

DIGITALA

UND

Lehrbuch
zum
Digitalen
Humanismus

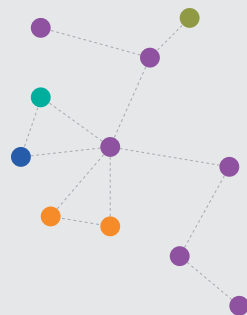
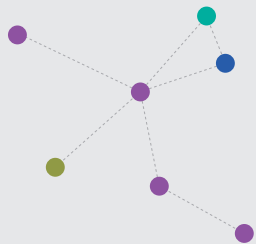
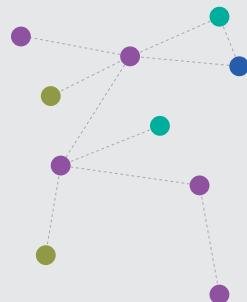
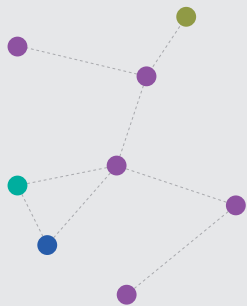
LISIEREA

UNG

mit
praktischen
Übungen

WIR

Anita Eichinger, Peter Knees
und Hannes Werthner (Hg.)



DIGITALI SIERUNG UND WIR

Lehrbuch zum Digitalen Humanismus
mit praktischen Übungen

Anita Eichinger, Peter Knees
und Hannes Werthner (Hg.)

Inhalt

Vorwort 6

Kapitel 1

**DIE DIGITALE
TRANSFORMATION 9**

Kapitel 2

**ENTWICKLUNG
DER INFORMATIK 25**

Kapitel 3

**EINFÜHRUNG IN DIE
KÜNSTLICHE INTELLIGENZ 39**

Kapitel 4

**AUSWIRKUNGEN DER
DATENINTENSIVEN
KÜNSTLICHEN INTELLIGENZ 51**

Kapitel 5

**KI-GESETZGEBUNG UND
IHRE HERAUSFORDERUNGEN 65**

Kapitel 6

**EMPFEHLUNGSSYSTEME UND
IHRE AUSWIRKUNGEN AUF
PLURALITÄT UND FAIRNESS 83**

Kapitel 7

DAS VERZERLTE NETZ:
VORURTEILE IM WEB 99

Kapitel 8

DIE NOTWENDIGKEIT
DES SCHUTZES
DER PRIVATSPHÄRE 111

Kapitel 9

ALGORITHMISCHE MODERATION
VON INHALTEN 123

Kapitel 10

DER MENSCH IM
MITTELPUNKT DER
SYSTEMENTWICKLUNG 139

Kapitel 11

PHILOSOPHISCHE
GRUNDLAGEN DES
DIGITALEN HUMANISMUS 155

Kapitel 12

DER DIGITALE HUMANISMUS 171

Glossar 190

Herausgeber:innen 198

Impressum 199

Die Digitalisierung betrifft uns alle – als Einzelne und als Gesellschaft. Die Veränderungen sind so weitreichend, dass eine Betrachtung allein durch die Linse der Informatik und der technischen Disziplinen nicht mehr ausreicht, um zu verstehen, was Digitalisierung bedeutet. Mit diesem Buch wollen wir dazu beitragen, diese Entwicklungen in ihrer Breite zu verstehen, und zeigen, dass wir eingreifen und diesen Prozess gestalten können.

Dieses Arbeitsbuch stellt in zwölf Kapiteln ausgewählte Schwerpunkte im Hinblick auf Digitalisierung vor. Es erläutert den technologischen Kontext und die Entwicklungen der Informatik sowie deren Auswirkungen auf die Gesellschaft auf Basis einer humanistischen Anschauung. Die Themen reichen von der digitalen Transformation bis zum Digitalen Humanismus. Ein Abschnitt behandelt die Geschichte der Informatik, ein weiterer bietet eine Einführung in die Künstliche Intelligenz (KI). Die Auswirkungen von KI auf Mensch und Natur werden ebenso aufgezeigt wie die Notwendigkeit einer Gesetzgebung zur Regulierung von KI. Zur Diskussion gestellt werden Empfehlungssysteme und ihr Einfluss auf Pluralität und Fairness sowie Vorurteile im Web. Der Notwendigkeit des Schutzes der Privatsphäre ist ein Kapitel gewidmet, ebenso den Problemen, die sich durch eine algorithmische Moderation von Inhalten ergeben können. All dies wird kontrastiert mit einer menschenzentrierten Systementwicklung auf Basis philosophischer Grundlagen, wie der Digitale Humanismus sie anstrebt.

Die Grundlage dieser Kapitel bildet das von namhaften internationalen Expert:innen auf Englisch verfasste interdisziplinäre Lehrbuch *Introduction to Digital Humanism*, herausgegeben von Hannes Werthner, Carlo Ghezzi, Jeff Kramer, Julian Nida-Rümelin, Bashar Nuseibeh, Erich Prem und Allison Stanger (Springer, 2024). Während dieses für den tertiären Bildungsbereich konzipiert ist, ist *Digitalisierung und wir* ein Arbeitsbuch mit praktischen Übungen, Anleitungen und Aktivitäten für höhere Schulen und die Erwachsenenbildung. Das Buch soll den Wissensvermittler:innen als Inspiration zur Gestaltung von Lehreinheiten dienen und Input dafür liefern. Es bildet einen Startpunkt für die Auseinandersetzung mit den vielschichtigen und fachübergreifenden Themenkomplexen. *Digitalisierung und wir* stellt einen altersgerechten und niederschweligen Zugang zu den Inhalten aus *Introduction to Digital Humanism* dar. Jedes Kapitel ist unterteilt in Abschnitte, ergänzt durch Wissensfragen, Fragen zur Diskussion, Aktivierungen und Gruppenarbeiten, jeweils farblich kodiert. Die Struktur des Arbeitsbuchs ist am Beginn in Form eines Netzwerks visualisiert.

Wir erheben hinsichtlich der Auswahl der Kapitel und Themen keinen Anspruch auf Vollständigkeit. Auch kann ein derartiges Buch aufgrund der Dynamik der Entwicklung niemals abgeschlossen sein. Vielmehr handelt es sich mehrheitlich um solche Themenkomplexe, anhand derer das Aufeinandertreffen technischer Lösungen und gesellschaftlicher Fragen diskutiert werden kann und die zeigen, wie der Stand der Technik Normen implementiert oder definiert. Wir erwarten, dass der Einsatz dieses Arbeitsbuchs konkreter aufzeigt, welche Themen vermehrt Hintergrundwissen erfordern bzw. einer genaueren Auseinandersetzung bedürfen. Es handelt sich somit um den Beginn einer Entwicklung und nicht um einen Endpunkt.

Bei der Ausgestaltung der Inhalte wurde auf Unterstützung durch Künstliche Intelligenz zurückgegriffen. Die englischen Inhalte wurden mit DeepL ins Deutsche übersetzt und teilweise mithilfe von ChatGPT vereinfacht. Die dabei gemachte Erfahrung wird einige, vor allem im Kontext dieses Buchs, beruhigen, andere enttäuschen: Eine maschinelle Erstellung der Lehrinhalte und pädagogischer Interventionen erwies sich als unzulänglich.

Unser Dank gilt den Autorinnen und Autoren der den einzelnen Kapiteln zugrunde liegenden Texten. Darüber hinaus Stefan Woltran für inhaltliches Feedback. Christian Mertens und Julia Friehs sei für Korrekturen und Lektorat gedankt. Lisa Ifsits für die wunderbare Grafik und inhaltliche Unterstützung. Ganz großer Dank gebührt Bettina K. Lechner für das didaktische Konzept.

Der Digitale Humanismus stellt die Menschen, die Gesellschaft und ihre Umwelt in den Mittelpunkt technischer Entwicklungen und zeigt auf, wie Technologie zu einer gerechteren, sozialeren Welt beitragen kann. Nicht zuletzt deshalb haben wir uns entschieden, dieses Buch *open access* zu publizieren, sodass eine Verwendung in anderen Ländern und Sprachen leicht möglich ist.

Anita Eichinger, Peter Knees und Hannes Werthner
Wien, November 2024



- KAPITELABSCHNITTE
- WISSENSFRAGEN
- FRAGE(N) ZUR DISKUSSION
- AKTIVIERUNG
- GRUPPENARBEIT

Kapitel 1

DIE DIGITALE TRANSFORMATION

In diesem Kapitel widmen wir uns der digitalen Transformation, die mittlerweile fast alle Lebensbereiche berührt. Wir betrachten die Auswirkungen von Informatik und Web auf Individuen und die Gesellschaft und zeigen, dass sie tiefgreifende Änderungen nach sich ziehen. Daher ist eine gezielte Gestaltung dieser Systeme notwendig.

Basis dieses Kapitels ist der Aufsatz „Digital Transformation, Digital Humanism: What Needs to Be Done“ von Hannes Werthner, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024. DOI: 10.1007/978-3-031-45304-5_8.

Einleitung

Wir befinden uns in einer spannenden Zeit. Die **digitale Transformation** verändert unser Leben. Diese Transformation begann mit der Erfindung des ersten elektronischen Computers vor etwa 80 Jahren und hat dazu geführt, dass Informationstechnologie (IT) heute das „Betriebssystem“ unserer Gesellschaft ist. Sichtbar wurde das vor allem zu Zeiten der COVID-19-Pandemie, in der viele von uns dank Homeoffice oder Homeschooling weiterarbeiten und weiterlernen konnten. Aber auch die Nachteile eines digitalen Lebens ohne physische Treffen und Zusammenkünfte wurden offenbar (Einsamkeit, Depression u. v. m.). Diese Veränderungen sind tiefgreifend und werden sich weiter verstärkt fortsetzen. Computer und Netzwerke beeinflussen jeden Teil unseres Lebens und haben zu einer engen Verbindung zwischen Menschen und Maschinen geführt. Mehr noch, wir erleben eine Ko-Evolution von Menschen und Technologie, wobei sich beide gegenseitig beeinflussen und weiterentwickeln.

Beispiele für die digitale Transformation in allen Lebensbereichen sind:

- **Kommunikation:** Digitale Kommunikationstechnologien erlauben uns, quasi in Echtzeit mit praktisch jedem Menschen auf der ganzen Welt auf verschiedene Arten zu interagieren.
- **Medien:** Aufbauend auf diesen Technologien findet Information online statt. Der Unterschied zwischen jenen, die Inhalte erstellen, und jenen, die sie konsumieren, verschwimmt. Physische Printmedien und klassischer Rundfunk werden als Massenmedien in Nischen abgedrängt.
- **Industrie:** Fertigungsprozesse und Logistik sind eng und weltweit vernetzt und werden verstärkt an individuelle Anforderungen angepasst.
- **Konsum und Wirtschaft:** Einkaufen in allen Formen verlagert sich in die Onlinewelt. Digitale Währungen und Transaktionen werden zur Normalität, und Bargeld verliert immer weiter an Bedeutung.
- **Unterhaltung:** Musik, Radio, Videos u. a. werden digital produziert und personalisiert vertrieben. Dabei wird das Konsum- und Nutzungsverhalten erfasst und zur Optimierung dieser Prozesse und zur Empfehlung anderer Inhalte eingesetzt.

Dies sind nur einige Beispiele, auch in Wissenschaft, Medizin und Gesundheit, Verwaltung, Landwirtschaft, Umwelt, Verkehr, Justiz, Polizei und Verbrechensbekämpfung – in der Kriminalität selbst – und auch in der Kriegsführung hat die Informationstechnologie die Welt verändert. In der Praxis greifen diese Bereiche ineinander und erlauben einen schnellen Austausch von Information und Anwendungen über ihre jeweiligen Grenzen hinaus. Noch nie gab es so viele Start-ups und Firmen mit neuen Ideen und Lösungen.

Nicht alles an dieser Entwicklung ist positiv. Obwohl der technologische Fortschritt viele Vorteile bietet, gibt es auch Nachteile und Herausforderungen. Diese reichen von wirtschaftlichen und gesellschaftlichen Problemen und Ungleichheiten bis hin zu politischen und ökologischen Krisen. Die Realität der digitalen Ökonomie zeigt zum Beispiel, dass neue Märkte vielfach unfair sind und gleichzeitig etablierte Wirtschaftstheorien auf die jetzige Situation nicht mehr direkt anwendbar sind. Ein neues, umfassendes Verständnis der digitalen Transformation ist daher notwendig, um unsere Gesellschaft in eine gerechtere, respektvollere und nachhaltigere Zukunft zu führen. Der Digitale Humanismus, den wir am Ende dieses Buchs erkunden (siehe Seite 171), versucht, Technologie so zu gestalten, dass eine sozialere, gerechtere, klimafreundlichere Welt entsteht.

AKTIVIERUNG

Digitale Transformation

Materialien:

- Kärtchen/Post-its
- Smartphone

Die Teilnehmer:innen legen ihr Smartphone vor sich auf den Tisch und notieren auf Kärtchen, welche positiven und negativen Auswirkungen der digitalen Transformation sie damit verbinden.

Das Web und seine Auswirkungen

Das Internet und das Web sind heute die wohl wichtigsten technologischen Errungenschaften. Ihre rasante Entwicklung hat weitreichende Auswirkungen auf unsere Gesellschaft und unser tägliches Leben. Was als einfaches Netzwerk begann, um Informationen zwischen Wissenschaftler:innen auszutauschen, hat sich zu einer globalen Infrastruktur entwickelt, die fast alle Aspekte des modernen Lebens beeinflusst.

- **Globale Vernetzung:** Das Internet verbindet Milliarden von Menschen weltweit. Dieser Zugang zu Informationen und Kommunikation hat die Art und Weise, wie wir lernen, arbeiten und interagieren, grundlegend verändert.
- **Digitale Kluft:** Trotz der weitreichenden Vorteile des Internets gibt es erhebliche Ungleichheiten beim Zugang. Während in Industrieländern fast jede:r online ist, haben in vielen Entwicklungsländern weniger als die Hälfte der Menschen einen Internetzugang. Diese digitale Kluft zeigt sich nicht nur zwischen Ländern, sondern auch innerhalb von Gesellschaften. Laut Schätzung der Internationalen Fernmeldeunion der Vereinten Nationen (ITU) hatte 2023 ein Drittel der Weltbevölkerung keinen Zugang zum Internet.
- **Wirtschaftlicher Wandel:** Das Web hat neue Geschäftsfelder und -modelle hervorgebracht und traditionelle Märkte massiv verändert. Onlineplattformen wie Amazon und Alibaba haben die Einzelhandelslandschaft revolutioniert, während Dienste wie Uber und Airbnb traditionelle Branchen wie Taxi- und Hotelgewerbe herausfordern.
- **Soziale Expansion:** Kommunikationstools wie Skype, Facebook, WhatsApp und Zoom haben neue Formen sozialer Interaktion geschaffen, die es Menschen ermöglichen, über große Entfernungen hinweg in Verbindung zu bleiben. Insbesondere in Zeiten von Krisen, wie während der COVID-19-Pandemie, spielten diese Tools eine entscheidende Rolle bei der Aufrechterhaltung sozialer und beruflicher Kontakte.
- **Psychologische Veränderungen:** Die ständige Verfügbarkeit und der allgegenwärtige Zugang zu digitalen Medien haben auch unsere Gewohnheiten und psychologischen Muster beeinflusst. Phänomene wie die „FOMO“ (Fear of Missing Out), also die Angst, etwas zu verpassen, spiegeln die neuen sozialen Zwänge wider, die durch das Web entstanden sind.

- **Veränderungen auf politischer und rechtlicher Ebene:** Das Web hat auch das politische Spektrum erweitert. Es ermöglicht schnelle und weitreichende Mobilisierungen für politische Kampagnen und Bewegungen, stellt uns gleichzeitig aber vor neue Herausforderungen in Bezug auf Fake News, Datenschutz und Cybersicherheit.
- **Veränderung physischer Räume:** Die Zunahme des elektronischen Handels (E-Commerce) hat nicht nur das Kaufverhalten verändert, sondern auch die physische Landschaft von Handel und Geschäften. Einkaufszentren und Ladengeschäfte verzeichnen einen Kundenschwund, während Logistikzentren und Lieferdienste boomen.

Das Netz von den Anfängen bis zur Gegenwart

Das Internet und das Web wurden ursprünglich auch von der US-amerikanischen Anti-Establishment-Bewegung und ihrer utopischen Vision einer offenen und dezentralisierten Kommunikation inspiriert. Frühe Formen des Internets wie Newsgroups und Bulletin-Boards

Kurzer Exkurs zu Informatik:

→ Siehe auch Kapitel 2, Seite 25

Informatik ist das Studium des Computers – einer Maschine, die für fast alles verwendet werden kann. Ein Computer kann nahezu jede Aufgabe erledigen, von der Wiedergabe von Musik bis hin zur Steuerung eines Kraftwerks.

- **Allzweckautomat:** Der Computer kann sein „Verhalten“ ändern und anpassen, basierend auf den Eingaben, die er erhält. Das macht ihn zu einer „General Purpose Technology“, die überall eingesetzt werden kann und schnell unsere Welt verändert.
- **Informatiker:innen** sind die Menschen hinter dieser Technologie. Sie entwickeln neue Software und Hardware, die unsere physischen und digitalen Welten formen.
- **Ubiquität der Software:** Software ist überall – sie steuert moderne Maschinen und sogar Dinge in unserem täglichen Leben, die wir nicht sehen können, wie z. B. in Smart Homes.

WISSENSFRAGEN

Das Web und seine Auswirkungen

Was versteht man unter der „digitalen Kluft“, und wie beeinflusst sie den Zugang zum Internet in verschiedenen Ländern?

Wie verändern Online-Plattformen wie Amazon und Airbnb traditionelle Geschäftsmodelle?

Inwiefern hat die COVID-19-Pandemie die Bedeutung digitaler Kommunikationstechnologien wie Skype und Zoom hervorgehoben?

In welcher Weise verstärken digitale Kommunikationsmittel, trotz ihrer Vorteile, soziale Isolation und psychologische Probleme, wie etwa „Fear of Missing Out“ (FOMO)?

AKTIVIERUNG

Gedankenexperiment – die Pandemie ohne IT

Ziel:

Ein Gedankenexperiment, das die Welt während der Pandemie ohne die Unterstützung durch Informationstechnologie (IT) betrachtet.

Materialien:

- Zettel und Stifte

1. Einführung und Gruppeneinteilung:

Beginnen Sie mit einer kurzen Einführung: Betonen Sie, wie stark unser Alltagsleben von der IT durchdrungen ist und wie dies während der Pandemie besonders deutlich wurde. Teilen Sie die Teilnehmer:innen in Kleingruppen mit je max. 3 bis 4 Personen. Sie dürfen für die Vorbereitung keine Smartphones oder Computer benutzen. Nur Zettel und Stift!

2. Anleitung – Die Welt ohne IT:

Anleitung des Gedankenexperiments mit folgenden Fragestellungen:

- ▶ Wie wäre die Pandemie verlaufen, wenn es keine IT gegeben hätte? Berücksichtigung von Homeoffice, Online-Unterricht, Informationsverbreitung über das Virus, Bestellen von Lebensmitteln und anderen Gütern etc.
- ▶ Wie hätten Regierungen, Unternehmen und Einzelpersonen reagiert?
- ▶ Welche alternativen Kommunikations- und Arbeitsmethoden hätten sich entwickeln können?

Die Teams diskutieren diese Fragen und entwickeln Szenarien, wie die Welt ohne die Unterstützung durch IT ausgesehen hätte.

3. Zusammenführung und Diskussion:

Jede Gruppe präsentiert kurz ihre Überlegungen und Szenarien. Gemeinsame Diskussion, welche Einblicke das Gedankenexperiment bietet und was dies über unsere Abhängigkeit von IT aussagt.

4. Reflexion und Schlussfolgerung:

Die Teilnehmer:innen reflektieren, was sie aus diesem Experiment gelernt haben.

ermöglichten einen freien Informationsaustausch und basierten auf dem Prinzip der basisdemokratischen Beteiligung. Schlüsselfiguren wie Doug Engelbart sahen in der Technologie ein Mittel, um menschliche Fähigkeiten zu erweitern und nicht, um sie zu ersetzen. Die Vernetzung von Menschen erlaubt weltweite Kooperationen. Das partizipative Online-Lexikon Wikipedia oder die Verbreitung von Open-Source-Software, die von allen genutzt und weiterentwickelt werden kann, sind nur zwei von vielen wunderbaren Beispielen in der vernetzten Online-Welt.

