil der Behandlung dieser Thematik im Unterricht ist die Sensibilisierung für ethische Fragestellungen. Deepfakes werfen grundlegende Fragen auf: Was ist Wahrheit? Wem kann man vertrauen? Welche Grenzen sollte es bei der Nutzung von Technologie geben? Diese Fragen sind nicht nur technisch, sondern auch philosophisch und gesellschaftlich relevant. Der Unterricht bietet einen geeigneten Rahmen, um solche Thematiken mit der Lehrperson und mit den Klassenkolleg*innen zu diskutieren und unterschiedliche Perspektiven kennenzulernen. Dadurch wird nicht nur Wissen vermittelt, sondern auch die Fähigkeit gestärkt, sich eine reflektierte Meinung zu bilden. In den geplanten Unterrichtsstunden

für diese Arbeit müssen Argumente von vorgegebenen Sichtweisen aus überlegt werden.

Zudem kann die Beschäftigung mit Deepfakes das Interesse an Technologie und Wissenschaft fördern. Wenn die Schüler*innen verstehen, wie künstliche Intelligenz funktioniert und welche Möglichkeiten sie bietet, könnte dies motivierend wirken und möglicherweise sogar berufliche Interessen wecken. Das kann zum Beispiel dazu führen, dass sich die*der ein*e oder andere Schüler*in nach der Oberstufe für ein technisches Studium entscheidet. Gleichzeitig lernen sie, dass technischer Fortschritt immer auch kritisch hinterfragt werden muss und nicht immer nur Vorteile mit sich bringt.

Aus den genannten Gründen ist die Forschung zu dieser Thematik von hoher Relevanz. Diese Arbeit setzt sich mit wichtigen Bereichen von KI-generierten und manipulierten Medien auseinander. Die genauere Beschreibung folgt im nächsten Kapitel, welches das Ziel dieser Arbeit aufzeigt und außerdem die genauen Forschungsfragen beinhaltet.

1.2 Zielsetzung und Forschungsfragen

Diese Arbeit verfolgt das Ziel, Literatur zu Deepfakes genau zu untersuchen. Ein besonderer Fokus liegt dabei auf den Risiken dieser Technologie. Es gibt zahlreiche Chancen und Möglichkeiten und es kann bei korrekter und sinnvoller Anwendung von künstlicher Intelligenz enorm viel Zeit gespart werden, jedoch bringen Deepfakes auch Probleme mit sich. Aus diesem Grund stellt sich an diese Arbeit folgende erste Forschungsfrage:

- Welche Risiken von Deepfakes werden in der Literatur beschrieben?

Somit kann aufgezeigt werden, welche Erkenntnisse bereits herausgefunden wurden. Sehr interessant sind allerdings auch die Gedanken von jungen Menschen zu diesem Thema und deshalb werden Jugendliche über Deepfakes befragt. Daraus ergibt sich folgende zweite Forschungsfrage.

- Welche Sichtweisen auf Deepfakes und deren Risiken zeigen sich aus der Perspektive von Oberstufenschüler*innen einer AHS?

Dadurch können wissenschaftliche Erkenntnisse mit Meinungen und Gedanken von Jugendlichen verglichen und Gemeinsamkeiten und wesentliche Unterschiede aufgezeigt werden. Mithilfe von Literaturrecherche und der Beantwortung eines anonymen Fragebogens von einigen Oberstufenschüler*innen können diese beiden Forschungsfragen beantwortet werden. Eine genauere Beschreibung der Vorgehensweise erfolgt im Kapitel 4 Methodik.

Außerdem werden auch Unterrichtsstunden zum Thema Deepfakes geplant und auch durchgeführt. Das Unterrichtskonzept wird dabei in vier verschiedenen Klassen ausprobiert und laufend überarbeitet und verbessert. Das symbolisiert auch eine gute Vorgehensweise als Lehrperson, wenn man eine Unterrichtsplanung öfters verwendet. Aus diesem Grund stellt sich folgende letzte und sehr wichtige Forschungsfrage für diese Arbeit.

- Inwiefern ist das im Rahmen der Masterarbeit erarbeitete spezifische Unterrichtskonzept für die Zielgruppe der Schüler*innen einer AHS-Oberstufe geeignet, zur Verbesserung von deren Medienkompetenz hinsichtlich Deepfakes beizutragen?

Die Beantwortung der drei Forschungsfragen erfolgt somit durch die Kombination verschiedener methodischer Zugänge. Einerseits bildet eine umfassende wissenschaftliche Literaturrecherche die theoretische Grundlage der Untersuchung, indem aktuelle Erkenntnisse zu Deepfakes, deren gesellschaftlichen Auswirkungen sowie den damit verbundenen Chancen und Risiken analysiert werden. Vor allem werden aber die Gefahren behandelt. Andererseits werden junge Personen im Rahmen einer empirischen Erhebung befragt, um ihre Kenntnisse und Einstellungen im Umgang mit Deepfakes zu erfassen. Ergänzend dazu werden praktische Erfahrungen aus dem schulischen Kontext einbezogen. Die Verbindung von theoretischer Literaturrecherche, empirischer Forschung und schulischer Praxis ermöglicht eine umfassende Betrachtung dieser wichtigen und derzeit sehr relevanten Thematik.

2 THEORETISCHER HINTERGRUND

Die rasante Entwicklung von künstlicher Intelligenz hat dazu geführt, dass KI-generierte oder manipulierte Medien immer realistischer und schwerer zu erkennen

werden. Forschung im Bereich der Deepfakes ist deshalb besonders wichtig, um diesen Herausforderungen gezielt begegnen zu können. Einerseits geht es darum, Technologien zu entwickeln, die KI-generierte und manipulierte Medien zuverlässig erkennen können, damit dem damit verbundenen Betrug und die entstehenden Desinformationen verringert werden können. Andererseits hilft die Forschung dabei, ein besseres Verständnis für die Funktionsweise und die Auswirkungen solcher Manipulationen zu gewinnen. Es gibt verschiedene Begriffe, die für diese Arbeit von besonderer Bedeutung sind. Aus diesem Grund sollen diese Begrifflichkeiten im folgenden Kapitel geklärt werden.

2.1 Begriffsklärungen

Sehr wichtige Begriffe für diese Arbeit sind *künstliche Intelligenz*, *Deepfakes*, *EduLARP* und *Design-Based-Research*.

Es ist nicht ganz einfach, künstliche Intelligenz zu definieren, da es keine einheitliche Meinung darüber gibt, was eigentlich *Intelligenz* bedeutet. Außerdem ist die Intelligenz einer Maschine nicht ganz vergleichbar mit der von Menschen (Kaplan et al., 2017, S.15). Dennoch existieren verschiedene Definitionsansätze, von denen der folgende eine geeignete Annäherung an den Begriff der künstlichen Intelligenz bietet: „John McCarthy, einer der Gründerväter dieses Wissensgebiets, hat 1955 über den zugehörigen Prozess gesagt, dass es darum geht, ‘Maschinen zu entwickeln, die sich verhalten, als verfügten sie über menschliche Intelligenz‘ “ (Kaplan et al., 2017, S.15).

Das Ziel besteht laut dieser Definition darin, Computerprogramme und technische Systeme zu entwickeln, die Aufgaben ausführen können, für die normalerweise menschliche Intelligenz erforderlich ist. Dazu gehören beispielsweise das Verstehen von Sprache, das Erkennen von Bildern, das Treffen von Entscheidungen oder das Lösen von Problemen. Das Erstaunliche daran ist, dass diese Maschinen dabei aber keine echte menschliche Intelligenz beinhalten, verhalten sich jedoch in bestimmten Situationen so, als könnten sie denken, lernen oder eigenständig handeln und das bringt viele Potentiale aber auch Probleme mit sich.

Auch die in dieser Arbeit behandelten Deepfakes nutzen künstliche Intelligenz, um Medien zu manipulieren oder generieren. Unter Deepfakes versteht man

Medieninhalte, die künstlich mittels KI erstellt oder verändert wurden. Das betrifft Bilder und Videos, aber auch reine Audioaufzeichnungen. Es wird zunehmend schwieriger, gefälschte und reale Medien zu unterscheiden (Kleemann, 2023). *Deepfakes* sind Medieninhalte, die zwar täuschend echt wirken, deren Inhalte aber in Wirklichkeit nie in der Form passiert sind. Diese Fälschungen werden laufend immer besser und sind deshalb teilweise schon sehr schwierig von realen Medien zu unterscheiden (Kleemann, 2023). Eine genauere Beschreibung dieser Technik erfolgt im Kapitel 2.2 Von der Bildbearbeitung zu Deepfakes.

Ein weiterer wichtiger Begriff für diese Arbeit ist das EduLARP. Diese Abkürzung steht für Educational Live Action Role Playing. Dabei begeben sich Schüler*innen in eine Rolle, die davor festgelegt wurde. Diese Methode kann sowohl in der Grundschule als auch in den Sekundarstufen eins und zwei angewendet werden (Zuzana et al., 2024, S.1). Für diese Arbeit wurden fünf unterschiedliche Rollen mit vorgegebenen Zielen erstellt und die Schüler*innen müssen an einer Gruppendiskussion über KI-generierte und manipulierte Medien teilnehmen. Die eigenen Meinungen und Gedanken sollen irrelevant sein, damit sich die Jugendlichen auf ihre zugewiesene Rolle konzentrieren können.

Zum Abschluss dieses Kapitels muss noch der Begriff Design-Based Research geklärt werden. In der Literatur werden auch die Synonyme „Design-Experiment“ und „Design Research“ verwendet. Dabei sollen Lernumgebungen gestaltet und parallel dazu auch Lerntheorien ausprobiert und weiterentwickelt werden (Beege, 2023, S.89). Im konkreten Fall dieser Arbeit werden zwei Unterrichtsstunden zum Thema Deepfakes geplant, durchgeführt und laufend verbessert. Dabei bleibt die Zielgruppe gleich und es handelt sich immer um Oberstufenschüler*innen einer allgemeinbildenden Schule. Nach jedem Durchgang wird das Feedback eingearbeitet und die verbesserten Unterrichtssequenzen werden in einer anderen Klasse durchgeführt.

Damit sind die für diese Arbeit zentralen Begriffe definiert. Im folgenden Kapitel erfolgt eine detaillierte Auseinandersetzung mit dem Phänomen der Deepfakes und ein Blick in die Geschichte der Bildbearbeitung.

2.2 Von der Bildbearbeitung zu Deepfakes

Deepfakes sind grundsätzlich fälschlicherweise echt wirkende Medien, wie bereits im vorigen Kapitel erläutert wurde: „[...] täuschend echt wirkende, künstlich erstellte oder veränderte Foto-, Video- oder Sprach-aufzeichnungen“ (Kleemann, 2023, S.1). Ein großes Problem ist dabei, dass es schon zahlreiche Möglichkeiten gibt, um solche Fälschungen zu erstellen und diese Tools von allen genutzt werden können:

Heute sind einfach nutzbare KI-Tools, mit denen sich hochwertige Fälschungen produzieren lassen, frei zugänglich. Deepfakes sind ein Massenprodukt und keine Seltenheit mehr. Dazu haben in den vergangenen Jahren vor allem drei Entwicklungen beigetragen: die kontinuierliche Verbesserung der KI, die stetige Zunahme verfügbarer Rechenleistung und der offene Zugang zu immer mehr Daten, mit denen die KI trainiert werden kann. Ein Ende und die volle Tragweite dieses Prozesses sind aktuell kaum absehbar (Kleemann, 2023, S.1-2).

An dieser Stelle muss allerdings gesagt werden, dass Medien schon viele Jahrzehnte davor manipuliert wurden. Die Manipulation von Bildern ist keineswegs ein neues Phänomen, sondern hat eine lange historische Tradition. Bereits im 20. Jahrhundert nutzten politische Machthaber wie Joseph Stalin, Mao Zedong, Adolf Hitler, Benito Mussolini, Fidel Castro und Leonid Breschnew gezielt bearbeitete Fotografien, um politische Botschaften zu verbreiten oder historische Ereignisse in ihrem Sinne darzustellen. Durch solche Eingriffe sollte die Wahrnehmung der Bevölkerung beeinflusst und die Geschichte nachträglich verändert oder neu interpretiert werden, sodass die politischen Machthaber, die zuvor genannt wurden, davon profitierten (Farid, 2020, S.9).

Mit dem Aufkommen der digitalen Bildbearbeitung Anfang der 1990er Jahre änderte sich die Situation grundlegend. Leistungsfähige Computer, digitale Kameras und benutzerfreundliche Bildbearbeitungsprogramme machten es zunehmend einfacher und kostengünstiger, Fotos zu verändern. Das war damals bereits ein wesentlicher Fortschritt, jedoch unterscheiden sich diese Medien von Deepfakes, indem sie ohne künstlicher Intelligenz bearbeitet wurden (Farid, 2020, S.9). Somit war das meistens aufwändiger und langwieriger.

Im Jahr 2026 werden somit keine besonderen Fachkenntnisse mehr benötigt, um die*den Bürgermeister*in einer großen Stadt beispielsweise in einem Video zurücktreten zu lassen, obwohl sie*er das nie getan hat, da es noch leichter wurde,

solche Medien zu erstellen. Wie diese Technologie funktioniert, wird im folgenden Kapitel genauer erläutert.

2.3 Funktionsweise – Deep Learning

Die Entwicklung von Verfahren der künstlichen Intelligenz hat in den vergangenen Jahren erhebliche Fortschritte bei der Verarbeitung von Sprache, Texten und Bildern ermöglicht. Eine zentrale Rolle spielt dabei das sogenannte Deep Learning. Darunter versteht man ein Teilgebiet des maschinellen Lernens (Folkers 2019, S. 3). Ziel dieser Methode ist es, dass Computersysteme komplexe Muster und Zusammenhänge in großen Datenmengen selbstständig erkennen und für bestimmte Aufgaben nutzbar machen.

Im Gegensatz zu klassischen Programmieransätzen werden die Regeln zur Problemlösung nicht explizit vorgegeben. Stattdessen lernt das System anhand von Beispieldaten, welche Merkmale für die jeweilige Aufgabe relevant sind. Eine Mustererkennung ist dafür sehr wichtig. Dazu kommen mehrschichtige neuronale Netze zum Einsatz, deren Struktur sich an der Funktionsweise des menschlichen Gehirns orientiert (Folkers 2019, S. 3). Diese Netze bestehen aus zahlreichen miteinander verbundenen Verarbeitungsebenen, die Informationen schrittweise analysieren und abstrahieren. Während eines Trainingsprozesses werden die internen Parameter des Modells fortlaufend angepasst, um die Genauigkeit der Ergebnisse zu verbessern. Dabei erfolgt schließlich die Abbildung komplexer Zusammenhänge durch die Verarbeitung einer ausreichend großen Anzahl an Datenpunkten (Folkers 2019, S. 3).

Im Bereich des maschinellen Lernens werden unterschiedliche Lernansätze unterschieden. Folkers nennt drei verschiedene Möglichkeiten:

1. Überwachtes Lernen

Beim überwachten Lernen erhält das Modell sowohl Eingabedaten als auch die dazugehörigen korrekten Ergebnisse, wodurch es lernt, zukünftige Daten richtig zuzuordnen (Folkers 2019, S. 3).

2. Unüberwachtes Lernen

Beim unüberwachten Lernen stehen hingegen ausschließlich Eingabedaten zur Verfügung, sodass das System eigenständig Strukturen oder Gemeinsamkeiten erkennen muss (Folkers 2019, S. 3).

3. Reinforcement Learning

Die dritte Methode stellt das Reinforcement Learning dar, bei dem das Modell durch Rückmeldungen auf seine Handlungen lernt, ohne dass eine konkrete Musterlösung vorgegeben wird (Folkers 2019, S. 3).

Die Mustererkennung mit anschließender Produktion ist auch für Deepfakes von großer Bedeutung. Diese Technologie ermöglicht es, dass viele Bild-, Video- oder auch reine Audiodateien ausgewertet und charakteristische Merkmale erkannt werden können. Darauf aufbauend können dann neue Inhalte erzeugt werden.

Die Auseinandersetzung mit Deep Learning und Deepfakes kann dazu beitragen, die Medienkompetenz der Schüler*innen zu verbessern. Das ist vor allem für die dritte Forschungsfrage sehr wichtig. Das folgende Kapitel beschreibt die Medienkompetenz genauer.

2.4 Medienkompetenz

Der Begriff Medienkompetenz umfasst grundsätzlich zwei Bereiche. Einerseits geht es um den Gegenstand, früher waren es hauptsächlich Massenmedien wie Zeitungen oder Zeitschriften und danach auch das Fernsehen. Andererseits geht es aber auch um die Mittel in der Pädagogik, früher gab es die Tafel und den Overhead-Projektor (Von Rein, 1996, S.7).

Diese Projektoren sind noch nicht ganz aus den österreichischen Schulen verschwunden, dennoch hat sich die Technologie ordentlich weiterentwickelt. Die zunehmende Verbreitung digitaler Medien hat die Anforderungen an den kompetenten Umgang mit Informationen erheblich verändert. Während Medien früher vor allem als Instrumente zur Informationsvermittlung dienten, sind Menschen heute mit einer Vielzahl unterschiedlicher Inhalte konfrontiert, deren Glaubwürdigkeit und Wahrheitsgehalt nicht immer eindeutig erkennbar sind. Es kursieren sehr viele Falschinformationen und diese Fake News nehmen immer neue Dimensionen an. Vor

diesem Hintergrund gewinnt das Konzept der Medienkompetenz zunehmend an Bedeutung. Medienkompetenz umfasst dabei nicht nur die Fähigkeit, Medien technisch bedienen und nutzen zu können, sondern auch die kritische Auseinandersetzung mit deren Inhalten und Wirkungsweisen und das gilt schon seit langer Zeit und nicht erst seit der Nutzung von künstlicher Intelligenz. Die Medienkompetenz beinhaltet die Fähigkeit, Informationen zu bewerten, Quellen einzuordnen und den Realitätsgehalt medial vermittelter Inhalte kritisch zu hinterfragen (Von Rein, 1996).

Diese Kompetenz ist für junge Menschen sehr wichtig, da viele Jugendliche erhebliche Zeit vor Bildschirmen und digitalen Medien verbringen. In der pädagogischen Diskussion wird Medienkompetenz daher als eine relevante Voraussetzung betrachtet, um sich in einer zunehmend digitalisierten Medienlandschaft orientieren und fundierte Urteile treffen zu können und Medien somit zu „beherrschen“. Das galt auch schon vor einigen Jahren, auch wenn sich die Medien immer weiter verändern:

Der Begriff der Medienkompetenz hat daher auch etwas Schillerndes. Er bezeichnet nicht nur die Notwendigkeit, Medien zu "beherrschen", sondern auch, ihren Einfluß [sic] abzuwehren, die vermittelten Informationen zu filtern, sich zu orientieren, den Realitätsgehalt einschätzen zu können. Medienkompetenz in der aktuellen Diskussion hat - zumindest aus pädagogischer Sicht - eher etwas Abwägendes, Kritisches, Qualifikatorisches (Von Rein, 1996, S.8).

Aus heutiger Sicht lässt sich das Erkennen von manipulierten und generierten Medien von der allgemeinen Medienkompetenz nicht mehr wegdenken. Es ist relevant, diese Kompetenz in der Schule zu erlernen, auch wenn in einigen Jahren die Methoden zur Erkennung mit großer Wahrscheinlichkeit schon wieder ganz anders sein werden. Somit zählt die Behandlung von Deepfakes derzeit zur allgemeinen Medienkompetenz: „Ähnlich wie bei der Erkennung von Desinformation wird allgemeiner Medienkompetenz zur Erkennung von Deepfakes eine positive Rolle zugesprochen“ (Karaboga et al., 2024, S.129).

Die Fähigkeit, Deepfakes zu erkennen, hängt nicht nur von spezifischem Wissen über diese Technologie ab, sondern auch stark von allgemeinen digitalen Kompetenzen. Menschen, die sicher mit dem Internet, sozialen Medien und digitalen Informationsquellen umgehen können, sind häufig besser in der Lage, manipulierte Inhalte kritisch zu hinterfragen und mögliche Fälschungen zu identifizieren (Karaboga

et al., 2024). Sie sind somit kritischer im Umgang mit solchen Medien und glauben nicht alles, was sie sehen. Das ist für die Medienkompetenz sehr wichtig.

Darüber hinaus leisten journalistische Medien einen wichtigen Beitrag zur Aufklärung, indem sie über die Entstehung, Verbreitung und Risiken von Deepfakes informieren. Die Resilienz der Personen wird diesbezüglich gestärkt. Sie lernen, sich an die neue Technik und somit auch an die neuen Probleme anzupassen. Dies kann das Bewusstsein für digitale Manipulationen schärfen und einen reflektierten Umgang mit Online-Inhalten fördern. Da Deepfakes insbesondere über soziale Netzwerke und Messenger-Dienste verbreitet werden, können auch persönliche Erfahrungen mit solchen Inhalten dazu beitragen, die eigene Fähigkeit zur Erkennung von Fälschungen zu verbessern (Karaboga et al., 2024). Viele Jugendliche verbringen viel Zeit auf sozialen Netzwerken wie Instagram oder TikTok und versenden täglich einige Nachrichten über Messenger-Dienste wie WhatsApp und sammeln dadurch auch persönliche Erfahrungen. Insgesamt spielen sowohl digitale Kompetenzen als auch die regelmäßige Auseinandersetzung mit Medieninhalten eine wichtige Rolle beim Erkennen von Deepfakes und das fördert dann auch die Medienkompetenz (Karaboga et al., 2024).

Für die geplanten Unterrichtsstunden im Rahmen dieser Arbeit ist auch das EduLARP von zentraler Bedeutung. Im letzten Kapitel des theoretischen Hintergrunds wird diese Unterrichtsmethode erläutert.

2.5 Spielerisches Lernen in der Schule – EduLARP

Im Educational Live Action Role Playing begeben sich Schüler*innen in bestimmte Rollen oder Personas und es wird ein interaktives Lernerlebnis geschaffen, das unterschiedlich lange dauern kann. Es hat sich durch die Forschung an dieser Methode gezeigt, dass das EduLARP dazu beitragen kann, auch mehr Frauen und unterrepräsentierte Gruppen im Bereich der Interaktion zwischen Mensch und Computern einzubinden (Robinson et al., 2023, S.1).

Bei einem Educational Live Action Role Playing müssen sich die Kinder bzw. Jugendlichen somit von den eigenen Gedanken und Einstellungen gegenüber bestimmter Themen verabschieden und für begrenzte Zeit neue Eigenschaften annehmen. Umso besser sich die Teilnehmer*innen in die vorgegebene Rolle

hineinversetzen können, umso authentischer und lehrreicher wird das EduLARP. Von hoher Relevanz ist außerdem die Nachbesprechung und ausführliche Reflexion eines Educational Larps. Wenn das Spiel mit dem restlichen regulären Unterricht gut miteinander verknüpft ist, dann ist die Methode erfolgreicher (Bowman et al., 2015, S.20). Die heutigen Generationen lernen leichter durch Interaktion und auch mithilfe von spielendem Lernen: „As our current generation continues to engage in a more interactive, game-focused manner than previous classroom models can support, educators are challenged to find innovative ways to engage students“ (Bowman et al., 2015, S.20).

Das nachfolgende Kapitel zeigt einen Überblick über den aktuellen Forschungsstand zu Gefahren und Chancen von Deepfakes.

3 FORSCHUNGSSTAND ZU GEFAHREN UND CHANCEN

In diesem Kapitel wird der aktuelle Forschungsstand zu Deepfakes in Bezug auf die Gefahren aufgearbeitet, jedoch wird auch auf die Chancen und Potentiale dieser Technik eingegangen.

3.1 Probleme und Risiken

Deepfakes gibt es noch nicht sehr lange, deshalb gibt es bislang nur begrenzte wissenschaftliche Erkenntnisse, insbesondere zur Wahrnehmung in der Bevölkerung. Für die Schweiz fehlen sogar umfassende Studien vollständig. Aus diesem Grund zielt eine Untersuchung aus dem Jahr 2024 zunächst darauf ab, grundlegende, beschreibende Daten zu erheben. Dabei wird analysiert, welche Erfahrungen Menschen in der Schweiz bereits mit Deepfakes gemacht haben: „Konkret erheben wir, ob Menschen in der Schweiz schon von Deepfakes gehört haben, ob sie schon Deepfakes gesehen haben und ob sie schon Deepfakes erstellt und weiterverbreitet haben“ (Vogler et al., 2024, S.126).

Es ist auch wichtig anzumerken, dass Kommunikation eine wichtige Rolle im Umgang mit neuen Technologien spielt. Diesbezüglich wird nicht nur von Deepfakes gesprochen, sondern auch von vielen anderen modernen Techniken. Der Austausch findet sowohl im persönlichen Umfeld, aber auch rein digital statt. Als Beispiele für den Kontakt mit Deepfakes werden Messenger-Dienste genannt, aber hauptsächlich geht

es um soziale Netzwerke. Viele Menschen kommen gerade über diese Kanäle erstmals mit Deepfakes in Berührung, etwa wenn ihnen entsprechende Inhalte zugeschickt werden und sie diese teilen (Vogler et al., 2024). Somit kann es dazu kommen, dass entsprechende KI-generierte oder manipulierte Medien sehr schnell verbreitet werden und das Löschen nahezu unmöglich wird.

Wer sich viel mit sozialen Medien und deren kritische und reflektierte Nutzung beschäftigt, kann laut Vogler auch leichter Deepfakes erkennen: „Je höher die Social-Media-Literacy, desto wahrscheinlicher werden Deepfakes und reale Videos richtig klassifiziert“ (Vogler et al., 2024, S.148). Für diese Erhebung in der Schweiz hat man sich als Forschungsfrage gestellt, ob Deepfakes eher als Chance oder aber als Risiko von der Bevölkerung eingeschätzt werden (Vogler, 2024, S.125). Es lässt sich feststellen, dass die gefährliche Seite von KI-generierten und manipulierten Medien überwiegt: „Deepfakes werden von der Schweizer Bevölkerung stark mit Risiken assoziiert. Es gibt dabei kaum Unterschiede zwischen soziodemografischen Gruppen“ (Vogler et al., 2024, S.150).

Ein Unterschied ist aber dennoch bei den individuellen Risiken zwischen Männern und Frauen erkennbar. Die weiblichen Teilnehmer*innen der Studie sehen sich selbst stärker gefährdet, dass die eigene Privatsphäre in Gefahr geraten könnte oder sie selbst Opfer eines KI-generierten oder manipulierten Videos werden könnten (Vogler et al., 2024, S.150). Generell lässt sich eine Assoziation mit den Problemen dieser Technik erkennen:

In einem weiteren Schritt haben wir gefragt, wie sehr sie Statements zum Risiko von Deepfake-Technologien für verschiedene Gesellschaftsbereiche einschätzen (1 = stimme überhaupt nicht zu bis 7 = stimme sehr stark zu). Deepfakes werden von Schweizerinnen und Schweizern stark mit Risiken assoziiert (Vogler 2024: S.136).

Es zeigt sich außerdem auch, dass bei gut erstellten Deepfakes sehr viele Personen der Teilnehmer*innen Deepfakes nicht von realen Medien unterscheiden können. Den Menschen, die eine hohe Social-Media-Literacy haben, fällt es aber wie zuvor erwähnt leichter, gefälschte Medien zu erkennen (Vogler et al., 2024, S.150). Das ist deshalb positiv hervorzuheben, da diese Personen auch mit mehr Deepfakes konfrontiert werden, da in den sozialen Medien wie TikTok oder Instagram viele solcher Bilder und Videos kursieren.

Vor diesem Hintergrund untersucht die Studie auch, über welche Medienkanäle Menschen mit Deepfakes in Kontakt kommen und ob ein Zusammenhang zwischen ihrer Mediennutzung und ihrem Umgang mit dieser Technologie besteht (Vogler et al., 2024). Diese Studie zeigt, dass soziale Medien wie beispielsweise Instagram oder TikTok die Hauptursache für die Verbreitung von *Deepfakes* sind. Es ist allerdings erstaunlich, dass Messenger-Dienste viel weniger davon betroffen sind. WhatsApp oder Telegram wurden vor allem während der Corona Pandemie mit der Verbreitung von Desinformation in Verbindung gebracht. Die Verbreitung von falschen Informationen oder Gerüchten nimmt mit Deepfakes neue Dimensionen an, allerdings werden über diese Messenger-Dienste weniger KI-generierte und manipulierte Medien verbreitet. Die Untersuchung hat nicht erfasst, welche konkreten Arten von Deepfakes Menschen auf verschiedenen Plattformen sehen. Deshalb ist unklar, ob es sich dabei tatsächlich um schädliche Inhalte handelt (Vogler et al., 2024).

Ein sehr heikles Thema ist aber auf jeden Fall die Erstellung von nicht einvernehmlichen Deepfake-Videos, in denen reale Personen zu sehen sind. Vor allem, wenn es um den Bereich Rufschädigung geht. Mit bestimmten Apps können Menschen sehr einfach solche Filme erstellen. „Vornehmliches Einsatzgebiet von Deepfake-Videos ist deren Verwendung für kriminelle Zwecke, insbesondere bei der Erstellung nicht einvernehmlicher Pornografie“ (Farid & Schindler, 2020, S.23). Das kann zu sehr unangenehmen Situationen für die Betroffenen führen.

Eine weitere Gefahr für die Gesellschaft ist die Verbreitung von Desinformation. Deepfakes könnten als Eskalation von Fake News beschrieben werden, da Desinformationskampagnen für realistische und authentische Fake News sorgen können: „Zudem wird die Demokratisierung ausgefeilter Techniken zur Erzeugung hochrealistisch gefälschter Audio-, Bild- und Videoaufnahmen die Bekämpfung betrügerischer Fehl- und Desinformationskampagnen voraussichtlich erschweren“ (Farid & Schindler, 2020, S.23). Als Beispiel wird genannt, dass Betrüger*innen im Jahr 2019 in einem Telefonat einen Vorsitzenden des Vorstandes in einem Unternehmen dazu gebracht haben, sehr viel Geld zu überweisen. Das funktionierte über eine synthetische Imitation der Stimme (Farid & Schindler, 2020). Es gibt somit nicht nur Betrug und Probleme mittels KI-generierte Bilder und Videos, sondern auch durch das „Fälschen“ von Stimmen.

Diesbezüglich muss der sogenannte „Enkeltrick“ erwähnt werden. Dabei handelt es sich um eine Betrugsmasche, von der meistens ältere Menschen betroffen sind. Personen werden zur Überweisung von Geld gedrängt und diese Form von Kriminalität wird immer authentischer, da mithilfe von künstlicher Intelligenz Audio-Deep-Fakes erzeugt werden können (Mößle et al., o.D., S.667). Das folgende Zitat beschreibt diese Methode genauer:

Ein Beispiel für die enorme Wandlungs- und Anpassungsfähigkeit bestimmter Deliktsbereiche ist dabei der sogenannte Enkeltrick - ein telefonisches Betrugsdelikt, das vor allem ältere Menschen betrifft und in den letzten Jahren einen deutlichen Zuwachs verzeichnen kann. Inzwischen werden Betrugsnachrichten auch über Messenger-Dienste versendet und zuweilen auch künstliche Intelligenz (KI) zur Durchführung des Betrugs verwendet - z.B. mittels Stimmenimitation/Audio-Deep-Fake (Mößle et al., o.D., S.667).

Diese Art von Telefonaten werden auch Schockanrufe genannt. Im Kontext dieses sogenannten Enkeltricks zeigt sich eine besonders auffallende Tendenz bezüglich der Anzeigen. Im Zusammenhang mit derartigen Betrugsfällen sprechen drei Gründe gegen eine Anzeigeerstattung. (Mößle et al., o.D., S.679). Es muss allerdings betont werden, dass die Fälle auch angezeigt werden könnten, wenn es nur einen Versuch gegeben hat. Am häufigsten wird die Tat als nicht ausreichend schwerwiegend wahrgenommen. Darüber hinaus gehen viele Menschen, die von solchen Enkeltricks betroffen sind, davon aus, dass eine Aufklärung durch die Polizei wenig erfolgversprechend ist. Ein weiterer wichtiger Faktor ist die Einschätzung, dass mit der Anzeige viel Bürokratie anfallen könnte. Diese Motive stellen die zentralen Hemmnisse für eine Anzeigeerstattung dar (Mößle et al., o.D., S.679). Das ist aber auch sehr problematisch, da die Täter*innen somit noch schwerer erwischt werden können, als das ohnehin bereits der Fall ist.

Es gibt außerdem eine große Angst, dass KI-generierte und manipulierte Inhalte das Vertrauen in die Medien verringern. Es kann zur Verbreitung von falschen Informationen kommen. Für die Politik spielen laut der Befragten in der Schweizer Bevölkerung die Risiken, die mit Deepfakes assoziiert werden, eine kleinere Rolle (Vogler, 2024).

Insgesamt gesehen zeigt sich aber schon ein größeres Problem im Zusammenhang mit politischen Themen. Deepfakes verstärken die politische Manipulation. In Amerika sprach man schon im Jahr 2019 von einer großen globalen Bedrohung vor allem für

die Politik: „Demnach werden Gegner und strategische Rivalen versuchen, mittels Deepfakes oder ähnlicher maschineller Lernverfahren generierte überzeugende – aber falsche – Bild-, Audio- und Videodateien gegen die USA und deren Verbündete und Partner einzusetzen“ (Farid et al., 2020, S.25).

Es kommt stark auf die Qualität der KI-generierten und manipulierten Medien an, welche Auswirkungen damit erzielt werden können. Bei den Parlamentswahlen in Großbritannien im Jahr 2019 wurden Deepfakes bereits eingesetzt, um die Erfolgsaussichten bestimmter Politiker*innen zu schmälern (Farid et al., 2020). Die Qualität war allerdings nicht optimal und dadurch können solche Medien schnell von vielen Menschen als *Fakes* identifiziert werden.

Falsche Informationen und Propaganda sind zwar keine neuen Begriffe, jedoch ermöglicht künstliche Intelligenz eine klare Verstärkung. Falschinformationen nehmen somit durch den Einsatz von Deepfakes ganz neue Dimensionen an, es handelt sich um eine Eskalation der Fake News (Von Gottberg, 2024). Wenn sich jemand kurz vor dem Wahlsonntag noch schnell Informationen beschaffen möchte, die dann darüber entscheiden, wo das Kreuz im Wahllokal gesetzt wird, dann kann KI Falschinformationen liefern (Von Gottberg, 2024, S.20).

Mittlerweile wissen aber auch schon viele Menschen, dass Chatbots nicht wie Suchmaschinen funktionieren, jedoch sind das leider noch zu wenig. Deepfakes hingegen können auch Leute mit hoher IT-Affinität täuschen, da sie immer authentischer wirken. Das kann, wie bereits erwähnt, auch bei politischen Themen zu einem Problem werden. Ein Beispiel dafür zeigt auch das Video, indem sich ein Präsident eines Kriegsgebiets zum Aufruf zur Kapitulation äußert, obwohl er das in Wirklichkeit nie gesagt hat (Vot Gottberg, 2024, S.21).

Weitere Probleme sind das Erstellen von Fake-Identitäten und der Identitätsdiebstahl. Beim Erstellen von Fake-Identitäten können Betrüger*innen nicht nur falsche Profile von Nutzer*innen von sozialen Medien erstellen, sondern es kann mithilfe von Deepfakes illegal erworbenes Geld auf Konten verschoben werden, die es eigentlich gar nicht geben dürfte. Der Einsatz von Deepfake-Technologien eröffnet somit neue Möglichkeiten für betrügerische Handlungen im Finanzsektor. Durch die realitätsnahe Nachbildung von Gesichtern und Stimmen können Betrüger*innen digitale Authentifizierungsverfahren umgehen und sich gegenüber Banken als andere

Personen ausgeben. Das Video-Ident-Verfahren kann durch diese Methode überlistet werden (Lukjantschuk et al., 2025, S.46).

Dadurch besteht die Gefahr, dass Konten unter falschen Identitäten eingerichtet werden und anschließend können sie für illegal erwirtschaftetes Geld genutzt werden. Darüber hinaus erleichtern Deepfakes den Missbrauch fremder Identitäten, indem sie eine glaubwürdige Täuschung gegenüber digitalen Bankdienstleistern ermöglichen. Wenn es keine physischen Filialen gibt, kann das gut funktionieren. Somit können Unbefugte Zugang zu bestimmten Bankkonten erhalten, obwohl sie dort gar nicht Kunden sind (Lukjantschuk et al., 2025, S.46).

Insgesamt gesehen gibt es sehr viele Probleme und Risiken, die Deepfakes mit sich bringen und einige davon wurden in diesem Kapitel genannt. An dieser Stelle muss aber gesagt werden, dass diese Technologie auch positive Aspekte mit sich bringt und Chancen und Potentiale von Deepfakes auch genannt werden sollten. Aus diesem Grund beschäftigt sich das folgende Kapitel genauer damit.

3.2 Chancen und Potentiale

Die Chancen von Deepfakes werden von der Schweizer Bevölkerung deutlich geringer bewertet als die Risiken. Positive Aspekte sehen die Teilnehmer*innen der Studie vor allem im persönlichen Bereich, zum Beispiel als Möglichkeit zur Förderung von Kreativität oder auch einfach nur zur Unterhaltung. Wirtschaftliche Vorteile, wie neue Arbeitsplätze für die Schweiz, werden hingegen eher skeptisch beurteilt (Vogler, 2024). Mit sehr wenigen Fachkenntnissen können somit viele Menschen kreative Ideen umsetzen und in sozialen Netzwerken wie Instagram, Facebook oder TikTok sorgen KI-generierte und manipulierte Medien für abwechslungsreiche Unterhaltung der Nutzer*innen.

Außerdem können Deepfakes auch zur Verringerung von Verzerrungen beitragen. Darunter können einige sensible Bereiche profitieren, wie beispielsweise das Bildungswesen, das Strafjustizsystem und die Arbeitswelt: „Similarly, in education and employment, our approach could critically evaluate biases in the admissions and hiring processes, where the visual representation of candidates often plays an important role“ (Liu et al., 2025, S.15).

Es lässt sich feststellen, dass es auch Chancen und Potentiale von Deepfakes gibt, jedoch sind diese Bereiche kein wesentlicher Teil dieser Arbeit. Dennoch ist eine kurze Gegenüberstellung von positiven Aspekten auch wichtig, damit eine differenzierte Betrachtung ermöglicht wird.

Das nachfolgende Kapitel berichtet über die Methodik und die Vorgehensweise bei der Forschung dieser Arbeit.

4 METHODIK

Diese Arbeit besteht aus zwei größeren Bereichen, die aber inhaltlich miteinander zusammenhängen. Der erste Teil dient der Beantwortung der ersten beiden Forschungsfragen und beinhaltet die Literaturrecherche und die Erstellung, Durchführung und Auswertung eines Fragebogens mithilfe der Inhaltsanalyse nach Mayring, der in zwei Klassen der zehnten Schulstufe ausgeteilt wird. Das Kapitel 4.1 beschreibt diesen Fragebogen genauer. Der zweite Teil dieser Arbeit benötigt die Planung, Durchführung und regelmäßige Überarbeitung von Unterrichtsstunden in Form von Design-Based Research.

4.1 Beschreibung des Fragebogens

Es handelt sich um einen selbst erstellten Fragebogen, da er genau an die Zielgruppe und an die Thematik angepasst wurde. Nach der sorgfältigen Überlegung der einzelnen Fragestellungen wurde außerdem noch Feedback von Studienkolleg*innen eingeholt und danach erfolgte eine Überarbeitung. Wichtig war es, dass auch Fragen zu sozialen Medien gestellt wurden, da es Zusammenhänge zwischen der Nutzung von Social Media und der Auseinandersetzung mit Deepfakes gibt. Außerdem war auch relevant, dass nicht nur inhaltliche Themen abgefragt wurden, sondern auch eigene Meinungen und Feedback zu den Unterrichtsstunden.

Mithilfe eines QR-Codes können die Schüler*innen den Fragebogen auf Microsoft Forms ausfüllen. Diese Möglichkeit wurde gewählt, da es eine unkomplizierte und anonyme Beantwortung der Fragen ermöglicht. Es gab einen Pretest in der Woche zuvor, bei dem zwei Jugendliche den Fragebogen ausfüllten, um die Verständlichkeit der Fragestellungen gewährleisten zu können. Die Befragung ist vollständig anonym. Zuerst wird abgefragt, welchem Geschlecht sich die*der Teilnehmende zugehörig

fühlt. Als Auswahlmöglichkeiten gibt es weiblich, männlich und divers. Anschließend starten bereits die inhaltlichen Fragen, um die Bearbeitungsdauer möglichst kurz zu halten. Die meisten Fragestellungen sind offene Fragen und erfordern eine selbst formulierte Antwort.

Wie viel Zeit verbringst du ungefähr pro Tag auf sozialen Medien? Möchtest du diese Zeit reduzieren? Hast du schon einmal Deepfakes erkannt? Woran hast du es bemerkt und auf welcher Plattform hast du dieses Medium entdeckt? Das sind die ersten Fragen, um herauszufinden, ob die*der Teilnehmende bereits viel persönliche Erfahrung mit sozialen Medien und KI-generierten und manipulierten Medien gemacht hat.

Anschließend wird das Gefühl beim Schauen eines Deepfake-Videos abgefragt, um erkennen zu können, ob es die Jugendlichen eher enttäuschend und gruselig oder aber lustig bzw. normal empfinden. Die darauffolgenden Fragen sollen Einschätzungen aus jugendlicher Perspektive aufzeigen, wie es in der Zukunft mit dieser Thematik weitergehen könnte: Wie werden sich Fake News durch die Möglichkeit von Deepfakes in Zukunft entwickeln? Eigene Einschätzung. Glaubst du, wird man wieder verstärkt auf gedruckte Informationsquellen wie eine Zeitung zurückgreifen? Warum/ Warum nicht?

Dann wird abgefragt, anhand welcher Merkmale die*der Teilnehmende Deepfakes erkennt und danach geht es in den folgenden beiden Fragen um die Problematik des Anstiegs der Anzahl von kursierenden Deepfakes: Welche Risiken entstehen, wenn immer mehr Deepfakes verbreitet werden? Wie hoch schätzt du den Anteil an KI-generierten bzw. manipulierten Medien ein, die du auf soziale Medien siehst? Abschließend können die Jugendlichen dann noch notieren, was sie sonst noch mitteilen möchten, falls ein relevanter Aspekt in den vorigen Fragestellungen noch gefehlt hat.

Dieser Fragebogen wird im Anschluss ausgewertet und darüber berichtet das folgende Kapitel.

4.2 Datenanalyse

Die Daten des Fragebogens werden mithilfe von Microsoft Forms erhoben und dann anschließend in einer Microsoft Excel Datei abgespeichert. Für jede Frage gibt es eine

Spalte an Antworten und für jede*n teilnehmende*n Schüler*in gibt es jeweils eine Zeile für die Antworten. Diese Daten werden danach anhand der qualitativen Inhaltsanalyse nach Mayring ausgewertet. Für die Kategorienbildung wird das induktive Verfahren verwendet. Dabei werden zuerst die Daten erhoben und anschließend die Kategorien anhand der erhaltenen Antworten gebildet: „Eine induktive Kategoriendefinition hingegen leitet die Kategorien direkt aus dem Material in einem Verallgemeinerungsprozess ab, ohne sich auf vorab formulierte Theorienkonzepte zu beziehen“ (Mayring, 2022, S.84).

Somit werden die Antworten einer Fragestellung betrachtet und danach mehrere Kategorien gebildet, die genauer definiert werden. Außerdem gibt es jeweils ein Beispiel und eine Kodierregel, die genauer beschreibt, welche Voraussetzungen eine Antwort erfüllen muss, um dieser Kategorie zugeordnet zu werden und möglicherweise werden auch unterschiedliche Teile einer Antwort verschiedenen Kategorien zugeordnet.

Die nachfolgende Tabelle zeigt ein Beispiel für die Darstellung einer Kategorie auf folgende Fragestellung: Wie würdest du dein Gefühl beim Schauen eines Deepfake-Videos beschreiben?

Kategorie	Definition	Ankerbeispiel	Kodierregel
3: Positive emotionale Reaktionen	Positive Gefühle wie Spaß, Interesse oder Humor beim Anschauen von Deepfake-Videos.	„es ist lustig“	Alle Antworten, in denen positive Emotionen wie lustig, interessant oder unterhaltsam genannt werden, werden dieser Kategorie zugeordnet.

Tabelle 1 Kategorienbildung Beispiel

In der ersten Spalte wird der Titel der Kategorie genannt, es geht um positive Reaktionen beim Schauen eines KI-generierten oder manipulierten Mediums. In der zweiten Spalte wird diese Kategorie genauer definiert, um die Zuordnung zu erleichtern. Die passende Definition dieser Kategorien ist für die Inhaltsanalyse sehr wichtig (Mayring, 2022, S.84). In der dritten Spalte steht das Ankerbeispiel. „es ist lustig“ kann definitiv als positive emotionale Reaktion beurteilt und somit dieser Kategorie zugeordnet werden. In der letzten Spalte werden die Regeln festgelegt, um

bei Antworten, die nicht optimal in diese Kategorie passen, dennoch entscheiden zu können, wo sie zugeordnet werden.

Anschließend kann ermittelt werden, wie viele Antworten pro Kategorie erhalten wurden und somit werden die Daten analysiert. Diese Auswertung und Analyse ist wesentlich für die Beantwortung der zweiten Forschungsfrage. Für die dritte Fragestellung ist die Methode Design-Based Research sehr wichtig, die im nachfolgenden Kapitel erläutert wird.

4.3 Design-Based Research

Es wird im Rahmen dieser Masterarbeit eine Doppelstunde zu Deepfakes für die AHS-Oberstufe geplant und ausprobiert. Bei der Methode Design-Based Research werden diese Unterrichtsstunden mehrfach durchgeführt und stets verbessert. Es gibt somit einen iterativen Prozess, bei dem immer wieder überarbeitet wird: „Forschung und Entwicklung werden als ein zirkulärer, iterativer Prozess konzipiert; Bildungsforschung und –praxis wirken kooperativ zusammen, wobei die Interessen und Ziele klar getrennt bleiben und die Handlungsschwerpunkte variieren können“ (Euler, 2014, S.8).

Dadurch werden die Unterrichtsstunden schrittweise optimiert und der Prozess wird dokumentiert: „The research and development process is realized in iterate cycles of design, testing, analysis, and redesign. An incremental optimization of the design is effected within these cycles, and the development processes and principles are simultaneously documented“ (Euler, 2014, S.19). Aus diesem Grund beschreibt das Kapitel 6 Deepfakes im Unterricht die einzelnen Durchgänge und Planungen der Doppelstunde detailliert.

Die nachfolgende Abbildung stellt den Ablauf von Design-Based Research dar. Anhand des kreisförmigen Pfeils erkennt man den iterativen Verlauf mit mehreren Wiederholungen.

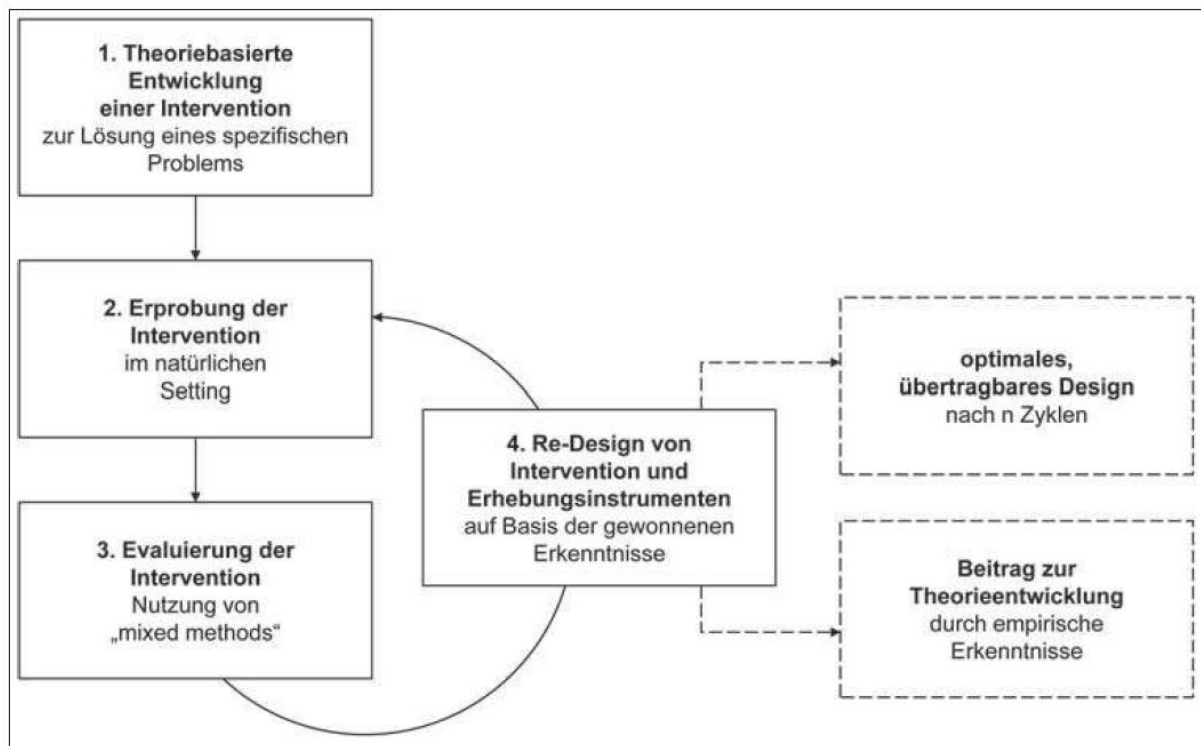


Abbildung 1 Design-Based Research Zyklen (Gess et al. 2014: S.12)

Zuerst werden die Unterrichtsstunden geplant und dann erprobt (erster Durchgang). Anschließend werden die gewonnen Erkenntnisse evaluiert und die Doppelstunde wird auf deren Basis verbessert (Schritt 4). Danach wird wieder bei Schritt 2 gestartet, die überarbeiteten Unterrichtsstunden werden wieder durchgeführt und der zweite Iterationszyklus wird gestartet. Es gibt keine allgemeingeltende optimale Anzahl an Zyklen, das kann individuell variieren (Gess et al., 2014, S.12).

Das folgende Kapitel beschreibt die Zielgruppe der geplanten Unterrichtsstunden genauer.

4.4 Zielgruppe

Die Zielgruppe der Fragebögen und auch der Unterrichtsstunden ist die Oberstufe von allgemeinbildenden höheren Schulen. Der Fragebogen wurde in zwei Klassen der zehnten Schulstufe der Gesamtschule in der Kandlgasse im siebten Wiener Gemeindebezirk ausgeteilt. Die Unterrichtsstunden wurden ebenfalls in dieser Schule durchgeführt und anschließend wurden die überarbeiteten Stunden im Gymnasium in der Klostersgasse im 18. Wiener Gemeindebezirk unterrichtet. Aufgrund von Sanierungsarbeiten wurden die beiden Klassen der neunten Schulstufe zum Zeitpunkt

dieser Durchführung im ehemaligen Ausweichquartier der Schule, im orthopädischen Spital Währing, unterrichtet.

5 PERSPEKTIVEN DER JUGENDLICHEN

Die vorliegende qualitative Erhebung untersucht Wahrnehmungen, Einstellungen und Einschätzungen von Jugendlichen über Deepfakes. Dabei wurden 35 Schüler*innen aus zwei Klassen der zehnten Schulstufe in der Gesamtschule in der Kandlgasse im siebten Wiener Gemeindebezirk befragt. Die Auswertung erfolgt in Anlehnung an die strukturierende Inhaltsanalyse nach Mayring und fokussiert zentrale inhaltliche Muster über alle offenen Fragen hinweg.

Dabei zeigt sich, dass viele der befragten Jugendlichen sehr viel Zeit vor Bildschirmen verbringen. Das wachsende Problembewusstsein zeigt sich aber auch bei den Antworten, die Schüler*innen sind teilweise schon sehr reflektiert bei dieser Problematik. Anschließend werden die Ergebnisse der Auswertung dem aktuellen Forschungsstand gegenübergestellt.

5.1 Auswertung der Fragebögen

Die Tabelle 3 zeigt die Aufteilung innerhalb der beiden Klassen. Es lässt sich feststellen, dass zum Zeitpunkt der Befragung überwiegend weibliche Schüler*innen anwesend waren.