Das Netz hat sich seit seiner Entstehung weiterentwickelt, nicht immer zum Vorteil der Nutzer:innen:

- **Wirtschaftsmodell:** Das heutige Web basiert überwiegend auf Werbung. Große Suchmaschinen und Onlineshops übernehmen das Modell traditioneller Zeitungen, erweitern es jedoch durch Personalisierung und Empfehlungssysteme. Dies führt dazu, dass das Nutzerverhalten zunehmend von der Optimierung ihrer Klicks bestimmt wird, was wiederum die Einnahmen maximiert.
- **Emotionale Anreize und Überwachung:** Die Bereitschaft, über soziale Medien und andere Online-Dienste emotional aufgeladene Inhalte zu präsentieren, die mehr Nutzerengagement erzeugen, sowie das Sammeln von Nutzerdaten haben zum „Überwachungs-kapitalismus“ geführt.
- **Von der Basisdemokratie zur Kontrolle:** Was einst als Mittel zur Förderung der freien Meinungsäußerung und Beteiligung begann, hat sich zu einer von wenigen großen Unternehmen kontrollierten Infrastruktur entwickelt.

Der Begriff „Überwachungs-kapitalismus“ wurde von der US-amerikanischen Wirtschaftswissenschaftlerin Shoshana Zuboff geprägt. Sie versteht darunter ein marktwirtschaftliches-kapitalistisches System, in dem die Daten der Nutzer:innen gesammelt werden, um daraus Verhaltensvorhersagen zu treffen und damit Gewinne zu erwirtschaften. Für Zuboff ist der Überwachungs-kapitalismus ein totalitäres Projekt.

→ **Buch-/Hörbuchempfehlung:** Shoshana Zuboff, Das Zeitalter des Überwachungs-kapitalismus. Campus Verlag, 2020.

WISSENSFRAGEN

Das Netz von den Anfängen bis zur Gegenwart

Wie hat sich das ursprüngliche Ideal der basisdemokratischen Beteiligung im Web verändert?

Was ist Überwachungskapitalismus, und wie wird dieser Begriff von Shoshana Zuboff beschrieben?

Welche sozialen und wirtschaftlichen Auswirkungen hat die zunehmende Kontrolle und Kommerzialisierung persönlicher Daten durch private Unternehmen?

Welche ethischen und gesellschaftlichen Herausforderungen ergeben sich aus der zentralen Rolle von Algorithmen bei der Steuerung unserer Interaktionen und Entscheidungen im Web? Aus der Rolle der Algorithmen oder auch der Tatsache, dass diese in der Hand weniger Konzerne liegen?

Illusion der Freiheit und algorithmische Steuerung

Im Web erleben wir eine scheinbare Freiheit, die jedoch durch Algorithmen eingeschränkt wird. Was wir sehen, welche Informationen wir erhalten und wie wir interagieren, wird zunehmend von Algorithmen bestimmt, die darauf ausgelegt sind, aus den gewonnenen/gespeicherten Daten unser Verhalten vorherzusagen und zu beeinflussen (algorithmische Filter). Unsere Entscheidungen scheinen individuell, sind jedoch oft das Ergebnis algorithmisch aufbereiteter, vorselektierter Informationen. Das Ergebnis ist eine Illusion von Individualität und Gemeinschaft, die tatsächlich eine Zusammenfassung individuell zugeschnittener Ansichten ist. Damit einher geht ein uns oft nicht bewusster Verlust der individuellen und kollektiven Autonomie.

Automatisierung, Effizienz und strukturelle wirtschaftliche Veränderungen

Die fortschreitende Entwicklung des Webs und digitaler Technologien hat vor allem die Automatisierung und die Steigerung der Effizienz in den Vordergrund gerückt. Ziel ist es, Systeme zu schaffen, die optimal und nahezu autonom funktionieren können. Dabei wurden jedoch oft die **Resilienz** der Systeme und der menschliche Aspekt vernachlässigt, was zu einer systematischen Verdrängung vordergründig unproduktiver Elemente führte.

Wichtige Entwicklungen auf struktureller wirtschaftlicher Ebene:

1. „Informatisierung“ der Wertschöpfungskette:

Die Wertschöpfungsketten in vielen Branchen wurden durch die

Resilienz bedeutet, dass die Funktionalität eines Systems unter störenden inneren oder äußeren Einflüssen aufrechterhalten oder nach einer Störung die Funktionalität schnellstmöglich wiederhergestellt werden kann. Ein Beispiel für nicht resiliente Systeme sind die Lieferketten, die während der COVID-19-Pandemie massiv beeinträchtigt waren. Ein Beispiel für ein resilientes System hingegen ist das Internet. Es ermöglicht, dass Informationen, wenn eine Teilverbindung gestört ist, auf anderen Verbindungen ans Ziel kommen.

verstärkte Nutzung von Informationstechnologie umgestaltet. Firmen können dadurch Tätigkeiten leichter auslagern, die Produktion wird über mehrere Unternehmen verteilt. Dies führte zu einer zunehmenden Spezialisierung und einer Neustrukturierung ganzer Industrien. Diese „Informatisierung“ ist auch ein Schlüsselfaktor für die Globalisierung, da sie Unternehmen ermöglicht, ihre Tätigkeiten dort durchzuführen, wo es am günstigsten ist.

2. Outsourcing an Kund:innen:

Viele Aufgaben, die traditionell von Unternehmen übernommen wurden, werden nun an Kund:innen ausgelagert. Dies reicht von der Selbstbedienung an Geldautomaten bis hin zu Online-Diensten, bei denen Kund:innen viele Schritte des Verkaufs- und Serviceprozesses selbst übernehmen, ohne dafür bezahlt zu werden. Dadurch verändern sich Berufsfelder, und Jobs gehen verloren.

3. Privatisierung öffentlicher Daten:

In der digitalen Welt werden persönliche und öffentliche Daten zunehmend von privaten IT-Unternehmen kontrolliert und kommerziell genutzt. Diese Unternehmen nutzen diese Daten, um ihre Produkte zu verbessern und zu personalisieren, was ihnen wiederum enormen Profit einbringt. Ein Beispiel dafür ist die Nutzung der Daten von Wikipedia oder all die digitalisierten Bücher in Bibliotheken, die wesentlich zur Verbesserung der von Google generierten Suchergebnisse beitragen.

Die Tatsache, dass öffentliche Daten in den Händen privater Unternehmen liegen, wirft wichtige Fragen zu Datenschutz, Eigentumsrechten und der gerechten Verteilung von durch Daten generiertem Wert auf. Es entsteht eine neue Form der Wirtschaft, in der Informationen und Daten zur Hauptwährung werden.

Entwicklung der Unternehmenslandschaft in der Plattformökonomie

In den letzten 25 Jahren hat sich die Unternehmenslandschaft dramatisch verändert, insbesondere durch das Aufkommen und den Erfolg von Plattformunternehmen wie Apple, Microsoft, Amazon, Google, YouTube, Facebook, Uber, Airbnb und Instagram.

WISSENSFRAGEN

Automatisierung, Effizienz und strukturelle wirtschaftliche Veränderungen

Was bedeutet die „Informatisierung“ der Wertschöpfungskette, und wie trägt sie zur Zerstörung traditioneller Industrien bei?

Welche Aufgaben wurden durch das Outsourcing an Kund:innen übertragen, und welche Auswirkungen hat dies auf die Effizienz von Unternehmen?

Wie nutzen Plattformunternehmen wie Google oder Facebook öffentliche Daten, um wirtschaftlichen Wert zu generieren?

Inwiefern beeinflusst das Outsourcing an Kund:innen die Verantwortung und den Verlust menschlicher Arbeitskraft, und welche sozialen Folgen ergeben sich daraus?

FRAGEN ZUR DISKUSSION

Automatisierung, Effizienz und strukturelle wirtschaftliche Veränderungen

Wie beeinflusst Automatisierung Arbeitsplätze und den Alltag? Welche Vor- und Nachteile werden wahrgenommen?

1

Mögliche Diskussion: Die Automatisierung bringt erhebliche Effizienzsteigerungen und entlastet Menschen von repetitiven Aufgaben, was die Produktivität steigert. Jedoch führt sie auch zu einem Verlust bestimmter Arbeitsplätze und stellt Herausforderungen an die Qualifikationen der Arbeitskräfte. Die Vorteile umfassen niedrigere Produktionskosten und höhere Präzision, während die Nachteile Arbeitsplatzverluste und erhöhte Anforderungen an die Umqualifizierung betreffen.

Inwiefern trägt das Outsourcing von Aufgaben an Kund:innen zur Verschiebung der Verantwortung und zum Verlust menschlicher Arbeitskraft bei, und welche sozialen Folgen hat dies?

2

Mögliche Diskussion: Durch das Outsourcing von Aufgaben wie Selbstbedienung an Kassen und Online-Check-ins tragen Kund:innen zunehmend Verantwortung, die früher bei Unternehmen lag. Diese Verlagerung führt zu einem Verlust an einfachen Arbeitsplätzen und verringert zwischenmenschliche Interaktionen. Sozial kann dies zu Vereinsamung und dem Gefühl führen, dass Unternehmen ihre Pflichten an die Konsument:innen weitergeben.

Sollten Nutzer:innen für die Bereitstellung ihrer Daten entlohnt werden? Wie könnte ein gerechtes Modell dafür aussehen?

3

Mögliche Diskussion: Die Entlohnung für Daten wäre eine Möglichkeit, die Wertschöpfung gerechter zu verteilen. Ein Modell könnte sein, dass Plattformen den Nutzer:innen die Wahl lassen, ihre Daten zu teilen, und dafür monetäre oder andere Anreize bieten. Es könnten auch Mikrozahlungen pro Nutzung oder Datensatz stattfinden, die dem jeweiligen Wert der Daten entsprechen.

GRUPPENARBEIT

Gestaltung fairer Plattformen

Ziel:

Teilnehmer:innen entwickeln Ideen, wie Plattformökonomien fairer und inklusiver gestaltet werden können.

Lernziele:

- Kreative Problemlösungskompetenz stärken.
- Bewusstsein für die Verantwortung von Unternehmen in der digitalen Wirtschaft schärfen.

Materialien:

- Notebook, Tablet bzw. Smartphone zur Recherche

1. Gruppenaufteilung:

Die Teilnehmer:innen werden in kleine Teams aufgeteilt.

2. Auswahl der Plattform und Analyse:

Jedes Team wählt eine spezifische Plattform (z. B. Uber, Facebook) und analysiert, welche sozialen oder wirtschaftlichen Probleme diese Plattform verursacht.

3. Lösungsfindung:

Die Teams entwickeln kreative Lösungen oder Geschäftsmodelle, die die negativen Auswirkungen adressieren, z. B. durch verbesserte Entlohnungsmodelle für Arbeiter:innen bei diesen Plattformen oder Datenschutzmaßnahmen für Nutzer:innen.

4. Abschluss:

Am Ende präsentiert jedes Team seine Lösung der Gruppe, gefolgt von einer Feedbackrunde.

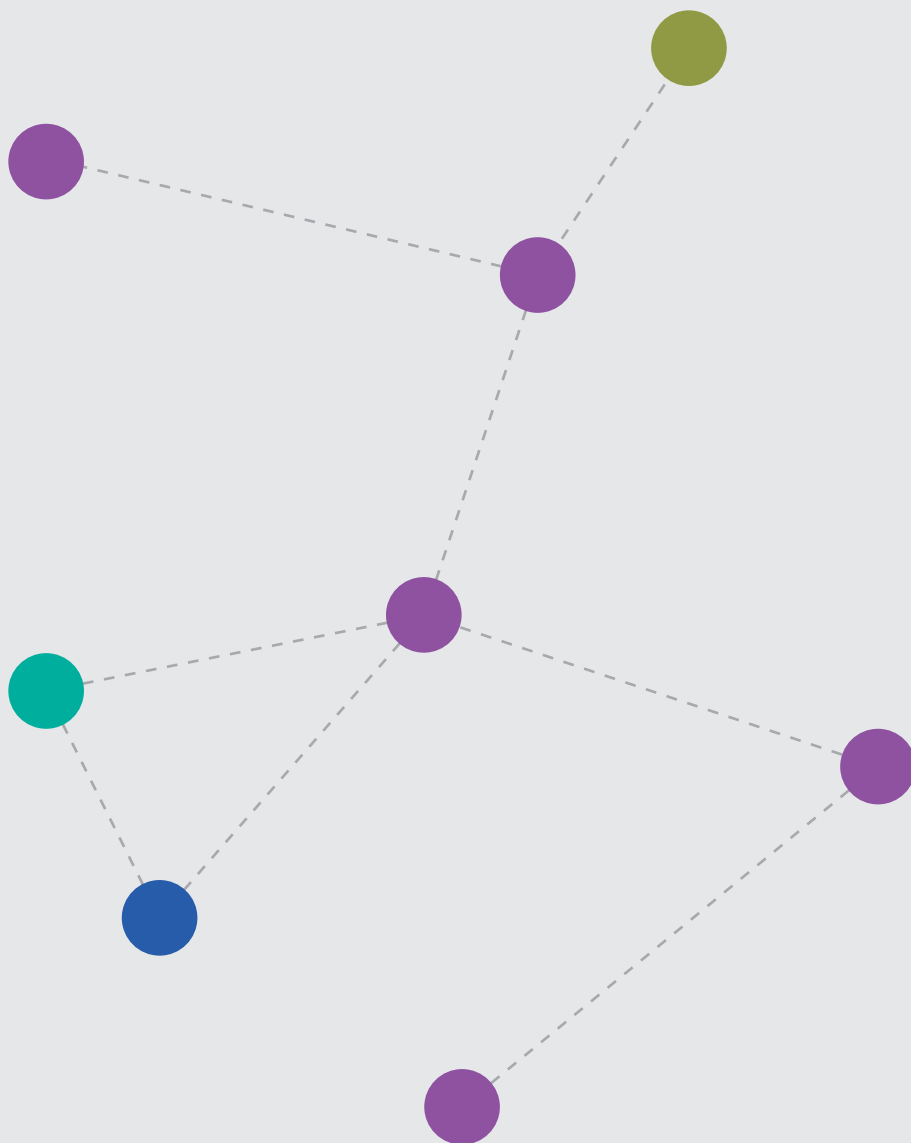
Diese Unternehmen haben gemeinsam, dass sie digitale Plattformen betreiben, die Angebot und Nachfrage auf innovative Weise zusammenbringen. Einige Schlüsselmerkmale dieser Entwicklung sind:

- **Virtualisierung des Geschäfts:** Diese Firmen haben Geschäftsmodelle, die vollständig digital sind, was zu einer hohen Skalierbarkeit führt, d. h., obwohl die Zahl der Benutzer:innen, Angebote und Funktionen steigt, erhöhen sich die Kosten des Betriebs kaum. Sie berühren nahezu alle wirtschaftlichen und gesellschaftlichen Bereiche und bieten technische sowie marktbezogene Dienstleistungen an.
- **Netzwerkeffekte und „Winner takes all“-Dynamik:** Die Plattformökonomie ist bekannt für ihre Netzwerkeffekte, bei denen der Nutzen der Plattform für jede:n Benutzer:in mit der Anzahl der Nutzer:innen steigt. Dies fördert eine Marktkonzentration, bei der wenige große Akteure dominieren.
- **Marktbeherrschung:** Plattformen wie diese dominieren den Markt oft mit Marktanteilen zwischen 50 bis 90 Prozent. Ihr Hauptwert liegt nicht in physischer Infrastruktur, sondern in den Netzwerken von Nutzer:innen und den Informationen, die sie generieren.

Die Dominanz der Plattformunternehmen führt auch zu Spannungen auf wirtschaftlicher und politischer Ebene. Sie bestimmen weitgehend die Regeln der Online-Welt, andere Firmen fühlen sich benachteiligt, nicht alle Bevölkerungsteile profitieren gleichermaßen von den Vorteilen der digitalen Welt. Diese zunehmende Macht der Plattformen wirft natürlich Fragen nach angemessener Regulierung und Kontrolle auf – eine beträchtliche Herausforderung für die Politik.

Schlussbetrachtungen

In unserer Diskussion haben wir die Entwicklung der Digitalisierung und des Internets von einer partizipativen Plattform zu einer zentralisierten Informationsinfrastruktur erörtert. Während diese Transformation zweifellos erhebliche Fortschritte und Vorteile mit sich gebracht hat, sind damit auch bedeutende sozioökonomische und politische Herausforderungen verbunden.



- KAPITELABSCHNITTE
- WISSENSFRAGEN
- FRAGE(N) ZUR DISKUSSION
- AKTIVIERUNG
- GRUPPENARBEIT

Kapitel 2

ENTWICKLUNG DER INFORMATIK

In diesem Kapitel befassen wir uns mit der Geschichte von Computern und elektronischer Datenverarbeitung. In wenigen Jahrzehnten hat die Informatik das menschliche Umfeld – Wirtschaft, Finanzwesen, soziale Beziehungen, Politik und Gesellschaft, um nur einige zu nennen – durch ihre bahnbrechenden Fortschritte wie den Personal Computer, das Internet, das World Wide Web, mobile Computer, maschinelles Lernen und Künstliche Intelligenz völlig verändert.

Basis dieses Kapitels ist der Aufsatz „Evolution of Computing“ von James R. Larus, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024. DOI: 10.1007/978-3-031-45304-5_3.

Vorgeschichte

Turing-Maschine

Die moderne Informatik, oft als Computerwissenschaft bezeichnet, kann man mit Alan Turing (1912–1954) beginnen, der 1936 als Student in Cambridge seine bahnbrechende Arbeit „On Computable Numbers, with an Application to the Entscheidungsproblem“ veröffentlichte. In diesem Aufsatz führte Turing das Konzept eines universellen Computers (der Turing-Maschine) ein und postulierte, dass dieser jeden **Algorithmus** ausführen kann.

Die Idee einer Rechenmaschine – eines Geräts, das in der Lage ist, eine Berechnung auszuführen – hatte mehrere Vorläufer. Bereits 1834 entwarf zum Beispiel der Engländer Charles Babbage (1791–1871) mit seiner Analytical Engine (die nie gebaut wurde) ein Modell eines Computers, das Rechenaufgaben ausführen kann. Seine Bekannte Lady Ada Lovelace (1815–1852) erkannte, dass eine derartige Maschine nicht nur rechnen oder mathematische Probleme lösen könnte, sondern man sie auch programmieren könnte. Ada Lovelace ist damit eine der wichtigen Mitbegründer:innen der Disziplin, die später Informatik heißen wird.

Turings Innovation bestand darin, die Anweisungen zur Steuerung des Computers (sein Programm) als Daten zu behandeln und damit ein unendlich veränderbares Gerät zu schaffen, das als Computer mit gespeicherten Programmen bekannt ist. Diese Innovation machte Computer zu universellen Automaten. Bis heute hat kein anderer Bereich menschlicher Erfindungen ein einziges Gerät hervorgebracht, das so mächtig ist. Sie können fast alles berechnen.

Ein **Algorithmus** ist eine eindeutige Handlungsanweisung zur Lösung eines Problems. Algorithmen bestehen aus endlich vielen, wohldefinierten Einzelschritten. Damit können sie zur Ausführung in ein Computerprogramm implementiert, aber auch in menschlicher Sprache formuliert werden.
(<https://de.wikipedia.org/wiki/Algorithmus>)

Eine Turing-Maschine ist jedoch eine mathematische Abstraktion und kein praktischer Computer. Die ersten elektronischen Computer wurden während des Zweiten Weltkriegs gebaut, um Probleme wie die Berechnung von der benötigten Artillerie und das Knacken von Geheimcodes zu lösen. Diese frühen Computer waren elektronisch und nicht mechanisch wie ihre unmittelbaren Vorgänger, aber sie folgten nicht Turings Weg und behandelten Programme als Daten; vielmehr wurden sie durch Umverdrahtung ihrer Schaltkreise programmiert.



AKTIVIERUNG

Besuch eines technischen Museums

Ziel:

Die Teilnehmer:innen besuchen ein technisches Museum, um alte Computer und deren Entwicklung zu betrachten und zu verstehen.

Museumswahl:

Auswahl eines technischen Museums in der Nähe, das eine Ausstellung zu alten Computern und deren Geschichte bietet. Wenn das Museum eine Online-Ausstellung anbietet, kann diese alternativ besucht werden.

Ergebnis:

Die Teilnehmer:innen sollten am Ende der Aktivierung ein besseres Verständnis für die Entwicklung der Computertechnologie haben und deren Auswirkungen auf die heutige Gesellschaft reflektieren können. Ein Museumsbesuch bietet eine praktische und anschauliche Möglichkeit, historische Technologiethemata zu erkunden und zu erleben.

1. Gruppenaufgaben: Einteilen der Teilnehmer:innen in kleine Gruppen und Verteilung spezifischer Aufgaben oder Fragen, die sie während des Museumsbesuchs beantworten sollen.