Befragte Schüler*innen (m/w/d)

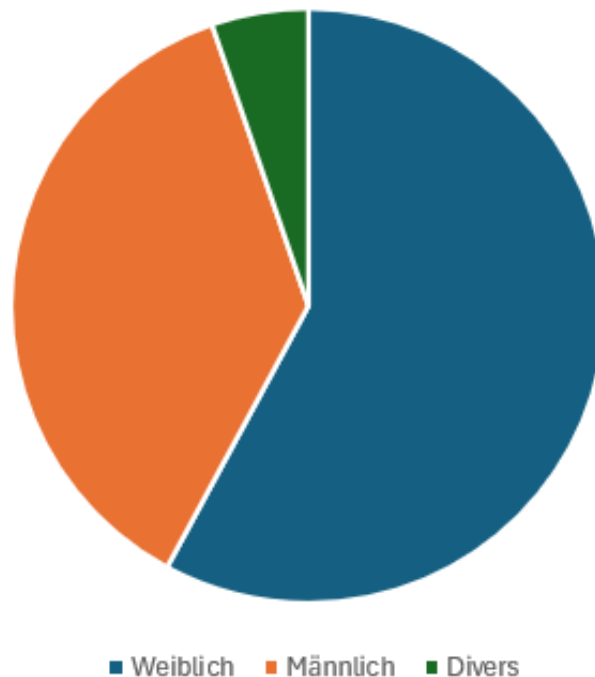


Abbildung 2 Befragte Schüler*innen (m/w/d)

Zunächst lässt sich hinsichtlich der täglichen Nutzung sozialer Medien eine deutliche Normalisierung hoher Nutzungszeiten aus Sicht der Jugendlichen feststellen. Viele Stunden beispielsweise auf TikTok zu verbringen, ist für viele junge Menschen zum ganz normalen Alltag geworden. Die Angaben reichen überwiegend von mehreren Stunden bis hin zu sehr intensiven Nutzungsformen, die nahezu den gesamten Tagesverlauf umfassen. Häufig genannte Werte wie „5 Stunden“ oder „6-7 Stunden“ verdeutlichen, dass soziale Medien für die Befragten ein großer Bestandteil ihres täglichen Alltags sind, vor allem, wenn man bedenkt, dass sie bis 13:50 Uhr meistens Unterricht haben und der Großteil dieser Bildschirmzeit erst nach der Schule beginnt. Es wird aber auch oft hinzugefügt, dass die Schüler*innen diese Zeit gerne verringern und die Stunden sinnvoller nutzen wollen. Somit zeigt sich etwa in der Aussage „4 Stunden Ja ich will die Zeit reduzieren“, dass Nutzungsdauer und Problembewusstsein gleichzeitig auftreten können, ohne dass daraus zwingend eine Verhaltensänderung resultiert. Viele greifen somit trotz des Bewusstseins über die Probleme lieber zum Smartphone als zum Buch.

In der Kategorie „hohe Nutzungsdauer“ wird die starke Integration sozialer Medien noch deutlicher. Aussagen wie „Um die 7 bis 12 std“ oder „Ich verbringe täglich

mehrere Stunden wenn nicht sogar den Großteil des Tages“ weisen darauf hin, dass Social Media in einigen Fällen nicht mehr nur als Freizeitaktivität, sondern als kontinuierlich begleitendes Medium verstanden werden kann. Das ist auch keine Seltenheit, da 13 von 35 Schüler*innen angegeben haben, dass sie täglich mehr als sieben Stunden auf sozialen Netzwerken verbringen. Diese extreme Nutzung wird dabei auch teilweise nicht negativ bewertet, sondern als normaler Bestandteil des Alltags dargestellt. Die folgende Abbildung zeigt in einem Kreisdiagramm die Anteile an Schüler*innen, die weniger als drei Stunden, zwischen drei und sechs Stunden und mehr als sechs Stunden täglich Zeit auf sozialen Medien verbringen.



Abbildung 3 Diagramm Tägliche Zeit auf sozialen Medien

Im Hinblick auf die Reduktion der Nutzung zeigt sich ein heterogenes Bild. Einerseits äußern viele Befragte den Wunsch, ihre Nutzungszeit zu verringern, wie in den Zitaten „5 Stunden mindestens. ja das will ich reduzieren“ oder „3 Stunden, ja ich möchte es reduzieren“ zu sehen ist. Diese Aussagen deuten auf ein grundsätzlich vorhandenes Problembewusstsein hin, das jedoch häufig nicht in konkrete Handlungsstrategien überführt, sondern einfach hingenommen wird. Andererseits existiert eine ebenso relevante Gruppe, die keine Veränderung anstrebt oder ihre Nutzung als unproblematisch betrachtet. Dies wird deutlich in Aussagen wie „nein ich wills nicht reduzieren“ oder „Ca 6/7 Stunden, nein ich möchte es nicht reduzieren“. Insgesamt

entsteht dadurch ein Spannungsfeld zwischen Reflexion und Akzeptanz intensiver Nutzung. Fakt ist aber, dass die Jugendlichen durchschnittlich bedenklich viel Zeit auf sozialen Netzwerken agieren und diese Erkenntnis ist wichtig für die Fragestellungen zu Deepfakes, da auf Social Media viele KI-generierte und manipulierte Medien verbreitet werden.

Die Analyse der Frage zur Erkennung von Deepfakes zeigt eine klare Dominanz der visuellen Wahrnehmungsstrategien. Die meisten Befragten orientieren sich bei der Beurteilung von Echtheit an konkrete visuelle Merkmale, die sie sehen oder eben nicht erkennen können. Besonders häufig werden dabei Gesichter, Hände, Finger, Augen und Bewegungsabläufe genannt. Typische Aussagen wie „Auf Veränderungen im Video achten, Fingeranzahl, Qualität, Farbverhältnisse“ oder „Auf die Augen, Zähne, Hintergrund, Farbe achten oder ob etwas sehr unrealistisch wirkt“ verdeutlichen, dass viele Schüler*innen auf visuelle Erkennungsmerkmale achten, auch wenn es sich um ein Video handelt. Diese Ergebnisse legen nahe, dass sich im Alltagswissen der Befragten bereits spezifische „Deepfake-Indikatoren“ etabliert haben. Es ist aber auch anzumerken, dass sehr viele Jugendliche die Finger und vor allem deren Anzahl genannt haben, jedoch muss man bei immer besser werdenden Deepfakes darauf achten, dass die Finger häufig schon gut dargestellt werden.

Neben dieser dominanten visuellen Strategie zeigt sich eine kleinere, jedoch methodisch interessante Gruppe von Antworten, die auf Intuition oder subjektives Gefühl zurückgreifen. Aussagen wie „Authentisch vom Gefühl her, eigentlich schwer zu erkennen“ oder „Man merkt es einfach“ verdeutlichen, dass ein Teil der Befragten keine expliziten Kriterien benennen, aber dennoch Deepfakes häufig erkennen kann. Diese Form der Einschätzung verweist auf implizites Medienwissen, das nicht bewusst verbalisiert wird, jedoch im Entscheidungsprozess eine Rolle spielt. Für die eigene Erkennung reicht das häufig aus, jedoch ist es dann schwierig, anderen Menschen Möglichkeiten zur Identifizierung zu erklären.

Eine weitere Ebene der Erkennungsstrategien bildet die Kontext- und Quellenprüfung. Hierbei werden externe Informationen herangezogen, um die Echtheit eines Videos zu verifizieren. Dies zeigt sich beispielsweise in der Aussage „Ich schaue mir die Kommentare an.“ oder „Man prüft die Quelle, achtet auf komische Fehler im Video und schaut, ob andere vertrauenswürdige Seiten es bestätigen.“ Das ist auch

eine sehr wesentliche Kenntnis, da bei immer besser werdenden KI-generierten bzw. manipulierten Medien eine forensische Prüfung oder eine sehr genaue Recherche der Herkunft der Medien nötig ist. Diese Antworten weisen auf eine erweiterte Form der Medienkompetenz hin, bei der nicht nur das audiovisuelle Material selbst, sondern dann zusätzlich auch dessen soziale Einbettung berücksichtigt wird. Dennoch bleibt diese Strategie im Vergleich zur visuellen Analyse weniger verbreitet bei den Schüler*innen in diesen Oberstufenklassen.

Die Einschätzungen zur zukünftigen Entwicklung von Fake News im Zusammenhang mit Deepfakes sind überwiegend pessimistisch geprägt. Viele Befragte erwarten eine zunehmende Verstärkung der Problematik, insbesondere hinsichtlich der Unterscheidbarkeit zwischen echten und manipulierten Inhalten. Das ist auch derzeit schon auf soziale Medien zu erkennen. Dies wird deutlich in Aussagen wie beispielsweise „Man weiß nicht mehr was man glauben soll“ oder „Es wird schwerer zu erkennen, was Fake ist und was nicht“. Diese Wahrnehmung weist darauf hin, dass Desinformationsprobleme verstärkt werden laut den Schüler*innen der Gesamtschule in der Kandlgasse.

Zudem wird häufig eine quantitative Zunahme von Fake News erwartet. In den Aussagen „Ich glaube es wird viel häufiger vorkommen, dass Fake News in die Welt gelangen“ oder „Es wird viel schlimmer werden“ lässt sich das erkennen. Dies deutet auf die Erwartung einer exponentiellen Ausbreitung digitaler Desinformation hin. Parallel dazu wird auch eine Erhöhung der Qualität der Fake News beschrieben, insbesondere durch realistischere und glaubwürdigere Inhalte. Dies zeigt sich in der Aussage „Deepfakes lassen Fake News durch Bild und Video „Beweise“ realistischer wirken“. Hier wird außerdem deutlich, dass visuelle Beweislogik als besonders problematisch wahrgenommen wird, da sie klassische Formen der Faktenprüfung verändert. Dieses große Problem wurde allerdings auch ausführlich in den Unterrichtsstunden davor besprochen.

Neben diesen Entwicklungen werden auch gesellschaftliche Risiken thematisiert. Diese reichen von politischer Manipulation über Betrug bis hin zu generellem Vertrauensverlust in digitale Medien. Aussagen wie „Es können Falschinformationen, Betrug, Rufschädigung und Vertrauensverlust in Medien entstehen.“ oder „Es werden vielleicht immer mehr politische Themen benutzt und es werden vielleicht immer mehr

Probleme dadurch auftreten“ zeigen, dass Deepfakes als gesellschaftliches Problem von den Jugendlichen verstanden werden. Besonders häufig wird der Verlust von Vertrauen betont, der als zentrale Konsequenz digitaler Manipulation wahrgenommen wird.

Im Hinblick auf den möglichen Anstieg von gedruckten Informationsquellen wie Zeitungen zeigt sich ein überwiegend skeptisches Bild gegenüber einer solchen Entwicklung. Viele Befragte gehen davon aus, dass digitale Medien weiterhin dominieren werden, insbesondere aufgrund ihrer Schnelligkeit und Allgegenwärtigkeit. Man erhält viele unterschiedliche Informationen in nur wenigen Minuten und die Algorithmen präsentieren den Menschen die Inhalte, die sie als interessant empfinden. Dass Zeitungen nicht wieder häufiger gelesen werden, wird in Aussagen wie „Nein, ich glaube, dass Soziale Medien in den Vordergrund rücken“ oder „Nein, weil Internet schneller ist und Zeitungen kaum noch genutzt werden“ deutlich.

Gleichzeitig existiert jedoch auf der anderen Seite auch eine andere Meinung, die gedruckten Medien einen höheren Vertrauensstatus zuschreibt. Dies zeigt sich in Aussagen wie „Ja höchstwahrscheinlich da dadurch News vertrauenswürdiger rüberkommen“ oder „Ich denke schon, weil die gedruckten Informationsquellen von Menschen selbst geschrieben werden“.

Das folgende Diagramm zeigt, dass der Anteil an KI-generierten Medien auf sozialen Netzwerken laut den Antworten der Schüler*innen sehr hoch ist. An der y-Achse lässt sich die Anzahl der Schüler*innen ablesen und somit kann ermittelt werden, wie viele Jugendliche die jeweilige Kategorie angekreuzt haben. 11 Schüler*innen haben angekreuzt, dass 10-20 % der Medien, die sie auf sozialen Netzwerken sehen, KI-generiert oder manipuliert sind. 12 Jugendliche haben sogar angegeben, dass der Anteil noch höher ist.

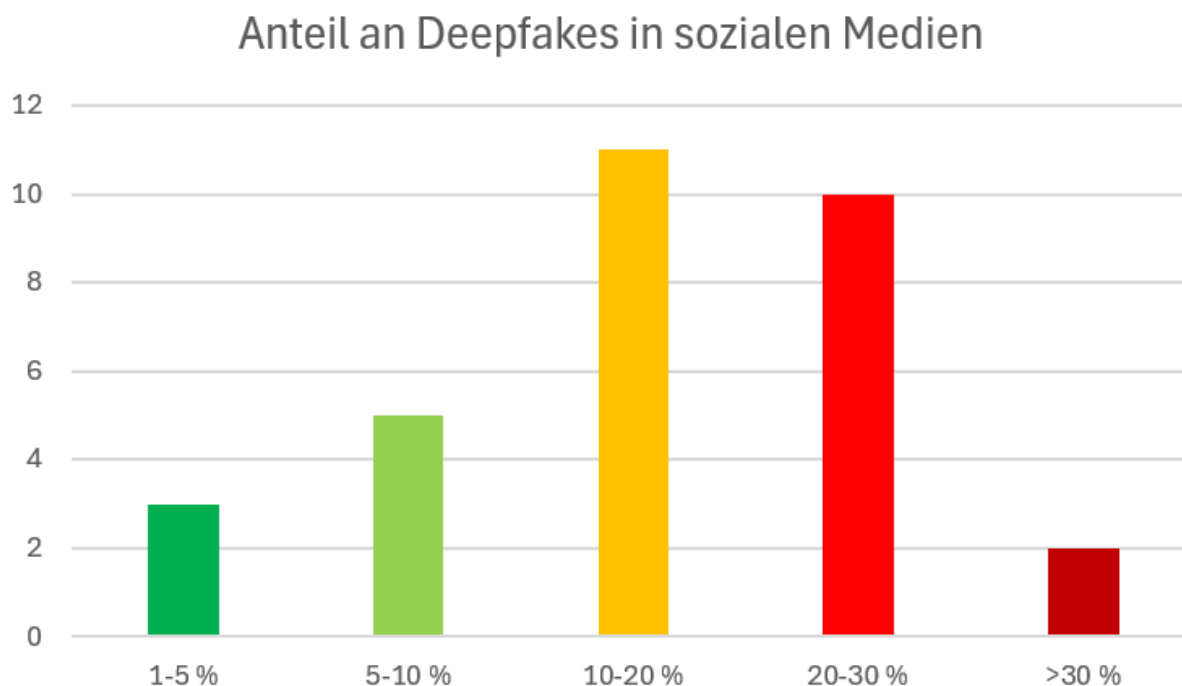


Abbildung 4 Anteil an Deepfakes in sozialen Medien

In den abschließenden offenen Stellungnahmen zeigt sich, dass sich die jungen Menschen bereits Sorgen über diese Thematik machen, bevor sie volljährig sind. Besonders häufig wird Unsicherheit und Angst thematisiert, etwa in den Antworten „Angst weil man es irgendwann nicht mehr erkennen kann“ oder „Ich denke, dass es problematisch für die Zukunft sein könnte“. Gleichzeitig werden auch Forderungen formuliert, insbesondere im Bereich der Medienkompetenz, wie in den Zitaten „Ich finde man sollte Sachen, bei denen man sich nicht sicher ist immer mit einer vertrauten und sicherern Quelle überprüfen“. Dies zeigt, dass neben Problembewusstsein auch ein gewisser Wunsch nach aktiver Bewältigung der Herausforderungen bei den Jugendlichen existiert.

Zusammenfassend lässt sich feststellen, dass die Schüler*innen Social Media intensiv nutzen und Deepfakes grundsätzlich als relevantes gesellschaftliches Problem wahrnehmen. Im folgenden Kapitel werden diese Perspektiven der Jugendlichen mit dem aktuellen Forschungsstand verglichen.

5.2 Vergleich mit aktuellem Forschungsstand

In den Antworten der Jugendlichen bezüglich der Risiken und Probleme von Deepfakes lassen sich einige Parallelen mit dem aktuellen Forschungsstand finden.

Es wird genannt, dass die immer schwieriger werdende Unterscheidbarkeit zwischen realen und manipulierten Medien ein großes Problem darstellt, was auch bei den Risiken im aktuellen Forschungsstand genannt wird. Viele Menschen verbringen täglich viel Zeit auf sozialen Medien und können bei guten Deepfakes kaum unterscheiden, ob das wirklich so passiert ist oder ob das die künstliche Intelligenz in wenigen Sekunden erschaffen hat. Die Abbildung 5 zeigt, dass der Anteil an Deepfakes in sozialen Medien bereits sehr hoch ist.

Thematisiert wird auch sowohl im aktuellen Forschungsstand als auch im vorherigen Kapitel der Perspektiven von Jugendlichen das Problem der Fake News und auch das Thema der Manipulation und Gerüchte, die verbreitet werden. Die folgenden Antworten dienen als Beispiele:

- „Ich glaube es werden sehr viele Fake News entstehen und ich finde es relativ gefährlich.“
- „Viel Manipulation viele Gerüchte“
- „Dadurch werden viel mehr fake news verbreitet und Lügen erzählt.“
- „Es können Falschinformationen, Betrug, Rufschädigung und Vertrauensverlust in Medien entstehen.“
- „Menschen fangen an, Sachen zu glauben und Videos/Fotos zu vertrauen, die nicht real sind“

Es lassen sich sowohl in der Forschung als auch im empirischen Teil dieser Arbeit viele Probleme und Risiken erkennen. Damit die Jugendlichen einen kritischen und reflektierten Blick auf Deepfakes haben, ist der Unterricht über das Thema sehr wichtig. Das folgende Kapitel befasst sich aus diesem Grund mit zwei konkreten Unterrichtsstunden, die das Thema behandeln.

6 DEEPFAKES IM UNTERRICHT

Im Rahmen der Masterarbeit wurden Unterrichtsstunden geplant, durchgeführt und laufend überarbeitet. Die erste Phase beschreibt das Kapitel 6.1 Planung.

6.1 Planung

Bereits im Herbst 2025 entwickelt der Autor dieser Arbeit im Rahmen eines Seminars über Fachdidaktik der Unterrichtsfächer Informatik und Digitale Grundbildung ein EduLARP zum Thema Deepfakes. Das Vorhaben wird zweimal präsentiert und Feedback kann somit eingeholt werden. Dadurch entstehen die Charakterkarten für die unterschiedlichen Rollen. Da sich damit in der Form aber keine zwei Unterrichtsstunden füllen lassen und Vorwissen in dieser Thematik notwendig ist, wird es in eine Doppelstunde integriert und am Ende der beiden Stunden in die Planung eingearbeitet. Zuerst wird das aktuelle Vorwissen abgefragt und die Schüler*innen müssen zwischen realen und KI-generierten Bildern unterscheiden können. Dieser Einstieg erhält positives Feedback von den Kolleg*innen im Seminar. Erst anschließend wird das vorhandene Vorwissen von einem theoretischen Vortrag, dem gemeinsamen Lesen eines Zeitungsartikels und dem Schauen eines Videos ergänzt.

Sehr positives Feedback erhält das Zeigen von eigenen erstellten Deepfakes. Konkret geht es um ein Bild, welches bearbeitet wurde und um zwei Videos, die mithilfe der App Grok erstellt wurden. Die Schüler*innen sollen anhand dieser Medien feststellen, was von der künstlichen Intelligenz gut gelungen ist und anhand welcher Merkmale man das generierte Medium erkennen kann.

Zum Abschluss der Unterrichtsstunden wird das zuvor erwähnte EduLARP geplant, bei dem die Jugendlichen in Kleingruppen über Deepfakes diskutieren sollen und jede*r Schüler*in erhält davor eine vorgegebene Rolle. Als Beispiel können ein*e besorgte*r Bürger*in und ein*e unwissende*r Jugendliche*r genannt werden. Eine Kollegin teilte als Feedback mit, dass die KI-generierten Bilder auf den Charakterkarten sowohl weibliche als auch männliche Personen im ungefähr gleichen Verhältnis darstellen sollen und diese Veränderung wurde anschließend auch durchgeführt. Somit müssen sich die Schüler*innen Argumente für und bzw. oder gegen Deepfakes überlegen, aber ausschließlich aus der Sicht der zugewiesenen Rolle. Die eigenen Meinungen sollen vernachlässigt werden. Danach werden im ersten Durchgang noch Fragen an die Schüler*innen gestellt, deren Ergebnisse im Kapitel 5 beschrieben wurden. Diesbezüglich gab es sehr viel Feedback, da manche Fragen nicht ganz klar formuliert waren und der Fragebogen wird deshalb überarbeitet.

Nachdem Feedback eingeholt wurde und die Planungen überarbeitet wurden, werden die Unterrichtsstunden gehalten. Konkreten Lehrplanbezug haben die beiden Sequenzen in der Kategorie Informatik, Mensch und Gesellschaft in folgendem Unterpunkt:

„Die Bedeutung von Informatik in der Gesellschaft beschreiben, die Auswirkungen auf die Einzelnen und die Gesellschaft einschätzen und Vor- und Nachteile an konkreten Beispielen abwägen können“ (Lehrplan RIS, entnommen am 12.06.2026).

Es gibt zwei konkrete Lernziele in Bezug auf Deepfakes, die in diesen Unterrichtsstunden erfüllt werden sollten:

- Schüler*innen (SuS) können den Begriff „Deepfake“ erklären.
- SuS können Probleme und Risiken von Deepfakes nennen.
- SuS können Erkennungsmerkmale von Deepfakes nennen und diese Strategien anwenden.
- Die SuS können aus vorgegebenen Rollen zum Thema Deepfakes argumentieren.

Genauere Informationen zu den Rahmenbedingungen der Stunden beinhaltet das nächste Kapitel.

6.2 Rahmenbedingungen der Unterrichtsstunden

Es gibt drei Durchgänge der Unterrichtseinheiten, wobei der erste Durchgang in zwei Klassen absolviert wird. Jedoch ist zwischen den beiden Terminen nur ein einzelner Tag und deshalb kann keine ausführliche Überarbeitung dazwischen stattfinden. Aus diesem Grund werden diese beiden Termine in einem Durchgang zusammengefasst. Es handelt sich dabei um zwei sechste Klassen in der Gesamtschule in der Kandlgasse. Sie befindet sich im siebten Wiener Gemeindebezirk und ist eine öffentliche Schule, somit muss kein Schulgeld von den Eltern bezahlt werden. Die sechste Klasse entspricht der zehnten Schulstufe. Die beiden weiteren Durchgänge werden im Gymnasium in der Klostergasse im 18. Wiener Gemeindebezirk gehalten. Es handelt sich jeweils um eine 5. Klasse, das entspricht der neunten Schulstufe und es ist auch eine öffentliche Schule. Da das Schulgebäude in diesem Jahr saniert wird, befinden sich die Klassen zum Zeitpunkt dieser Durchgänge im Ausweichquartier, dem ehemaligen orthopädischen Spital Währing.

6.3 Durchgang 1

Der erste Durchgang wurde zweimal durchgeführt. Sowohl am 15. als auch am 16. Dezember 2025 wurde das Unterrichtskonzept ausprobiert. Der erste Versuch war in der Klasse 6B in der Kandlgasse. Das Klassenzimmer ist sehr lang und deshalb gibt es insgesamt sechs Reihen, dafür ist der Raum nicht breit. Insgesamt gesehen ergibt sich somit keine optimale Ausgangssituation, da viele Schüler*innen sehr weit von der Lehrperson weg sitzen, wenn an der Tafel etwas geschrieben wird.

An dem betreffenden Tag waren insgesamt neunzehn Schüler*innen anwesend, die am Unterrichtsgeschehen teilnahmen. Der Informatikunterricht fand in den beiden ersten Stunden des Vormittags statt, wodurch der Tag für die Lernenden direkt mit einem fachlich anspruchsvollen, aber zugleich praxisorientierten Fach begann. Der Unterricht startete pünktlich um 8 Uhr und von 8:50 bis 9 Uhr gab es eine kurze Pause, da es in dieser Schule nach jeder Stunde am Vormittag zehn Minuten Pause gibt. Die Tabelle 2 zeigt die geplante Unterrichtsmatrix der Doppelstunde im Detail:

Nr.	Thema, Aufgabenstellung	Medien, Hilfsmittel	Sozialform, Aktivität	Zeit
1	Begrüßung , Vorstellung des heutigen Programms	PowerPoint Präsentation, Beamer	Vorstellung der Lehrperson	0-5 Min.
2	KI-generierte Bilder von realen Fotos unterscheiden Die SuS stimmen mittels Handzeichen ab, welche Bilder real sind. Anschließend wird analysiert, woran man die Unterschiede erkennt.	PowerPoint Präsentation, Beamer	Fragen an die Schüler*innen, Vortrag der Lehrperson	5-10 Min.
3	Mindmap erstellen Die Schüler*innen melden sich durch Handzeichen und schreiben alle Gedanken zum Thema Deepfakes an die Tafel.	Tafel	Aufschreiben an die Tafel	10-20 Min.
4	Vortrag Die Lehrperson erklärt ausführlich das Thema, die Erkennungsmerkmale und die Technologie und ergänzt somit die Gedanken der Schüler*innen.	PowerPoint Präsentation, Beamer	Vortrag der Lehrperson	20-30 Min.

5	Video – Deepfakes Am Beamer wird ein Video gezeigt, in dem beschrieben wird, wie Menschen in der Politik Probleme aufgrund von Deepfakes bekommen können. Anschließend wird im Plenum über das Video gesprochen.	Beamer, Lautsprecher	Aufmerksames Anschauen eines Videos, Diskutieren im Plenum	45-55 Min.
6	Deepfakes analysieren Die Schüler*innen analysieren KI-manipulierte bzw. generierte Videos und sprechen über die Erkennungsmerkmale.	PowerPoint Präsentation, Beamer	Diskutieren im Plenum	55-65 Min.
7	EduLARP Vorbereitung Die Schüler*innen überlegen sich selbstständig oder mithilfe des Smartphones und künstlicher Intelligenz Argumente für die Diskussion.	Smartphone, Künstliche Intelligenz	Gruppenarbeit	65-75 Min.
8	EduLARP Diskussion In einer Diskussionsrunde diskutieren Personen aus unterschiedlichen Branchen und Lebensphasen über Deepfakes. Alle wollen ihren Standpunkt und ihre Meinung zu dem Thema einbringen und versuchen, die anderen davon zu überzeugen. Jede*r Schüler*in begibt sich dazu in eine vorgegebene Rolle mit vorgegebenen Zielen und überlegt sich vor der Diskussion passende Argumente.	Zettel	Gruppendiskussion	75-90 Min.
9	Resümee und Beantwortung des Fragebogens Schüler*innen melden sich durch Handzeichen, um über die Erfahrungen während des EduLARPs zu berichten. Anschließend scannen die Jugendlichen den QR-Code, der zum Fragebogen führt und die Schüler*innen beantworten die Fragen zum Thema Deepfakes.	PowerPoint Präsentation, Handy	Besprechen im Plenum, Ausfüllen desw Fragebogens	90-100 Min.

Tabelle 2 Unterrichtsmatrix erster Durchgang

Die Unterrichtseinheiten begannen mit einer Einführung in das Thema Deepfakes. Dazu unterscheiden die Schüler*innen zunächst zwischen realen und KI-generierten Bildern und es erkannten alle Jugendlichen die KI-generierten Fotos. Anschließend wird das Vorwissen in Form einer Mindmap gesammelt. Die*Der Schüler*in, die*der sich für einen Beitrag meldete, musste die Wortmeldung selbst an der Tafel in die Mindmap eintragen. Das führte leider dazu, dass sich nicht viele Leute meldeten und nur sehr offensichtliche Einträge wie künstliche Intelligenz oder Manipulation aufgeschrieben wurden. Danach wurden Grundlagen zu Deepfakes, deren Funktionsweise und die möglichen Erkennungsmerkmale mithilfe der vorbereiteten PowerPoint Präsentation erklärt. Mithilfe eines Videos sowie der Analyse von manipuliertem Bild- und Videomaterial vertieften die Lernenden ihr Verständnis und reflektierten mögliche gesellschaftliche Auswirkungen.

Darauf aufbauend bereiteten sich die Schüler*innen auf das EduLARP vor, in dem sie unterschiedliche gesellschaftliche Rollen einnehmen und die Chancen und Risiken von Deepfakes diskutieren sollen. Dazu gibt es folgende Aufgabenstellung:

„In einer Diskussionsrunde werden Personen aus unterschiedlichen Branchen und Lebensphasen über Deepfakes diskutieren. Alle wollen ihren Standpunkt und ihre Meinung zu dem Thema einbringen und versuchen, die anderen davon zu überzeugen. Jede*r Schüler*in begibt sich dazu in eine vorgegebene Rolle mit vorgegebenen Zielen und überlegt sich vor der Diskussion passende Argumente.“

Die Vorbereitung verlief leider sehr chaotisch und nur wenige Schüler*innen machten sich Notizen zu den überlegten und recherchierten Argumenten. Das verpflichtende Mitschreiben würde die Vorbereitung wahrscheinlich verbessern. Die Reduzierung auf Vierergruppen würde außerdem die Sprechanteile erhöhen. Es gibt beim ersten Durchgang fünf Charakterkarten für das EduLARP, auf denen die Fotos immer KI-generiert sind. Die Bilder wurden mithilfe von ChatGPT erstellt. Folgender Prompt zeigt die Eingabe für das Foto der ersten Charakterkarte:

„Erstelle ein Bild einer jungen Frau für eine Charakterkarte einer Diskussionsrunde über Deepfakes. Sie soll eine Influencerin darstellen und ein Smartphone in der Hand halten“.

Die erste Charakterkarte ist die*der Influencer*in, die*der mithilfe von KI-generierten und manipulierten Medien interessante Beiträge erstellen möchte:



Influencer*in

Motivation: Reichweite & Einnahmen steigern

Haltung: eher positiv zu Deepfakes

Pro: kreativ, viral

Kontra: eigener Ruf möglicherweise gefährdet (Identitätsdiebstahl...)

Satz: „Hauptsache es bringt Views!“

Abbildung 5 Charakterkarte Influencer*in

Die zweite Rolle, die vergeben wurde, ist die*der besorgte Bürger*in, bei der es im Wesentlichen darum geht, die Gefahren und Risiken dieser Technologie zu betonen.



Besorgte*r Bürger*in

Motivation: Risiken betonen

Haltung: ganz klar gegen Deepfakes

Pro: nichts

Kontra: Manipulation, Betrug, Unsicherheit

Satz: „Man weiß bald gar nicht mehr, was echt ist!“

Abbildung 6 Charakterkarte Besorgte*r Bürger*in

Die dritte Charakterkarte zeigt die*den unwissenden Jugendliche*n, die*der ganz unreflektiert viel Zeit auf sozialen Medien verbringt und Deepfakes als lustig empfindet.



Unwissende*r Jugendliche*r

Motivation: Spaß, Unterhaltung, möchte coole Videos schauen

Haltung: positiv, unreflektiert

Pro: witzig, cool

Kontra: versteht Risiken kaum

Satz: „Ist doch lustig, wenn Katzen Teig kneten.“

Abbildung 7 Charakterkarte Unwissende*r Jugendliche*r

Die vierte Rolle, die anschließend vergeben wurde, zeigt die*den Geschäftsführer*in, die*der sich als einzige Person sowohl Argumente dafür als auch dagegen überlegen sollte.



Geschäftsführer*in

Motivation: viele Nutzer*innen, Vertrauen halten

Haltung: neutral zu Deepfakes

Pro: kreative Tools

Kontra: Fakes gefährden Glaubwürdigkeit

Satz: „Innovation ja, aber kontrolliert.“

Abbildung 8 Charakterkarte Geschäftsführer*in

Die letzte Charakterkarte zeigt die*den Bürgermeister*in einer großen Stadt, die*der betonen sollte, dass Deepfakes für die Politik sehr gefährlich werden können.



Bürgermeister*in

Motivation: Ruf der Stadt schützen,
Vertrauen der Bürger behalten

Haltung: eher gegen Deepfakes

Pro: evtl. für gekennzeichnete Satire

Kontra: Gefahr für öffentliche Ordnung,
persönliche Rufschädigung

Satz: „Gefälschte Videos können eine ganze
Stadt verunsichern.“

Abbildung 9 Charakterkarte Bürgermeister*in

Den Abschluss bildeten eine gemeinsame Reflexion der Diskussion sowie die Beantwortung eines digitalen Fragebogens zur Evaluierung der beiden Unterrichtseinheiten.

Am darauffolgenden Tag wurden diese Unterrichtseinheiten in der Parallelklasse nochmal durchgeführt. Das Klassenzimmer war, genauso wie am Vortag, länglich und schmal und es waren 17 Schüler*innen anwesend. Da die Zeit zwischen den beiden Doppelstunden zu kurz war, um genau zu evaluieren, wurde die gleiche Sequenz nochmal unterrichtet mit nur einer Änderung bei der dritten Aufgabe. Aus diesem Grund zeigt die folgende Unterrichtsmatrix nur die dritte Zeile.

Nr.	Thema, Aufgabenstellung	Medien, Hilfsmittel	Sozialform, Aktivität	Zeit
3	Mindmap erstellen Die Schüler*innen tippen über Mentimeter ihre Vorerfahrungen ins Smartphone und über den Beamer werden die Antworten an die Wand projiziert.	Mentimeter, Beamer	Eintippen ins Smartphone	10-20 Min.

Tabelle 3 Änderung erster Durchgang

Das hatte den Vorteil, dass viele unterschiedliche Gedanken eingetippt wurden. Als Beispiele können Manipulation, Betrug und Fake News genannt werden. Der wesentliche Nachteil war aber, dass es dann schwierig war, genauer auf die Begriffe einzugehen, da die Begriffe oder Phrasen anonym eingetippt wurden.

Die Antworten im Fragebogen lieferten sehr viel positives Feedback. „Ich fand die Informationen nützlich und es war geschickt ein KI generiertes Video zu zeigen um zu schauen woran man es erkennt“. Dieses Beispiel zeigt, dass die praktische Anwendung der Erkennungsmerkmale des theoretischen Vortrags anhand von KI-generierten Videos gut angekommen ist. Außerdem hat auch vielen Schüler*innen das Hineinversetzen in eine bestimmte Rolle gefallen: „Mir hat die Gruppenarbeit mit den Rollen gefallen es wurde alles sehr gut erklärt und es war verständlich“ Auf der anderen Seite gab es aber auch negative Kritik, da dieses Thema eventuell schon in anderen Fächern behandelt wurde: „Deepfakws und fake news werden in fast jeder klasse unterrichtet, weshalb dies langweilig wird und man alles schon kennt“. Im Großen und Ganzen haben die Unterrichtsstunden den Schüler*innen aber gut gefallen. Außerdem erkennt man am Beispiel „Gut gefallen hat mir, dass man lernt, wie man Deepfakes erkennt und warum sie gefährlich sind.“, dass auch die Medienkompetenz hinsichtlich Deepfakes gefördert werden kann, da dabei die kritische Auseinandersetzung mit den Wirkungsweisen und auch mit den Inhalten sehr wichtig ist. Das wurde im Kapitel 2.4 Medienkompetenz genauer erläutert.

Es werden folgende Veränderungen am Ende des ersten Durchgangs vorgenommen:

- Die Erklärung des Dagstuhl-Dreiecks zu Beginn der Unterrichtseinheiten wird ergänzt, um den Schüler*innen unterschiedliche Sichtweisen auf technologische Themen zu zeigen.
- Mindmap wird an der Tafel und nicht über Mentimeter erstellt, damit genauer auf Wortmeldungen eingegangen werden kann.
- Ein kleiner Ausschnitt eines Zeitungsartikels wird durch einen längeren sehr aktuellen ersetzt.
- Die Anzahl an Rollen wird von fünf auf vier Charakterkarten reduziert und die*der Geschäftsführer*in entfernt. Damit werden die Sprechanteile erhöht und die*der Firmenchef*in, die*der eine sehr schwierige neutrale Haltung gegenüber Deepfakes einnehmen musste, wird gestrichen.

6.4 Durchgang 2

Zwischen dem ersten und dem zweiten Durchgang ist viel Zeit vergangen und die Unterrichtsstunden wurden dazwischen anhand des erhaltenen Feedbacks in der Universität, aber auch von den Lehrpersonen und Schüler*innen in der Schule, verändert. Anschließend wurde am 22. Mai 2026 im Gymnasium im 18. Wiener Gemeindebezirk in der Klostergasse unterrichtet. Allerdings wurden die Unterrichtsstunden während des Umbaus der Schule absolviert und aus diesem Grund befanden sich die Klassen im ehemaligen orthopädischen Spital an der Grenze zwischen Hernals und Währing. Eine Studienkollegin ermöglichte mir das Unterrichten in einer fünften Klasse, das entspricht der 9. Schulstufe.

Es handelte sich um eine Doppelstunde am Freitagnachmittag ohne offizielle Pause dazwischen. Zuerst wurde das in der Informatikdidaktik sehr wichtige Dagstuhl-Dreieck erklärt, um den Schüler*innen drei wesentliche Perspektiven auf das Thema Deepfakes zu zeigen. Wie funktioniert diese Technologie? Wie nutzt man diese Technologie? Welche gesellschaftlichen und kulturellen Auswirkungen bringt diese Technologie mit sich? Bei Deepfakes reicht es nämlich nicht zu verstehen, *wie* künstliche Intelligenz Videos, Bilder oder Stimmen manipuliert. Es sollte auch erkannt werden, welche Folgen das für Menschen, Medien, Politik oder Vertrauen in Informationen hat. Gleichzeitig spielt die praktische Nutzung eine Rolle, zum Beispiel wie man Deepfakes erkennt oder verantwortungsvoll mit KI umgeht. Die Inhalte der

beiden Unterrichtsstunden, vor allem aber das EduLARP, konzentrieren sich hauptsächlich auf die dritte Frage.

Anschließend wurden wieder auf drei PowerPoint Folien reale und KI-generierte Bilder gezeigt und die Schüler*innen mussten erkennen, welches keine reale Abbildung darstellt. Es haben alle Jugendliche mittels Handzeichen die generierten Bilder erkannt. Um die Einführung in das Thema abzuschließen, fertigten wir eine Mindmap an der Tafel an. In die Mitte wurde der Begriff „Deepfakes“ geschrieben und dann konnten die Schüler*innen alle Gedanken zu diesem Thema aufschreiben. Nach dem Aufzeigen durften sie zur Tafel kommen und konnten ihren Kommentar selbst in Stichworten notieren. Folgende Gedanken wurden dabei an die Tafel geschrieben:

- Cyber Mobbing
- Unrealistisch
- Social Media
- KI
- Fake News
- Spezifisch: Grok

Sehr interessant war, dass die App „Grok“ genannt wurde. Damit wurden auch Videos für diese Unterrichtsstunden erstellt und die App selbst stand in den letzten Monaten immer wieder in unterschiedlichen Zeitungen unter Kritik. Die App wurde weniger zensiert und es gibt viele unpassende Antworten, aber auch sehr problematische Videos können mithilfe von Grok erstellt werden.

Nach dem kurzen theoretischen Vortrag wurde der Artikel „Deepfakes aufdecken: So arbeiten IT-Forensiker“ aus der Zeitung „Der Standard“ vom 29.03.2026 gemeinsam gelesen. Das wurde zwischen dem ersten und zweiten Durchgang der Unterrichtsstunden verändert, da dieser Artikel sehr aktuell und wichtig in diesem Zusammenhang ist. Wenn Deepfakes immer besser werden, können sie vielleicht bald nur noch durch forensische Prüfung oder genaue Recherche der Herkunft identifiziert werden. Dann helfen Erkennungsmerkmale wie eine monotone Stimme eventuell nicht mehr.

Anschließend wurde das Video über die Probleme der Politik aufgrund von Deepfakes angesehen. Es gab dazu nur eine einzige Wortmeldung. Ganz anders war

das allerdings bei den KI-generierten Bildern und Videos von der Lehrperson selbst. Die Medien wurden über den Beamer gezeigt und sehr viele Schüler*innen meldeten sich zu Wort. Sie analysierten die Stärken und Schwächen dieser Deepfakes und waren fasziniert von der Technologie. Als Beispiel kann dabei genannt werden, dass eine goldene Schere in der Hand sofort zu einer Gabel wurde, da sich die Szene in einem Restaurant abspielt und die Künstliche Intelligenz weiß, dass in den meisten Fällen beim Essen einer Pizza eine Gabel in der Hand sein sollte.

Das EduLARP zum Abschluss war eine sehr gelungene Übung. Im Vergleich zum ersten Durchgang setzten sich die einzelnen Rollen der Diskussionsrunde zuerst zusammen und alle mussten sich gemeinsam Argumente überlegen. Im zweiten Schritt durften sie auch das Handy und künstliche Intelligenz zur Recherche verwenden. Die Schüler*innen haben auch während der Diskussionsphasen viel miteinander gesprochen und ihre Rollen vertreten. Es war vermutlich auch ein Vorteil, dass diesmal nur vier statt fünf Rollen vergeben wurden und es aus diesem Grund eine kleinere Gruppengröße gab. Außerdem war es Freitagnachmittag, die Jugendlichen hatten keine Sorgen über anstehende Schularbeiten oder Tests am gleichen Tag und zwei Gruppen konnten wir auch außerhalb des Klassenzimmers diskutieren lassen. Somit gab es keine Störungen oder Ablenkungen zwischen den einzelnen Debatten.

Die folgende Tabelle zeigt die Stundenmatrix dieser Einheiten:

Nr.	Thema, Aufgabenstellung	Medien, Hilfsmittel	Sozialform, Aktivität	Zeit
1	Begrüßung, Vorstellung des heutigen Programms, Erklärung des Dagstuhl-Dreieck Die SuS lernen drei wesentliche Perspektiven auf technische Themen kennen.	PowerPoint Präsentation, Beamer	Fragen an die Schüler*innen, Vortrag der Lehrperson	0-5 Min.
2	KI-generierte Bilder von realen Fotos unterscheiden Die SuS stimmen mittels Handzeichen ab, welche Bilder real sind. Anschließend wird analysiert, woran man die Unterschiede bemerkt.	PowerPoint Präsentation, Beamer	Fragen an die Schüler*innen, Vortrag der Lehrperson	5-10 Min.

3	Mindmap erstellen Die Schüler*innen melden sich durch Handzeichen und schreiben alle Gedanken zum Thema Deepfakes an die Tafel.	Tafel	Aufschreiben an die Tafel	10-20 Min.
4	Vortrag Die Lehrperson erklärt ausführlich das Thema, die Erkennungsmerkmale und die Technologie und ergänzt somit die Gedanken der Schüler*innen.	PowerPoint Präsentation, Beamer	Vortrag der Lehrperson	20-30 Min.
5	Wie arbeiten IT-Forensiker*innen Es wird gemeinsam ein Zeitungsartikel aus der Zeitung „Der Standard“ gelesen. Ein*e Schüler*in liest den Text laut vor und anschließend wird der Abschnitt besprochen.	Zeitungsartikel, Beamer	Gemeinsames Lesen, Diskutieren im Plenum	30-45 Min.
6	Video – Deepfakes Am Beamer wird ein Video gezeigt, in dem beschrieben wird, wie Menschen in der Politik Probleme aufgrund von Deepfakes bekommen können. Anschließend wird im Plenum über das Video gesprochen.	Beamer, Lautsprecher	Aufmerksames Anschauen eines Videos, Diskutieren im Plenum	45-55 Min.
7	Deepfakes analysieren Die Schüler*innen analysieren KI-manipulierte bzw. generierte Videos und sprechen über die Erkennungsmerkmale.	PowerPoint Präsentation, Beamer	Diskutieren im Plenum	55-65 Min.
8	EduLARP Vorbereitung Schüler*innen, die gleiche Rollen zugeteilt bekommen haben, setzen sich in Kleingruppen zusammen und überlegen sich Argumente bezüglich Deepfakes. Das Handy und künstliche Intelligenz dürfen sie nur in den letzten Minuten verwenden.	Zettel, Stift, Handy, Künstliche Intelligenz	Gruppenarbeit	65-80 Min.
9	EduLARP Diskussion In einer Diskussionsrunde diskutieren Personen aus unterschiedlichen Branchen	Zettel	Gruppendiskussion	80-95 Min.

	und Lebensphasen über Deepfakes. Alle wollen ihren Standpunkt und ihre Meinung zu dem Thema einbringen und versuchen, die anderen davon zu überzeugen. Jede*r Schüler*in begibt sich dazu in eine vorgegebene Rolle mit vorgegebenen Zielen und überlegt sich vor der Diskussion passende Argumente.			
10	Resümee und Auftrag für nächste Einheit Zum Abschluss werden Erfahrungen aus den einzelnen Diskussionen im Plenum gesammelt und das Ausfüllen des Fragebogens in der nächsten Einheit wird erklärt.	-	Besprechen im Plenum	95-100 Min.

Tabelle 4 Unterrichtsmatrix zweiter Durchgang

Die Schüler*innen wurden eine Woche später zu den Unterrichtsstunden befragt und sowohl Wissensfragen als auch Fragen als Feedback mussten die Jugendlichen ausfüllen. Es wurden sieben Tage zwischen dem Unterricht und dem Ausfüllen des Fragebogens eingeplant, damit festgestellt werden kann, was die Jugendlichen eine Woche später noch wissen.

Die Auswertung des Fragebogens zeigt insgesamt gesehen, dass soziale Medien intensiv genutzt werden, aber auch das Problembewusstsein gegenüber digitalen Inhalten wie Deepfakes steigt. Die Schüler*innen nutzen soziale Medien überwiegend mehrere Stunden täglich, teilweise sehr intensiv, reflektieren diese Nutzung jedoch häufig kritisch. Antworten wie beispielsweise „4 Stunden Ja ich will die Zeit reduzieren“ zeigen, dass zwar ein Problembewusstsein vorhanden ist, dieses aber nicht immer zu konkreten Veränderungen führt, sondern die Jugendlichen dennoch sehr viel Zeit vor den Bildschirmen verbringen.

Bei den möglichen Erkennungsmerkmalen von KI-generierten oder manipulierten Medien dominiert vor allem die visuelle Strategie. Die Jugendlichen achten besonders auf Merkmale wie Fingeranzahl, Augen, Haut oder Hintergrund und bewerten Inhalte als falsch, wenn sie ganz perfekt oder eher unnatürlich wirken. Außerdem wird auch die Stimme als Merkmal genannt, zum Beispiel wenn sie besonders monoton wirkt.

Bei den Gefahren von Deepfakes werden von den Schüler*innen vor allem Betrug, Rufschädigung von Personen und Fake News bzw. Desinformationen genannt. Viele gehen davon aus, dass das Erkennen von Deepfakes immer schwieriger wird. Diese Erkenntnisse, die sich die Schüler*innen gemerkt haben, zeigen, dass die Medienkompetenz der Jugendlichen in den beiden Unterrichtsstunden gefördert wurde.

Die beiden Einheiten werden insgesamt sehr positiv bewertet. Besonders hervorgehoben werden interaktive Elemente wie Gruppenarbeit und Diskussionen. Die meisten würden den Unterricht weiterempfehlen, vor allem aufgrund der Aktualität und Relevanz des Themas sowie des persönlichen Lerngewinns. Kritikpunkte beziehen sich nur vereinzelt auf die Komplexität der Inhalte oder auch auf die Dauer, da sich ein*e Schüler*in weniger als zwei Unterrichtsstunden mit diesem Thema auseinandersetzen wollte. Die meisten Jugendlichen reagierten aber positiv auf die Stunden, wie zum Beispiel folgende Antwort zeigt: „guter Aufbau: mit Gruppenarbeit, MindMap, Fake oder echt,... / verändern: eigentlich gar nichts“. Somit müssen nur noch Kleinigkeiten angepasst werden. Folgende Veränderungen werden am Ende des zweiten Durchgangs vorgenommen.

- Das Dagstuhl Dreieck wird durch das Frankfurt Dreieck ersetzt, da es detaillierter ist.
- Die Schüler*innen müssen bei der Vorbereitung auf das EduLARP alle Argumente aufschreiben, damit sie während der Diskussion auf ihre Notizen als Unterstützung zurückgreifen können.

6.5 Durchgang 3

Zwischen den Durchgängen zwei und drei verging nicht viel Zeit, bereits am 18. Juni 2026 erfolgte die Durchführung der finalen Unterrichtsstunden. Es wurde wieder in einer fünften Klasse der allgemeinbildenden höheren Schule in der Klostergasse im 18. Bezirk in Wien unterrichtet. Es war sehr spät, da der Unterricht erst um 15:25 begonnen und somit die Doppelstunde bis 17:05 gedauert hat und es waren nur acht Schüler*innen anwesend, von denen haben sechs Jugendliche eine Woche später den Fragebogen ausgefüllt. Es wurden nur noch zwei kleinere Aspekte verändert und der Großteil der Planungen vom zweiten Durchgang wurden übernommen.

Das Dagstuhl Dreieck wurde durch das Frankfurt Dreieck ersetzt, da dieses sehr detailliert drei wesentliche Blickwinkel auf das Thema Deepfakes beschreibt. Die Abbildung 11 zeigt das Dreieck, welches mit den Schüler*innen zu Beginn der Einheiten besprochen wurde. Diese beiden Unterrichtsstunden befassen sich vor allem mit der grünen Seite des Dreiecks: gesellschaftliche und kulturelle Wechselwirkungen.

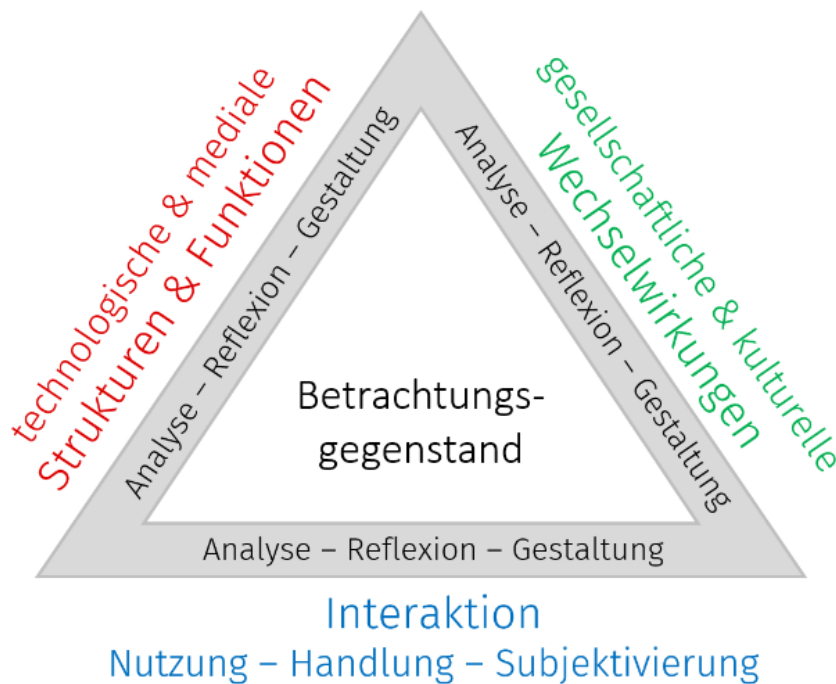


Abbildung 10 Frankfurt Dreieck (Brinda, T. et al. 2019: 27)

Die nachfolgende Tabelle zeigt die beiden Veränderungen genauer. Neben dem Frankfurt Dreieck wurde auch die EduLARP Vorbereitung verändert. Die Schüler*innen mussten ihre Argumente, die sie sich während der Vorbereitungszeit überlegten und anschließend recherchierten, aufschreiben. Das sorgte für eine bessere Diskussion, da alle Jugendlichen immer wieder auf ihre Stichwortzettel blickten und somit konnten alle Argumente in der Debatte des EduLARPs verwendet werden.

Nr.	Thema, Aufgabenstellung	Medien, Hilfsmittel	Sozialform, Aktivität	Zeit
1	Begrüßung, Vorstellung des heutigen Programms, Erklärung des Frankfurt Dreieck Die SuS lernen drei wesentliche Perspektiven auf technische Themen kennen.	PowerPoint Präsentation, Beamer	Fragen an die Schüler*innen, Vortrag der Lehrperson	0-5 Min.
8	EduLARP Vorbereitung Schüler*innen, die gleiche Rollen zugeteilt bekommen haben, setzen sich in Kleingruppen zusammen, überlegen sich Argumente bezüglich Deepfakes und müssen sie notieren. Das Handy und künstliche Intelligenz dürfen sie nur in den letzten Minuten verwenden.	Zettel, Stift, Handy, Künstliche Intelligenz	Gruppenarbeit	65-80 Min.

Tabelle 5 Veränderungen dritter Durchgang

Bei der Analyse der erhaltenen Antworten und des Feedbacks eine Woche nach den Unterrichtsstunden wird deutlich, dass die Schüler*innen den Unterricht häufig als abwechslungsreich und interaktiv wahrgenommen haben. Das EduLARP wird dabei besonders positiv hervorgehoben. Aussagen wie „Ich fand die abwechslungsreiche Gestaltung des Unterrichts sehr gut“ oder „Ja, weil er mit der Mindmap am Anfang und dem Roleplay am Ende eine gute Abwechslung hatte“ verdeutlichen, dass den Jugendlichen die unterschiedlichen Methoden besonders positiv in Erinnerung geblieben sind.

Bei den Erkennungsstrategien haben viele Schüler*innen visuelle Merkmale genannt, wie beispielsweise die Anzahl der Finger. Die folgende Antwort zeigt ein Beispiel für eine positive Rückmeldung diesbezüglich: „Das uns gezeigt wurde wie wir erkennen wenn es KI generiert ist“. Somit konnte die Medienkompetenz hinsichtlich Deepfakes verbessert werden. Das Thema wurde von den Jugendlichen auch als sehr relevant gesehen. Probleme wie Falschinformationen, Mobbing und Rufschädigung wurden genannt.