- ▶ Technische Entwicklung: Notieren der wichtigsten technischen Fortschritte, die anhand der ausgestellten Computer sichtbar sind.
- ▶ Anwendungen: Die Teilnehmer:innen sollen herausfinden, für welche Zwecke die alten Computer genutzt wurden und wie sie das Leben der Menschen beeinflussten.
- ▶ Persönliche Geschichten: Recherche über Geschichten oder Anekdoten über die Personen, die diese Computer entwickelt oder benutzt haben.

2. Interaktive Elemente: Die Teilnehmer:innen sollen ermutigt werden, gegebenenfalls vorhandene interaktive Stationen auszuprobieren. Viele Museen bieten Nachbildungen alter Computerspiele oder -anwendungen an, die den Besucher:innen ein Gefühl dafür geben, wie die Technologie früher genutzt wurde.

3. Präsentation der Ergebnisse: Nach dem Museumsbesuch treffen sich die Teilnehmer:innen wieder, um ihre Erfahrungen und Erkenntnisse zu teilen. Jede Gruppe präsentiert kurz ihre Beobachtungen und beantwortet die gestellten Fragen.

4. Diskussion: Diskussion über die Bedeutung der historischen Entwicklung der Computertechnologie.

Fragestellungen:

- ▶ Was hat die Teilnehmer:innen an den alten Computern am meisten überrascht?
- ▶ Welche technischen Fortschritte haben die größten Auswirkungen auf die heutige Technologie?
- ▶ Wie hat sich die Nutzung von Computern im Laufe der Zeit verändert?

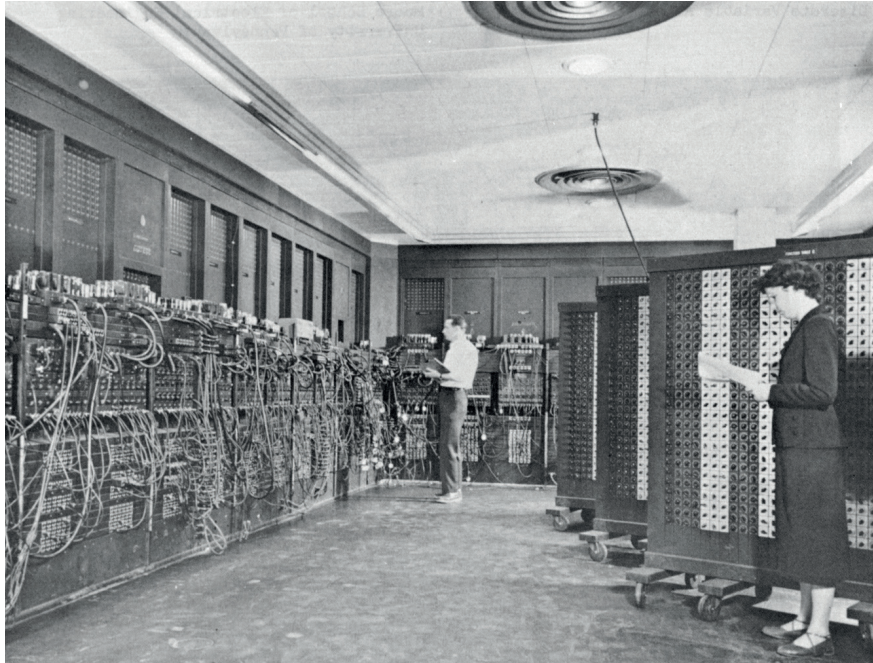


Abb. 1: ENIAC (1947), der erste elektronische Allzweckcomputer.
Quelle: Wikipedia (Public domain). <https://en.wikipedia.org/wiki/ENIAC>

ENIAC

Der erste elektronische Allzweckcomputer in den USA war der ENIAC. Er war ein bedeutender Schritt in der Entwicklung der Computertechnologie.

Von-Neumann-Architektur

Doch schon bald nach dem Krieg schrieb der ungarisch-amerikanische Mathematiker John von Neumann, aufbauend auf den Arbeiten vieler anderer, eine Arbeit, die Turings Erkenntnisse mit praktischer Technik vereinte. Er beschrieb eine Architektur für speicherprogrammierbare Computer, die den Grundstein für die Computerindustrie legte. Diese sogenannte Von-Neumann-Architektur ist immer noch die Blaupause für heutige Computer.

Computer als programmierbare Maschinen

Die gesellschaftlichen Auswirkungen der ersten Computer waren zunächst gering. Die neuen Maschinen verdrängten kaum menschliche Arbeitskräfte. Wenn, dann meist Frauen, die wissenschaftliche Berechnungen von Hand oder mit mechanischen Rechenmaschinen durchführten. Gleichzeitig wurden Frauen zu den ersten Programmierern, wie beispielsweise im Jet Propulsion Laboratory der NASA (siehe Abb. 2). Auf diesem Foto von 1959 arbeitet eine Programmiererin mit einem frühen Computer namens IBM 704.

Trotz ihrer begrenzten Nutzungsmöglichkeiten gab es aber gleichzeitig große intellektuelle Aufregung über das Potenzial dieser „denkenden Maschinen“.

Turing-Test

In einem weiteren bahnbrechenden Beitrag stellte Alan Turing die Frage, ob eine Maschine „denken“ kann, mit seinem berühmten Turing-Test, der besagt, dass eine Maschine als intelligent bezeichnet werden kann, wenn Menschen nicht unterscheiden können, ob sie sich mit einer Maschine oder einem anderen Menschen unterhalten (siehe auch Infobox auf Seite 166).

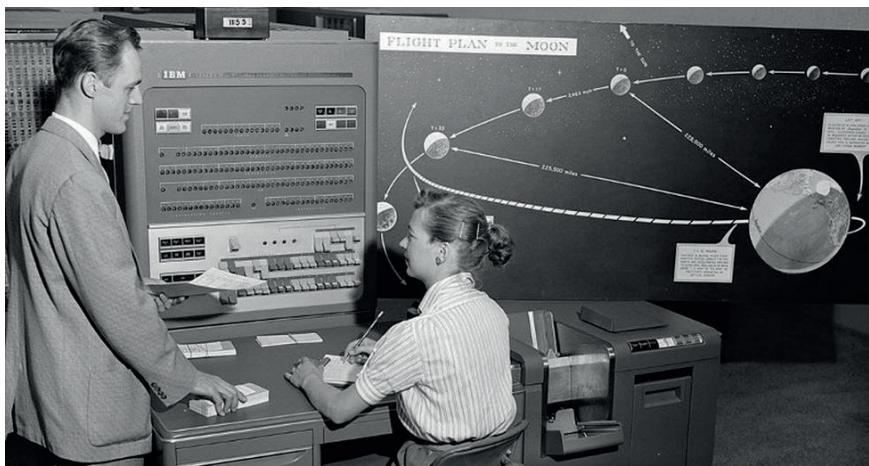


Abb. 2: Eine Programmiererin arbeitet an einem frühen Maschinencomputer.

Quelle: <https://www.flickr.com/photos/nasamarshall/30683267841/in/photostream>. Lizenz: CC BY-NC 2.0; Credits: NASA/JPL-Caltch

Siebzig Jahre später, mit dem Aufkommen von ChatGPT, ist Turings Formulierung immer noch aufschlussreich und relevant, auch wenn dieser eine Vielzahl anderer Definitionen und Ansichten gegenüberstehen.

Entwicklung der Mikroprozessoren

In den 1970er Jahren ermöglichte die zunehmende Dichte integrierter Schaltkreise die Herstellung eines „Computers auf einem Chip“. Diese Mikroprozessoren verwandelten Computer von teuren, schwer zu konstruierenden Maschinen für Unternehmen in kleine, preiswerte Geräte, die Unternehmer für die Herstellung innovativer Produkte nutzen konnten.

Aufstieg der Personal Computer (PCs)

Mikroprozessoren ermöglichten den Bau von Personal Computern, die so klein und preiswert waren, dass sie für den privaten Gebrauch erschwinglich wurden. Die rasch wachsende Anzahl von Computern machte Software zu einem profitablen, eigenständigen Geschäft.

Bill Gates, einer der Mitbegründer von Microsoft, war einer der Ersten, die erkannten, dass sich der Wert von Computern durch handelsübliche Mikroprozessoren dramatisch von den physischen Computern auf die Software zur Erledigung von Aufgaben verlagert. IBM beschleunigte diesen Wandel, indem es seinen legendären PC mit Standardkomponenten baute und andere Unternehmen nicht daran hinderte, „IBM-kompatible“ Computer zu bauen. Dies führte zu einer breiten Verfügbarkeit und sinkenden Preisen, was wiederum der aufstrebenden Softwareindustrie zugutekam.

Grafische Benutzeroberflächen (GUIs)

Die Interaktion mit den frühen Computern erfolgte in Textform. Programme wurden in einer Programmiersprache geschrieben, und der Computer wurde mit Textbefehlen gesteuert. In den späten 1960er und 1970er Jahren wurden grafische Benutzeroberflächen (Graphical User Interfaces, kurz: GUIs) entwickelt. Diese Änderung machte Computer für viel mehr Menschen zugänglich und nützlich. Computer haben auch

andere menschliche Interaktionen übernommen. Spracherkennung und Spracherzeugung sind seit Langem etablierte Techniken für die Interaktion. Mittlerweile ist es nicht ungewöhnlich, ein Smartphone oder ein anderes Gerät zu steuern, indem man mit ihm spricht.

Fazit aus diesem Abschnitt:

Die frühe Geschichte der Informatik ist geprägt von bahnbrechenden Arbeiten und Innovationen, die den Grundstein für die heutigen Computer legten. Alan Turings Konzept der Turing-Maschine und John von Neumanns Von-Neumann-Architektur sind zentrale Meilensteine, die die moderne Computerwissenschaft ermöglichten und bis heute prägen. Die Entwicklung der Computertechnologie, von den ersten Rechenmaschinen bis hin zu Personal Computern und Mikroprozessoren, hat die Art und Weise, wie Menschen arbeiten und kommunizieren, revolutioniert. Die Einführung von GUIs und Spracherkennung hat Computer noch zugänglicher und vielseitiger gemacht.

Computer und Kommunikation

Die Anfänge des Internets

Computer wären nur wenig spannender als die heutigen Taschenrechner, wenn sie nur mathematische Berechnungen durchführen könnten. Aber es wurde schnell klar, dass zwischen Computern Informationen ausgetauscht werden können, sodass sie nicht nur rechnen, sondern auch kommunizieren und zusammenarbeiten können. Eine der grundlegenden Innovationen war das ARPANET, der direkte Vorläufer des Internets. Das US-Verteidigungsministerium schuf das ARPANET Ende der 1960er/Anfang der 1970er Jahre als dezentrales und resilientes Kommunikationssystem, das einen Atomangriff überstehen können sollte.

Kommerzialisierung des Internets

Das Internet begann als Forschungsprojekt und Infrastruktur der US-Regierung. Zunächst hatten nur das Militär, einige wenige Universitäten und einige regierungsnahe Unternehmen Zugang. In den frühen 1990er Jahren beschloss die National Science Foundation (NSF), das Internet zu kommerzialisieren. In einer wenig beachteten, aber enorm erfolgreichen Aktion übergab sie das Internet der technischen Gemeinschaft, die es aufgebaut hatte, und den privaten Unternehmen, die die einzelnen Netze betreiben, aus denen das heutige Internet besteht.

WWW – Das World Wide Web

Während seiner Arbeit am **CERN**, einem physikalischen Forschungslabor in der Schweiz, entwickelte **Tim Berners-Lee** ein vernetztes Hypertext-System, das er optimistisch „World Wide Web (WWW)“ nannte. Das CERN gab seinen Entwurf und seine Software 1991 für die Öffentlichkeit frei. Die akademische Gemeinschaft, die bereits mit dem Internet vertraut war, stürzte sich schnell auf das Web.

Eine entscheidende Veränderung brachte das Aufkommen des World Wide Web. **Tim Berners-Lee** entwickelte das Web an der Forschungseinrichtung der Europäischen Organisation für Kernforschung (**CERN**) in der Schweiz. 1991 wurden der Entwurf und die Software des Webs veröffentlicht. Schnell begannen Unternehmen Websites zu erstellen, und die breite Bevölkerung kaufte PCs, um Zugang zum Internet zu erhalten. Über sogenannte Hyperlinks war es nun, ohne großes Vorwissen, möglich, sich im Internet zu bewegen. Das im Internet vorhandene Wissen war allen zugänglich. Dies sollte zu einer großen Demokratisierung des Wissens führen. In einem späteren Abschnitt werden wir sehen, wie sich dieser freie Zugang zu Information entwickelt hat.

Smartphones

Die nächste wichtige und radikale Veränderung war das mobile Computing, das anwendbar wurde, als Computer ausreichend stromsparend wurden, um als Smartphones verpackt zu werden. Der entscheidende Durchbruch für die mobile Datenverarbeitung war die Einführung des iPhones durch Apple im Jahr 2007. Es vereinte in einem taschengroßen Gehäuse eine Touchscreen-Benutzeroberfläche, die für ein kleines Gerät ohne Tastatur oder Maus geeignet ist, und eine kontinuierliche Konnektivität über das drahtlose Telefonnetz.

Für den Großteil der Weltbevölkerung sind Smartphones heute der wichtigste Zugang zum Internet. Innerhalb weniger Jahre wurde das

Smartphone zu einem ständigen Begleiter, den die meisten Menschen immer bei sich tragen.

Smartphones haben auch die Art der Datenverarbeitung verändert, indem sie Kameras und GPS-Empfänger (zur Standortbestimmung) mit dem Computer verbunden haben. Wie viele andere Funktionen ist auch die Standortbestimmung ein zweischneidiges Schwert: Sie bietet einerseits ausgefeilte Karten- und Navigationsmöglichkeiten, ermöglicht andererseits aber auch das Aufspüren von Personen durch Werbetreibende und Kriminelle.

Auswirkungen des Internets und WWW

Das Internet und das WWW ermöglichten in den letzten drei Jahrzehnten ein bemerkenswertes Spektrum gesellschaftlicher und persönlicher Veränderungen. Sie machten es möglich, Informationen zu finden, Handel zu treiben und überall zu kommunizieren – scheinbar ohne Kosten. Tatsächlich hat sich das Werbegeschäft als grundlegendes Geschäftsmodell des Internets etabliert.

Fazit aus diesem Abschnitt:

Die Entwicklung der Computertechnologie war ein Wendepunkt, der die Kommunikation und Zusammenarbeit zwischen Menschen und Maschinen ermöglichte. Das ARPANET, als Vorläufer des Internets, legte den Grundstein für ein dezentrales Kommunikationssystem.

Die Einführung des World Wide Web machte das Internet allgemein zugänglich und führte zu einem enormen öffentlichen Interesse und wirtschaftlichen Investitionen. Mit der Verbreitung von Smartphones wurde das Internet zu einem allgegenwärtigen Begleiter im Alltag der Menschen. Diese Entwicklungen haben die Art und Weise, wie wir Informationen austauschen und kommunizieren, grundlegend verändert und sind ein wesentlicher Bestandteil der digitalen Transformation unserer Gesellschaft. Das Werbegeschäft ist das bestimmende Geschäftsmodell des Internets.

WISSENSFRAGEN

Computer und Kommunikation

Warum war die Entwicklung des World Wide Webs so wichtig? War damit tatsächlich eine Demokratisierung des Wissens verbunden?

Welche zwei entscheidenden Veränderungen trugen in den frühen 1990er Jahren zur breiten öffentlichen Nutzung des Internets bei?

Wie hat das iPhone die mobile Datenverarbeitung und die Nutzung des Internets verändert?

Warum wurde das Werbegeschäft das grundlegende Geschäftsmodell des Internets?

Wie hat die Entwicklung der grafischen Benutzeroberflächen (GUIs) die Nutzung und Zugänglichkeit von Computern für eine breitere Öffentlichkeit verändert?

1 **Computer bezeichneten menschliche Arbeitskräfte, vor allem Frauen, die Berechnungen per Hand durchführten. Nach und nach wurden Frauen aus dem aufstrebenden Bereich der Informatik verdrängt. Gibt es abseits der Automatisierung noch andere Gründe, warum Frauen ab den 1950er/1960er Jahren verdrängt wurden?**

Mögliche Diskussion: Neben der Automatisierung spielte auch die Geschlechterdiskriminierung eine Rolle, da Informatik als „technische Männerdomäne“ definiert wurde. Die zunehmende Professionalisierung und der Fokus auf Hardware sowie Mathematik wurden als „männlich“ betrachtet, was Frauen den Zugang zu vielen Positionen erschwerte. Ebenso trugen gesellschaftliche Vorurteile und eine fehlende Förderung von Frauen in technischen Berufen zur Verdrängung bei.

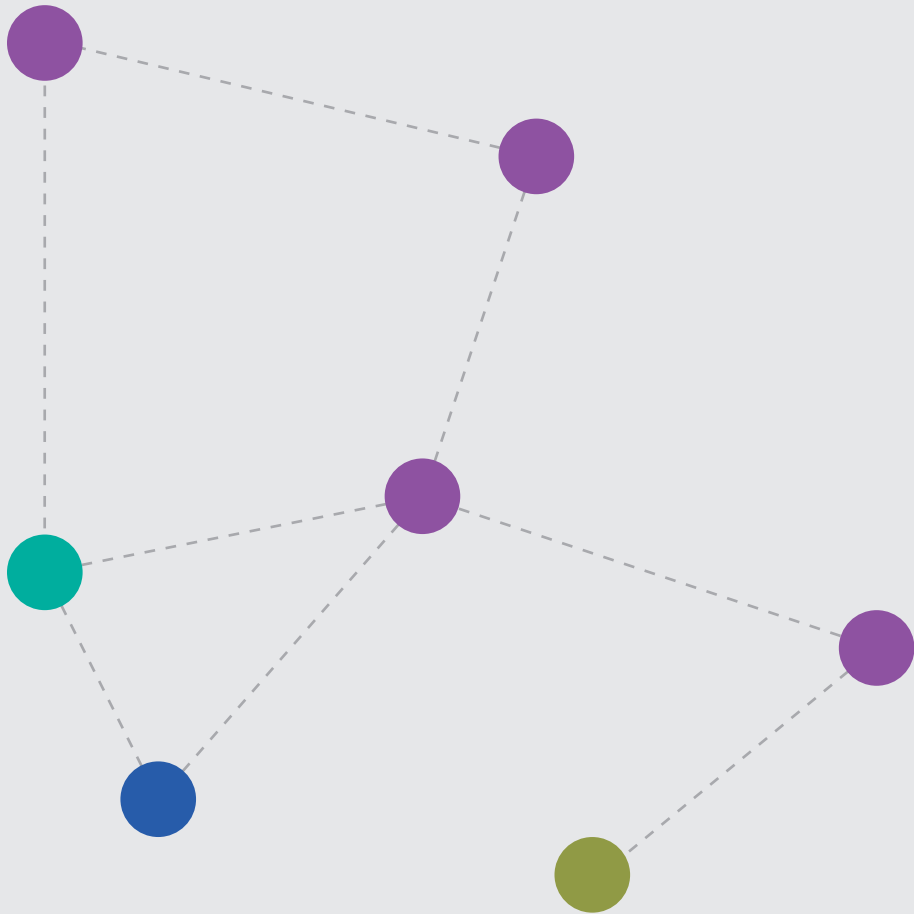
2 **Das Internet ist ein sehr resilientes System. Inwiefern entspricht das Bild eines Straßennetzes dem des Internets? Und inwiefern nicht? Trifft das Bild eher auf die westliche Welt zu?**

Mögliche Diskussion: Die Funktionsweise des Internets entspricht der eines weitverzweigten Straßennetzes, in dem Autos auf unterschiedlichen Wegen von A nach B kommen können, bei Stau, Sperrungen etc. ausweichen können.

Weitere Entwicklungen und Auswirkungen

Die Entwicklungen in der Informationstechnologie führten zu grundlegenden Veränderungen in allen Lebensbereichen – der sogenannten *digitalen Transformation* der Gesellschaft. Daten spielen eine zentrale Rolle und sind von großem Interesse zur Erschließung neuer Anwendungen. Dabei sind die drei Bereiche *Künstliche Intelligenz und maschinelles Lernen*, *Big Data und Cloud-Computing* und *Sicherheit und Datenschutz* eng miteinander verbunden. Sie werden in anderen Kapiteln dieses Buchs tiefergehend behandelt.

Abseits der technischen Fortschritte und Anwendungen sollte noch ein anderer Aspekt der Informatik betrachtet werden, der großen Einfluss auf die Gesellschaft hat, nämlich jener, Handlungen und intellektuelle Fähigkeiten als einen formalen und analysierbaren Prozess zu betrachten. Diese Sichtweise des „Computational Thinking“ hat auch andere Bereiche der Bildung, Wissenschaft und Wirtschaft beeinflusst. Über Jahrhunderte hinweg boten wissenschaftliche und technische Errungenschaften informelle, natürlichsprachliche Beschreibungen, wie eine Aufgabe zu bewältigen ist. Die Informatik brachte Strenge und Formalismus in die Beschreibung von Lösungen in Form von Algorithmen.