Ein*e Schüler*in hat die Unterrichtsstunden als langwierig empfunden. Das könnte daran liegen, dass eine Gruppe beim EduLARP früher fertig war und dann aber noch Zeit zur Verfügung hatten. Auch wird vereinzelt erwähnt, dass das Thema bereits

teilweise bekannt war, wodurch weniger neue Impulse von den Jugendlichen wahrgenommen wurden. Insgesamt gesehen, fallen die Empfehlungen des Unterrichts allerdings positiv aus, wie folgende Antworten zeigen:

- „Ja, weil diese Unterrichtsstunden mich viel über dieses Thema unterrichtet haben“
- „Ja, weil er mit der minimal am Anfang und dem roleplay am Ende eine gute Abwechslung hatte und es nicht langweilig war.“

Somit konnte im Zuge dieser drei Durchgänge die Doppelstunde entsprechend überarbeitet und verbessert werden und es wurde versucht, die Medienkompetenz der Jugendlichen hinsichtlich Deepfakes zu verbessern. Insgesamt legen die Ergebnisse der Befragung nahe, dass das entwickelte Unterrichtskonzept geeignet ist, die Medienkompetenz von Schüler*innen einer AHS-Oberstufe hinsichtlich Deepfakes nachhaltig zu fördern. Der zeitliche Abstand zwischen Unterricht und Befragung unterstreicht dabei die Aussagekraft der Ergebnisse, da wesentliche Inhalte und Kompetenzen auch eine Woche nach den Unterrichtseinheiten noch reflektiert und angewendet werden konnten. Die Schüler*innen können Deepfakes nach den Unterrichtseinheiten sehr präzise definieren und ihre Entstehung erklären. Die Antworten „Das sind mit KI erstellte Bilder, Videos oder Audios, die echt wirken.“ Und „Die KI lernt aus vielen Daten und erstellt neue Inhalte.“ zeigen diese Erkenntnis. Außerdem lernten die Schüler*innen auch besonders kritisch zu sein, bevor man Inhalte auf sozialen Medien glaubt.

Durch das Redesign der entwickelten Unterrichtsstunden wurden Probleme an der Planung erkannt, analysiert, verändert und erneut ausprobiert. Am Ende sind zwei Unterrichtsstunden entstanden, die mithilfe von Feedback von vielen Schüler*innen und mehreren Lehrpersonen und Studienkolleg*innen so gut wie möglich optimiert wurden und die sehr wichtige Thematik der KI-generierten bzw. manipulierten Medieninhalte abwechslungsreich behandeln.

7 DISKUSSION

Im Theorieteil dieser Arbeit wurden viele Gefahren und Risiken von Deepfakes genannt. Darüber wurde in den letzten Jahren viel geforscht. Das könnte an der rasanten technologischen Entwicklung liegen. Einigen Menschen gehen die Fortschritte der künstlichen Intelligenz zu schnell und das birgt viele Gefahren. Es wurden in dieser Arbeit einige Risiken genannt, jedoch ist die Technologie nicht mehr aufzuhalten. Man muss daran forschen, wie man sie nutzen kann. Auf der anderen Seite sollte man aber nicht weniger daran arbeiten, die Gefahren zu minimieren.

Auch viele Jugendliche erkennen bereits die großen Gefahren von Deepfakes. Einige der befragten Schüler*innen zeigten sich sehr reflektiert und empfinden es nicht als positiv, kreative Videos schnell erstellen zu können. Ein möglicher Grund dafür könnte sein, dass viele Menschen die KI-generierten bzw. manipulierten Medien auf sozialen Netzwerken nicht erkennen. Beispielsweise werden den Jugendlichen Videos von älteren Verwandten zugeschickt, die fasziniert vom Inhalt sind und in Wirklichkeit hat das Video nie stattgefunden, da es sich um einen Deepfake handelt. Das könnte bei den Schüler*innen Bedenken hervorrufen. Aber auch die Gefahren von Identitätsdiebstahl und das Verbreiten von Fake News besorgen die junge Generation. Der Einfluss von KI-generierten und manipulierten Medien auf die Politik könnte darauf einen geringeren Einfluss haben, da es sich bei der Befragung ausschließlich um Jugendliche handelt und sie noch bei keiner größeren politischen Wahl mitentscheiden durften.

Bei der Durchführung der Unterrichtsstunden sind sowohl die Analyse der selbst erstellten Deepfakes als auch das EduLARP besonders gut bei den Schüler*innen angekommen. Das könnte daran liegen, dass in vielen Unterrichtsstunden vermehrt Theorie ohne interaktive Übungen vermittelt wird und die beiden Übungen eine Abwechslung dazu darstellten. Vor allem die Schüler*innen des zweiten Durchgangs am Freitagnachmittag wollten die kurzen Videos immer wieder sehen und erkannten jedes Mal neue Erkennungsmerkmale der KI-Generierung. Auch beim EduLARP gab es in dieser Doppelstunde hitzige und ausführliche Diskussionen. Das könnte daran liegen, dass es die letzten beiden Stunden vor dem Wochenende waren und sich die Schüler*innen somit besser in die Rolle begeben konnten. Wenn am gleichen Tag eine Schularbeit oder ein Test geschrieben werden muss, dann könnte es um einiges

schwieriger sein, sich auf das EduLARP und die zugeteilte Rolle vollständig einzulassen.

8 FAZIT UND AUSBLICK

Diese Arbeit setzt sich mit der Beantwortung von drei Forschungsfragen auseinander. Dadurch soll das Thema Deepfakes von drei wichtigen Blickpunkten aus betrachtet werden. Neben der Auseinandersetzung mit wissenschaftlicher Literatur und den Perspektiven von Jugendlichen einer AHS-Oberstufe in Wien befasste sich diese Arbeit auch mit zwei umgesetzten Unterrichtsstunden zu diesem Thema. Zum Abschluss werden im Fazit die Antworten auf die drei Forschungsfragen zusammengefasst:

Welche Risiken von Deepfakes werden in der Literatur beschrieben?

Wesentliche Probleme von Deepfakes sind die Verbreitung von Falschinformation, Identitätsdiebstahl, Betrug und Manipulation von Wahlen. Beim Betrug ist der Einzeltrick zu erwähnen, bei dem Stimmen nachgeahmt werden können. Vor allem Frauen bemerken, dass ihre eigene Privatsphäre durch Deepfakes in Gefahr sein könnte. In wissenschaftlichen Literaturquellen werden somit viele unterschiedliche Probleme aufgrund der Möglichkeit von Deepfakes beschrieben.

Welche Sichtweisen auf Deepfakes und deren Risiken zeigen sich aus der Perspektive von Oberstufenschüler*innen einer AHS?

Viele der befragten Jugendlichen sehen Deepfakes bereits aus einer sehr reflektierten Perspektive. Mehr als die Hälfte gaben an, dass mindestens zehn Prozent der Bilder und Videos in den sozialen Netzwerken mit künstlicher Intelligenz generiert oder manipuliert sind. Die Oberstufenschüler*innen, die für diese Forschung befragt worden sind, gaben an, dass Betrug, Vertrauensverlust und die Verbreitung von Fake News ihrer Meinung nach ein großes Problem werden und somit eine kritische Sichtweise auf Deepfakes haben und nicht unreflektiert solche Medien erstellen und ansehen.

Inwiefern ist das im Rahmen der Masterarbeit erarbeitete spezifische Unterrichtskonzept für die Zielgruppe der Schüler*innen einer AHS-Oberstufe geeignet, zur Verbesserung von deren Medienkompetenz hinsichtlich Deepfakes beizutragen?

Die Ergebnisse der Evaluierung deuten darauf hin, dass das im Rahmen der Masterarbeit entwickelte Unterrichtskonzept einer Doppelstunde geeignet ist, die Medienkompetenz von Schüler*innen der AHS-Oberstufe im Hinblick auf Deepfakes zu fördern. Die Proband*innen waren eine Woche später in der Lage, Deepfakes fachlich korrekt zu definieren, Erkennungsmerkmale von KI-generierten und manipulierten Medien und deren Gefahren und Risiken zu nennen. Außerdem gab es viele Antworten, die auf eine sehr reflektierte Herangehensweise auf diese heikle Thematik haben.

Aus den Unterrichtserfahrungen der drei Durchgänge können Erkenntnisse abgeleitet werden. Zur Ermittlung des Vorwissens der Schüler*innen eignete sich das gemeinsame Verfassen einer Mindmap an der Tafel am besten. Mentimeter hatte zwar den Vorteil, dass es mehr Meldungen gab, jedoch kam es dann zu keinen genaueren Gesprächen. Eine positive Überraschung gab es beim Besprechen der selbst erstellten Deepfakes. Die Schüler*innen konnten einige Erkennungsmerkmale feststellen und es haben viele Jugendliche für Wortmeldungen aufgezeigt. Wesentlich war auch die Gruppengröße bei der Diskussion über Deepfakes. Bei diesem EduLARP war es besser, wenn weniger Schüler*innen miteinander diskutierten und dafür mehr Gruppen gebildet worden sind. Aus diesem Grund wurden die Charakterkarten von fünf auf vier unterschiedliche fiktive Personen reduziert. Es wäre vermutlich besser gewesen, gleich von Beginn an kleinere Gruppen zu bilden. Somit konnten einige Erkenntnisse gemacht werden und wertvolle Erfahrungen wurden gesammelt.

Es ist sehr wichtig, dass in diesem Bereich weiterhin geforscht wird, da die Potentiale von Deepfakes vermutlich noch bei weitem nicht alle bekannt sind. Dabei handelt es sich sowohl um nützliche als auch um schädliche Verwendung. Täuschend echte Medien stehen jetzt auf jeden Fall genauso wie früher das Internet den Menschen als neue Technologie zur Verfügung und die Auseinandersetzung mit deren Vor- und Nachteilen wird deshalb auch in Zukunft von großer Relevanz sein.

9 LITERATURVERZEICHNIS

- Beege, B. (2023). Hochschulelehre lernen: Förderung von Lerntransfer durch authentische Aktivitäten – eine Design-Based-Research-Studie (1st ed.). Münster: Waxmann Verlag.
- Bowman, S. L., & Standiford, A. (2015). Educational larp in the middle school classroom-a mixed method case study. *International Journal of Role-Playing*.
- Brinda, T., Brüggem, N., Diethelm, I., Knaus, T., Kommer, S., Kopf, C., ... & Weich, A. (2019). Frankfurt-Dreieck zur Bildung in der digital vernetzten Welt. In *Informatik für alle* (S. 25-33). Gesellschaft für Informatik.
- Farid, H., & Schindler, H. J. (2020). Eine Bedrohung für Demokratie und Gesellschaft.
- Folkers, A. (2019). Steuerung eines autonomen Fahrzeugs durch Deep Reinforcement Learning. Springer Spektrum.
- Gess, C., Rueß, J. & Deicke, W. (2014). Design-based Research als Ansatz zur Verbesserung der Lehre an Hochschulen – Einführung und Praxisbeispiel. In: *Qualität der Wissenschaft*.
- Gosch, N. (2026). KI-Bilder: Wer haftet für Bikini-Fotos? In: *Kleine Zeitung 1.Februar 2026*.
- Kaplan, J., & Lenz, G. (2017). Künstliche Intelligenz (1. Auflage.). Frechen: mitp Verlags GmbH & Co. KG.
- Karaboga, M., Frei, N., Puppis, M., Vogler, D., Raemy, P., Ebbers, F., ... & Rovelli, S. (2024). Deepfakes und manipulierte Realitäten. *Technologiefolgenabschätzung und Handlungsempfehlungen für die Schweiz*. TA-SWISS Publikationsreihe (Hrsg.): TA, 81, 2024.
- Kleemann, A. (2023). Deepfakes – Wenn wir unseren Augen und Ohren nicht mehr trauen können. SWP- Aktuell.
- Lehrplan RIS (2026)
<https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10008568>. (entnommen am 12.06.2026).

- Liu, Y., Padmanabhan, B., & Viswanathan, S. (2025). From Deception to Perception: The Surprising Benefits of Deepfakes for Detecting, Measuring, and Mitigating Bias.
- Lukjantschuk, J., Dambacher, N., & Klenk, N. (2025). Wie kann in Zeiten von Deepfakes das Original von Fälschungen unterschieden werden?. In: *Smarte Unternehmen, digitale Gesellschaft–Wie KI & Co. das Leben und Arbeiten der Zukunft gestalten können* (t) en, 45.
- Kusche, I. (2024). Politische Öffentlichkeit, Desinformation und das Problem von Deepfakes. In A. Bahr & G. Fröhlich (Hrsg.), *Ain't Nothing Like the Real Thing? Formen und Funktionen medialer Artefakt-Authentifizierung* (S. 149–168). transcript.
- Mayring, P. (2022) Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim: Beltz Verlagsgruppe.
- Möble, T., Fleps, T., Gerstner, D., Keil, V., & Wachter, E. (o.D.) Kriminologische Forschung Baden-Württemberg (KriFoBW). Erste Ergebnisse der landesweiten Sicherheitsbefragung-vom Enkeltrick bis KI.
- Robinson, R. B., Johansson, K., Fey, J. C., Márquez Segura, E., Back, J., Waern, A., Lynne Bowman, S. & Isbister, K. (2023). Edu-larp@ CHI. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (S. 1-5).
- Vogler, D., Rauchfleisch, A. & De Seta, G. (2024). Wahrnehmung von Deepfakes in der Schweizer Bevölkerung. In: Deepfakes und manipulierte Realitäten. Technologiefolgenabschätzung und Handlungsempfehlung für die Schweiz. Universität Zürich.
- Von Gottberg, Joachim (2024) Deepakes Fiktive und alternative Wahrheiten. In: *JMS-Report 6/2024*.
- Von Rein, A. (Ed.). (1996). Medienkompetenz als Schlüsselbegriff. Bad Heilbrunn: Klinkhardt.
- Zuzana, V., Roman, M., Kateřina, T., Martin, Ž., Pavel, S., Tsitouras, C. & Simos, T. (2024). Computer environment of virtual eduLARP game. *AIP Conference Proceedings*.

10 Hilfsmittel und ethische Aspekte

Im Rahmen der Erstellung dieser Masterarbeit wurde die KI-gestützte Softwareanwendung ChatGPT-5 unterstützend als Hilfsmittel eingesetzt. Der Fragebogen und dessen Struktur wurde ohne künstliche Intelligenz erstellt, jedoch bei der Auswertung wurden die Kategorien und Zuordnungen von ChatGPT vorgeschlagen. Danach wurden die Ergebnisse, wenn es nötig war, überarbeitet. Hilfreich waren die Prompts zur Auswertung des Fragebogens, zur Erstellung der Bilder der Charaktere für das EduLARP und zur sprachlichen Formulierung mancher Sätze. Auch bei der Interpretation der Daten hat KI geholfen. ChatGPT war nicht hilfreich für die Recherche von wissenschaftlicher Literatur, diesbezüglich wurden u:search und Google Scholar verwendet. Die inhaltliche Verantwortung der Arbeit liegt vollständig bei dem Verfasser.

Die Fragebögen wurden anonym ausgefüllt und es wurde eine Einverständniserklärung der Direktor*innen eingeholt, damit die Befragung durchgeführt werden konnte. Auch die Schüler*innen wurden über die Freiwilligkeit und die Anonymität der Antworten aufgeklärt.

11 ANHANG

Einverständniserklärung für Direktor*innen



Einverständnis Fragebogen

Sehr geehrte Frau*Herr Direktor*in!

Im Rahmen meiner Masterarbeit an der Universität Wien wurde ein Fragebogen für Schüler*innen der Oberstufe erstellt, der sich mit Deepfakes (KI-generierte und manipulierte Medien) befasst. Die Teilnahme ist freiwillig und die Daten werden anonym erhoben und in der wissenschaftlichen Arbeit werden keine Namen der Schüler*innen genannt, sondern nur der Name der Schule und die Schulstufe beschrieben. Die erhobenen Daten werden ausschließlich für wissenschaftliche Zwecke verwendet. Darüber werden auch die Schüler*innen informiert.

Damit die Schüler*innen Ihrer Schule an dieser Studie teilnehmen können, benötige ich Ihr Einverständnis:

- ☐ Ja, ich stimme zu.
- ☐ Nein, ich stimme nicht zu.

Unterschrift:

Information für Schüler*innen

Im Rahmen meiner Masterarbeit an der Universität Wien wurde ein Fragebogen für Schüler*innen der Oberstufe erstellt, der sich mit Deepfakes (KI-generierte und manipulierte Medien) befasst. Die Teilnahme ist freiwillig und die Daten werden anonym erhoben und in der wissenschaftlichen Arbeit werden keine Namen genannt, sondern nur der Name der Schule und die Schulstufe beschrieben. Die erhobenen Daten werden ausschließlich für wissenschaftliche Zwecke verwendet.

Vielen Dank für die Teilnahme!

Antworten und Kategorien der Fragebögen 1. Durchgang

1. Wie viel Zeit verbringst du ungefähr pro Tag auf sozialen Medien? Möchtest du diese Zeit reduzieren?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Geringe Nutzungsdauer	Die Person verbringt bis zu 3 Stunden pro Tag auf sozialen Medien.	„2-3 Stunden, Ja“	Alle Antworten mit einer Nutzungsdauer von bis einschließlich 3 Stunden pro Tag werden dieser Kategorie zugeordnet.
2: Mittlere Nutzungsdauer	Die Person verbringt mehr als 3 bis einschließlich 6 Stunden pro Tag auf sozialen Medien.	„4 Stunden Ja ich will die Zeit reduzieren“	Alle Antworten mit einer Nutzungsdauer von mehr als 3 bis einschließlich 6 Stunden pro Tag werden dieser Kategorie zugeordnet.
3: Hohe Nutzungsdauer	Die Person verbringt mehr als 6 Stunden pro Tag auf sozialen Medien.	„7-9 Stunden, ja“	Alle Antworten mit einer Nutzungsdauer von mehr als 6 Stunden pro Tag werden dieser Kategorie zugeordnet. Bei qualitativen Angaben ohne genaue Zeitangabe (z. B. „Großteil des Tages“) erfolgt die Zuordnung zu dieser Kategorie, wenn eine sehr intensive Nutzung beschrieben wird.

Kategorie 1: Geringe Nutzungsdauer

- „1 stunde, nein“
- „3 Stunden, ja ich möchte es reduzieren“
- „2-3 Stunden, Ja“
- „Ca.3 Stunden pro Tag, es könnte schlimmer sein.“
- „3h“
- „So 2h und ich finde persönlich, dass es zu viel ist.“
- „2-3 Stunden“
- „2 stunden“
- „2 stunden“

Kategorie 2: Mittlere Nutzungsdauer

- „4 Stunden Ja ich will die Zeit reduzieren“
- „5 stunden“
- „5 stunden“
- „Ich verbringe um die 4 Stunden auf Social Modus und ja ich will es reduzieren“
- „Ich verbringe viel auf Social Media, meine Bildschirmzeit liegt so zwischen 5-7 Stunden. Ich möchte es auf jeden Fall reduzieren.“
- „5 stunden ja ich möchte diese zeit reduzieren“
- „5 Stunden mindestens. ja das will ich reduzieren“
- „4“
- „Ungefähr4 Stunden. Ist mir p3rsönlich zu viel so würd ich schon gern sie reduzieren 😞“
- „ca 6 Stunden“
- „4-5 , am wochendende mehr (10) ,nein“
- „4 ja“

Kategorie 3: Hohe Nutzungsdauer

- „12.3 nein ich wills nicht reduzieren“

- „Um die 7 bis 12 std“
- „7-8h“
- „Ich verbringe täglich mehrere Stunden wenn nicht sogar den Großteil des Tages“
- „9.30 std ich würd schon gern es etwas reduzieren“
- „10 stunden“
- „Ca 6/7 Stunden, nein ich möchte es nicht reduzieren“
- „8 stunden, nicht wircklich“
- „7-9 Stunden, ja“
- „Ca. 7“
- „So 6-7 Stunden. Eigentlich schon aber es ist schwer.“
- „6-7 Stunden“
- „~6-7 stunden ,ja ich will weniger haben“

2. Hast du schon einmal Deepfakes erkannt? Woran hast du es bemerkt und auf welcher Plattform hast du dieses Medium entdeckt?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Erkennung durch visuelle Auffälligkeiten	Der Deepfake wird aufgrund unrealistischer Darstellungen oder sichtbarer Bildfehler erkannt.	„Ja weil sie unrealistisch aussehen“	Alle Antworten, in denen unrealistische Bewegungen, Gesichter, Hände, Augen, Farben oder andere sichtbare Auffälligkeiten als Erkennungsmerkmal genannt werden, werden dieser Kategorie zugeordnet.
2: Erkennung durch technische Merkmale oder Qualität	Der Deepfake wird anhand von KI-typischen Merkmalen, schlechter Qualität, Stimme oder Mimik erkannt.	„Öfters. Deepfakes sind meist an der stimmt oder der Mimik erkennbar“	Alle Antworten, in denen die Erkennung auf Qualität, KI-Generierung, Stimme, Mimik, Kontrast, Schattierung oder ähnliche technische Merkmale zurückgeführt wird, werden dieser Kategorie zugeordnet.
3: Erkennung durch Vergleich, Recherche oder Kontextwissen	Der Deepfake wird durch Vergleiche mit echten Inhalten, Recherche oder logische Unstimmigkeiten erkannt.	„Tiktok, bei dem Video wusste man es, weil es einen Vergleich gab zum echten“	Alle Antworten, in denen die Erkennung auf Vergleiche, Googeln, Hinweise anderer Personen oder inhaltliche Unstimmigkeiten zurückgeführt wird, werden dieser Kategorie zugeordnet.

Kategorie 1: Erkennung durch visuelle Auffälligkeiten

- „Ja, Instagram. Entweder es stand dabei, Fotos oder Videos waren unrealistisch (z.B. komische Bewegungen)“
- „Ja weil sie unrealistisch aussehen“
- „Ja hab es auf TikTok bekommen. Es war unrealistisch“
- „Die sind zu unrealistisch und fühlen sich Fake an“
- „Man merkt es oft anhand des komischen Aussehens, meist sind Hände oder Augen die ein bisschen komisch wirken“
- „Auf Tiktok habe ich oft erkannt, dass Sachen nicht stimmen können“
- „Ja meistens. Man bemerkt es meistens, weil die Bilder nicht perfekt genug sind, also sind meistens verschwommen oder Hände haben zum beispiel 7 finger statt 5. Auf Tiktok Insta usw.“
- „Ja man erkennt es an der Haut und so“
- „Ja schon des Öfteren, ich finde solche inhalte kann man gut erkennen, da sie nicht echt aussehen“

- „Auf Tik Tok hab ich ein Video gesehen wo die Kleidung eines jungen eine Make-70 Tutorial gemacht hat. Man hat gesehen wie plötzlich Sachen sich ändern oder verschwinden“
- „Verschiedene Farben, sah unrealistisch aus“
- „Ja, ich sehe es immer auf TikTok oder Instagram. Ja an den Bewegungen und an den Augen und Gesichtsausdrücke.“
- „Ja ich hab's erkannt. An den Augen und der Farbe. Auf TikTok“

Kategorie 2: Erkennung durch technische Merkmale oder Qualität

- „Ja, es hat entweder zu perfekt ausgeschaltet oder einfach komisch, tiktok und insta“
- „Ein Kollege hat mir ein Deepfake geschickt. Es war sehr deutlich dass es mit KI generiert wurde.“
- „Auf TikTok gibt es viele KI erstellte Videos die erkennt man eigentlich direkt.“
- „an der Qualität“
- „TikTok, dumme Kommentare“
- „Öfters. Deepfakes sind meist an der Stimme oder der Mimik erkennbar“
- „Ja. Auf TikTok. Bewegungen von dieser Person/diesem Teil in dem Video, der Kontrast, die Schattierung und die Struktur schienen nicht echt zu sein.“
- „Hände, Hintergrund, Gesicht, monotone Stimme. Auf TikTok, Instagram, Twitter“

Kategorie 3: Erkennung durch Vergleich, Recherche oder Kontextwissen

- „Ja auf insta.“
- „Ich hab's immer bemerkt ich weiß nicht wieso“
- „TikTok, bei dem Video wusste man es, weil es einen Vergleich gab zum echten“
- „Ja, YouTube“
- „Man sieht's oft auf Instagram und ich finde es einfach zu erkennen.“
- „hab es gegoogelt“
- „Ja, ich habe dann diese Nachricht gegooglet und sie war Fake. Ich habe es auch Tik-Tok gesehen“
- „Ja, auf tiktok.“
- „Ja, auf tiktok“
- „Ja, auf Instagram gibt es viele“
- „Ja, Instagram, er sagt was kann nicht sein“
- „Nein“
- „ich habe das nicht gesehen“
- „Nein“

3. Wie würdest du dein Gefühl beim Schauen eines Deepfake-Videos beschreiben?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Negative emotionale Reaktionen	Gefühle wie Angst, Ekel, Grusel, Nervosität oder Unbehagen beim Anschauen von Deepfake-Videos.	„Uncanny, gruselig“	Alle Antworten, in denen negative Emotionen (z. B. gruselig, unangenehm, verängstigt, enttäuscht, genervt) dominieren, werden dieser Kategorie zugeordnet.
2: Neutrale / emotionslose Reaktionen	Keine starke emotionale Reaktion oder Gleichgültigkeit gegenüber Deepfake-Videos.	„Neutral“	Alle Antworten, in denen keine klare emotionale Bewertung erfolgt oder Gleichgültigkeit ausgedrückt wird (z. B. neutral, egal, normal, fühle nichts), werden dieser Kategorie zugeordnet.

Kategorie	Definition	Ankerbeispiel	Kodierregel
3: Positive emotionale Reaktionen	Positive Gefühle wie Spaß, Interesse oder Humor beim Anschauen von Deepfake-Videos.	„es ist lustig“	Alle Antworten, in denen positive Emotionen wie lustig, interessant oder unterhaltsam genannt werden, werden dieser Kategorie zugeordnet.
4: Ambivalente Reaktionen	Gemischte Gefühle, z. B. gleichzeitig Faszination und Unsicherheit oder wechselnde Emotionen.	„Manchmal sind sie lustig aber dann denke ich mir wie krass es eigentlich ist so etwas leichtes zu erstellen.“	Alle Antworten, in denen sowohl positive als auch negative oder widersprüchliche Emotionen genannt werden, werden dieser Kategorie zugeordnet.

Kategorie 1: Negative emotionale Reaktionen

- „Genervt“
- „Uncanny, gruselig“
- „Es ist seltsam und gruselig was damit alles gemacht werden kann“
- „Disgust“
- „Ich empfinde nicht wirklich etwas wobei es manchmal nervt“
- „Ich fühle mich angelogen oder halt gescammt“
- „Ist komisch“
- „Furchteinflößend. Wenn es so etwas schon machen kann(tanzende Früchte, etc.), was kann es noch?“
- „verängstigt“
- „unangenehm“

Kategorie 2: Neutrale / emotionslose Reaktionen

- „Fake.“
- „Egal“
- „Normal“
- „Neutral“
- „desinteressiert, wenn ich weiß das es fake ist würd ich nicht weiterschauen wollen“
- „Neutral fühle nichts“
- „komisch“
- „Neutral, ist mir eigentlich egal“
- „Ja ich würde es mir anschauen aber ich glaube ich würde wissen das es Fake ist“
- „/“

Kategorie 3: Positive emotionale Reaktionen

- „Es fühlt sich so komisch an, Ich lache gefühlt dabei weil ziemlich erkennbar ist“
- „es ist lustig“
- „lustig“
- „Lustig“
- „Ich hab's lustig gefunden“
- „Interessiert“

Kategorie 4: Ambivalente Reaktionen

- „Ein komisches Gefühl, weil so etwas einfach erstellt werden kann und so real wirkt“

- „Enttäuscht, das man KI für sowas nutzen würde.“
- „Manchmal sind sie lustig aber dann denke ich mir wie krass es eigentlich ist so etwas leichtes zu erstellen.“

4. Wie werden sich Fake News durch die Möglichkeit von Deepfakes in Zukunft entwickeln?

Eigene Einschätzung

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Zunahme von Täuschung und Realismus	Fake News werden durch Deepfakes realistischer, glaubwürdiger und schwerer von echten Inhalten zu unterscheiden.	„Deepfakes lassen Fake News durch Bild und Video „Beweise“ realistischer wirken“	Alle Antworten, die eine steigende Glaubwürdigkeit, Realitätsnähe oder Täuschungswirkung durch Deepfakes beschreiben, werden dieser Kategorie zugeordnet.
2: Zunahme von Verbreitung und Häufigkeit	Fake News treten häufiger auf und verbreiten sich stärker in Zukunft.	„Ich glaube es werden sehr viele Fake News entstehen und ich finde es relativ gefährlich.“	Alle Aussagen, die eine erhöhte Menge, Häufigkeit oder stärkere Verbreitung von Fake News beschreiben, werden dieser Kategorie zugeordnet.
3: Zunahme von Unsicherheit und Erkennbarkeitsschwierigkeit	Es wird schwieriger, echte und gefälschte Inhalte zu unterscheiden.	„Es wird schwerer zu erkennen, was Fake ist und was nicht.“	Alle Aussagen, die eine steigende Unsicherheit oder sinkende Erkennbarkeit von Fake News thematisieren, werden dieser Kategorie zugeordnet.
4: Negative gesellschaftliche Folgen	Deepfakes führen zu Problemen wie Manipulation, Gefahr, Politikmissbrauch oder kriminellen Folgen.	„Es wird viel Erpressung und Nötigung geben.“	Alle Aussagen, die konkrete negative Konsequenzen für Gesellschaft, Politik oder Sicherheit beschreiben, werden dieser Kategorie zugeordnet.
5: Technologischer Fortschritt (neutral/unkonkret)	Entwicklung von KI/Deepfakes wird allgemein als Fortschritt oder Weiterentwicklung beschrieben, ohne konkrete Bewertung.	„Sie werden sich sehr entwickeln“	Alle allgemeinen Aussagen über technische Weiterentwicklung ohne klare inhaltliche Ausarbeitung werden dieser Kategorie zugeordnet.
6: Keine oder unsichere Einschätzung	Die Person kann keine klare Prognose geben oder äußert Unsicherheit.	„keine ahnung“	Alle Aussagen ohne inhaltliche Zukunftseinschätzung oder mit expliziter Unsicherheit werden dieser Kategorie zugeordnet.

Kategorie 1: Zunahme von Täuschung und Realismus

- „Deepfakes lassen Fake News durch Bild und Video „Beweise“ realistischer wirken“
- „Sie werden immer genauer und realistischer. Dadurch wird es sicher zu Problemen kommen“
- „Es wird schwerer zu erkennen, was Fake ist und was nicht.“
- „Schwerer zu erkennen“
- „Ich denk es werden später viel mehr deep Fakes geben die wir nicht mal erkennen werden“

- „Werden schweren zu unterscheiden“
- „Es wird sehr schwer zu unterscheiden ob etwas real oder KI generiert ist“
- „Ich glaube sie werden viel glaubwürdiger und beängstigender“
- „Ich denke fast alles wir mit der Ki generiert und man wir das gar nicht erkennen das Nachrichten Fake sind“

Kategorie 2: Zunahme von Verbreitung und Häufigkeit

- „Es wird immer schlimmer und schlimmer.“
- „Es wird viel schlimmer werden“
- „Es wird Zukünftig zu einem Massiven Problem“
- „Ich glaube es wird viel häufiger vorkommen, dass Fake News in die Welt gelangen“
- „Die werden sich sehr entwickeln und zwar sehr, aber nicht positiv wie zb durch social media. Glaube es wird viel mehr Fake news geben.“
- „Ich glaube das Fake News viel mehr geglaubt werden.“
- „Es werden vielleicht mehr Fake News bevorzugt und denen mehr geglaubt als echten Begebenheiten“

Kategorie 3: Zunahme von Unsicherheit und Erkennbarkeitsschwierigkeit

- „Man weiß nicht mehr was man glauben soll“
- „Es wird immer schwerer zum erkennen“
- „Schwerer zu erkennen“
- „Ich denke es wird sehr schwer zu unterscheiden ob etwas real oder KI generiert ist“

Kategorie 4: Negative gesellschaftliche Folgen

- „Ich glaube es werden sehr viele Fake News entstehen und ich finde es relativ gefährlich.“
- „Es werden vielleicht immer mehr politische Themen benutzt und es werden vielleicht immer mehr Probleme dadurch auftreten“
- „Es könnte gefährlich werden“
- „Es wird viel Erpressung und Nötigung geben.“

Kategorie 5: Technologischer Fortschritt (neutral/unkonkret)

- „Sie werden besser“
- „Starke Fortschritte durch KI. KI wird jetzt über den Jahren sich verbessern von der Qualität her.“
- „Sie werden zu normalen Sachen die man täglich sieht“
- „Sie werden sich sehr entwickeln z.B dich Social Media oder generell“
- „Ich glaube deepfakes wird sich sehr weite Fortschritte machen es werden viele mehrere technologische Möglichkeiten geben“
- „Sehr schnell und fortgeschrittener“

Kategorie 6: Keine oder unsichere Einschätzung

- „keine ahnung“
- „-“
- „-“
- „schnell“

5. Glaubst du, wird man wieder verstärkt auf gedruckte Informationsquellen wie eine Zeitung zurückgreifen? Warum/ Warum nicht?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Nein – digitale Medien bleiben dominant	Die Befragten glauben nicht, dass gedruckte Medien wichtiger werden; soziale Medien/Internet bleiben führend.	„Nein, ich glaube, dass Soziale Medien in den Vordergrund rücken“	Alle Aussagen, in denen klar verneint wird, dass Zeitungen wichtiger werden, und stattdessen digitale Medien als dominant beschrieben werden, werden dieser Kategorie zugeordnet.
2: Ja – Rückkehr zu gedruckten Medien (Vertrauensgewinn)	Gedruckte Medien werden wieder wichtiger, da sie als glaubwürdiger oder vertrauenswürdiger wahrgenommen werden.	„Ja höchstwahrscheinlich da dadurch News vertrauenswürdiger überkommen.“	Alle Aussagen, in denen Zeitungen wegen höherer Glaubwürdigkeit, Wahrheit oder Kontrolle als bevorzugte Quelle genannt werden, werden dieser Kategorie zugeordnet.
3: Vielleicht / unsichere Entwicklung	Es wird eine mögliche, aber nicht sichere Rückkehr zu Zeitungen gesehen.	„Vielleicht wenn ki zu stark ist und sie runter genommen werden muss“	Alle Antworten mit Unsicherheit, Einschränkungen oder hypothetischen Szenarien („vielleicht“, „könnte sein“) werden dieser Kategorie zugeordnet.
4: Digitale Überlegenheit durch Geschwindigkeit/Bequemlichkeit	Digitale Medien bleiben wichtiger, da sie schneller, praktischer und stärker verbreitet sind.	„Nein, weil Internet schneller ist und Zeitungen kaum noch genutzt werden.“	Alle Aussagen, die digitale Medien aufgrund von Geschwindigkeit, Verfügbarkeit oder Gewohnheit als überlegen darstellen, werden dieser Kategorie zugeordnet.
5: Misstrauen gegenüber allen Informationsquellen	Sowohl digitale als auch gedruckte Quellen werden als potenziell unsicher oder manipulierbar angesehen.	„Nein, KI wird immer besser und besser daher wird man es nicht mehr erkennen ob es fake oder Wahr ist.“	Alle Aussagen, die generelles Misstrauen gegenüber allen Medien oder Informationsquellen ausdrücken, werden dieser Kategorie zugeordnet.

Kategorie 1: Nein – digitale Medien bleiben dominant

- „Nein, ich glaube, dass Soziale Medien in den Vordergrund rücken“
- „Nein“
- „Nein, im Notfall vielleicht, weil niemand der neuen Generation diese lesen würde“
- „Nein, weil heutzutage immer mehr Sachen im Internet passieren“
- „Nein. Mittlerweile hat fast jeder ein Handy und benutzt es häufig und ich glaube nicht das auf einmal alle oder viele aufhören viel Handy zu schauen.“
- „Nein“
- „Denke nicht weil social media schon so weit verbreitet ist“
- „Nein, weil Internet schneller ist und Zeitungen kaum noch genutzt werden.“
- „Nein, weil man auf Sozialen Medien schneller Information findet“
- „Nein, ich glaube die meisten werden Informationen aus den Sozialen Medien.“
- „Ich denke nicht, weil es altmodisch ist“

Kategorie 2: Ja – Rückkehr zu gedruckten Medien (Vertrauensgewinn)

- „Ja höchstwahrscheinlich da dadurch News vertrauenswürdiger überkommen.“
- „Ja“
- „Ja“
- „Ja wird man weil man glaubwürdige Quellen sucht/braucht“
- „Tatsächlich schon, weil dies keine gefälschten quellen sind.“
- „Zeitungen werden von Menschen auf wahrheit überprüft, deswegen würde man mehr auf Zeitungen zurückgreifen“
- „Ich denke schon, weil die gedruckten Informationsquellen von Menschen selbst geschrieben werden.“
- „Ja ,weil das internet immer unglaubwürdig wirkt“

Kategorie 3: Vielleicht / unsichere Entwicklung

- „Schon, weil nur bestimmte Menschen Zeitungen veröffentlichen dürfen aber blind sollte man es trotzdem nicht vertrauen“
- „Vielleicht wenn KI zu stark ist und sie runter genommen werden muss“
- „Könnte sein da es nicht mehr sicher sein wird was echt oder Fake ist“
- „Vllt, weil man gerade in den Schuhen versucht den neueren Generationen das beizubringen.“
- „Vielleicht nur paar Menschen , Weil es dann später vielleicht irgendwann zu viel wird“

Kategorie 4: Digitale Überlegenheit durch Geschwindigkeit/Bequemlichkeit

- „Ich glaube allgemein die Welt wird digitalisiert, ob mit Absicht oder nicht keiner wird mehr auf Zeitungen und sowas achten.“
- „Ich glaub nicht, da die Technologie erweitert werden kann und verbessert werden kann“
- „Ich denke nicht, weil die meisten Leute haben jetzt einfach keine Lust jeden Tag eine Zeitung zu lesen, wenn man sowieso schon in den Sozialen Medien sehr viel erfährt“
- „Ich denke nicht, weil es altmodisch ist“

Kategorie 5: Misstrauen gegenüber allen Informationsquellen

- „Nein, KI wird immer besser und besser daher wird man es nicht mehr erkennen ob es fake oder Wahr ist.“
- „Ich glaub nicht weil sie auch Fake sein könnten“
- „Weil man sich sicher sein will ob es stimmt“

- „Ich denke schon, denn im Internet behaupten die Leute oft Dinge, um Aufmerksamkeit zu erregen, und die Medien werden ihnen dann eher Glauben schenken als Bloggern und Influencern.“

6. Welche Merkmale würdest du heranziehen, um herauszufinden, ob ein Video echt ist?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Visuelle/inhaltliche Merkmale im Video	Es werden konkrete visuelle Hinweise im Video genutzt (z. B. Gesicht, Hände, Finger, Bewegungen, Hintergrund, Farben, Schatten, Qualität).	„Auf Veränderungen im Video achten, Fingeranzahl, Qualität, Farbverhältnisse“	Alle Antworten, in denen konkrete visuelle oder inhaltliche Video-Merkmale zur Echtheitsprüfung genannt werden, werden dieser Kategorie zugeordnet.
2: Subjektives Gefühl / Intuition	Die Einschätzung basiert auf einem allgemeinen Gefühl oder subjektiver Wahrnehmung statt konkreter Merkmale.	„Authentisch vom Gefühl her, eigentlich schwer zu erkennen“	Alle Aussagen, in denen „Gefühl“, Intuition oder allgemeines Erkennen ohne konkrete Kriterien genannt werden, werden dieser Kategorie zugeordnet.
3: Kontext- und Quellenprüfung	Die Echtheit wird durch externe Informationen wie Kommentare, andere Quellen oder Verbreitung auf mehreren Plattformen überprüft.	„Ich schaue mir die Kommentare an.“	Alle Antworten, in denen externe Quellen, Kommentare oder Verifikation außerhalb des Videos genutzt werden, werden dieser Kategorie zugeordnet.
4: Allgemeine oder unspezifische Wahrnehmung	Die Echtheit wird allgemein oder sehr vage beschrieben, ohne konkrete Kriterien.	„Man merkt es einfach“	Alle Antworten, die keine klaren Prüfkriterien nennen und nur allgemeines „Erkennen“ beschreiben, werden dieser Kategorie zugeordnet.

Kategorie 1: Visuelle/inhaltliche Merkmale im Video

- „Auf Veränderungen im Video achten, Fingeranzahl, Qualität, Farbverhältnisse“
- „Bewegungen sachen verschwinden“
- „Körperhaltung, schauen ob sich Objekte oder andere Sachen verändern“
- „Man schaut auf den Übergang“
- „Ich schau auf den Hintergrund und Körper“
- „Genauer hinschauen, bei einem Menschen die Gesichtsausdrücke, die Finger, usw.“
- „Haut, Augen, Finger“
- „Auf die Augen, Zähne, Hintergrund, Farbe achten oder ob etwas sehr unrealistisch wirkt.“
- „Schattierung, Anzahl von Fingern, Textur“
- „Die Videos haben meistens Stellen die sich plötzlich ändern und Gegenstände und Leute sehen meist zu echt oder einfach nur komisch aus“
- „auf die kleinen details achten die normalerweise menschen machen“
- „Die Zähne, Augen und andere Merkmale anschauen“
- „Farben, Bewegungen, Gesicht, Haut, Hintergrund,...“
- „Gesicht, Bewegung“
- „An der quali“
- „Augen, Finger, Haut, Farbe“
- „Ob es natürlich aussieht oder nicht“
- „Die Augen, ob es zu weich ausschaut, Farben, Sachen die verachwinden“
- „Gesicht, Hände, Bewegung“
- „einfach daran schauen on etwas falsch abgebildet ist (schatten)“

- „der Hintergrund ist meistens schlecht gemacht, dinge verschwinden oder verändern sich“
- „Ich würde es mir genauer anschauen und genau auf die Farben oder schauen ob es realistisch ist“
- „Wie die Person aussieht, Finger, Hintergrund“
- „Schatten, Gesichtsausdrücke(wenn vorhanden), Bewegungen, etc.“
- „Ich würde als erstes auf das Aussehen des Videos achten und schauen ob etwas nicht passt“
- „Augen, Hände, Gang und Bewegung, kleine Details und Perspektivenwechsel“
- „Finger Sachen die Eif so verschwinden“

Kategorie 2: Subjektives Gefühl / Intuition

- „Authentisch vom Gefühl her, eigentlich schwer zu erkennen“

Kategorie 3: Kontext- und Quellenprüfung

- „Ich schaue mir die Kommentare an.“
- „Andere Quellen wo das Video vorkommt und auch Details achten“
- „Man prüft die Quelle, achtet auf komische Fehler im Video und schaut, ob andere vertrauenswürdige Seiten es bestätigen.“

Kategorie 4: Allgemeine oder unspezifische Wahrnehmung

- „Man merkt es einfach“
- „Ich schau auf die Augen und Körper“
- „.“

7. Welche Risiken entstehen, wenn immer mehr Deepfakes verbreitet werden?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Verlust von Wahrheit / Realitätsunsicherheit	Es wird beschrieben, dass nicht mehr zwischen echt und falsch unterschieden werden kann.	„Man weiß irgendwann nicht mehr was echt ist und was nicht“	Alle Aussagen, in denen Unsicherheit über Realität vs. Fälschung im Vordergrund steht, werden dieser Kategorie zugeordnet.
2: Vertrauensverlust in Medien und Gesellschaft	Deepfakes führen zu sinkendem Vertrauen in Informationen, Medien oder andere Menschen.	„Vertrauen verschwindet, Missverständnisse“	Alle Aussagen, die Vertrauensverlust in Medien, Internet oder zwischen Menschen beschreiben, werden dieser Kategorie zugeordnet.
3: Verbreitung von Fake News, Desinformation und Manipulation	Deepfakes führen zu mehr Fake News, Lügen, Gerüchten und gezielter Manipulation.	„Dadurch werden viel mehr fake news verbreitet und Lügen erzählt.“	Alle Aussagen, die Desinformation, Manipulation oder Verbreitung falscher Inhalte nennen, werden dieser Kategorie zugeordnet.
4: Individuelle und gesellschaftliche Schäden	Konkrete negative Folgen für Personen oder Gesellschaft (z. B. Betrug, Angst, Konflikte, Panik).	„Man kann leichter gescammt werden“	Alle Aussagen, die konkrete Schäden wie Betrug, Angst, Streit, Panik oder Rufschädigung beschreiben, werden dieser Kategorie zugeordnet.

Kategorie	Definition	Ankerbeispiel	Kodierregel
5: Überforderung / kognitive und soziale Folgen	Menschen sind überfordert, glauben alles oder verlieren Orientierung.	„Menschen werden anfangen zu glauben das alles fake ist“	Alle Aussagen, in denen Überforderung, Orientierungslosigkeit oder falscher Glaube beschrieben wird, werden dieser Kategorie zugeordnet.
6: Gesellschaftlicher Systemverlust / extreme Folgen	Sehr starke oder systemische Konsequenzen wie gesellschaftlicher Zerfall oder „Leben in einer Lüge“.	„Die Menschen werden einfach in einer Lüge leben.“	Alle Aussagen mit extremen, systemweiten oder apokalyptisch wirkenden Folgen werden dieser Kategorie zugeordnet.

Kategorie 1: Verlust von Wahrheit / Realitätsunsicherheit

- „Man weiß irgendwann nicht mehr was echt ist und was nicht“
- „Nichtmehr erkennbar was echt ist“
- „Man wird nicht mehr wissen was echt ist oder nicht.“
- „Man kann nicht sagen was echt ist“
- „Man kann nicht mehr echt von Fake auseinander scheiden und man kann einfacher getäuscht werden“
- „das man nicht mehr weiß was wahr und was falsch ist“
- „Man kann nicht mehr echt von Fake auseinander scheiden und man kann einfacher getäuscht werden“
- „Dass man nicht mehr das echte vom gefälschten erkennen kann“
- „Viel mehr Menschen werden anfangen Sachen zu glauben keiner wird mehr erkennen was real und was Fake ist“
- „Dass Menschen nicht mehr wissen was wahr oder was Fake ist“

Kategorie 2: Vertrauensverlust in Medien und Gesellschaft

- „Vertrauen verschwindet, Missverständnisse“
- „Fake News und Missverständnisse zwischen Menschen.“
- „Das Internet wird immer weniger glaubwürdig“
- „Kein Vertrauen und Angst“
- „Es kann Streit entstehen“
- „Menschen können sich verarscht fühlen“
- „Das Internet wirkt unglaubwürdiger. man greift wieder auf gedruckte Informationsquellen“

Kategorie 3: Verbreitung von Fake News, Desinformation und Manipulation

- „Fake news“
- „Fake news“
- „Desinformation der Bevölkerung“
- „Viel Manipulation viele Gerüchte“
- „Dadurch werden viel mehr fake news verbreitet und Lügen erzählt.“
- „Lauchtgläubige Menschen glauben dies und verbreiten es immer weiter“
- „Es können Falschinformationen, Betrug, Rufschädigung und Vertrauensverlust in Medien entstehen.“
- „Menschen fangen an, Sachen zu glauben und Videos/Fotos zu vertrauen, die nicht real sind“

Kategorie 4: Individuelle und gesellschaftliche Schäden

- „Man kann leichter gescammt werden“

- „Falsche Informationen, ältere Bevölkerung macht sich Sorgen“
- „Es kann Streit entstehen“
- „Sehr viele Menschen vor allem Jugendliche und Kinder die noch nicht mit diesem Thema bekannt sind, glauben dann häufig daran“
- „Menschen werden versorgt“
- „Fake News, Menschen werden versorgt“
- „Betrüger rufen Sie möglicherweise mit der Stimme Ihrer Eltern an und bitten um Geld.“
- „Fake News, Menschen werden versorgt“
- „Man kann leichter gescammt werden“

Kategorie 5: Überforderung / kognitive und soziale Folgen

- „Man könnte irgendwann echte Videos als Fake labeln.“
- „das man nicht mehr weiß was wahr und was falsch ist“
- „Man wird nicht mehr wissen was echt ist oder nicht.“
- „Menschen werden anfangen zu glauben das alles Fake ist und wenn dann wirklich etwas passiert z.B. Krieg werden Sie das dann auch nicht glauben“
- „Viel mehr Menschen werden anfangen Sachen zu glauben keiner wird mehr erkennen was real und was Fake ist“
- „Menschen fangen an, Sachen zu glauben und Videos/Fotos zu vertrauen, die nicht real sind“

Kategorie 6: Gesellschaftlicher Systemverlust / extreme Folgen

- „Massenpanik“
- „Wir werden dümmer“
- „Das die Gesellschaft Sachen glaubt die nicht echt sind“
- „Die Menschen werden einfach in einer Lüge leben.“

8. Wie hoch schätzt du den Anteil an KI-generierten bzw. manipulierten Medien ein, die du auf soziale Medien siehst?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Sehr geringer Anteil	Einschätzung, dass nur 1–5 % der gesehenen Inhalte KI-generiert/manipuliert sind.	„1 - 5 %“	Alle Angaben, die den Anteil auf 1–5 % schätzen, werden dieser Kategorie zugeordnet.
2: Geringer Anteil	Einschätzung von etwa 5–10 % KI-generierter/manipulierter Inhalte.	„5 - 10 %“	Alle Angaben mit einer Schätzung von 5–10 % werden dieser Kategorie zugeordnet.
3: Mittlerer Anteil	Einschätzung von etwa 10–20 % KI-generierter/manipulierter Inhalte.	„10-20 %“	Alle Angaben mit einer Schätzung von 10–20 % werden dieser Kategorie zugeordnet.
4: Erhöhter Anteil	Einschätzung von etwa 20–30 % KI-generierter/manipulierter Inhalte.	„20-30 %“	Alle Angaben mit einer Schätzung von 20–30 % werden dieser Kategorie zugeordnet.
5: Sehr hoher Anteil	Einschätzung von über 30 % KI-generierter/manipulierter Inhalte.	„über 30 %“	Alle Angaben mit einer Schätzung von mehr als 30 % werden dieser Kategorie zugeordnet.

Kategorie 1: Sehr geringer Anteil (1–5 %)

- „1 - 5 %“
- „1 - 5 %“
- „1 - 5 %“

Kategorie 2: Geringer Anteil (5–10 %)

- „5 - 10 %“
- „5 - 10 %“
- „5 - 10 %“
- „5 - 10 %“
- „5 - 10 %“

Kategorie 3: Mittlerer Anteil (10–20 %)

- „10-20 %“
- „10-20 %“
- „10-20 %“
- „10-20 %“
- „10-20 %“
- „10-20 %“
- „10-20 %“
- „10-20 %“
- „10-20 %“
- „10-20 %“
- „10-20 %“

Kategorie 4: Erhöhter Anteil (20–30 %)

- „20-30 %“
- „20-30 %“
- „20-30 %“
- „20-30 %“
- „20-30 %“
- „20-30 %“
- „20-30 %“
- „20-30 %“
- „20-30 %“
- „20-30 %“

Kategorie 5: Sehr hoher Anteil (>30 %)

- „über 30 %“
- „über 30 %“

9. Was hat dir an den beiden Unterrichtseinheiten über Deepfakes gut gefallen? Was hat dir nicht gut gefallen? Falls etwas für dich unverständlich war, beschreibe es bitte kurz!

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Positives Feedback (Inhalt, Methode, Sozialform)	Aussagen, in denen Aspekte der Unterrichtseinheiten als gut, interessant, verständlich oder gelungen bewertet werden.	„Die Diskussion hat mir gut gefallen und das Video der rest nicht.“	Alle positiven Bewertungen zu Inhalt, Erklärung, Materialien, Gruppenarbeit oder Methoden werden dieser Kategorie zugeordnet.
2: Negatives Feedback (Kritik / Verbesserung)	Aussagen, in denen Kritik, Langeweile, Unverständlichkeit oder Verbesserungsvorschläge geäußert werden.	„War etwas langweilig“	Alle Aussagen mit negativer Bewertung oder Kritik an Unterricht, Material oder Durchführung werden dieser Kategorie zugeordnet.
3: Neutral / keine konkrete Bewertung	Aussagen ohne klare positive oder negative Bewertung oder nur sehr allgemeine Aussagen.	„-“	Alle Antworten ohne inhaltliche Bewertung (z. B. nur Strich, sehr allgemein oder unklar) werden dieser Kategorie zugeordnet.
4: Ambivalentes Feedback (positiv und negativ kombiniert)	Aussagen, die sowohl positive als auch negative Aspekte enthalten.	„Mir hat gefallen das es verständlicher gemacht worden ist, ich finde die Gefahr der KI hätte man mehr betonen können“	Alle Antworten, in denen Lob und Kritik gleichzeitig vorkommen, werden dieser Kategorie zugeordnet.