● KAPITELABSCHNITTE ● WISSENSFRAGEN ● FRAGE(N) ZUR DISKUSSION
● AKTIVIERUNG ● GRUPPENARBEIT

Kapitel 3

EINFÜHRUNG IN DIE KÜNSTLICHE INTELLIGENZ

Dieses Kapitel gibt einen Überblick über die wichtigsten Methoden der Künstlichen Intelligenz (KI). Die Methoden der symbolischen KI sind in der Logik verwurzelt, während beim maschinellen Lernen (ML) Algorithmen Informationen aus Daten extrahieren und Muster zur Anwendung auf neue Fälle erstellen.

Basis dieses Kapitels sind die Aufsätze „A Short Introduction to Artificial Intelligence: Methods, Success Stories, and Current Limitations“ von Clemens Heitzinger und Stefan Woltran, „Trustworthy Artificial Intelligence: Comprehensible, Transparent and Correctable“ von Ute Schmid und „Evolution of Computing“ von James R. Larus, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024. DOIs: 10.1007/978-3-031-45304-5_9, 10.1007/978-3-031-45304-5_10 und 10.1007/978-3-031-45304-5_3.

Einleitung

Künstliche Intelligenz ist das Forschungsgebiet der Informatik, das sich mit der Simulation menschlichen Denkens beschäftigt. Der Begriff wurde 1956 auf der Dartmouth-Konferenz zur Künstlichen Intelligenz geprägt. Allerdings gibt es keine eindeutige Definition dieses Begriffs. Er definiert sich eher über die Verwendung von zwei großen Methodenklassen.

Diese beiden Methodenklassen sind (1) symbolische oder wissensbasierte Methoden und (2) maschinelles Lernen. Beide Bereiche existieren seit den Anfängen, wobei der Durchbruch des maschinellen Lernens erst ab den 2010er Jahren stattfand. Die erste Implementierung eines maschinellen Lernprogramms war ein Programm zum Erlernen einer Strategie für das Spiel Dame und wurde 1952 von Arthur Samuel (1901–1990) realisiert. Zu den frühen Ansätzen gehörten auch das Perceptron als Versuch, die Struktur des menschlichen Gehirns zu imitieren und damit kognitive Prozesse in der Maschine simulieren zu können. Perceptron wurde von Frank Rosenblatt (1928–1971) 1957 entwickelt. Es besteht aus einer „Sensoreinheit“ mit Fotozellen und Antworteinheiten, die die Mustererkennung anzeigen. Die Maschine wird trainiert, indem Buchstaben in das Sichtfeld der Fotozelle gelegt werden. Nach Abschluss des Trainings kann sie die Buchstaben korrekt erkennen. Bei komplexeren Aufgaben, wie der Erkennung verschiedener Schriftbilder, erreichte die Maschine eine Genauigkeit von 85 Prozent.

Im Allgemeinen sollten KI-Ansätze nur für Probleme eingesetzt werden, die mit Standardalgorithmen, also klassischer Programmierung von Computern, nicht gelöst werden können. Während Standardalgorithmen, zumindest im Prinzip, Korrektheit (eine Eingabe ergibt die beabsichtigte Ausgabe) und Vollständigkeit (für alle möglichen Eingaben kann eine Ausgabe berechnet werden) garantieren, gilt dies für datenbasierte KI im Allgemeinen nicht. Für viele sicherheitskritische Bereiche kommen KI-Algorithmen daher in der Regel eigentlich nicht infrage.

Symbolische KI – wissensbasierte Systeme

Als symbolische KI werden jene Methoden bezeichnet, die auf der expliziten Beschreibung der Probleme bzw. der notwendigen Lösungsschritte für den Computer beruhen. Zur Beschreibung der Probleme werden Logik oder verwandte formale Sprachen verwendet; das eigentliche Finden möglicher Lösungen („Suche“) ist ein zentraler Aspekt der symbolischen KI. In diesem Modell ist es relativ einfach, die Erklärbarkeit einer Lösung zu gewährleisten. Zudem ist die Korrektheit einer Lösung im Allgemeinen eindeutig und nicht von Wahrscheinlichkeiten abhängig. Das „intelligente“ Verhalten ergibt sich hier einfach durch die Rechenleistung. Es sei an dieser Stelle darauf hingewiesen, dass in diesem Modell die „Erklärbarkeit“ einer Lösung konzeptionell leicht zu erhalten ist (bei größeren Spezifikationen werden die Erklärungen für den Menschen allerdings tendenziell unverständlich).

Die 1980er Jahre waren die Blütezeit dieser wissensbasierten Methoden, besonders im Zusammenhang mit Expertensystemen. Man hoffte, dass KI-Systeme menschliche Expert:innen in vielen Bereichen unterstützen könnten – von der medizinischen Diagnostik über die Produktionsplanung bis hin zu intelligenten Tutorsystemen in der Lehre. Dabei entstanden effiziente Algorithmen zum Ziehen von

Das Polanyi-Paradoxon bezieht sich auf die Erkenntnis des Philosophen Michael Polanyi, dass Menschen oft mehr wissen, als sie ausdrücken oder formalisieren können. Es beschreibt das Phänomen, dass ein Großteil des menschlichen Wissens implizit ist und nicht leicht in explizite Regeln oder Anweisungen gefasst werden kann. Polanyi fasste dies in dem Satz zusammen: „Wir wissen mehr, als wir sagen können.“

Ein Beispiel für das Polanyi-Paradoxon ist das Fahrradfahren. Die meisten Menschen können Fahrrad fahren, aber nur wenige können die physikalischen und motorischen Prozesse, die dabei ablaufen, genau beschreiben. Dieses implizite Wissen ist oft unbewusst und wird durch Erfahrung und Praxis erworben, nicht durch formale Instruktionen.

In der Künstlichen Intelligenz und insbesondere in der symbolischen KI wird das Polanyi-Paradoxon zu einer Herausforderung. Während einige Aspekte des Wissens explizit kodiert und in Expertensystemen verwendet werden können, bleibt ein großer Teil des menschlichen Wissens außerhalb des Bereichs dessen, was formalisierbar ist. Dieses implizite Wissen lässt sich nicht leicht in logische Regeln oder formale Sprachen übersetzen, was die Entwicklung umfassender wissensbasierter Systeme erschwert.

Schlussfolgerungen. Symbolische KI-Systeme erzielten besonders im Bereich von Spielen Erfolge, da Brettspiele wie Schach klare Regeln haben, die den Spieler:innen ihre Zugmöglichkeiten vorgeben.

Trotz großer Hoffnungen konnten Expertensysteme nur teilweise überzeugen. Ein Hauptproblem war der sogenannte Knowledge-Engineering-Engpass: Menschliches Wissen ist nur zum Teil explizit verfügbar und formal darstellbar. Große Bereiche des menschlichen Wissens, insbesondere Wahrnehmungswissen und hochautomatisierte Handlungsroutinen, sind implizit und können mit Methoden der Wissenserfassung nicht oder nur unzureichend erfasst werden. Dieses Phänomen ist auch als **Polanyi-Paradoxon** bekannt: „Wir wissen mehr, als wir sagen können.“

Subsymbolische KI – maschinelles Lernen

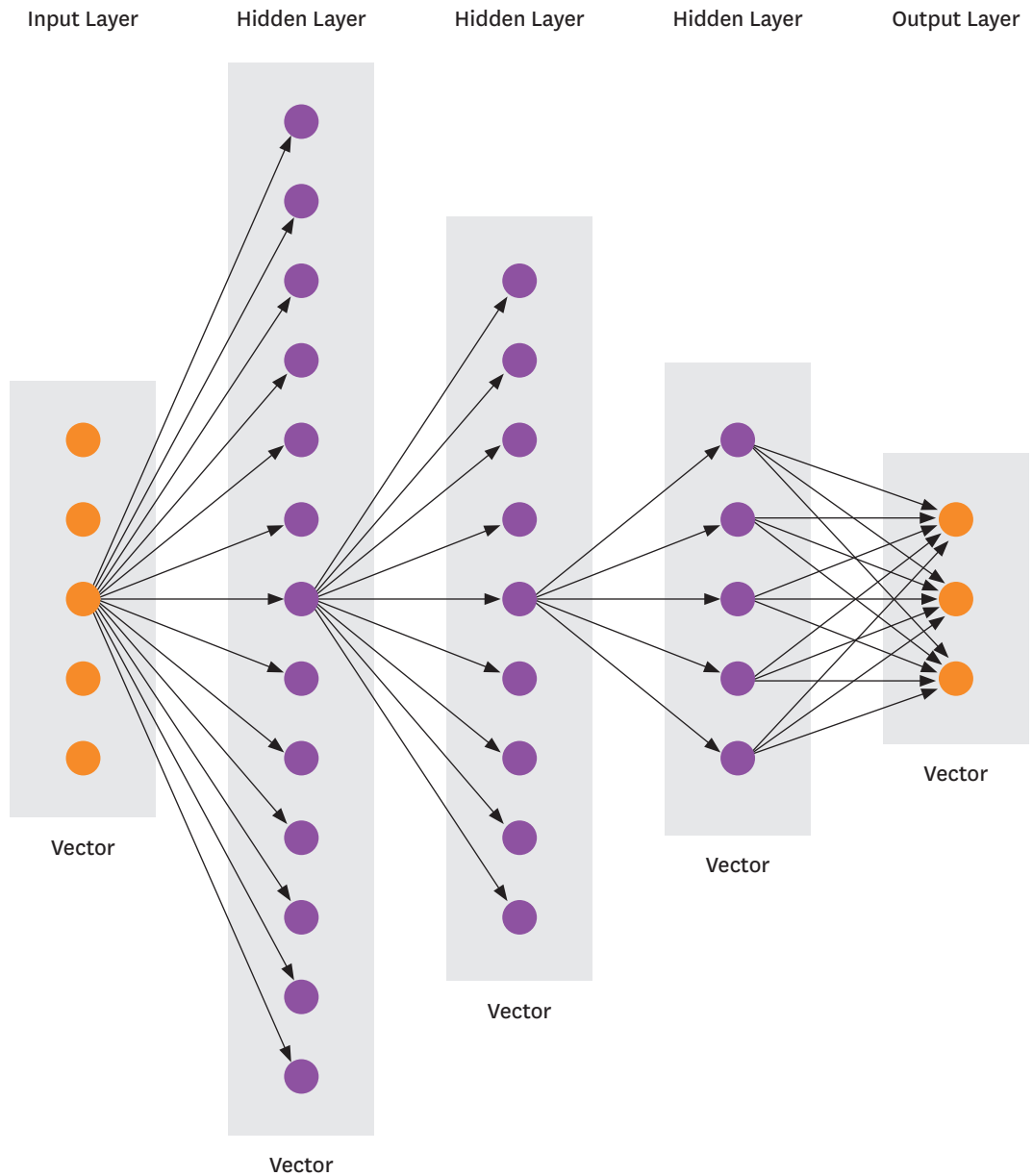
Beeindruckende Erfolge bei der Anwendung von tiefen neuronalen Netzen hat das Interesse an KI seit Beginn der 2010er Jahre neu entfacht, diesmal mit dem Fokus auf maschinelles Lernen. Der Hauptgrund dafür ist, dass es nun möglich war, ohne komplexe Vorverarbeitung direkt aus verschiedenen Datenarten wie Bildern oder Texten zu lernen (End-to-End-Lernen). Früher mussten Daten oft in tabellarischer Form vorliegen, um maschinell verarbeitet werden zu können. Heute können neuronale Netze direkt aus Rohdaten wie Fotos oder Röntgenbildern lernen.



Abb. 3: Bei der Bildererkennung klassifiziert der Algorithmus die Eingaben (Pixel des Bildes) und identifiziert das Objekt im Bild. Dies ist eine überwachte Lernaufgabe, bei der der Algorithmus aus Trainingsdaten lernt. Das Eingabebeispiel hier ist ein Foto eines Hundes, und die Ausgabe ist das Label „Hund“.
Quelle: KI-generiert

Beim maschinellen Lernen geht es oft um Klassifikationsaufgaben, bei denen der Algorithmus Funktionen findet, die Eingaben korrekt klassifizieren oder vorhersagen. Ein Beispiel ist die Bildererkennung: Die Eingaben sind die Pixel eines Bildes, und die Ausgabe kann die Identifikation eines Objekts im Bild sein. Dies nennt man überwachte Lernaufgaben, bei denen der Lernalgorithmus aus Trainingsdaten lernt, die aus Eingaben und definierten Ausgaben bestehen. Die Ausgabedaten werden oft als Labels bezeichnet; ein Eingabebeispiel kann z. B. ein Foto eines Hundes sein, und die entsprechende Ausgabe kann das Label „Hund“ sein.

Es gibt viele Algorithmen für das überwachte Lernen, aber künstliche neuronale Netze sind derzeit am bekanntesten und wichtigsten.



Grafik 1: Diese schematische Darstellung veranschaulicht, wie ein künstliches neuronales Netz (Artificial Neural Network, ANN) Eingabedaten schrittweise verarbeitet und transformiert, um eine Klassifikation oder Vorhersage zu treffen. Links befindet sich der Eingabevektor (Input Layer), der durch mehrere versteckte Schichten transformiert wird. In jeder versteckten Schicht (Hidden Layer) werden die Eingabedaten durch Gewichtungen und Aktivierungsfunktionen verarbeitet. Rechts ist der Ausgangsvektor (Output Layer) zu sehen, der das Ergebnis darstellt. Die Pfeile zeigen die Verbindungen zwischen den Neuronen an. Der Übersichtlichkeit halber werden jeweils nur die Verbindungen eines Neurons in die nächste Schicht dargestellt.

Quelle: Clemens Heitzinger, DOI: 10.1007/978-3-031-45304-5_9

Kriterium	Symbolische KI	Subsymbolische KI (maschinelles Lernen)
Beschreibung	Methoden basieren auf expliziten Regeln und Logik zur Problemlösung	Algorithmen lernen Muster aus Daten
Kernprinzip	Explizite Beschreibung der Problemlösungsstrategien	Lernen aus Beispielen und Daten
Erklärbarkeit	Lösungen sind im Prinzip leicht erklärbar	Lösungen sind oft schwieriger zu erklären (Blackbox-Charakter)
Anwendung	Geeignet für gut definierte Probleme mit klaren Regeln	Geeignet für komplexe Probleme, bei denen explizite Regelbeschreibung schwierig ist
Beispiel	Expertensysteme, Schachprogramme	Bild- und Spracherkennung, Empfehlungssysteme
Daten	Benötigt spezifisches Wissen und Regeln	Benötigt große Mengen an Daten zum Training
Flexibilität	Weniger flexibel bei Veränderungen der Problemstruktur	Hohe Flexibilität und Anpassungsfähigkeit an neue Daten
Korrektheit und Vollständigkeit	Kann Korrektheit und Vollständigkeit garantieren	Kann dies im Allgemeinen nicht garantieren
Herausforderungen	Knowledge-Engineering-Engpass, explizite Wissensdarstellung schwierig	Bedarf an großen und repräsentativen Datensätzen, Erklärbarkeit

Tabelle 1: Vergleich symbolische KI mit subsymbolischer KI. Quelle: Bettina K. Lechner

Neuronale Netze sind sehr flexible Datenstrukturen. Sie bestehen aus sogenannten Neuronen, einfachen Einheiten, die Eingabewerte anhand von Gewichtungen und Aktivierungsfunktionen in Ausgabewerte transformieren und die in Schichten (Layers) angeordnet sind. Die Neuronen der Eingabeschicht (Input Layer) repräsentieren die Eingabedaten in Form eines Vektors, die Neuronen der Ausgabeschicht (Output Layer) jene der Ausgabedaten. Dazwischen befindet sich eine variable Anzahl von verborgenen Schichten (Hidden Layer). Die Neuronen aufeinanderfolgender Schichten sind über Gewichte verbunden, die im Zuge des Trainings angepasst werden, um die Trainingsdaten möglichst gut abbilden zu können. Ein bedeutender Durchbruch hierfür war der Backpropagation-Algorithmus.

Im Jahr 2012 gewann ein tiefes neuronales Netz namens AlexNet die ImageNet Challenge, bei der Bilder aus 1.000 Kategorien klassifiziert werden sollen. AlexNet konnte direkt aus den Bildern lernen, was ein großer Fortschritt war. Ähnliche Entwicklungen gibt es auch bei der Verarbeitung natürlicher Sprache, wie maschineller Übersetzung (DeepL) oder Textgenerierung (GPT-3).

Eine bahnbrechende neue Entwicklung sind Transformatoren, eine bestimmte Art von neuronalen Netzen, die den sogenannten Aufmerksamkeitsmechanismus (Attention Mechanism) nutzen, um auch Beziehungen zwischen Wörtern über große Entfernungen in einem Text zu lernen. Transformatoren wurden ursprünglich in der maschinellen Übersetzung eingesetzt und lieferten zu diesem Zeitpunkt die beste und schnellste maschinelle Übersetzung. Sie wurden für die Verwendung in großen Sprachmodellen (Large Language Models, LLMs) wie InstructGPT, ChatGPT und GPT-4 angepasst und sind ein Meilenstein in der Verarbeitung natürlicher Sprache.

Denn bislang basierte maschinelles Lernen auf Klassifikationen. Das Revolutionäre an diesen großen Sprachmodellen ist, dass sie „das nächst Wahrscheinliche vorhersagen“ und damit tatsächlich „Neues“ kreieren.

Der wichtigste Faktor für den Erfolg lernbasierter KI-Systeme ist oft ein großer Trainingssatz. Zum Beispiel werden große Sprachmodelle wie ChatGPT mit Milliarden von Dokumenten aus dem Internet trainiert. Diese Modelle können auf allgemeine Fragen gut antworten und realistische Dialoge führen, obwohl sie oft inhaltliche Fehler machen, die ihren Mangel an tiefem Verständnis zeigen.

WISSENSFRAGEN

Einleitung, subsymbolische und symbolische KI

Was ist der grundlegende Unterschied zwischen symbolischer KI und subsymbolischer KI?

Was ist Künstliche Intelligenz (KI), und wer prägte den Begriff?

In welchen Bereichen zeigt die symbolische KI ihre Stärken, und wo stößt sie an ihre Grenzen?

Welche zwei Hauptrichtungen gibt es in der Künstlichen Intelligenz (KI), und wie unterscheiden sie sich?

Welche Vorteile bietet maschinelles Lernen im Vergleich zu wissensbasierten Systemen?

Warum sind KI-Algorithmen notwendig, und in welchen Bereichen kommen sie in der Regel nicht zum Einsatz?

FRAGEN ZUR DISKUSSION

Symbolische bzw. subsymbolische KI

1 Sollte mehr Wert auf die Erklärbarkeit von KI-Systemen gelegt werden, auch wenn das die Flexibilität und Anpassungsfähigkeit einschränkt?

Diskutiert werden könnte, ob es wichtiger ist, dass KI-Systeme nachvollziehbar und transparent sind (wie bei symbolischer KI) oder ob die Flexibilität und Leistungsfähigkeit von maschinellem Lernen Vorrang haben sollte, auch wenn die Systeme schwieriger zu erklären sind.

2 In welchen Anwendungsbereichen könnte eine Kombination aus symbolischer und subsymbolischer KI besonders nützlich sein?

Diese Frage könnte eine Diskussion darüber anregen, wie die Stärken beider Ansätze kombiniert werden könnten, um vielseitigere und robustere KI-Systeme zu entwickeln.

3 Wie kann verhindert werden, dass beim Training von großen Sprachmodellen (LLMs) hauptsächlich Texte aus der männlichen, westlichen, englischsprachigen Welt verwendet werden?

Diskutiert werden könnte, was sich verändern würde, wenn vermehrt kleinere Sprachen, Daten aus dem Globalen Süden etc. zum Einsatz kämen.

Schlussbetrachtungen

Maschinelles Lernen (ML) hat die Art und Weise verändert, wie Computer programmiert werden. Früher gaben Programmierer:innen explizite Anweisungen, jetzt lernen Computer durch Beobachtung. ML basiert auf Wahrscheinlichkeiten, da es auf Daten angewiesen ist, die immer begrenzt sind. Dadurch können wichtige Daten fehlen oder nicht repräsentativ sein. ML-Systeme sind oft schwer nachvollziehbar und gelten als „Blackboxes“.