Kategorie 1: Positives Feedback (Inhalt, Methode, Sozialform)

- „Selber argumentieren“
- „Hat mir gut gefallen“
- „Es war selbstständiges lernen dabei und es hat Spaß gemacht in der Gruppe zu diskutieren, als auch die eigene Freizeit, wenn man die Aufgabe schnell erledigte“
- „Der Workshop war ziemlich interessant. Es war informativ“
- „Die Diskussionen haben mir gut gefallen.“
- „Visuelle Beispiele“
- „Mir hat die Gruppenarbeit mit den Rollen gefallen es wurde alles sehr gut erklärt und es war verständlich.“
- „In Gruppen arbeiten macht immer spaß“
- „Die Folien“
- „Ich finde alles war sehr gut, am besten hat mir diese Rollen Spiel gefallen“
- „es war unterhaltsam“
- „Die Videos die uns gezeigt wurden“
- „Hat mir gut gefallen, alles gut rübergebracht“
- „Die Diskussion war cool👍“
- „Es war interessant“
- „Mir hat es gefallen das man sich in Rollen verteilen kann und diskutiert werden kann es war oke ,und es wurde erklärt das war gut“
- „Gut gefallen hat mir, dass man lernt, wie man Deepfakes erkennt und warum sie gefährlich sind.“
- „Ich fand die Informationen nützlich und es war geschickt ein KI generiertes Video zu zeigen um zu schauen woran man es erkennt“
- „Wo man schauen musste was real und was Fake ist“

- „Das wir in Gruppen arbeiten konnten“
- „Es war alles sehr ausführlich“
- „interessante spiel mit fotos und thema ist geil“

Kategorie 2: Negatives Feedback (Kritik / Verbesserung)

- „ich finde es sollte immer mehr interessanter gemacht werden, Schüler nehmen den Unterricht nicht mehr ernst da es einfach uninteressant geworden ist“
- „war nicht verständlich genug und nicht gut erklärt.“
- „Deepfakws und fake news werden in fast jeder klasse unterrichtet, weshalb dies langweilig wird und man alles schon kennt“

Kategorie 3: Neutral / keine konkrete Bewertung

- „“
- „/“
- „“

Kategorie 4: Ambivalentes Feedback (positiv und negativ kombiniert)

- „Mir hat gefallen das es verständlicher gemacht worden ist, ich finde die Gefahr der KI hätte man mehr betonen können“
- „Alles war gut, einbisschen langweilig“

10. Was möchtest du abschließend zu diesem Thema noch sagen? (positive Aspekte, Ängste, wenig Kontakt mit der Thematik,...)

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Positive oder konstruktive Abschlusss Aussagen	Abschließende Aussagen mit positiver Bewertung, Lerngewinn oder konstruktiven Empfehlungen.	„War sehr informativ“	Alle Aussagen, die einen positiven Eindruck, Nutzen oder sinnvolle Empfehlungen enthalten, werden dieser Kategorie zugeordnet.
2: Negative/ängstliche oder kritische Aussagen	Aussagen mit Sorgen, Kritik, Angst oder negativer Bewertung von KI/Deepfakes.	„Angst weil man es irgendwann nicht mehr erkennen kann“	Alle Aussagen, die Angst, Ablehnung, Kritik oder problematische Zukunftserwartungen ausdrücken, werden dieser Kategorie zugeordnet.
3: Neutral / kein Interesse / keine Meinung	Aussagen ohne klare Bewertung oder mit Gleichgültigkeit bzw. fehlender Meinung.	„Es ist eigentlich egal“	Alle Antworten ohne erkennbare emotionale oder inhaltliche Position werden dieser Kategorie zugeordnet.
4: Reflexive / regulative Aussagen (Umgang & Umgangsregeln)	Aussagen, die Verhaltensempfehlungen oder Umgangsstrategien mit KI/Deepfakes enthalten.	„Ich finde man sollte Sachen, bei denen man sich nicht sicher ist immer mit einer vertrauten und sicherern Quelle überprüfen“	Alle Aussagen, die Empfehlungen zum Umgang, Medienkompetenz oder Schutzstrategien enthalten, werden dieser Kategorie zugeordnet.
5: Unsichere / unklare oder irrelevante Antworten	Sehr kurze, unklare oder inhaltlich nicht verwertbare Antworten.	„1“	Alle nicht interpretierbaren, sehr kurzen Antworten werden dieser Kategorie zugeordnet.

Kategorie 1: Positive oder konstruktive Abschlusss Aussagen

- „War sehr informativ“
- „War okay“
- „Man soll Kinder bei bringen was Fake ist und was echt oder zumindest Merkmale“
- „Es hat mir sehr gefallen, der Unterricht wurde gut erklärt“
- „alles gut“

Kategorie 2: Negative/ängstliche oder kritische Aussagen

- „Deepfakes sind schlecht, wenn man sie benutzt um Leute zu manipulieren“
- „Ich glaube es ist wichtig, dass jeder erkennt was Deep Fakes sind und hoffentlich gerät es nicht außer Kontrolle bei den sozialen Medien“
- „Ich habe keine starke Meinung dazu.“
- „keine ahnung“
- „Angst weil man es irgendwann nicht mehr erkennen kann“
- „Ich denke, dass es problematisch für die Zukunft sein könnte.“
- „Man wollte aufpassen wenn man was ki generiertes sieht und nicht direkt alles von den medien glauben“
- „Viele Menschen sind versorgt weil sie nicht mehr wissen was real und was unrealistisch ist aber ki wird immer fortgeschrittener“

- „Viele Zeichnungen sind heutzutage KI und deswegen werden viele Künstler beschuldigt dass ihre Zeichnungen KI sind was ich einbisschen dumm finde“
- „Ich bin kategorisch gegen neuronale Netze, obwohl ich sie auch benutze. Die Leute vergessen, wie man Texte korrekt schreibt oder Gedankenstriche verwendet.“

Kategorie 3: Neutral / kein Interesse / keine Meinung

- „Es ist eigentlich egal“
- „-“
- „Nichts“
- „Nichts“
- „Nichts“
- „/“
- „Nichts“
- „-“
- „Nichts“

Kategorie 4: Reflexive / regulative Aussagen (Umgang & Umgangsstrategien)

- „Ich finde man sollte Sachen, bei denen man sich nicht sicher ist immer mit einer vertrauten und sicherern Quelle überprüfen“
- „Ja man soll aufpassen“

Kategorie 5: Unsichere / unklare oder irrelevante Antworten

- „“
- „1“
- „ich werde an das nicht glauben“

Antworten inklusive Feedback 2. Durchgang

1. Welche Charakteristika kennzeichnen Deepfakes? Wie würdest du Deepfakes definieren? Erkläre in etwa drei Sätzen, wie die Technologie hinter Deepfakes funktioniert.

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Deepfakes als KI-generierte oder manipulierte Medien	Deepfakes werden als künstlich erzeugte oder veränderte Bilder, Videos oder Audios beschrieben.	„Deepfakes sind gefälschte Videos, Bilder oder Stimmen, die durch KI entstanden sind.“	Alle Aussagen, die Deepfakes als durch KI erzeugte, bearbeitete oder manipulierte Medieninhalte definieren, werden dieser Kategorie zugeordnet.
2: Realistische Wirkung und Täuschung	Deepfakes werden als besonders echt wirkende Inhalte beschrieben, die Menschen täuschen oder irreführen können.	„Deepfakes sind KI-generierte oder manipulierte Medieninhalte, die echt wirken.“	Alle Aussagen, die die realistische Wirkung, Täuschung, Irreführung oder Verbreitung von Falschinformationen hervorheben, werden dieser Kategorie zugeordnet.
3: Technische Funktionsweise von Deepfakes	Die Entstehung von Deepfakes wird über KI-Training, Datensätze, Deep Learning oder Mustererkennung erklärt.	„Die KI wird mit vielen Daten trainiert und lernt dadurch, Personen realistisch zu ersetzen/nachzumachen.“	Alle Aussagen, die beschreiben, wie Deepfakes technisch erzeugt werden, werden dieser Kategorie zugeordnet.
4: Erkennungsmerkmale und Qualitätsmängel	Deepfakes werden anhand typischer Fehler, Ungenauigkeiten oder visueller Auffälligkeiten beschrieben.	„Bei schlechteren Deepfakes sind oft genaue Details ungenau. Z.B. Hände, der Hintergrund, die Schatten usw.“	Alle Aussagen, die auf Fehler, Unstimmigkeiten oder typische Merkmale von Deepfakes eingehen, werden dieser Kategorie zugeordnet.

Kategorie 1: Deepfakes als KI-generierte oder manipulierte Medien

- „Deepfakes sind mit KI erzeugte oder veränderte Videos, Bilder oder Stimmen, die sehr echt wirken.“
- „Deepfakes sind gefälschte Videos, Bilder oder Stimmen, die durch KI entstanden sind.“
- „Deepfakes sind unechte, mit KI erstellte Bilder/Videos und nicht immer gleich zu erkennen.“
- „Manipulierte oder hergestellte Bilder, die gefälscht sind.“
- „Deepfakes sind KI-generierte oder manipulierte Medieninhalte, die echt wirken.“
- „Deepfakes sind mit KI erstellte oder manipulierte Videos, Bilder oder Stimmen. Sie wirken oft echt, obwohl sie gefälscht sind.“
- „es ist KI, die versucht eine Person nachzuahmen.“
- „Ich würde sagen, dass Deepfakes gefälschte Videos im Internet sind. Meistens mit KI erstellt.“
- „Deepfakes sind meist von KI erstellte oder von Menschen editierte Bilder.“
- „Deep fakes sind Fotos oder Videos meist von KI erstellt, und enthalten oft falsche oder unangemessene Inhalte.“

Kategorie 2: Realistische Wirkung und Täuschung

- „Deepfakes sind mit KI erzeugte oder veränderte Videos, Bilder oder Stimmen, die sehr echt wirken.“
- „Danach kann sie neue, täuschend echte Inhalte erstellen.“
- „Deepfakes sind KI-generierte oder manipulierte Medieninhalte, die echt wirken.“
- „Deepfakes sind durch Technologie (Photoshop, KI) geänderte Bilder, die versuchen, zu misinformieren, zu verführen oder zu schockieren.“
- „Deepfakes sehen "zu echt" aus.“
- „Wenn Falsche Informationen zum Beispiel auf sozialen Medien verbreitet werden. Es sind meistens schockierende Nachrichten.“
- „Man will möglichst viele Aufrufe bekommen.“
- „Deep fakes sind Fotos oder Videos meist von KI erstellt, und enthalten oft falsche oder unangemessene Inhalte.“

Kategorie 3: Technische Funktionsweise von Deepfakes

- „Die KI wird mit vielen Daten trainiert und lernt dadurch, Personen realistisch zu ersetzen/nachzumachen.“
- „Die KI lernt aus vielen Aufnahmen einer Person deren Gesicht oder Stimme.“
- „Sie entstehen mithilfe von KI, der zuerst viele Bilder gezeigt werden, damit die KI dann selbst Dinge erstellen kann.“
- „Sie entstehen durch Deep-Learning-Modelle, die mit vielen Bildern, Videos oder Stimmen trainiert werden.“
- „Die KI lernt mit vielen Daten Gesicht, Stimme und Bewegungen einer Person nachzuahmen“
- „Es gibt Videos, Audioaufnahmen und so weiter von der Person.“
- „Die KI ermöglicht es jedem leichter und schneller als je zuvor Deepfakes zu erstellen und dann zu verbreiten.“
- „Diese verwenden ein Muster, um eben das Bild nach Belieben zu ändern.“

Kategorie 4: Erkennungsmerkmale und Qualitätsmängel

- „Aber sie sind nicht immer perfekt.“
- „Tunnelblick: Sie konzentrieren sich auf einzelne Stellen, lassen bei anderen aber nach. Monoton.“
- „Bei schlechteren Deepfakes sind oft genaue Details ungenau. Z.B. Hände, der Hintergrund, die Schatten usw.“
- „Es gibt bei bestimmten Dingen Ungenauigkeiten bei Erstellung von Bildern.“

2. Welche Methoden existieren zur Erkennung von Deepfakes? Teile zu jeder Methode mit, inwieweit sie noch zuverlässig ist.

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Visuelle und auditive Analyse von Inhalten	Deepfakes werden anhand von Bild-, Video- oder Stimmfehlern erkannt.	„Bild/Video analysieren: KI macht oft Fehler (z.B. falsche Schatten, "zu perfekte" Bilder...)“	Alle Aussagen, die auf Auffälligkeiten in Bild, Video, Bewegung oder Stimme zur Erkennung von Deepfakes verweisen, werden dieser Kategorie zugeordnet.
2: Quellen- und Kontextprüfung	Die Überprüfung von Quellen, Metadaten, Faktenchecks oder Kontextinformationen dient zur Erkennung von Deepfakes.	„Quelle prüfen: sehr zuverlässig“	Alle Aussagen, die auf die Prüfung der Herkunft, Glaubwürdigkeit oder zusätzlicher Informationen verweisen, werden dieser Kategorie zugeordnet.

Kategorie	Definition	Ankerbeispiel	Kodierregel
3: Technische Erkennung durch KI oder Experten	Deepfakes werden mithilfe spezieller Programme, KI-Systeme oder Fachpersonen erkannt.	„KI Programme, die Deepfakes erkennen können - relativ zuverlässig“	Alle Aussagen, die technische Erkennungssysteme oder Expert:innen als Erkennungsmethode nennen, werden dieser Kategorie zugeordnet.
4: Grenzen und sinkende Zuverlässigkeit der Erkennung	Die Befragten thematisieren, dass Deepfake-Erkennung durch den technologischen Fortschritt zunehmend schwieriger wird.	„die meisten sind nicht mehr so zuverlässig weil es immer besser wird“	Alle Aussagen, die auf die sinkende Zuverlässigkeit von Erkennungsmethoden aufgrund besser werdender KI hinweisen, werden dieser Kategorie zugeordnet.

Kategorie 1: Visuelle und auditive Analyse von Inhalten

- „Bildfehler erkennen: unnatürliche Bewegungen oder Beleuchtung, zu "perfekt"“
- „1. Bild/Video analysieren: KI macht oft Fehler (z.B. falsche Schatten, "zu perfekte" Bilder, undeutlicher Hintergrund, unmögliche Handbewegungen, Mundbewegungen passen nicht zum Sprechen, ...)“
- „2. Stimme analysieren: KI generierte Stimmen klingen oft monoton und unnatürlich“
- „Auf Körperteile achten. Die Stimme untersuchen.“
- „Ob das Bild auch echt genommen wurde, was sehr hilfreich sein kann, da KI sich verbessern kann. Bei KI: Mehr Finger, monotone Stimme, verschwommene Bilder.“
- „man kann sie durch monotonen Stimmen und den tunnel blick erkennen.“
- „Man hält Ausschau für die in Frage 1. erwähnten Merkmale.“
- „Genau Analysieren - gut bei ungenauen Deepfakes. Stimme anhören - auch gut bei ungenauen Deepfakes.“
- „Bewegungen, Stimme passt zu Bild, monoton, Finger, Augen“

Kategorie 2: Quellen- und Kontextprüfung

- „Quelle prüfen: sehr zuverlässig“
- „Metadaten prüfen: zuverlässig, wenn Originaldaten vorhanden sind.“
- „Bildfehler erkennen: teilweise zuverlässig KI-Erkennungsprogramme: hilfreich, aber nicht perfekt
Quellen prüfen: sehr zuverlässig Faktenchecks/Metadaten: meist zuverlässig“
- „Wenn die Quelle nicht seriös ist.“
- „Wenn es einen schockierenden Titel hat.“
- Kategorie 3: Technische Erkennung durch KI oder Experten
- „3. KI Programme, die Deepfakes erkennen können - relativ zuverlässig, aber es gibt einen Wettkampf zwischen den KIs“
- „KI-Erkennungstools: aktuell recht zuverlässig, aber nicht perfekt.“
- „Digitale Wasserzeichen: sehr zuverlässig, aber selten genutzt.“
- „Man verwendet KI, die das erkennen können.“
- „Die KI analysieren lassen- KI vs KI, jede Seite verbessert sich.“
- „Andere KIs können Genauigkeiten sehen, die Menschen nicht auf den ersten Blick erkennen würde.“
- „Es gibt auch Experten zum Thema“

Kategorie 4: Grenzen und sinkende Zuverlässigkeit der Erkennung

- „Fehler erkennen: nicht mehr so zuverlässig“
- „wird unzuverlässiger, da die KI immer besser wird“
- „eigentlich gut erkennbar, aber manche KIs können das auch schon sehr gut“

- „heute nur noch begrenzt zuverlässig“
- „hilfreich, aber nicht perfekt“
- „teilweise zuverlässig“
- „jede Seite verbessert sich.“
- „die aber unzuverlässiger werden, je besser die KI wird.“
- „die meisten sind nicht mehr so zuverlässig weil es immer besser wird“

3. Welche Gefahren bringen Deepfakes mit sich? Nenne jeweils 1-3 Gefahren, die

* Einzelpersonen

* die Politik

* die Gesellschaft

Betreffen.

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Desinformation, Fake News und Manipulation	Deepfakes werden als Mittel zur Verbreitung falscher Informationen und zur Manipulation von Meinungen oder Entscheidungen beschrieben.	„Politik: Fake-News“	Alle Aussagen, die Fake News, Desinformation, Propaganda oder Manipulation von Informationen thematisieren, werden dieser Kategorie zugeordnet.
2: Rufschädigung, Betrug und persönliche Schäden	Deepfakes werden als Gefahr für Einzelpersonen durch Rufschädigung, Identitätsmissbrauch, Erpressung oder Betrug beschrieben.	„Einzelpersonen: Rufschädigung, Betrug, Identitätsmissbrauch“	Alle Aussagen, die persönliche, finanzielle oder psychische Schäden für Einzelpersonen nennen, werden dieser Kategorie zugeordnet.
3: Beeinflussung von Politik und Wahlen	Deepfakes werden als Gefahr für demokratische Prozesse und politische Entscheidungen beschrieben.	„Politik: Beeinflussung von Wahlen“	Alle Aussagen, die Wahlmanipulation, politische Einflussnahme oder Falschinformationen über Politiker und Parteien thematisieren, werden dieser Kategorie zugeordnet.
4: Vertrauensverlust und gesellschaftliche Unsicherheit	Deepfakes führen zu Misstrauen, Unsicherheit und Schwierigkeiten bei der Unterscheidung von Wahrheit und Fälschung.	„Gesellschaft: Unsicherheit (Was ist echt? Was nicht?)“	Alle Aussagen, die gesellschaftliche Verunsicherung, Vertrauensverlust oder Orientierungslosigkeit beschreiben, werden dieser Kategorie zugeordnet.

Kategorie 1: Desinformation, Fake News und Manipulation

- „Einzelpersonen: Betrug/Politik: Fake-News/Gesellschaft: Verbreitung von falschen Informationen“
- „Politik: falsche Aussagen“
- „Politik: Desinformation“
- „Sie liefern falsche Informationen. Sie können auch als Propaganda dienen.“
- „Politik: Fake-News“
- „Gesellschaft: Verbreitung von Desinformation“
- „Politik: Fake News“
- „Die Gesellschaft: Propaganda kann verbreitet werden“
- „Andere Wahlergebnisse, wenn z.B. Fakeinfos von Politikern veröffentlicht werden.“

- „1) Falsch informiert sein“
- „2) Falsche Werbung für Wahlen zum Beispiel oder Falsche Infos über Politiker“
- „Gesellschaft: Anschließend von Gruppen oder einzelnen, fake news über Produkte und Leute“

Kategorie 2: Rufschädigung, Betrug und persönliche Schäden

- „Einzelpersonen: Täuschung, schlechte Gerüchte, gefälschte Videos“
- „Einzelpersonen: Rufschädigung, Erpressung/Betrug (Enkeltrick)“
- „Einzelpersonen: Rufschädigung, Betrug, Identitätsmissbrauch“
- „Pornographische Medien von echten Personen“
- „Einzelpersonen: Rufschädigung Betrug Cybermobbing“
- „Einzelpersonen: der Ruf einer Person kann geschadet werden, durch faken Videos.“
- „Enkeltrick bei Einzelpersonen(vor allem alte Menschen).“
- „Wenn Fakenews dich im Internet bloßstellen, kann das einen mental schaden.“
- „Bedrohende Inhalte für Personen(gruppen), falsche Investitionen verkauft,“
- „Einzelpersonen: Ruf geschädigt, Betrug, Unsicherheit.“

Kategorie 3: Beeinflussung von Politik und Wahlen

- „Politik: Beeinflussung von Wahlen“
- „Politik: Manipulation von Wahlen“
- „Politik: Wahlmanipulation“
- „man kann fake news über eine Partei im Internet verbreiten, weshalb manche Personen sie vielleicht nicht mehr wählen wollen“
- „Ruf von Politikern.“
- „Gesellschaft leidet, wenn Wahlen damit beeinflusst werden.“
- „können falsche Inhalte über Parteien oder einzelne Individuen verbreiten und die Wahlen dadurch beeinflussen.“

Kategorie 4: Vertrauensverlust und gesellschaftliche Unsicherheit

- „Gesellschaft: falsche Informationen, weniger Vertrauen in Dinge“
- „Gesellschaft: Unsicherheit (Was ist echt? Was nicht?), Verbreitung von Fake News, Vertrauensverlust (auch auf Social Media)“
- „Vertrauensverlust“
- „Gesellschaft: Unsicherheit, Misstrauen gegenüber Medien“
- „Politik: Vertrauensverlust“
- „Gesellschaft: Misstrauen gegenüber Medien“
- „Panik oder Konflikte“
- „Keiner weiß mehr was stimmt und was nicht.“
- „Einzelpersonen: Unsicherheit.“

4. Welche neuen Erkenntnisse hast du aus den beiden Unterrichtsstunden gewonnen?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Wissen über Deepfakes und deren Funktionsweise	Die Befragten geben an, neues Wissen über Entstehung, Funktionsweise oder Verbreitung von Deepfakes gewonnen zu haben.	„Deepfakes werden mit KI erstellt“	Alle Aussagen, die technische Grundlagen, Entstehung oder Verbreitung beschreiben, werden dieser Kategorie zugeordnet.
2: Erkennung von Deepfakes und Medienkompetenz	Die Befragten berichten über neu erlernte Fähigkeiten zur Erkennung oder kritischen Bewertung von Deepfakes.	„Außerdem weiß ich jetzt besser, wie Deepfakes erkannt werden können.“	Alle Aussagen zu Erkennungsmethoden, Prüfstrategien oder kritischem Umgang mit Medien werden dieser Kategorie zugeordnet.
3: Risiken und gesellschaftliche Auswirkungen von Deepfakes	Die Befragten nennen Gefahren, Manipulation oder gesellschaftliche Folgen von Deepfakes.	„Gefahr von Falschinformationen und Manipulation“	Alle Aussagen zu Risiken, Täuschung, Missbrauch oder gesellschaftlichen Auswirkungen werden dieser Kategorie zugeordnet.
4: Reflexion und Perspektivenerweiterung	Die Befragten reflektieren unterschiedliche Sichtweisen oder neue Einsichten zum Thema.	„jede Gruppe an Menschen (Jugendliche, Bürgermeister,...) hat eine andere Sichtweise was Deep Fakes betrifft.“	Alle Aussagen, die Perspektivwechsel, Reflexion oder Überraschung ausdrücken, werden dieser Kategorie zugeordnet.
5: Antworten ohne inhaltliche Aussage / Leerantworten	Antworten ohne verwertbaren inhaltlichen Bezug zum Thema Deepfakes (z. B. Leerzeichen, Striche oder fehlende Aussagen).	„-“	Alle Einträge ohne inhaltliche Aussage oder ohne Bezug zur Fragestellung werden dieser Kategorie zugeordnet.

Kategorie 1: Wissen über Deepfakes und deren Funktionsweise

- „Deepfakes werden mit KI erstellt“
- „Wie Deepfakes funktionieren und welche Gefahren & Probleme sie mit sich bringen“
- „Ich habe gelernt, wie Deepfakes erstellt werden“
- „Dass KI wirklich leider sehr viel besser geworden ist“
- „Tunnelblick von KI“
- „Das KI gegen KI benutzt wird.“

Kategorie 2: Erkennung von Deepfakes und Medienkompetenz

- „wie man sie erkennen kann“
- „Außerdem weiß ich jetzt besser, wie Deepfakes erkannt werden können.“
- „Sehr viel tiefgründiger als gedacht, es gibt noch viele Erkennungsarten für Deepfakes.“
- „Nicht sicher selbst posten.“

Kategorie 3: Risiken und gesellschaftliche Auswirkungen von Deepfakes

- „Deepfakes werden mit KI erstellt Sie können echt wirken und Menschen täuschen Gefahr von Falschinformationen und Manipulation“
- „Welche Gefahren & Probleme sie mit sich bringen“
- „das es sehr leicht ist, Leute auf dem Internet nach zu armen und schlimme Sachen damit anstellen kann“
- „Das Deepfakes öfter vor kommen als ich gedacht hätte.“

Kategorie 4: Reflexion und Perspektivenerweiterung

- „jede Gruppe an Menschen (Jugendliche, Bürgermeister,..) hat eine andere Sichtweise was Deep Fakes betrifft.“
- „Ich habe gelernt, wie realistisch Deepfakes geworden sind und dass man Inhalte immer kritisch prüfen sollte.“
- „Sehr viel tiefgründiger als gedacht, es gibt noch viele Erkennungsarten für Deepfakes.“
- „Das Deepfakes öfter vor kommen als ich gedacht hätte.“
- „Dass KI wirklich leider sehr viel besser geworden ist“

Kategorie 5: Antworten ohne inhaltliche Aussage / Leerantworten

- „-“
5. Mit welchen 3 - 5 Adjektiven (= Eigenschaftswörtern wie langweilig, spannend, relevant, (un)interessant ...) würdest du die beiden Unterrichtsstunden beschreiben?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Positive Bewertung (interessant/spannend/motivierend)	Die Unterrichtsstunden werden insgesamt als interessant, spannend oder ansprechend beschrieben.	„spannend, interessant, informativ“	Alle Aussagen, die die Stunden positiv als interessant, spannend, cool oder ansprechend bewerten, werden dieser Kategorie zugeordnet.
2: Informations- und Wissensbezogenheit	Die Unterrichtsstunden werden als informativ, lehrreich oder wissensvermittelnd beschrieben.	„interessant, interaktiv, informativ“	Alle Aussagen, die den Informationsgehalt oder Lernwert betonen (z. B. informativ, relevant), werden dieser Kategorie zugeordnet.
3: Aktualität und gesellschaftliche Relevanz	Die Unterrichtsstunden werden als aktuell, relevant oder wichtig für den Alltag bzw. die Gesellschaft beschrieben.	„spannend, relevant, informativ, modern“	Alle Aussagen, die Aktualität, gesellschaftliche Bedeutung oder Relevanz hervorheben, werden dieser Kategorie zugeordnet.

Kategorie	Definition	Ankerbeispiel	Kodierregel
4: Methodisch-didaktische Wahrnehmung (interaktiv/abwechslungsreich)	Die Unterrichtsstunden werden hinsichtlich ihrer Gestaltung als interaktiv, abwechslungsreich oder angenehm beschrieben.	„interessant, interaktiv, informativ“	Alle Aussagen, die Unterrichtsform, Interaktivität oder Abwechslung thematisieren, werden dieser Kategorie zugeordnet.