Trotz dieser Herausforderungen hat maschinelles Lernen Computern menschenähnliche Fähigkeiten verliehen, die schwer zu programmieren sind. Dieser Wandel lässt einige Menschen befürchten, dass Computer „intelligent“ werden und die menschlichen Fähigkeiten bald übertreffen könnten und sich in Folge der menschlichen Kontrolle entziehen könnten. Aber eigentlich geht es darum, KI und ML durch verantwortungsvollen Einsatz nutzbar zu machen und mit konkreten Problemen und Auswirkungen umzugehen.

Es gibt drei Hauptfaktoren, die zum heutigen Stand der KI geführt haben. Der erste ist die Verfügbarkeit von riesigen Datensätzen und Datenbanken für Lernzwecke, auch dank des Internets. Dazu gehören alle Modalitäten, z. B. markierte Bilder für überwachtetes Lernen und große Sammlungen hochwertiger Texte für die maschinelle Übersetzung. Der zweite Faktor ist die Verfügbarkeit riesiger Rechenressourcen für das Lernen, insbesondere Grafikkarten (GPUs) und GPU-Cluster für Berechnungen mit neuronalen Netzen. Der dritte Faktor sind algorithmische Fortschritte und Software-Tools. Während neuronale Netze im Laufe der Geschichte der Künstlichen Intelligenz immer wieder auftauchten, waren und sind neue Strukturen und neue Algorithmen für die Anwendungen von großer Bedeutung. Die jüngsten Durchbrüche werden zu einer gewissen Konsolidierung in dem Sinne führen, dass die Lernalgorithmen effizienter und besser verständlich werden.

Die Forschung im Bereich KI wird mit hohem Tempo weitergehen, immer mehr menschliche Fähigkeiten werden erreicht und übertroffen. Es liegt an uns, diese Co-Evolution von Mensch und Maschine zu gestalten.

AKTIVIERUNG

Werden KI-Roboter die Weltmacht übernehmen und menschliche Fähigkeiten übertreffen?

Ziel:

Die Teilnehmer:innen sollen über die rasante Entwicklung der Künstlichen Intelligenz (KI) reflektieren und die möglichen zukünftigen Auswirkungen auf Gesellschaft und Alltag diskutieren.

Materialien:

- Stifte
- Moderationskarten

1. Brainstorming:

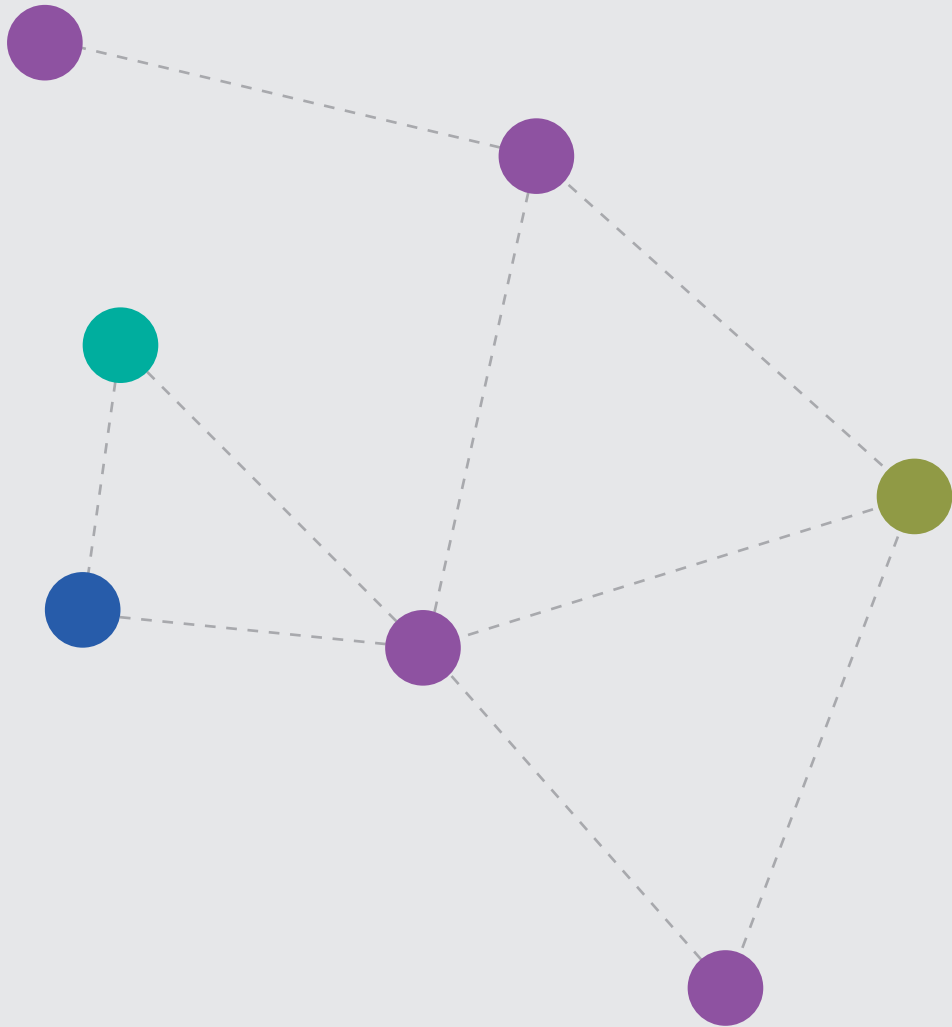
- ▶ Die Teilnehmer:innen werden gebeten, spontan zu nennen, welche menschlichen Fähigkeiten Computer ihrer Meinung nach in den nächsten zehn Jahren erreichen könnten.
- ▶ Die Antworten werden auf Moderationskarten geschrieben und an einer Pinnwand gesammelt.

2. Gruppenarbeit:

- ▶ Die Teilnehmer:innen werden in kleine Gruppen aufgeteilt, und jede Gruppe wählt eine der aufgelisteten Fähigkeiten aus.
- ▶ Die Gruppen diskutieren, wie es wäre, wenn Computer diese Fähigkeit erreichen würden. Sie sollen mögliche Vorteile und Nachteile skizzieren und sich überlegen, welche ethischen oder gesellschaftlichen Fragen damit verbunden sein könnten.

3. Präsentation und Diskussion:

- ▶ Jede Gruppe stellt ihre Ergebnisse im Plenum vor.
- ▶ Im Anschluss an jede Präsentation gibt es eine kurze Diskussion, in der die anderen Teilnehmer:innen Feedback geben und weitere Gedanken einbringen.



● KAPITELABSCHNITTE ● WISSENSFRAGEN ● FRAGE(N) ZUR DISKUSSION
● AKTIVIERUNG ● GRUPPENARBEIT

Kapitel 4

AUSWIRKUNGEN DER DATENINTENSIVEN KÜNSTLICHEN INTELLIGENZ

Dieses Kapitel beleuchtet die Auswirkungen von datenintensiver Künstlicher Intelligenz, insbesondere des maschinellen Lernens. Es geht um die Herausforderungen, die mit der Nutzung großer Datenmengen verbunden sind, und die Folgen dieser Technologien, insbesondere im Zusammenhang mit der Klimakrise. Es zeigt auf, wie KI sinnvoll eingesetzt werden kann, um sowohl die positiven Effekte zu maximieren als auch die negativen Auswirkungen zu minimieren.

Basis dieses Kapitels sind die Aufsätze „Artificial Intelligence and Large-Scale Threats to Humanity“ von Guglielmo Tamburrini, „A Short Introduction to Artificial Intelligence: Methods, Success Stories, and Current Limitations“ von Clemens Heitzinger und Stefan Woltran und „Evolution of Computing“ von James R. Larus, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024.
DOIs: 10.1007/978-3-031-45304-5_16, 10.1007/978-3-031-45304-5_9, 10.1007/978-3-031-45304-5_3.

Einleitung

Die Nutzung großer Datenmengen zum Trainieren von KI-Systemen bringt Herausforderungen mit sich, vor allem in Bezug auf die notwendige Infrastruktur für Datenverwaltung und Rechenleistung. Diese Anforderungen haben weitreichende Folgen, insbesondere im globalen Kontext der Klimakrise. Während KI dabei helfen kann, Klimadaten zu analysieren und Lösungen zur Eindämmung der Erderwärmung zu entwickeln, ist sie gleichzeitig selbst ein großer Energieverbraucher. Das Training von KI-Modellen erfordert viel Energie und führt zu einem erhöhten Ausstoß von Treibhausgasen, was die globale Erwärmung weiter vorantreibt. KI hat somit das Potenzial, sowohl zur Lösung als auch zur Verschärfung der Klimakrise beizutragen.

Diese ambivalente Rolle von KI verdeutlicht die Komplexität der Thematik und unterstreicht die Notwendigkeit, sorgfältig abzuwägen, wie und wo KI eingesetzt wird, um die bestmöglichen Ergebnisse für die Gesellschaft und die Umwelt zu erzielen.

Auswirkungen von Big Data

Der Fortschritt im maschinellen Lernen und anderen Bereichen wie beispielsweise den Large Language Models (LLMs) basiert auf der Fähigkeit, riesige Datenmengen, die als „Big Data“ bezeichnet werden, zu sammeln und zu analysieren. Diese Datenmengen wurden ursprünglich für Webanwendungen wie Suchmaschinen entwickelt, die Big Data Centers mit tausenden von Computern nutzen, um Internetseiten zu durchsuchen und schnell auf Anfragen der Nutzer:innen zu reagieren. Jedes Mal, wenn eine Suche durchgeführt wird, sind Tausende von Computern involviert, um die Daten zu verarbeiten. Diese anspruchsvolle Art der Datenverarbeitung wird als Parallelverarbeitung bezeichnet.

Die Internetsuche und die Speicherung großer Datenmengen wurden durch Fortschritte in der Informatik ermöglicht, wie z. B. durch die Entwicklung von Netzwerken mit hoher Geschwindigkeit und kostengünstigen Speichermedien. Zunächst war diese Infrastruktur nur wenigen Unternehmen zugänglich, aber Amazon stellte sie mit

Cloud-Computing auf eine breitere Basis. *Cloud-Computing* ermöglicht den kostengünstigen Zugang zu leistungsstarken Computern, die über das Internet verbunden sind, wodurch Unternehmen große Datenmengen sammeln, speichern und analysieren können.

Diese Entwicklungen haben jedoch auch schwerwiegende Folgen für den Datenschutz, da viele Websites jede Benutzerinteraktion aufzeichnen und speichern. Diese Daten werden zudem genutzt, um Systeme für maschinelles Lernen zu trainieren. Die Idee dahinter ist, dass Daten als das neue „Öl“ bezeichnet werden, also eine wertvolle Ressource, die profitabel genutzt werden kann.

Gleichzeitig wird oft übersehen, dass die Verarbeitung großer Datenmengen hohe Kosten, einen hohen Energiebedarf und CO₂-Ausstoß verursacht. Die Idee, dass unendliches Wachstum Wohlstand schafft, passt gut zur digitalen Welt, in der Systeme theoretisch unendlich erweitert werden können. In der Realität sind Ressourcen jedoch begrenzt, und Big Data und KI nehmen einen bedeutenden Teil dieser Ressourcen in Anspruch.

Fazit aus diesem Abschnitt:

Die Fähigkeit, riesige Datenmengen zu sammeln und zu analysieren, hat das maschinelle Lernen und andere Bereiche der Informatik revolutioniert. Während diese Entwicklungen enormes Potenzial bieten und zur Entstehung profitabler Geschäftsfelder geführt haben, gehen sie auch mit erheblichen Herausforderungen einher. Dazu gehören hohe Kosten, erheblicher Energieverbrauch und ein erhöhter CO₂-Ausstoß, der die Umwelt belastet. Darüber hinaus sind die Datenschutzrisiken erheblich, da persönliche Daten oft ohne ausreichende Transparenz oder Zustimmung genutzt werden. Diese Entwicklungen zeigen, dass die Nutzung von Big Data sorgfältig abgewogen werden muss, um negative Auswirkungen auf die Gesellschaft und die Umwelt zu minimieren.

KI und die Klimakrise

KI beeinflusst die Klimakrise auf zwei unterschiedliche Arten. Einerseits unterstützen KI-Modelle Klimaforscher:innen bei der Entwicklung präziser Klimamodelle und Vorhersageinstrumente, was dabei hilft, die Klimaerwärmung besser zu verstehen und diese einzudämmen. Daher spielen KI-Anwendungen eine zunehmend wichtige Rolle bei der Reduzierung von Treibhausgasemissionen, indem sie beispielsweise effizientere Energieverbrauchsmuster in Bereichen wie Produktion, Transport und Heizung erkennen und fördern. Solche Technologien wurden auch zur Unterstützung des europäischen Green Deal vorgeschlagen, und die „AI for Good“-Plattform (→ <https://aiforgood.itu.int>) der Vereinten Nationen hilft, KI-basierte Lösungen zur Bekämpfung des Klimawandels zu skalieren. Technologisch fortgeschrittene Länder sollten dabei weniger fortgeschrittenen Ländern den Zugang zu KI-Ressourcen erleichtern, um globale Klimaschutzbemühungen zu unterstützen.

KI als Treiber von Treibhausgasemissionen

Andererseits gibt es jedoch auch Anwendungen von KI, die den Klimawandel verschärfen. Insbesondere Öl- und Gasunternehmen setzen KI ein, um die Effizienz bei der Förderung und Verarbeitung fossiler Brennstoffe zu verbessern. Diese Technologien helfen dabei, mehr fossile Brennstoffe zu fördern, zu verarbeiten und zu vermarkten, was letztendlich zu einem höheren CO₂-Ausstoß führt. Dies steht im Widerspruch zu den Zielen vieler großer Unternehmen, Kohlenstoffneutralität zu erreichen, wie ein Bericht von Greenpeace 2016¹ zeigte.

Energieverbrauch und CO₂-Fußabdruck der KI

Neben den spezifischen KI-Anwendungen müssen auch die allgemeinen Auswirkungen der KI als Forschungs-, Industrie- und Handelsbereich auf den Klimawandel berücksichtigt werden. Die wachsende

¹ Greenpeace (2016), Oil in the Cloud. How Tech Companies are Helping Big Oil Profit from Climate Destruction (Greenpeace Report), <https://www.greenpeace.org/usa/reports/oil-in-the-cloud/> (abgerufen am: 4.9.2024).

Natural Language

Processing (NLP) ist ein Bereich der Künstlichen Intelligenz, der sich mit der menschlichen Sprache beschäftigt. Es geht darum, Computern die Fähigkeit zu geben, unsere natürliche Sprache zu verstehen, zu interpretieren und zu erzeugen. Anwendungen von NLP umfassen etwa maschinelle Übersetzung, Spracherkennung, Textanalyse, Chatbots und vieles mehr.

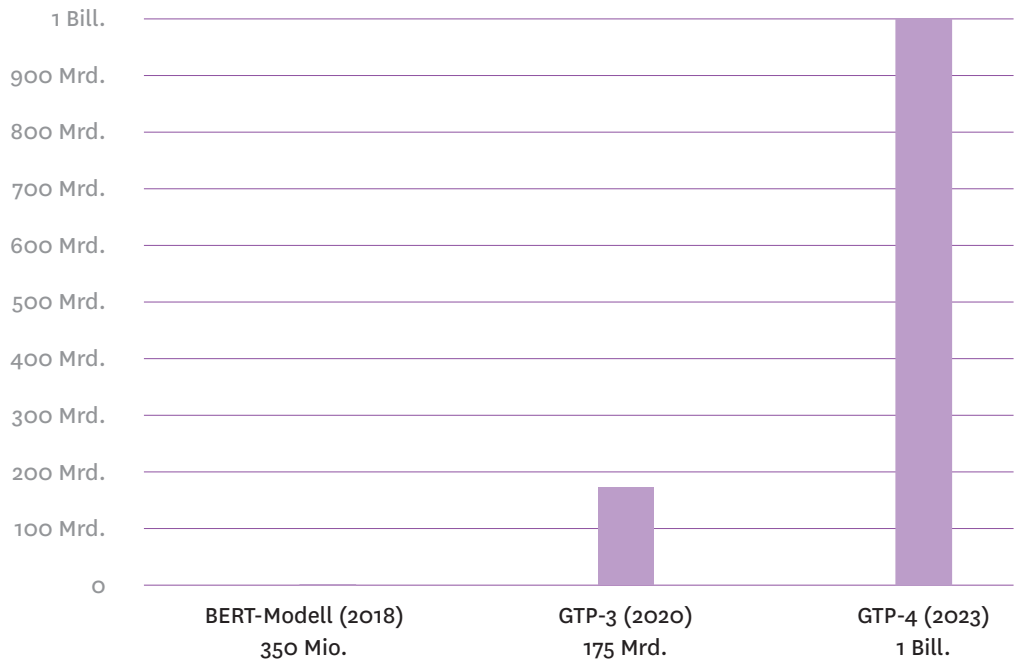
Besorgnis über den Energieverbrauch, der für Datenverarbeitung und die benötigte Hardwareinfrastruktur anfällt, hat sich insbesondere durch alarmierende Schätzungen verstärkt. Beispielsweise verbrauchen Bitcoin-Transaktionen jährlich so viel Energie wie ein ganzes Land in der Größe von Argentinien. Rechenzentren und Datenübertragungsnetze sind für etwa ein Prozent der globalen energiebedingten Treibhausgasemissionen verantwortlich.

Speziell für KI-Modelle zeigt sich ein ähnliches Bild: Das Training großer Modelle für die Verarbeitung natürlicher Sprache (**Natural Language Processing, NLP**) hatte anfangs einen Kohlenstoff-Fußabdruck, der vergleichbar mit dem von fünf durchschnittlichen Autos über ihren gesamten Lebenszyklus war. Neuere Schätzungen haben diese Zahl etwas reduziert, aber die Umweltbedenken bestehen weiterhin. Beispielsweise wird geschätzt, dass bis zu 15 Prozent des gesamten Stromverbrauchs von Google in den Jahren 2019 bis 2021 auf die Entwicklung und Nutzung von KI-Modellen zurückzuführen sind. Nur zehn Prozent dieses Verbrauchs entfallen auf das Training der Modelle, während die restlichen 90 Prozent für die Anwendung der trainierten Modelle genutzt werden.

Steigender Stromverbrauch durch größere KI-Modelle

Es ist unsicher, wie sich der Stromverbrauch von KI-Systemen in Zukunft entwickeln wird, aber er hängt direkt von der Größe der KI-Modelle ab. Es werden immer größere Modelle entwickelt. Die Größe dieser Netze wird durch die Anzahl der Verbindungen zwischen den neuronalen Einheiten bestimmt. Im Bereich der Verarbeitung natürlicher Sprache (NLP) hat sich die Anzahl dieser Parameter stark erhöht: von 350 Millionen bei einem **BERT-Modell** im Jahr 2018 auf 175 Milliarden bei GPT-3 im Jahr 2020 und bis zu mehreren Billionen bei GPT-4 im Jahr 2023. Auch in anderen KI-Bereichen verfolgen Forscher:innen das Ziel, die Genauigkeit durch immer größere Modelle zu verbessern. Dies führt jedoch zu einem höheren Stromverbrauch und einem größeren CO₂-Fußabdruck der KI-Branche. Zudem können KI-Anwendungen, die den Kohlenstoff-Fußabdruck von Prozessen verringern, paradoxerweise deren Nutzung steigern und somit den Gesamt-CO₂-Ausstoß erhöhen.

BERT (Bidirectional Encoder Representations from Transformers) ist der Name eines Algorithmus, den Google 2018 vorgestellt hat.



Grafik 2: Die Anzahl der neuronalen Verbindungen beim BERT-Modell von 2018 in Höhe von 350 Millionen ist im wahrsten Sinne des Wortes verschwindend im Vergleich mit dem aktuellen Modell. Quelle: Bettina K. Lechner

Maßnahmen zur Reduzierung des CO₂-Fußabdrucks der KI

Um den CO₂-Fußabdruck der KI erheblich zu reduzieren, wäre eine Stromversorgung aus erneuerbaren Energiequellen ideal. Im Jahr 2020 jedoch stammten fast zwei Drittel des weltweiten Stroms aus fossilen Energiequellen, was deutlich macht, dass es noch viele Jahre kontinuierlicher Bemühungen bedarf, um eine global grüne Stromversorgung zu erreichen.

In der Zwischenzeit können KI-Forscher:innen und Unternehmen einige bewährte Methoden anwenden, um den Energieverbrauch zu senken und den CO₂-Ausstoß der KI zu verringern:

- 1. Energieeffiziente Architekturen:** Die Auswahl von KI-Modellen, die weniger Energie verbrauchen, kann einen großen Unterschied machen.

- 2. Optimierte Hardware:** Die Verwendung von Prozessoren, die speziell für das Training von KI-Modellen entwickelt wurden, kann den Energieverbrauch erheblich reduzieren.
- 3. Effizientere Rechenzentren:** Berechnungen sollten in Rechenzentren durchgeführt werden, die sauberere Energiequellen nutzen und energieeffizienter sind.