Kategorie 1: Positive Bewertung (interessant/spannend/motivierend)

- „interessant, aktuell, wichtig“
- „wichtig, interessant, gut“
- „interessant, interaktiv, informativ“
- „spannend, interessant, informativ“
- „Spannend, interessant, informativ, relevant, modern“
- „spannend, relevant, cool!“
- „spannend, relevant, informativ, chillig“
- „spannend, interessant, relevant, informativ, lustig“
- „Interessant, Informativ, Relevant“
- „spannend, interessant, neu, sehr alltagsnahe“
- „ansprechend, angenehm, persönlich“
- „Interessant, relevant, spannend, obwohl ich das meiste es wusste war trotzdem cool.“

Kategorie 2: Informations- und Wissensbezogenheit

- „interessant, interaktiv, informativ“
- „spannend, interessant, informativ“
- „Spannend, interessant, informativ, relevant, modern“
- „Interessant, Informativ, Relevant“
- „spannend, interessant, neu, sehr alltagsnahe“
- „Interessant, relevant, spannend, obwohl ich das meiste es wusste war trotzdem cool.“

Kategorie 3: Aktualität und gesellschaftliche Relevanz

- „interessant, aktuell, wichtig“
- „Spannend, interessant, informativ, relevant, modern“
- „spannend, relevant, cool!“
- „spannend, relevant, informativ, chillig“
- „spannend, interessant, relevant, informativ, lustig“
- „Interessant, Informativ, Relevant“
- „Interessant, relevant, spannend, obwohl ich das meiste es wusste war trotzdem cool.“

Kategorie 4: Methodisch-didaktische Wahrnehmung (interaktiv/abwechslungsreich)

- „interessant, interaktiv, informativ“
- „spannend, interessant, informativ“
- „Spannend, interessant, informativ, relevant, modern“
- „spannend, relevant, informativ, chillig“
- „spannend, interessant, relevant, informativ, lustig“
- „ansprechend, angenehm, persönlich“

6. Was hat dir an den Einheiten gut gefallen? Was hat dir nicht gut gefallen und wie würdest du es verändern?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Interaktive und soziale Lernformen (positiv bewertet)	Besonders positiv werden Gruppenarbeit, Diskussionen, Mindmaps und Rollenspiele hervorgehoben.	„Ich fand es cool, dass wir dann auch noch in kleinen Gruppen diskutieren konnten in verschiedenen Rollen.“	Alle Aussagen, die Gruppenarbeit, Austausch, Diskussion oder interaktive Methoden positiv bewerten, werden dieser Kategorie zugeordnet.
2: Inhaltliche Qualität (Beispiele, Verständlichkeit, Informativität)	Die Inhalte und Beispiele werden als hilfreich, interessant oder informativ bewertet, teilweise auch mit Kritik an Verständlichkeit.	„Gut gefallen haben mir die Beispiele und die Erklärung der Deepfakes.“	Alle Aussagen, die sich auf Inhalte, Beispiele, Erklärungen oder Informationsgehalt beziehen, werden dieser Kategorie zugeordnet.
3: Kritik an Umfang und Verständlichkeit	Einzelne Inhalte werden als zu lang, zu kompliziert oder schwer verständlich beschrieben.	„Weniger gut war, dass manche Inhalte etwas kompliziert waren.“	Alle Aussagen, die Überforderung, Unklarheit oder zu großen Umfang thematisieren, werden dieser Kategorie zugeordnet.
4: Verbesserungs- und Veränderungsvorschläge	Konkrete Vorschläge zur Verbesserung des Unterrichts werden genannt.	„Vllt bei der Mind map beim Brainstormen an der Tafel die Leute eif rausgehen lassen“	Alle Aussagen, die konkrete didaktische oder organisatorische Änderungen vorschlagen, werden dieser Kategorie zugeordnet.
5: Gesamtzufriedenheit / keine Änderungswünsche	Die Befragten äußern insgesamt Zufriedenheit und sehen keinen Änderungsbedarf.	„verändern: eigentlich gar nichts“	Alle Aussagen, die keine Kritik oder Änderungswünsche enthalten, werden dieser Kategorie zugeordnet.

Kategorie 1: Interaktive und soziale Lernformen (positiv bewertet)

- „guter Aufbau: mit Gruppenarbeit, MindMap, Fake oder echt,... / verändern: eigentlich gar nichts“
- „Ich fand es cool, dass wir dann auch noch in kleinen Gruppen diskutieren konnten in verschiedenen Rollen.“
- „gut: interaktiv (Mindmap, Diskussion)“
- „Gut gefallen haben mir die Beispiele und Diskussionen.“
- „Ich mag die Abwechslung.“
- „Es war sehr erkenntnisreich. Ich fand die Gruppenarbeit gut, sehr abwechslungsreich.“
- „Ich fand es sehr gut, auch das es interaktiv war.“

Kategorie 2: Inhaltliche Qualität (Beispiele, Verständlichkeit, Informativität)

- „Gut gefallen haben mir die Beispiele und Diskussionen.“
- „Gut gefallen haben mir die Beispiele und die Erklärung der Deepfakes.“
- „mir hat es sehr gut gefallen, weil die Lernperson nett und geduldig war. Es war auch ziemlich informativ“
- „Dass sie informativ war.“
- „Es war sehr erkenntnisreich.“

Kategorie 3: Kritik an Umfang und Verständlichkeit

- „Weniger gut war, dass manche Inhalte etwas kompliziert waren.“
- „Dass sie informativ war. Vielleicht ein wenig zu lang.“

Kategorie 4: Verbesserungs- und Veränderungsvorschläge

- „gut: interaktiv (Mindmap, Diskussion); könnte besser gemacht werden: teilweise deutlicher, kürzer erklären“
- „Ich würde mehr praktische Beispiele zeigen.“
- „Vllt bei der Mind map beim Brainstormen an der Tafel die Leute eif rausgehen lassen ohne aufzeigen dann wird mehr so ein flow vllt.“

Kategorie 5: Gesamtzufriedenheit / keine Änderungswünsche

- „guter Aufbau: mit Gruppenarbeit, MindMap, Fake oder echt,... / verändern: eigentlich gar nichts“
- „Finde ich schon, das Thema finde ich generell sehr interessant“
- „Ich mag die Abwechslung. Ich hätte nichts geändert.“
- „Spannend“

7. Würdest du den Unterricht einem Freund / einer Freundin empfehlen? Wenn ja, warum, wenn nein warum nicht?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Empfehlung aufgrund von Relevanz und Aktualität	Der Unterricht wird empfohlen, weil das Thema als aktuell, wichtig oder gesellschaftlich relevant wahrgenommen wird.	„Ja, weil das Thema aktuell und wichtig ist und man lernt, Medien kritischer zu hinterfragen.“	Alle Aussagen, die Empfehlung mit gesellschaftlicher Relevanz, Aktualität oder Wichtigkeit begründen, werden dieser Kategorie zugeordnet.
2: Empfehlung aufgrund von Interesse und Spannung	Der Unterricht wird empfohlen, weil das Thema als spannend, interessant oder unterhaltsam wahrgenommen wird.	„Ja, weil er insgesamt sehr spannend und lustig war.“	Alle Aussagen, die Empfehlung mit persönlichem Interesse, Spannung oder Spaß begründen, werden dieser Kategorie zugeordnet.
3: Lern- und Kompetenzgewinn als Begründung	Der Unterricht wird empfohlen, weil neue Inhalte gelernt werden oder Medienkompetenz gefördert wird.	„Ja, man lernt viele neue Sachen.“	Alle Aussagen, die Lernen, Erkenntnisgewinn oder Medienkritik als Begründung nennen, werden dieser Kategorie zugeordnet.
4: Interessengebundene Empfehlung	Der Unterricht wird nur empfohlen, wenn Interesse am Thema besteht oder bestimmte Bedingungen erfüllt sind.	„Ja, wenn er sich für das Thema interessiert!“	Alle Aussagen, die eine eingeschränkte oder bedingte Empfehlung ausdrücken, werden dieser Kategorie zugeordnet.

Kategorie 1: Empfehlung aufgrund von Relevanz und Aktualität

- „Ja, das Thema ist sehr wichtig heutzutage.“
- „Ja, weil das Thema aktuell und wichtig ist und man lernt, Medien kritischer zu hinterfragen.“
- „Ja, weil das Thema aktuell und wichtig ist.“
- „Ja, damit sie mehr über die Gefahren des Internets wissen und nicht alles sofort glauben“
- „Ja, es schadet nicht über KI gelehrt zu werden.“
- „Ja, da es ein wichtiges Thema ist.“
- „Ja schon weil es wir immer wichtiger“

Kategorie 2: Empfehlung aufgrund von Interesse und Spannung

- „ja, es ist ein spannendes Thema“
- „Ja, weil er insgesamt sehr spannend und lustig war.“
- „Ja für jede Person, die in dem Thema eine Interesse“

Kategorie 3: Lern- und Kompetenzgewinn als Begründung

- „Ja, man lernt viele neue Sachen.“
- „Ja, weil das Thema aktuell und wichtig ist und man lernt, Medien kritischer zu hinterfragen.“
- „Ja, damit sie mehr über die Gefahren des Internets wissen und nicht alles sofort glauben“
- „Ja wahrscheinlich schon, weil alles sehr verständlich war.“

Kategorie 4: Interessengebundene Empfehlung

- „Ja für jede Person, die in dem Thema eine Interesse“
- „Ja, wenn er sich für das Thema interessiert!“

Antworten inklusive Feedback 3. Durchgang

1. Welche Charakteristika kennzeichnen Deepfakes? Wie würdest du Deepfakes definieren? Erkläre in etwa drei Sätzen, wie die Technologie hinter Deepfakes funktioniert.

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Deepfakes als KI-generierte oder manipulierte Medien	Deepfakes werden als künstlich erzeugte oder manipulierte Bilder, Videos oder Audios beschrieben.	„Das sind mit KI erstellte Bilder, Videos oder Audios, die echt wirken.“	Alle Aussagen, die Deepfakes als KI-generierte oder manipulierte Medien definieren, werden dieser Kategorie zugeordnet.
2: Realistische Wirkung und gesellschaftliche Folgen	Deepfakes werden als täuschend echt beschrieben und mit Manipulation, Desinformation oder negativen Auswirkungen in Verbindung gebracht.	„Es werden Informationen täuschend realistisch dargestellt.“	Alle Aussagen, die die täuschende Wirkung oder gesellschaftliche Folgen wie Desinformation oder Rufschädigung thematisieren, werden dieser Kategorie zugeordnet.
3: Technische Funktionsweise und Erkennungsmerkmale	Die Entstehung oder Erkennung von Deepfakes wird anhand der Arbeitsweise von KI oder typischer Fehler erklärt.	„KI lernt aus vielen Daten.“	Alle Aussagen, die beschreiben, wie Deepfakes entstehen oder woran sie erkannt werden können, werden dieser Kategorie zugeordnet.

Kategorie 1: Deepfakes als KI-generierte oder manipulierte Medien

- „Das sind mit KI erstellte Bilder, Videos oder Audios, die echt wirken.“
- „Deepfakes sind von KI veränderte Bilder.“
- „Deepfakes sind von KI dargestellte Inhalte.“
- „KI manipulierte Bilder oder Videos“
- „KI-generierte Fälschungen“
- „Bilder, Videos, Stimmen“

Kategorie 2: Realistische Wirkung und gesellschaftliche Folgen

- „Es werden Informationen täuschend realistisch dargestellt.“
- „Diese verbreiten sich schnell im Internet und lösen oft starke Emotionen aus.“
- „Außerdem können sie Leuten, wie zum Beispiel Politikern sehr schaden.“
- „Wirken echt“

Kategorie 3: Technische Funktionsweise und Erkennungsmerkmale

- „Man kann sie an kleinen Fehlern oder einem unnatürlichen Aussehen erkennen.“
- „Bilder sind einfacher als Videos zu faken.“
- „Man kann Deepfakes erkennen in dem oft im Hintergrund nicht alles passt.“
- „Auf die Hände sollte man auch achten da diese nicht korrekt von der KI generiert werden kann.“
- „Plötzliche Bewegungen sehen oft komisch aus.“
- „KI lernt aus vielen Daten.“
- „Erstellt neue Inhalte“

2. Welche Methoden existieren zur Erkennung von Deepfakes? Teile zu jeder Methode mit, inwieweit sie noch zuverlässig ist.

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Visuelle und auditive Merkmale analysieren	Deepfakes werden anhand von Auffälligkeiten im Bild, Video oder Ton erkannt.	„Bei Videos: Hand und Mund Bewegungen, Art des Sprechens“	Alle Aussagen, die Bild-, Video- oder Tonmerkmale zur Erkennung von Deepfakes nennen, werden dieser Kategorie zugeordnet.
2: Quellen- und Faktenprüfung	Die Überprüfung von Quellen, Autor:innen, Veröffentlichungsdatum oder Fakten dient der Erkennung von Deepfakes.	„Quellen überprüfen → sehr zuverlässig“	Alle Aussagen, die die Herkunft oder Glaubwürdigkeit eines Mediums überprüfen, werden dieser Kategorie zugeordnet.
3: Technische Erkennung durch KI-Tools	Spezielle Programme oder KI-Systeme werden zur Erkennung von Deepfakes eingesetzt.	„KI Bilderkenner“	Alle Aussagen, die KI-Erkennungstools oder technische Hilfsmittel nennen, werden dieser Kategorie zugeordnet.
4: Grenzen der Erkennung und sinkende Zuverlässigkeit	Die Befragten weisen darauf hin, dass Erkennungsmethoden durch die Weiterentwicklung von KI zunehmend unzuverlässig werden.	„durch die Besserung der KI: immer unzuverlässiger“	Alle Aussagen, die die begrenzte oder abnehmende Zuverlässigkeit von Erkennungsmethoden thematisieren, werden dieser Kategorie zugeordnet.

Kategorie 1: Visuelle und auditive Merkmale analysieren

- „Bei Videos: Hand und Mund Bewegungen, Art des Sprechens, Inhalt; Bilder: Hände(Finger), Mund, Haut, unnatürlich“
- „selber das Bild oder Video nach Fehlern untersuchen“
- „selber nach ungewöhnlichen Sachen schauen“
- „Schauen ob es ungewöhnliche Merkmale gibt“
- „Bild-/Tonfehler → mittel“

Kategorie 2: Quellen- und Faktenprüfung

- „Quellen überprüfen → sehr zuverlässig“
- „Autor/Datum überprüfen → auch sehr zuverlässig“
- „Quellen/Fakten prüfen → sehr zuverlässig“

Kategorie 3: Technische Erkennung durch KI-Tools

- „Ki erkennungstools → nicht immer sicher“
- „KI Bilderkenner“
- „KI-Erkennungstools → gut“

Kategorie 4: Grenzen der Erkennung und sinkende Zuverlässigkeit

- „Ki erkennungstools → nicht immer sicher“
- „durch die Besserung der KI: immer unzuverlässiger“
- „solange die KI sich nicht extremst bessert ist diese Methode bisher zuverlässig“
- „Bild-/Tonfehler → mittel“

3. Welche Gefahren bringen Deepfakes mit sich? Nenne jeweils 1-3 Gefahren, die

* Einzelpersonen

* die Politik

* die Gesellschaft

betreffen.

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Desinformation und Manipulation	Deepfakes werden als Gefahr für die Verbreitung falscher Informationen und die gezielte Manipulation von Personen oder Meinungen beschrieben.	„Politik: Verbreitung von Falschinformationen, Manipulation von Wahlen“	Alle Aussagen, die Fake News, Desinformation, Manipulation oder Beeinflussung thematisieren, werden dieser Kategorie zugeordnet.
2: Persönliche Schäden für Einzelpersonen	Deepfakes werden als Ursache für persönliche, finanzielle oder psychische Schäden beschrieben.	„Einzelpersonen: Rufschädigung, Identitätsdiebstahl, Erpressung“	Alle Aussagen, die Rufschädigung, Betrug, Mobbing, Identitätsdiebstahl oder persönliche Nachteile nennen, werden dieser Kategorie zugeordnet.
3: Gesellschaftliche Folgen und Vertrauensverlust	Deepfakes führen zu Angst, Unsicherheit, Misstrauen oder	„Gesellschaft: Verbreitung von Fake News, Angst“	Alle Aussagen, die gesellschaftliche Auswirkungen wie

Kategorie	Definition	Ankerbeispiel	Kodierregel
	beeinflussen das gesellschaftliche Zusammenleben.		Verunsicherung, Misstrauen oder Vertrauensverlust beschreiben, werden dieser Kategorie zugeordnet.
4: Politische Einflussnahme	Deepfakes werden als Gefahr für Politik, demokratische Prozesse und Wahlen beschrieben.	„Wahlmanipulation“	Alle Aussagen, die sich auf Wahlen, Politiker oder politische Entscheidungen beziehen, werden dieser Kategorie zugeordnet.

Kategorie 1: Desinformation und Manipulation

- „Politik: Verbreitung von Falschinformationen“
- „Sie liefern falsche Informationen.“
- „Falschinformation“
- „Politik: die Meinung eines Politikers kann zum Beispiel in einem Video verändert werden.“
- „Manipulation“
- „Politik: Desinformation“

Kategorie 2: Persönliche Schäden für Einzelpersonen

- „Einzelpersonen: Rufschädigung, Identitätsdiebstahl, Erpressung“
- „Mobbing“
- „Einzelpersonen: beeinflusst dich den falschen Inhalt“
- „Rufschädigung, Betrug“
- „Einzelpersonen: Betrug, Rufschädigung, Identitätsdiebstahl“

Kategorie 3: Gesellschaftliche Folgen und Vertrauensverlust

- „Gesellschaft: Verbreitung von Fake News, Angst“
- „Angst, Verwirrung“
- „Gesellschaft: genereller Einfluss auf das Handeln der Menschen auch gegenüber einander“
- „Gesellschaft: Falschinformationen, Misstrauen, Verunsicherung“

Kategorie 4: Politische Einflussnahme

- „Politik: Manipulation von Wahlen“
- „Politik: die Meinung eines Politikers kann zum Beispiel in einem Video verändert werden.“
- „Wahlmanipulation“
- „Politik: Wahlbeeinflussung, Vertrauensverlust“

4. Welche neuen Erkenntnisse hast du aus den beiden Unterrichtsstunden gewonnen?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Wissen über Deepfakes und deren Entstehung	Die Befragten berichten, dass sie neues Wissen über die Entstehung oder Funktionsweise von Deepfakes gewonnen haben.	„Entstehung von Deepfakes“	Alle Aussagen, die die Entstehung oder Funktionsweise von Deepfakes als neue Erkenntnis nennen, werden dieser Kategorie zugeordnet.
2: Erkennung und kritischer Umgang mit Medien	Die Unterrichtseinheiten vermittelten Wissen zur Erkennung von Deepfakes sowie zum kritischen Hinterfragen von Medieninhalten.	„Wie man deepfakes erkennt“	Alle Aussagen, die Erkennungsmöglichkeiten oder einen kritischeren Umgang mit Informationen beschreiben, werden dieser Kategorie zugeordnet.
3: Risiken und Auswirkungen von Deepfakes	Die Befragten geben an, die Gefahren und Folgen von Deepfakes besser verstanden zu haben.	„Die Auswirkungen der Deepfakes wurden mir noch deutlicher.“	Alle Aussagen, die Gefahren, Risiken oder gesellschaftliche Auswirkungen thematisieren, werden dieser Kategorie zugeordnet.

Kategorie 1: Wissen über Deepfakes und deren Entstehung

- „Wie man deepfakes erkennt und aus welchen Gründen sie entstehen.“
- „Entstehung von Deepfakes“

Kategorie 2: Erkennung und kritischer Umgang mit Medien

- „Wie gefährlich Deep fakes sind und das man immer mehr aufpassen sollte was man auch weiter erzählt und wovon man sich nicht hundert prozent sicher ist, weil sich sowas schnell verbreitet“
- „Wie man deepfakes erkennt“
- „Das man nicht immer alles glauben soll“
- „Sie werden immer besser und schwieriger zu unterscheiden“
- „Erkennungsmöglichkeiten“

Kategorie 3: Risiken und Auswirkungen von Deepfakes

- „Wie gefährlich Deep fakes sind und das man immer mehr aufpassen sollte was man auch weiter erzählt und wovon man sich nicht hundert prozent sicher ist, weil sich sowas schnell verbreitet“
- „Die Auswirkungen der Deepfakes wurden mir noch deutlicher“
- „Sie werden immer besser und schwieriger zu unterscheiden“
- „Gefahren verstanden“

5. Mit welchen 3 - 5 Adjektiven (= Eigenschaftswörtern wie langweilig, spannend, relevant, (un)interessant ...) würdest du die beiden Unterrichtsstunden beschreiben?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Positive Gesamtbewertung	Die Unterrichtsstunden werden insgesamt als spannend, interessant oder positiv erlebt.	„Großartig spannend cool“	Alle Aussagen, die den Unterricht allgemein positiv bewerten (z. B. spannend, interessant, cool, großartig), werden dieser Kategorie zugeordnet.
2: Informations- und Lerngewinn	Die Unterrichtsstunden werden als informativ, lehrreich oder erkenntnisreich beschrieben.	„Spannend, Informativ, Interessant, Relevant, Lehrreich“	Alle Aussagen, die den Wissensgewinn oder Informationsgehalt hervorheben, werden dieser Kategorie zugeordnet.
3: Unterrichtsgestaltung und Methodik	Die Gestaltung der Unterrichtsstunden wird als abwechslungsreich, strukturiert oder motivierend beschrieben.	„Abwechslungsreich, spannend, strukturiert, informierend“	Alle Aussagen, die sich auf den Aufbau oder die methodische Gestaltung des Unterrichts beziehen, werden dieser Kategorie zugeordnet.
4: Kritische oder eingeschränkte Bewertung	Die Befragten äußern einzelne Einschränkungen oder Kritikpunkte trotz insgesamt positiver Bewertung.	„schon bekannte infos, bissi unspektakulär“	Alle Aussagen, die den Unterricht als teilweise wenig neu, unspektakulär oder eingeschränkt bewerten, werden dieser Kategorie zugeordnet.

Kategorie 1: Positive Gesamtbewertung

- „Interessant“
- „lustig“
- „spannend“
- „neugierig“
- „interessant“
- „wichtig“
- „Großartig spannend cool“
- „Spannend interessant“
- „Spannend“
- „Interessant“
- „Relevant“

Kategorie 2: Informations- und Lerngewinn

- „informierend“
- „Informativ“
- „Lehrreich“
- „Relevant“

Kategorie 3: Unterrichtsgestaltung und Methodik

- „Abwechslungsreich“
- „abwechslungsreich“
- „strukturiert“

Kategorie 4: Kritische oder eingeschränkte Bewertung

- „schon bekannte infos“
- „bissi unspektakulär“

6. Was hat dir an den Einheiten gut gefallen? Was hat dir nicht gut gefallen und wie würdest du es verändern?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Positive Wahrnehmung von Inhalt und Gestaltung	Der Unterricht wird insgesamt positiv bewertet, insbesondere hinsichtlich Abwechslung, Beispielen, Gruppenarbeit oder interessanten Elementen.	„Ich fand die abwechslungsreiche Gestaltung des Unterrichts sehr gut“	Alle Aussagen, die Unterricht, Inhalte oder Methoden grundsätzlich positiv bewerten, werden dieser Kategorie zugeordnet.
2: Kritik an Ablauf und Methodik	Einzelne Unterrichtsbestandteile werden als zu lang, unnötig oder verbesserungswürdig beschrieben.	„Vielleicht nicht den langen Text am Anfang vorlesen.“	Alle Aussagen, die konkrete Kritik am Ablauf, an Materialien oder an der Durchführung äußern, werden dieser Kategorie zugeordnet.
3: Wünsche nach Verbesserung oder mehr Praxisbezug	Es werden konkrete Vorschläge gemacht, z. B. mehr Praxis oder realistischere Beispiele.	„mehr Praxis, mehr echte Deepfakes“	Alle Aussagen, die konkrete Verbesserungs- oder Änderungswünsche enthalten, werden dieser Kategorie zugeordnet.

Kategorie 1: Positive Wahrnehmung von Inhalt und Gestaltung

- „Es war sehr abwechslungsreich.“
- „Sonst war es gut gestaltet und auch lustig mit dem edu lab.“
- „Ich fand die abwechslungsreiche Gestaltung des Unterrichts sehr gut“
- „Das uns gezeigt wurde wie wir erkennen wenn es KI generiert ist“
- „Gut: Beispiele, Gruppenarbeit, Diskussion“
- „Gruppenarbeit j“

Kategorie 2: Kritik an Ablauf und Methodik

- „Ich finde nur, dass wir den zeitungsbericht nicht alle zusammen lesen sollten, weil das echt lange gedauert hat.“
- „Vielleicht nicht den langen Text am Anfang vorlesen.“
- „Das thema war uns halt schon bekannt weshalb wir auch wneig zu sagen hatten“

Kategorie 3: Wünsche nach Verbesserung oder mehr Praxisbezug

- „mehr Praxis, mehr echte Deepfakes“

7. Würdest du den Unterricht einem Freund / einer Freundin empfehlen? Wenn ja, warum, wenn nein warum nicht?

Kategorie	Definition	Ankerbeispiel	Kodierregel
1: Positive Empfehlung aufgrund von Lerngewinn und Relevanz	Der Unterricht wird empfohlen, weil er Wissen vermittelt oder als nützlich bzw. aktuell bewertet wird.	„Ja, weil diese Unterrichtsstunden mich viel über dieses Thema unterrichtet haben“	Alle Aussagen, die Empfehlung mit Lerngewinn, Aktualität oder Nützlichkeit begründen, werden dieser Kategorie zugeordnet.
2: Positive Empfehlung aufgrund von Gestaltung und Interesse	Der Unterricht wird empfohlen, weil er als abwechslungsreich, interessant oder gut gestaltet wahrgenommen wird.	„Ja, weil er mit der Mindmap am Anfang und dem Roleplay am Ende eine gute Abwechslung hatte und es nicht langweilig war.“	Alle Aussagen, die Empfehlung mit Spaß, Abwechslung oder Gestaltung begründen, werden dieser Kategorie zugeordnet.
3: Eingeschränkte oder negative Empfehlung	Der Unterricht wird nur teilweise oder nicht eindeutig empfohlen, z. B. wegen Langeweile oder fehlender Motivation.	„Geht so, da es lustig war jedoch langweilig und in die Länge gezogen“	Alle Aussagen, die Kritik oder eine eingeschränkte Empfehlung enthalten, werden dieser Kategorie zugeordnet.

Kategorie 1: Positive Empfehlung aufgrund von Lerngewinn und Relevanz

- „Ja, weil diese Unterrichtsstunden mich viel über dieses Thema unterrichtet haben“
- „Ja“
- „Jap Aufjedenfall“
- „Ja“
- „Aktuelles Thema“
- „Nützlich“
- „Deepfakes besser erkennen“

Kategorie 2: Positive Empfehlung aufgrund von Gestaltung und Interesse

- „Ja, weil er mit der minimal am Anfang und dem roleplay am Ende eine gute Abwechslung hatte und es nicht langweilig war.“

Kategorie 3: Eingeschränkte oder negative Empfehlung

- „Geht so, da es lustig war jedoch langweilig und in die länge gezogen“