Diese Praktiken sind besonders im Bereich der Informations- und Kommunikationstechnologie von Bedeutung und können unabhängig von der Entwicklung des globalen Strommix umgesetzt werden.

Die Rolle von Forschungs- und Berufsverbänden

Forschungs- und Berufsverbände können ebenfalls eine Schlüsselrolle dabei spielen, eine „grünere“ KI zu fördern. Sie könnten die Kriterien für ein „gutes“ Forschungsergebnis neu definieren, indem sie nicht nur die Genauigkeit von KI-Modellen bewerten, sondern auch deren Energieeffizienz berücksichtigen.

Durch diese neuen Maßstäbe kann die KI-Forschung zu intelligenteren und ressourcenschonenderen Systemen führen, die den Energieverbrauch minimieren. Diese Bemühungen sollten parallel zu den notwendigen Änderungen in der Stromerzeugung und -versorgung vorangetrieben werden.

Es liegt in der ethischen, sozialen und wissenschaftlichen Verantwortung sowohl von KI-Akteur:innen als auch von politischen Einrichtungen, sicherzustellen, dass KI zur Eindämmung der Klimaerwärmung beiträgt und nicht dazu, die Klimakrise weiter zu verschärfen.

Insgesamt bieten KI-Technologien viele Möglichkeiten zur Effizienzsteigerung, doch es bleibt eine kollektive Entscheidung, ob diese Technologien genutzt werden, um die Klimakrise zu bekämpfen oder zu verschärfen.

AI for Good – „gute“ KI-Anwendungen

Neben dem Thema Klimawandel kann KI auch auf andere Weise wertvolle Unterstützung bieten, um schwierige Probleme zu lösen. Diese Anwendungen, oft als „AI for Good“ bezeichnet, zeigen, wie KI

der Gesellschaft zugutekommen kann. Beispiele dafür finden sich in vielen Bereichen, wie der Medizin, in der KI zur Behandlung, Diagnose, Früherkennung, Krebsvorsorge und Medikamentenentwicklung eingesetzt wird. Auch bei der Erkennung von Hassreden, Fake News und in Hilfsmitteln für Menschen mit Behinderungen spielt KI eine wichtige Rolle.

Fazit aus diesem Abschnitt:

Künstliche Intelligenz (KI) birgt sowohl ein enormes Potenzial zur Bekämpfung der Klimakrise als auch das Risiko, diese zu verschärfen. KI kann helfen, effizientere Energieverbrauchsmuster zu identifizieren und Klimamodelle zu verbessern. Gleichzeitig tragen die energieintensiven Prozesse großer KI-Modelle erheblich zum CO₂-Ausstoß bei. Um diese negativen Auswirkungen zu minimieren, müssen bewährte Verfahren angewendet und der Übergang zu erneuerbaren Energiequellen beschleunigt werden. Aber es gibt auch KI-Anwendungen, sogenannte „AI for Good“, die zeigen, wie nützlich KI sein kann.

Drängende Fragen und Reflexion

In den Medien wird die Gefahr von KI häufig mit apokalyptischen Szenarien, wie etwa rebellischen Robotern, dargestellt. Das eigentliche Problem sehen wir jedoch in der ständig zunehmenden Nutzung großer Datenmengen, um immer leistungsfähigere KI-Modelle zu trainieren. Dies bringt zahlreiche Herausforderungen mit sich, darunter Fragen des Urheberrechts, des Schutzes der Privatsphäre und der Auswirkungen personalisierter Systeme. Zudem spielt die kostengünstige Arbeit für die Datenerfassung und das Training der Modelle eine große Rolle. Weitere wichtige Themen, wie die Auswirkungen von KI auf Bildung und Arbeitswelt, können hier aus Platzgründen nicht behandelt werden.

Ein konkretes Beispiel ist die Entwicklung des Programms ChatGPT, das mithilfe einer Vielzahl von Quellen wie Büchern, Artikeln, Websites und sogar Social-Media-Kommentaren trainiert wurde. Dabei

WISSENSFRAGEN

KI und die Klimakrise

Wie kann Künstliche Intelligenz (KI) zur Eindämmung der Klimaerwärmung beitragen?

Welchen Anteil am weltweiten energiebezogenen CO₂-Ausstoß haben Rechenzentren und Datenübertragungsnetze?

Welche Maßnahmen können ergriffen werden, um den CO₂-Fußabdruck von KI-Anwendungen zu reduzieren?

In welchen Bereichen kann „AI for Good“ zur Lösung gesellschaftlicher Herausforderungen beitragen, und welche ethischen Überlegungen sollten dabei berücksichtigt werden?

Wie können Forschungs- und Berufsverbände dazu beitragen, Standards für die verantwortungsvolle Entwicklung und Anwendung von KI festzulegen und zu fördern?

FRAGEN ZUR DISKUSSION

KI und die Klimakrise

Sollten KI-Unternehmen gesetzlich dazu verpflichtet werden, ihren CO₂-Fußabdruck offenzulegen und Maßnahmen zur Reduzierung zu ergreifen? Warum oder warum nicht?

1

Mögliche Diskussion: Die Verpflichtung zur Offenlegung könnte zu mehr Transparenz und Verantwortungsbewusstsein führen, aber es könnte auch argumentiert werden, dass dies die Innovationskraft und Wettbewerbsfähigkeit einschränken könnte.

In welchem Maß sollten KI-Entwickler:innen und -Forscher:innen ihre Prioritäten ändern, um mehr Wert auf die Energieeffizienz ihrer Modelle zu legen?

2

Mögliche Diskussion: Eine stärkere Priorisierung der Energieeffizienz könnte den CO₂-Fußabdruck von KI-Anwendungen erheblich reduzieren, jedoch könnte dies auch zu Kompromissen bei der Genauigkeit oder Leistungsfähigkeit der Modelle führen.

wurden die Urheber:innen dieser Inhalte nicht um ihre Zustimmung gebeten. Diese Art von Urheberrechtsproblemen besteht schon länger im Internet, wird aber besonders relevant, wenn persönliche Daten für KI-Assistenten genutzt werden. Social-Media-Plattformen sind darauf ausgelegt, Nutzer:innen durch die Nutzung ihrer Vorlieben und ihres Browserverlaufs möglichst lange zu binden, was zu Phänomenen wie Filterblasen und Echokammern führt. Diese tragen zur politischen Polarisierung bei und gefährden langfristig demokratische Prozesse.

Diese Probleme könnten sich verschärfen, wenn KI-Systeme gezielt das Wissen über ihre Nutzer:innen verwenden, um Antworten zu geben, die deren bestehende Überzeugungen bestätigen oder verstärken. Es ist daher wichtig, vorsichtig zu sein, wie wir KI-Systeme mit persönlichen Daten füttern. Darüber hinaus wird oft übersehen, dass das Training von KI-Modellen oft auf menschliche Arbeit angewiesen ist, die häufig unter schlechten Bedingungen, oft im Globalen Süden, stattfindet. Diese Arbeitskräfte sind damit beschäftigt, problematische Inhalte wie Hassreden oder Bilder von Gewalt zu kennzeichnen. So wurde bekannt, dass die Anpassung von ChatGPT zur Vermeidung toxischer Antworten von unterbezahlten kenianischen Arbeiter:innen durchgeführt wurde, was erstmals breite mediale Aufmerksamkeit erregte. Dieses Problem ist jedoch nicht neu und scheint fest in der aktuellen Praxis der Entwicklung fortschrittlicher KI-Systeme verankert zu sein.

Fazit aus diesem Abschnitt:

Die Nutzung von Künstlicher Intelligenz ist mit zahlreichen ethischen und sozialen Herausforderungen verbunden. Die zunehmende Abhängigkeit von großen Datenmengen zur Verbesserung von KI-Modellen birgt Risiken wie den Verlust der Privatsphäre, Probleme im Urheberrecht und die Ausbeutung kostengünstiger Arbeitskräfte, insbesondere im Globalen Süden. Zudem besteht die Gefahr, dass KI-Systeme durch die Nutzung persönlicher Daten den Nutzer:innen nur das bieten, was sie hören wollen, was zu einer Verstärkung von Filterblasen und einer Bedrohung demokratischer Prozesse führen kann. Es ist daher entscheidend, dass der Einsatz von KI verantwortungsvoll und mit Bedacht erfolgt, um negative Auswirkungen auf Gesellschaft und Individuen zu minimieren.

AKTIVIERUNG

Verschiedene Szenarien der Auswirkungen von KI

Ziel:

Die Teilnehmer:innen sollen erkennen, wie komplex und weitreichend die Auswirkungen von KI sind und dass es viele Perspektiven zu berücksichtigen gibt. Ziel ist es, Lösungsansätze zu entwickeln, die sowohl die Chancen verantwortungsvoll nutzen als auch die Risiken der KI-Technologien minimieren.

Beispielszenarien:

1. KI und Klimaschutz:

Eine Gruppe könnte darüber diskutieren, wie KI effektiv zur Bekämpfung des Klimawandels eingesetzt werden kann. Welche KI-Technologien könnten in ihrem Szenario genutzt werden? Welche potenziellen Herausforderungen und Risiken sehen die Gruppenmitglieder dabei?

2. Datenschutz und Ethik:

Eine andere Gruppe könnte ein Szenario bearbeiten, in dem sie herausfinden muss, wie Unternehmen ihre KI-Systeme trainieren und gleichzeitig den Datenschutz wahren können. Welche Maßnahmen sollten ergriffen werden, um sicherzustellen, dass KI-Systeme ethisch vertretbar bleiben?

3. Auswirkungen auf den Globalen Süden:

Eine dritte Gruppe könnte sich mit den ethischen Fragen rund um die Nutzung von kostengünstigen Arbeitskräften im Globalen Süden für die Datenverarbeitung und -kennzeichnung befassen. Wie könnte die Ausbeutung vermieden und faire Arbeitsbedingungen geschaffen werden?

4. KI zur Unterstützung von Umweltüberwachung

In diesem Szenario könnte eine Gruppe diskutieren, wie KI-Systeme zur Überwachung von Umweltbedingungen eingesetzt werden können. Sie könnten sich auf Technologien konzentrieren, die Luft- und Wasserqualität analysieren, den Energieverbrauch optimieren oder den Erhalt von Ökosystemen unterstützen. Die Gruppe sollte dabei auch die Herausforderungen und potenziellen Risiken betrachten, wie etwa den Energiebedarf von KI-Systemen selbst und mögliche Datenschutzprobleme bei der Erfassung von Umweltdaten.

NOTIZEN



● KAPITELABSCHNITTE ● WISSENSFRAGEN ● FRAGE(N) ZUR DISKUSSION
● AKTIVIERUNG ● GRUPPENARBEIT

Kapitel 5

KI-GESETZGEBUNG UND IHRE HERAUS- FORDERUNGEN

Dieses Kapitel untersucht, wie rechtliche und politische Institutionen auf die Herausforderungen der digitalen Transformation, insbesondere der KI, reagieren können. Es wird betont, dass ein System von Gesetzen notwendig ist, um Grundrechte zu schützen. Das Kapitel beschreibt außerdem die Entwicklung des Digital Services Act (DSA) und des Digital Markets Act (DMA) sowie verschiedener weiterer KI-Gesetze. Abschließend werden aktuelle und zukünftige Herausforderungen der KI-Governance sowie deren Einfluss auf die Gesellschaft und demokratische Werte diskutiert.

Basis dieses Kapitels ist der Aufsatz „Human Rights Alignment: The Challenge Ahead for AI Lawmakers“ von Marc Rotenberg, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024. DOI: 10.1007/978-3-031-45304-5_38.

Einleitung

Die Reaktionen von Regierungen und internationalen Organisationen auf die Herausforderungen der KI-Governance, also der Regelung und Steuerung von Künstlicher Intelligenz, sind ein entscheidender Test dafür, ob wir neue Technologien zum Vorteil und Nutzen der Gesellschaft einsetzen können. Diese Bemühungen basieren auf der Annahme, dass Gesetze und Regulierungen notwendig sind, um Grundrechte zu schützen.

Gesetze sind jedoch oft unvollkommen, da Sprache unpräzise ist und sich Technologien schnell weiterentwickeln. Große Unternehmen werden sich gegen Beschränkungen wehren, und es besteht die Gefahr, dass etablierte Menschenrechte wie der Schutz der Menschenwürde und der Privatsphäre verändert werden, um die technologische Entwicklung zu ermöglichen.

Deshalb ist es wichtig, Prinzipien für die Steuerung von KI zu formulieren, damit die Gesellschaft in der Lage ist, die technologische Entwicklung zu kontrollieren und sicherzustellen, dass digitale Technologien auf den Menschen ausgerichtet bleiben.

Darüber hinaus braucht es aber auch rechtliche Rahmen, um die Grundrechte der Nutzer:innen im digitalen Raum generell sicherzustellen – unabhängig von KI-Gesetzgebungen. Mit dem Digital Services Act (DSA) und dem Digital Markets Act (DMA) hat die EU Maßnahmen gesetzt, um diesen sicheren digitalen Raum zu schaffen, in dem vor allem Nutzerrechte sichergestellt werden. Das Gesetz über digitale Dienste (DSA) und das Gesetz über den digitalen Markt (DMA) bilden ein einheitliches Regelwerk, das in der gesamten EU gilt.

EU-Maßnahmen für einen sicheren digitalen Raum

Digital Services Act (DSA)

Der DSA zielt auf die Schaffung eines sichereren digitalen Raums, in dem die Grundrechte aller Nutzer:innen digitaler Dienste geschützt sind, ab. Er soll die Pflichten von digitalen Diensten regeln, die als Vermittler zwischen Endnutzer:innen einerseits und den Anbieter:innen von Waren, Dienstleistungen und Inhalten andererseits fungieren. Ein großes Anliegen des DSA ist die Transparenz gegenüber Nutzer:innen. Diese sollen nicht nur in den Nutzungsbedingungen der Plattformen transparent über ihre Rechte und Pflichten informiert werden, sondern sollen beispielsweise auch besser über ihnen angezeigte Werbung informiert werden. Dies betrifft insbesondere die Transparenz von Algorithmen. Auf individualisierte Werbung gestützte Geschäftsmodelle bleiben aber dem Grunde nach zulässig.

Der DSA regelt die Haftung von digitalen Diensten für illegale Inhalte, die Nutzer:innen hochgeladen haben, und führt Sorgfaltspflichten der Anbieter:innen ein, um die negativen Auswirkungen von Online-Diensten einzudämmen. Er zielt dabei auf sehr große Online-Plattformen mit mehr als 45 Millionen Nutzer:innen ab. Den Betreiber:innen digitaler Plattformen wird im DSA daher eine verantwortungsvolle Rolle im Kampf gegen illegale Online-Inhalte zugeschrieben.

Eine wesentliche Neuerung ist, dass der Geltungsbereich des DSA an den Ort der Nutzer:innen gebunden ist, nicht an die Niederlassung der Anbieter:innen.

Wenn gegen den DSA verstoßen wird, dürfen Strafen von bis zu sechs Prozent der Jahreseinnahmen oder des Jahresumsatzes des betreffenden Anbieters verhängt werden.

Digital Markets Act (DMA)

Der DMA zielt auf die Schaffung eines wettbewerbsfähigen und fairen Markts im Digitalsektor ab, innerhalb dessen auch kleinen und neuen Akteur:innen Wachstumschancen eröffnet werden und sich Unternehmen und Verbraucher:innen nicht auf das Akzeptieren unfairer

Bedingungen einlassen müssen, die von etablierten und marktmächtigen Anbietern diktiert werden.

Der digitale Binnenmarkt in der EU soll frei, offen und fair bleiben. Um das zu gewährleisten, sollen Anbieter:innen mit großem wirtschaftlichen und daher auch großem gesellschaftlichen Einfluss klaren und vor allem strengeren Regeln unterliegen als bislang. Dabei geht es sowohl um aktive Handlungspflichten als auch um Unterlassungspflichten für Gatekeeper-Plattformen.

Als ein Gatekeeper ist ein Dienst dann zu qualifizieren, wenn er

- a) erheblichen Einfluss hat,
- b) einen Dienst von zentraler Bedeutung für das Erreichen von Verbraucher:innen im Internet betreibt und
- c) dabei eine gefestigte und dauerhafte Position hat.

Dabei kann es um Suchmaschinen, soziale Netzwerke, Video-Sharing-Plattformen, Messaging-Dienste, Betriebssysteme, aber auch um ganz andere Online-Vermittlungsdienste gehen. Für all diese Eigenschaften gibt es definierte spezifische Schwellenwerte.

Für Nutzer:innen ist wichtig, dass Gatekeeper durch den DMA insbesondere von der Zusammenführung von personenbezogenen Daten aus ihren zentralen Diensten mit Daten aus anderen Diensten abgehalten werden. Sie dürfen Nutzer:innen auch nicht mehr daran hindern, vorinstallierte Software oder Apps zu deinstallieren. Auch dürfen sie die Nutzung ihrer Dienste nicht von einer Registrierung bei einem anderen Gatekeeper-Dienst abhängig machen. Es muss auch möglich sein, dass Nutzer:innen ihre eigenen Daten einfach an einen anderen Anbieter übertragen können (Datenportabilität). Dies sind nur einige konkrete Beispiele, die der DMA regelt.

Kommt ein Gatekeeper seinen Pflichten nicht nach, dann drohen empfindliche Strafen in Höhe von bis zu zehn Prozent des gesamten weltweiten Jahresumsatzes. Verhängen kann diese Strafen nach dem DMA-Vorschlag die Europäische Kommission.

Maßnahmen zur Regulierung von KI

In diesem Abschnitt geht es um die Entwicklung von Rechtsrahmen zur Steuerung von Künstlicher Intelligenz (KI). Zu den Governance-Rahmenwerken gehören wichtige politische und globale Vereinbarungen wie

1. die KI-Grundsätze der OECD, die KI-Leitlinien der G20,
2. die UNESCO-Empfehlung zur KI-Ethik und
3. das KI-Gesetz der Europäischen Union.

Diese Rahmenwerke bilden die Grundlage für Rechtsnormen und internationale Vereinbarungen und beeinflussen das Verhalten derjenigen, die KI-Systeme entwickeln und nutzen.

In diesen Rahmenwerken tauchen mehrere Schlüsselbegriffe immer wieder auf: „Fairness“, „Genauigkeit“, „Transparenz“, „menschenzentriert“ und „vertrauenswürdig“. Diese Begriffe könnten als Bausteine des KI-Rechts betrachtet werden, da sie grundlegende Werte darstellen, auf die sich spezifischere Vorgaben beziehen.

1. Die KI-Grundsätze der OECD und die KI-Leitlinien der G20

Die OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung) hat **fünf Grundsätze** für den verantwortungsvollen Umgang mit vertrauenswürdiger KI aufgestellt:

1. Integratives Wachstum, nachhaltige Entwicklung und Wohlstand
2. Menschenzentrierte Werte und Fairness
3. Transparenz und Erklärbarkeit
4. Robustheit, Sicherheit und Schutz
5. Rechenschaftspflicht

Die OECD hat auch **fünf Empfehlungen** für nationale Politik und internationale Zusammenarbeit im Bereich der vertrauenswürdigen KI formuliert:

1. Investitionen in die KI-Forschung und -Entwicklung
2. Förderung eines digitalen Ökosystems für KI
3. Gestaltung eines günstigen politischen Umfelds für KI
4. Aufbau von Humankapazitäten und Vorbereitung auf den Wandel auf dem Arbeitsmarkt
5. Internationale Zusammenarbeit für vertrauenswürdige KI

Diese Prinzipien wurden 2019 von den G20-Staaten bestätigt, und mehr als 50 Länder haben entweder die OECD-Leitprinzipien oder die G20-Leitlinien für Künstliche Intelligenz bestätigt.

2. UNESCO-Empfehlung zur Ethik der Künstlichen Intelligenz

Im November 2021 verabschiedeten die 193 Mitgliedstaaten der UNESCO die Empfehlung zur Ethik der Künstlichen Intelligenz. Sie soll nicht nur die Menschenrechte und die Menschenwürde schützen, sondern diese auch fördern. Zu den wichtigsten Punkten gehören:

1. **Schutz von Daten:** Die UNESCO-Empfehlung fordert zusätzliche Maßnahmen über das hinaus, was Technologieunternehmen und Regierungen bisher tun, um dem:der Einzelnen mehr Schutz zu bieten. Dies beinhaltet mehr Transparenz, Kontrolle und Handlungsfähigkeit in Bezug auf persönliche Daten.
2. **Verbot von Social Scoring und Massenüberwachung:** Die Empfehlung verbietet den Einsatz von KI-Systemen für **Social Scoring** und Massenüberwachung ausdrücklich.

Social Scoring: Das soziale, wirtschaftliche und politische Verhalten von Privatpersonen, Unternehmen und anderen Organisationen (wie z. B. Nichtregierungsorganisationen) soll der Ermittlung einer „sozialen Reputation“ dienen. Ein niedriger Punktestand kann zu Einschränkungen im Alltag führen, etwa beim Zugang zu sozialen Diensten oder bei der Arbeitsplatz- und Ausbildungssuche. (vgl. <https://de.wikipedia.org/wiki/Sozialkredit-System>)

3. **Überwachung und Bewertung:** Es werden neue Instrumente vorgeschlagen, um die Umsetzung der Empfehlungen zu unterstützen, darunter ethische Folgenabschätzungen.
4. **Schutz der Umwelt:** KI soll als wichtiges Instrument im Kampf gegen den Klimawandel und Umweltprobleme eingesetzt werden.

Die Empfehlung konzentriert sich nicht nur auf Werte und Grundsätze, sondern auch auf deren praktische Umsetzung durch konkrete politische Maßnahmen. Die UNESCO fordert die Mitgliedstaaten auf, Rahmen für ethische Folgenabschätzungen und Aufsichtsmechanismen einzuführen.

3. Das KI-Gesetz der EU

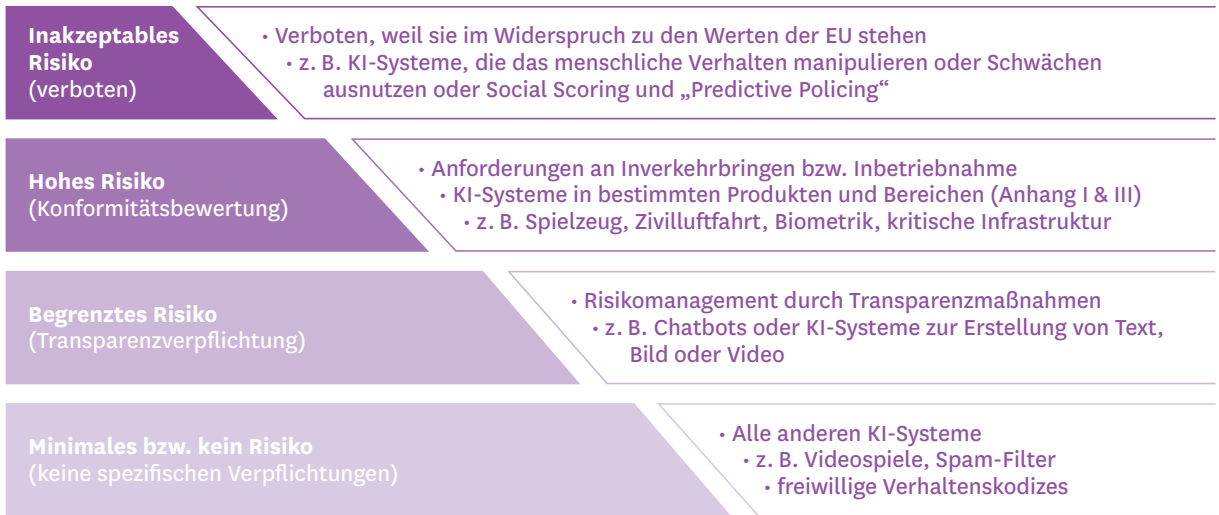
Mit dem KI-Gesetz will das Europäische Parlament sicherstellen, dass KI-Systeme

- sicher,
- transparent,
- nachvollziehbar,
- nicht diskriminierend und
- umweltfreundlich sind.

Das Gesetz legt Verpflichtungen fest, die auf das Risiko ausgerichtet sind, das von dem jeweiligen KI-System ausgeht. Je höher das Risiko, desto strenger die Regelung. KI-Systeme, die als Bedrohung für Menschen gelten, werden verboten.

Wichtige Links zum Governance-Rahmenwerk:

- OECD: Künstliche Intelligenz in der Gesellschaft:
https://www.oecd.org/de/publications/kunstliche-intelligenz-in-der-gesellschaft_6b89dea3-de/full-report.html.
- G20-Leitlinien für Künstliche Intelligenz:
<https://wp.oecd.ai/app/uploads/2021/06/G20-AI-Principles.pdf>
- UNESCO Recommendation on the Ethics of Artificial Intelligence:
<https://unesdoc.unesco.org/ark:/48223/pf0000380455>
- KI-Gesetz (AI-Act) der Europäischen Union:
https://commission.europa.eu/news/ai-act-enters-force-2024-08-01_de



Grafik 3: KI-Systeme werden entsprechend ihrem Risikopotenzial nach inakzeptablem, hohem, begrenztem und minimalem/keinem Risiko kategorisiert © RTR (CC BY 4.0)

Die Grafik 3 zeigt die Risikokategorien des KI-Gesetzes. In der obersten Risikokategorie werden KI-Systeme als „inakzeptabel“ eingestuft, etwa wenn sie den freien Willen der Nutzer:innen beeinflussen oder Social Scoring beinhalten, also die KI-gestützte Bewertung des Verhaltens einzelner Bürger:innen durch staatliche Stellen. Gemäß dem Anwendungsbereich des im August 2024 in Kraft getretenen KI-Gesetzes ist ihre Verwendung in der EU dann verboten.

Für die nächste Gruppe, „Hohes Risiko“, ist der Verwendungsbereich von größerer Bedeutung, das heißt, nicht das KI-System selbst wird als risikoreich angesehen, sondern der Bereich, in dem es eingesetzt wird. Hier werden verschiedene Bereiche genannt, z. B. bei Entscheidungen über den Zugang zu Bildung oder Beschäftigung oder bei Strafverfolgung oder Migration. Fällt ein KI-System in diese Kategorie, müssen Hersteller:innen und Nutzer:innen eine Reihe von Verpflichtungen einhalten – wie die Einrichtung von Risiko- und Qualitätsmanagementsystemen –, und das System muss bei der Kommission registriert werden.

Bei der dritten Gruppe, „Begrenztes Risiko“, sind „nur“ Transparenzpflichten vorgesehen. Die Anbieter:innen müssen ihre Nutzer:innen darüber informieren, dass der Inhalt von einer KI erzeugt wurde. „Risikofreie“ KI-Systeme werden nicht durch das KI-Gesetz geregelt. Dazu

gehören zum Beispiel Spam-Filter für E-Mail-Programme. Hier wird das Risiko für die Nutzer:innen als so gering eingeschätzt, dass keine Regelungen vorgesehen sind.

Beispiele für verbotene KI-Systeme sind:

- **kognitive Verhaltensmanipulation**, zum Beispiel sprachgesteuertes Spielzeug, das gefährliches Verhalten bei Kindern fördert
- **Social Scoring**: Klassifizierung von Menschen auf Grundlage ihres Verhaltens
- **biometrische Kategorisierung** von Menschen bezüglich sensibler Merkmale
- **biometrische Echtzeit-Fernidentifizierungssysteme** wie zum Beispiel Gesichtserkennung

Das EU-Parlament verabschiedete das Gesetz über Künstliche Intelligenz im März 2024, und der Europäische Rat erteilte seine Zustimmung im Mai 2024. Das Gesetz wird 24 Monate nach seinem Inkrafttreten in vollem Umfang anwendbar sein. Systeme mit hohem Risiko werden mehr Zeit haben, um die Anforderungen zu erfüllen; die sie betreffenden Verpflichtungen werden erst 36 Monate nach dem Inkrafttreten des Gesetzes gelten.

Fazit aus diesem Abschnitt:

Die Gesetzgebung zur Künstlichen Intelligenz entwickelt sich rasch und umfasst eine Vielzahl von Prinzipien und Regelungen, die darauf abzielen, KI-Systeme sicher, fair und menschenzentriert zu gestalten. Die UNESCO, die OECD und die EU setzen sich für den Schutz der Menschenrechte, für Transparenz und Nachhaltigkeit in der KI-Nutzung ein. Trotz dieser Fortschritte bleiben Herausforderungen bestehen, insbesondere in Bezug auf die Durchsetzung von Verboten und die Bewältigung der rasanten technologischen Entwicklungen. Eine effektive internationale Zusammenarbeit und eine ständige Anpassung der Gesetze sind unerlässlich, um sicherzustellen, dass KI-Systeme den gesellschaftlichen Werten und Grundrechten entsprechen.

WISSENSFRAGEN

Maßnahmen zur Regulierung von KI

Was sind wiederkehrende Elemente verschiedener KI-Gesetze und -Empfehlungen?

Was sind die fünf Empfehlungen der OECD für nationale Politik und internationale Zusammenarbeit im Bereich der vertrauenswürdigen KI?

Was fordert die UNESCO-Empfehlung zum Schutz persönlicher Daten?

Worauf zielt der Digital Services Act ab?

Was bedeutet Datenportabilität?

Was ist das Ziel des KI-Gesetzes der EU?

Was sind wichtige Herausforderungen bei der Governance von KI-Systemen?

Welche Strategien können eingesetzt werden, um die Einhaltung von KI-Gesetzen weltweit sicherzustellen?

Welche Maßnahmen können ergriffen werden, um sicherzustellen, dass Menschen die Kontrolle über KI-Systeme behalten?

Wie können die Prinzipien Fairness, Genauigkeit, Transparenz, Menschenzentriertheit und Vertrauenswürdigkeit in der Praxis umgesetzt werden, um sicherzustellen, dass KI-Systeme den gesellschaftlichen Erwartungen entsprechen?

1

Mögliche Beiträge:

- **Fairness:** Sicherstellen, dass die Datensätze, auf denen KI-Modelle trainiert werden, vielfältig und repräsentativ sind, um Verzerrungen zu minimieren. Implementierung von Audits und regelmäßigen Überprüfungen der Algorithmen, um Diskriminierung zu erkennen und zu beheben.
- **Genauigkeit:** Entwicklung strenger Testverfahren und Validierungstechniken, um die Genauigkeit der Modelle zu gewährleisten. Implementierung von Mechanismen zur kontinuierlichen Überwachung und Anpassung der Modelle basierend auf neuen Daten.
- **Transparenz:** Bereitstellung verständlicher Erklärungen darüber, wie die KI-Systeme ihre Entscheidungen treffen. Das könnte durch die Entwicklung und Anwendung von „Explainable AI“-Technologien (XAI-Technologien) erreicht werden, die die Entscheidungsprozesse von KI-Modellen verständlicher machen.
- **Menschenzentriertheit:** Einbeziehung von Benutzer:innen und Betroffenen in den Entwicklungsprozess von KI-Systemen, um sicherzustellen, dass die Systeme ihren Bedürfnissen und Erwartungen entsprechen. Dies könnte durch partizipative Designansätze und Benutzerfeedback-Schleifen erfolgen.
- **Vertrauenswürdigkeit:** Aufbau und Aufrechterhaltung von Vertrauen durch die Implementierung von Datenschutzmaßnahmen und Sicherheitsprotokollen, die den Schutz der Benutzerdaten gewährleisten. Regelmäßige Sicherheitsüberprüfungen und Zertifizierungen können ebenfalls zur Vertrauensbildung beitragen.

Welche Rolle sollten internationale Organisationen wie die OECD und die UNESCO bei der Gestaltung und Durchsetzung von KI-Gesetzen spielen, und wie kann eine effektive internationale Zusammenarbeit sichergestellt werden? Welche Rolle könnte die UNO spielen?

2

Mögliche Antworten:

- **Rolle der OECD und UNESCO:** Diese Organisationen sollten globale Standards und Leitlinien entwickeln, die als Grundlage für nationale Gesetzgebungen dienen können. Sie können als Plattformen für den Austausch bewährter Praktiken und Erfahrungen fungieren und dabei helfen, ein gemeinsames Verständnis der Herausforderungen und Lösungen im Bereich der KI-Governance zu entwickeln.
- **Internationale Zusammenarbeit:** Eine effektive Zusammenarbeit kann durch die Schaffung internationaler Foren und Arbeitsgruppen gefördert werden, in denen Regierungen, Unternehmen und Zivilgesellschaft zusammenarbeiten können. Multilaterale Abkommen und Memoranden können helfen, gemeinsame Ziele und Verpflichtungen festzulegen.
- **Herausforderungen und Lösungen:** Unterschiede in den nationalen Gesetzgebungen und Interessen können Herausforderungen darstellen. Diese können durch diplomatischen Dialog und Verhandlungen überwunden werden. Beispiele für erfolgreiche internationale Kooperationen, wie das Pariser Abkommen zum Klimaschutz, könnten als Modelle dienen. Initiativen zur Harmonisierung der Vorschriften und zur gegenseitigen Anerkennung von Standards können ebenfalls zur Förderung der Zusammenarbeit beitragen.

AKTIVIERUNG

Freiheit unter Beobachtung – Die unsichtbaren Fesseln der KI-Überwachung

Ziel:

Die Teilnehmer:innen sollen durch das Rollenspiel erkennen, wie wichtig KI-Gesetze sind, auch wenn sie denken, dass sie nichts zu verbergen haben. Dabei wird verdeutlicht, wie Überwachung und Bewertung durch KI-Systeme die Freiheit und Rechte der Bürger:innen einschränken können.

Materialien:

- Rollenbeschreibung für politische Entscheidungsträger:innen und Bürger:innen
- Karten oder Marker zur Darstellung der verschiedenen Überwachungsmaßnahmen
- Ein Whiteboard oder Flipchart für die Diskussion und Zusammenfassung

1. Einführung und Aufteilung:

- ▶ Die Teilnehmer:innen werden in zwei Gruppen aufgeteilt: politische Entscheidungsträger:innen und Bürger:innen.
- ▶ Eine kurze Einführung in die Thematik wird gegeben, inklusive einer Erklärung der verbotenen KI-Werkzeuge wie Social Scoring und biometrische Überwachung.

2. Szenario:

- ▶ Das Szenario beginnt damit, dass die politischen Entscheidungsträger:innen neue Überwachungsmaßnahmen ankündigen, um „die Sicherheit und Effizienz zu steigern“.
- ▶ Die Bürger:innen erfahren von diesen Maßnahmen und diskutieren, wie sich diese auf ihr tägliches Leben auswirken könnten.

3. Interaktion:

- ▶ Die politischen Entscheidungsträger:innen setzen verschiedene Überwachungsmaßnahmen ein, wie z. B. das Erfassen von Daten aus sozialen Medien, die Nutzung von Gesichtserkennung und das Bewerten der Bürger:innen basierend auf ihrem Verhalten.
- ▶ Die Bürger:innen müssen sich in ihrem Alltag damit auseinandersetzen und versuchen, normale Aktivitäten wie das Beantragen von Jobs, das Reisen oder das Nutzen von sozialen Dienstleistungen zu bewältigen, während sie unter ständiger Beobachtung stehen.

4. Diskussion:

Nach der Interaktion kommen beide Gruppen zusammen und diskutieren ihre Erfahrungen.

Fragen zur Reflexion:

- ▶ Wie haben sich die Bürger:innen unter der ständigen Überwachung gefühlt?

- ▶ Welche Herausforderungen haben die politischen Entscheidungsträger:innen bei der Durchsetzung der Überwachungsmaßnahmen erlebt?
- ▶ Welche Rechte wurden durch die Überwachung eingeschränkt?
- ▶ Warum sind Gesetze, die solche Überwachungsmaßnahmen verbieten, wichtig?

5. Schlussfolgerung:

- ▶ Eine kurze Zusammenfassung der Bedeutung von KI-Gesetzen und deren Schutzmechanismen.
- ▶ Hervorhebung der Rolle der internationalen Zusammenarbeit und der ständigen Anpassung von Gesetzen, um die Rechte der Bürger:innen zu schützen.

GRUPPENARBEIT

KI-Regulierungsgesetz

Materialien:

Notebook, Tablet bzw.
Smartphone zur Recherche

Erstellen Sie eine Simulation, in der die Teilnehmer:innen in Gruppen eingeteilt werden, um ein KI-Regulierungsgesetz zu entwerfen. Die Gruppen müssen Aspekte wie Transparenz, Rechenschaftspflicht, Datenschutz und ethische Überlegungen einbeziehen. Anschließend präsentieren die Gruppen ihre Vorschläge und diskutieren die Vor- und Nachteile der verschiedenen Ansätze.

Aktuelle und künftige Herausforderungen

Eine der größten Herausforderungen bei der Regulierung von KI besteht darin, bestimmte Anwendungen zu verbieten und die Einhaltung dieser Verbote durchzusetzen. Viele Länder unterstützen trotz Verboten weiterhin den Einsatz von KI-Technologien für Zwecke wie Social Scoring und Massenüberwachung.

Der Einsatz von KI zur biometrischen Fernidentifizierung bleibt auch in anderen Ländern problematisch, da diese Systeme Bilder mit Personenprofilen verknüpfen können, was zu polizeilichen Eingriffen führen kann, bevor es zu einer Straftat kommt. In demokratischen Gesellschaften erfordert die wirksame Steuerung von KI die Einführung von Beschränkungen für solche Systeme.

Die rechtlichen Rahmenbedingungen, die derzeit von der EU, dem Europarat und vielen Regierungen weltweit erarbeitet werden, stehen vor dem Problem, nicht unmittelbar wirksam zu werden. Von der Verabschiedung der Gesetze bis zur tatsächlichen Umsetzung ver-

gehen Monate bzw. Jahre. Der Zeitraum von der Verabschiedung des EU-Gesetzes über Künstliche Intelligenz bis zur tatsächlichen Anwendung desselben umfasst zwei Jahre, wodurch sich eine Lücke auftut, die den Einsatz unregulierter KI-Systeme ermöglicht. Der Vorschlag der Industrie, einen vorläufigen „KI-Pakt“ oder „Verhaltenskodex“ zu entwickeln, wird jedoch von zivilgesellschaftlichen Organisationen abgelehnt, da er die demokratische Entscheidungsfindung untergraben könnte.

Gesamt betrachtet haben zurzeit nur die Europäische Kommission und die Volksrepublik China gültige Gesetze und Regulierungen von Künstlicher Intelligenz, die auch durchgesetzt werden können. In allen anderen Regionen der Welt gibt es aktuell lediglich Empfehlungen.

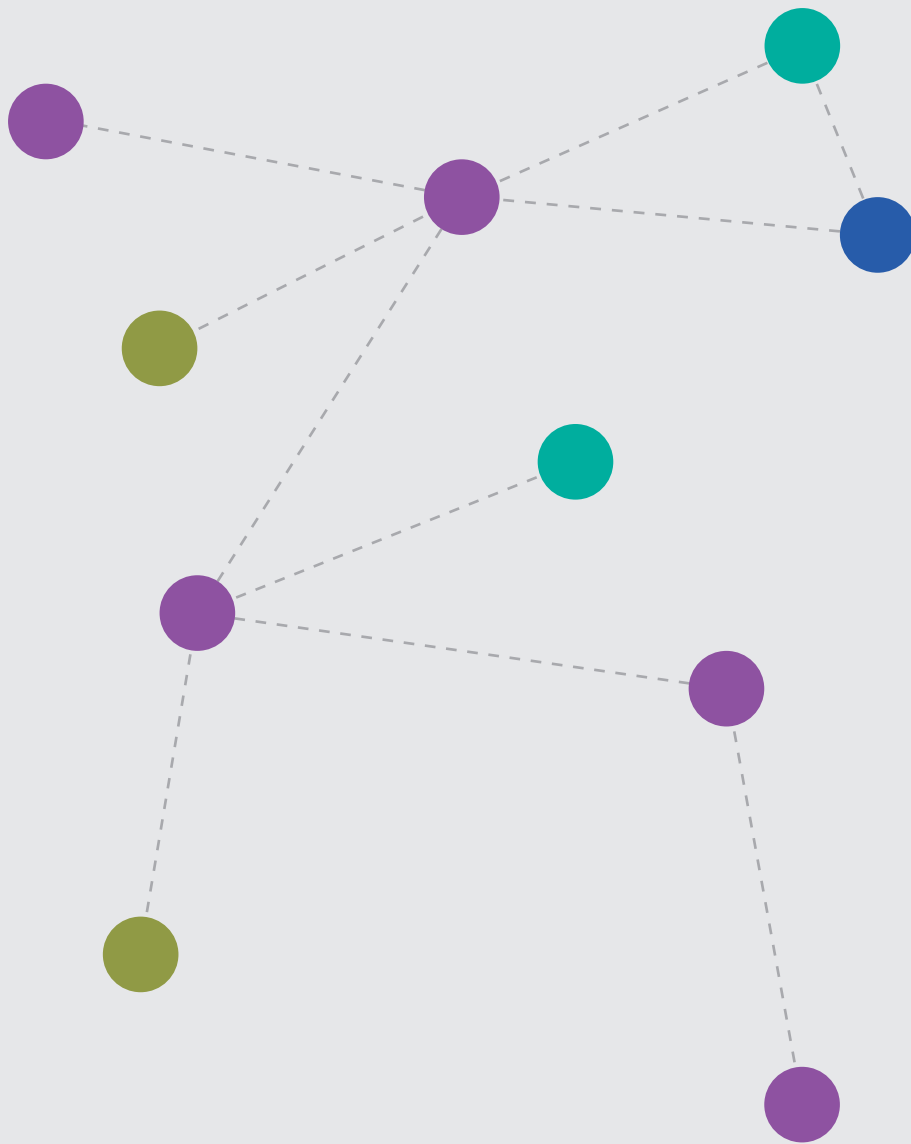
Die Governance hat zwei große Herausforderungen:

1. Durchsetzen von Gesetzen,
2. weltweite Gültigkeit.

Schließlich gibt es die existenzielle Herausforderung: Wird der Mensch die Kontrolle über die von ihm geschaffenen KI-Systeme behalten? Diese Frage ist in den letzten Jahren zunehmend in den Fokus gerückt. Die Fähigkeit, wirksame Gesetze zur Regulierung der KI zu entwickeln, wird dadurch noch wichtiger.

Fazit aus diesem Abschnitt:

Die Herausforderung besteht darin, effektive Gesetze zu schaffen und diese durchzusetzen, um gefährliche KI-Anwendungen zu verhindern. Die internationale Zusammenarbeit ist notwendig, um einheitliche Standards zu setzen und die demokratischen Werte zu schützen. Es muss sichergestellt werden, dass der Mensch die Kontrolle über die KI-Systeme behält und diese im Einklang mit demokratischen Werten und Grundrechten stehen.



- KAPITELABSCHNITTE
- WISSENSFRAGEN
- FRAGE(N) ZUR DISKUSSION
- AKTIVIERUNG
- GRUPPENARBEIT

Kapitel 6

EMPFEHLUNGSSYSTEME UND IHRE AUSWIRKUNGEN AUF PLURALITÄT UND FAIRNESS

Empfehlungssysteme sind in verschiedenen Anwendungen weitverbreitet, z. B. beim Onlineshopping, in sozialen Medien und bei der Personalisierung von Nachrichten. Dieses Kapitel gibt eine kurze Einführung in das Thema Empfehlungssysteme, erklärt ihre grundlegenden Mechanismen, ihre Auswirkungen und erörtert Ansätze zur Berücksichtigung von Vielfalt und Fairness in diesen Systemen.

Basis dieses Kapitels ist der Aufsatz „Recommender Systems: Techniques, Effects, and Measures Toward Pluralism and Fairness“ von Peter Knees, Julia Neidhardt und Irina Nalis, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024. DOI: 10.1007/978-3-031-45304-5_27.

Einleitung

Empfehlungssysteme sind Software-Tools, die Daten über das Verhalten und die Vorlieben der Nutzer:innen sammeln. Diese Daten werden dann verwendet, um den Nutzer:innen künftig Inhalte vorzuschlagen, die ihnen aufgrund des bisher Gesehenen/Gelesenen gefallen könnten. Diese Empfehlungen können sich auf eine Vielzahl von Produkten beziehen – von Büchern, Filmen und Musik über Reise- und Jobangebote bis zu Finanzdienstleistungen. Empfehlungssysteme sind heute weitverbreitet und werden von großen Onlineplattformen wie Google, Amazon, Facebook, YouTube und Netflix eingesetzt, um ihre Dienste an die Bedürfnisse und Vorlieben der Nutzer:innen anzupassen.

Empfehlungssysteme spielen eine zentrale Rolle im modernen Internet, da sie nicht nur den Zugang zu digitalen Inhalten erleichtern, sondern auch beeinflussen, welche Inhalte den Nutzer:innen präsentiert werden. In gewisser Weise sind sie zu Gatekeepern des Zugangs zu Wissen und Information geworden. Sie definieren, welche digitalen Inhalte uns im Internet angeboten werden. Sie sind in gewisser Weise aber auch notwendig, damit wir uns in der immens großen Menge an Information zurechtfinden.

In diesem Kapitel konzentrieren wir uns auf zwei wichtige Aspekte von Empfehlungssystemen, die gesellschaftliche Auswirkungen haben können: Diversität und Fairness. Wir zeigen, dass Empfehlungssysteme diese Werte fördern, aber auch bedrohen können. Diese Diskussion wird in den folgenden Kapiteln vertieft, wo wir auch Themen wie Verzerrungen und Vorurteile im Web und die automatische Moderation von Inhalten behandeln.

Empfehlungssysteme: Konzepte und Praktiken

Empfehlungssysteme sind spezialisierte Softwareanwendungen, die darauf abzielen, Nutzer:innen gezielt Artikel, Produkte oder Inhalte vorzuschlagen, die auf ihren individuellen Vorlieben und Verhaltensweisen im Web basieren. Diese Systeme nutzen eine Vielzahl von Datenquellen, um die personalisierten Empfehlungen zu erstellen.

Zu den wichtigsten Daten, die Empfehlungssysteme verwenden, gehören:

- **Informationen über die Artikel:** Dies umfasst Eigenschaften, mit denen ein Produkt im Netz verknüpft ist, wie Textbeschreibungen, Schlüsselwörter, Kategorien (z. B. Buchgenre oder Produkttyp), Veröffentlichungsdatum und Preis. Diese Daten helfen dem System, die Eigenschaften der Artikel zu verstehen und sie miteinander zu vergleichen.
- **Informationen über die Nutzer:innen:** Um relevante Empfehlungen aussprechen zu können, benötigen die Systeme auch Daten über die Nutzer:innen. Dazu gehören demografische Informationen (z. B. Alter, Geschlecht), Daten zum bisherigen Verhalten des Nutzers, der Nutzerin (wie frühere Käufe, Klicks, Bewertungen) und manchmal auch Informationen aus den sozialen Netzwerken der Nutzer:innen, um zu sehen, was Freunde oder Kontakte mögen.

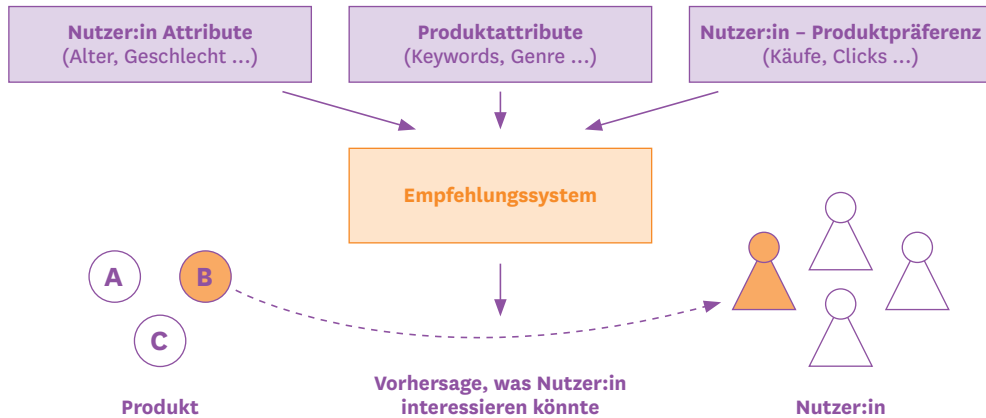
Die Art und Weise, wie Empfehlungssysteme diese Informationen verwenden, sind ganz unterschiedlich. Hier gibt es zum Beispiel folgende Ansätze:

- **Inhaltsbasiertes Filtern:** Hierbei werden Artikel empfohlen, die ähnliche Merkmale aufweisen wie Artikel, die die Nutzerin bzw. der Nutzer zuvor konsumiert hat.
→ **Beispiel:** Auf einer Streaming-Plattform wird Lena, die regelmäßig Dokumentationen über Höhlenmalereien anschaut, eine weitere Dokumentation über das Leben steinzeitlicher Menschen vorgeschlagen, weil diese ähnliche Genremerkmale wie die vorher angesehenen Dokus hat.
- **Kollaboratives Filtern:** Dieser Ansatz basiert auf den Präferenzen anderer Nutzer:innen mit ähnlichen Interessen. Das System analysiert, welche Artikel von Nutzer:innen mit ähnlichem Verhalten positiv bewertet wurden, und empfiehlt diese Artikel.

- **Beispiel:** Eine große Einkaufsplattform schlägt Tom vor, zusätzlich zum Kauf eines Mixers ein Kochbuch zu erwerben, weil andere Kund:innen, die den Mixer gekauft haben, auch dieses Kochbuch erworben haben.
- **Demografiebasiertes Filtern:** Empfehlungen basieren auf demografischen Daten, die allgemeinere Nutzergruppen betreffen. So könnten beispielsweise bestimmte Filme bevorzugt jüngeren Nutzer:innen vorgeschlagen werden, basierend auf dem, was diese Altersgruppe im Allgemeinen bevorzugt.
 - **Beispiel:** Eine Musikplattform empfiehlt Mia, die im Jahr 2000 geboren wurde, Playlists wie „Teen Party“ oder „Pop Rising“, die sich also an ihre Altersgruppe richten.
- **Beliebtheitsbasierte Empfehlungen:** Artikel, die aktuell im Trend liegen oder besonders häufig angesehen oder gekauft werden, werden bevorzugt empfohlen. Dies sorgt dafür, dass Nutzer:innen die „angesagtesten“ Inhalte zu sehen bekommen.
 - **Beispiel:** Eine Videoplattform zeigt Markus Videos an, die seinen Freunden gefallen haben, obwohl er diese Seiten oder Inhalte nicht gelikt hat.
- **Hybride Empfehlungssysteme:** Diese kombinieren mehrere der oben genannten Ansätze, um die Empfehlungsqualität weiter zu verbessern. Ein hybrides System könnte zum Beispiel sowohl inhaltsbasierte als auch kollaborative Filtermethoden verwenden, um die am besten passenden Artikel zu empfehlen.
 - **Beispiel:** Sarah bekommt auf dem Online-Marktplatz „BuyMore“ den neuen Thriller „Dark Shadows“ empfohlen, basierend auf der Kombination von ihrem bisherigen Kaufverhalten, den Vorlieben ähnlicher Nutzer:innen und aktuellen Trends.

Damit ein Empfehlungssystem diese personalisierten Vorschläge machen kann, muss es ein Modell oder Profil der Nutzerin bzw. des Nutzers erstellen. In diesem Profil wird gesammelt, welche Vorlieben der:die Nutzer:in hat, welche Artikel sie oder er in der Vergangenheit konsumiert hat und welche demografischen Daten bekannt sind. Dies gibt Anlass zu Bedenken hinsichtlich des Datenschutzes, da umfangreiche persönliche Daten gesammelt und gespeichert werden.

Traditionell wurden Empfehlungssysteme hauptsächlich danach bewertet, wie präzise sie vorhersagen können, ob ein:e Nutzer:in einen bestimmten Artikel oder Inhalt mögen wird. Dies förderte die Wahrscheinlichkeit eines Klicks bzw. Kaufs. Diese Fokussierung auf Genauigkeit hat jedoch auch Nachteile: Sie kann zu Filterblasen und



Grafik 4: Schematische Darstellung eines Empfehlungssystems.

Quelle: Hannes Werthner

Echokammern führen. Das bedeutet, dass Nutzer:innen zunehmend nur noch Inhalte sehen, die ihre bestehenden Interessen bestätigen, und andere Perspektiven oder neue Inhalte sie kaum noch erreichen. Dadurch werden Vielfalt und Unterschiedlichkeit der konsumierten Informationen eingeschränkt, was besonders bei Nachrichten oder sozialen Medien problematisch ist. Die Qualität eines Empfehlungssystems hängt also auch davon ab, wie vielfältig und ausgewogen die vorgeschlagenen Inhalte sind.

Fazit aus diesem Abschnitt:

Dieser Abschnitt zeigt die grundlegenden Mechanismen von Empfehlungssystemen auf, die durch die Analyse von Nutzerdaten personalisierte Vorschläge machen. Es wird deutlich, dass diese Systeme eine Balance zwischen Genauigkeit und Vielfalt finden müssen, um die Bildung von Filterblasen und Echokammern zu vermeiden. Genauigkeit ist wichtig, um den Nutzer:innen den für sie optimalen Inhalt zu empfehlen. Vielfalt der Empfehlungen hingegend ist entscheidend, um den Nutzer:innen eine breite Palette von Inhalten anzubieten und eine Fülle von Meinungen zu fördern. Dies unterstreicht die Notwendigkeit, bei der Entwicklung von Empfehlungssystemen nicht nur technische Überlegungen zu berücksichtigen.

WISSENSFRAGEN

Empfehlungssysteme: Konzepte und Praktiken

Welche Datenarten verwenden Empfehlungssysteme, um personalisierte Vorschläge zu erstellen?

Was versteht man unter einem inhaltsbasierten Ansatz bei Empfehlungssystemen?

Welche Herausforderungen entstehen durch Filterblasen in Empfehlungssystemen, und wie können diese das Weltbild der Nutzer:innen beeinflussen?

Welche Risiken birgt die ausschließliche Fokussierung auf die Genauigkeit in der Vorhersage des Nutzerverhaltens/-geschmacks bei Empfehlungssystemen?

Wie können Ansätze zur Erhöhung der Diversität und Fairness in Empfehlungssystemen die Bildung von Echokammern verhindern?

1

Wie können Empfehlungssysteme gestaltet werden, um eine Balance zwischen Genauigkeit und Vielfalt zu erreichen?

Man könnte auch darauf achten, dass bei der Bewertung von Empfehlungen nicht nur geschaut wird, wie gut sie passen, sondern auch darauf, wie unterschiedlich die Vorschläge sind. Das sorgt für mehr Abwechslung.

Eine Möglichkeit wäre, unterschiedliche Methoden zu kombinieren, sodass man nicht nur das zeigt, was man mag, sondern auch einmal etwas Neues oder anderes zum Entdecken anbietet. Das könnte helfen, den eigenen Horizont zu erweitern und nicht immer nur in der eigenen „Blase“ zu bleiben.

Eine weitere Idee wäre, das System so zu gestalten, dass es einem gezielt ab und zu neue oder weniger bekannte, eventuell sogar konträre Inhalte vorschlägt, um die Auswahl vielfältiger zu machen.

2

Welche ethischen Herausforderungen ergeben sich durch die Nutzung von Nutzerdaten in Empfehlungssystemen, und wie könnten diese Herausforderungen adressiert werden?

Ein großes Problem ist der Datenschutz. Daten, die über einen gesammelt werden, könnten ohne Zustimmung genutzt oder missbraucht werden. Eine Lösung wäre, transparent zu zeigen, wofür Daten verwendet werden.

Ein weiteres Problem ist, dass die Empfehlungen unbewusst bestimmte Gruppen benachteiligen könnten. Das könnte man verhindern, indem die Algorithmen regelmäßig überprüft und angepasst werden, um sicherzustellen, dass sie fair sind.

Außerdem besteht die Gefahr, dass die Systeme einen zu stark beeinflussen und vielleicht Inhalte zeigen, die gefährlich oder polarisierend sind. Hier könnte es helfen, strenge ethische Standards für die Algorithmen einzuführen und unabhängige Prüfungen durchzuführen, um sicherzustellen, dass die Inhalte in Ordnung sind.

AKTIVIERUNG

Meine persönlichen Empfehlungen

Ziel:

Das Bewusstsein der Teilnehmer:innen für die Mechanismen von Empfehlungssystemen stärken und sie dazu anregen, ihre eigenen Erfahrungen mit personalisierten Vorschlägen zu reflektieren.

Materialien:

- Moderationskärtchen und Stifte
- Pinnwand und Pins

Dauer:

Abhängig von der Anzahl der Gruppen, mindestens jedoch rund 20 Minuten

1. Einteilung der Teilnehmer:innen in Teams:

Jedes Team erhält mehrere Moderationskärtchen und Stifte.

2. Jedes Team erhält folgende Aufgabe:

- ▶ eine eigene Empfehlung für ein fiktives System zu entwickeln. Die Teilnehmer:innen sollen sich vorstellen, dass sie selbst der Algorithmus sind und eine Empfehlung für eine:n Nutzer:in erstellen.

3. Die Teams überlegen:

- ▶ auf welcher Grundlage sie die Empfehlung erstellen (z. B. frühere Vorlieben, aktuelle Trends, Alter, Geschlecht).
- ▶ wie sie sicherstellen können, dass die Empfehlung divers ist und nicht zu einer Filterblase führt.
- ▶ wie sie Fairness in ihre Empfehlung integrieren können.

4. Präsentation und Diskussion:

- ▶ Jede Gruppe präsentiert den anderen Teilnehmer:innen ihre Erkenntnisse.

5. Abschließende Fragen zur Reflexion:

- ▶ Wie beeinflussen Empfehlungssysteme den Alltag der Teilnehmer:innen?
- ▶ Welche Verantwortung sollten Entwickler:innen und Nutzer:innen solcher Systeme tragen?

Empfehlungssysteme als Bedrohung für Vielfalt und Fairness?

Ein großes Problem bei Empfehlungssystemen ist, dass sie oft die Vielfalt der Inhalte einschränken, indem sie ähnliche Artikel basierend auf früheren Entscheidungen bevorzugen und andere, davon abweichende gar nicht erst präsentieren. Eli Pariser (*1980) prägte dafür 2011 den Begriff „Filterblasen“. Wenn jemand beispielsweise in sozialen Medien Beiträge zum Thema Migration liest, könnte das System immer mehr Inhalte zu diesem Thema vorschlagen, sodass andere interessante Themen verdrängt werden. Das führt dazu, dass das Thema Migration überrepräsentiert wird und die Vielfalt der Inhalte abnimmt. Pariser argumentierte, dass solche Filter und personalisierten Dienste in direktem Zusammenhang mit Isolation und sozialer Spaltung stehen, was letztlich zu Polarisierung führt und eine Gefahr für demokratische Gesellschaften darstellt.

Mehrere Studien haben untersucht, wie sich das Nutzerverhalten bei der Verwendung von Empfehlungssystemen auf die Vielfalt der konsumierten Inhalte auswirkt. Sie kamen jedoch zu unterschiedlichen Ergebnissen. Daher wurde vorgeschlagen, weitere Unterscheidungen einzuführen und zwischen den folgenden Arten von Filterblasen zu differenzieren:

- **Technologische Filterblasen:** Diese entstehen durch Algorithmen, die dazu führen, dass die Vielfalt der Empfehlungen, die ein:e Nutzer:in erhält, im Laufe der Zeit abnimmt.
- **Gesellschaftliche Filterblasen:** Diese sind die gesellschaftlichen Auswirkungen technologischer Filterblasen, bei denen sich die Gesellschaft in Gruppen mit stark unterschiedlichen Meinungen aufspaltet. Dieser Prozess wird auch als Personalisierungs-Polarisierung-Hypothese bezeichnet.

Die Bedeutung von Vielfalt ist nicht nur eine technische Frage, sondern auch eine gesellschaftliche, da in einer demokratischen Gesellschaft Bürger:innen Informationen aus einer Vielzahl von Quellen mit unterschiedlichen Sichtweisen benötigen, um fundierte Entscheidungen treffen zu können.

Fairness und unterschiedliche Interessen bei Empfehlungssystemen

Bei Empfehlungssystemen gibt es verschiedene Beteiligte, die unterschiedliche Interessen haben, die manchmal im Widerspruch zueinander stehen können:

- **Nutzer:innen:** Sie möchten sicherstellen, dass die empfohlenen Artikel für sie relevant sind und dass sie fair behandelt werden, unabhängig von Merkmalen wie Ethnie oder Geschlecht.
- **Anbieter:innen von Produkten, Dienstleistungen, Inhalten etc.:** Sie möchten sicherstellen, dass ihre Inhalte gleichberechtigt mit denen anderer Anbieter:innen empfohlen werden.
- **Die Betreiber:innen der Empfehlungssysteme:** Sie möchten den Nutzen für alle Beteiligten maximieren, eigene Vorteile sichern, zum Beispiel durch Empfehlungen, die besonders profitabel sind.

Ein typisches Beispiel für diese Fairnessprobleme sind Jobempfehlungen auf Plattformen wie LinkedIn. Hier stellen sich einige Fragen wie zum Beispiel:

- Sind die Empfehlungen von Stellenangeboten gerecht auf die Nutzer:innen verteilt?
- Sind die Bewertungen die Eignung von Bewerber:innen betreffend fair, oder werden bestimmte Bewerber:innen unter- oder überbewertet?
- Haben die Nutzer:innen eine faire Chance, in den Ergebnislisten zu erscheinen, wenn Personalverantwortliche nach Kandidatinnen bzw. Kandidaten für eine offene Stelle suchen?

Darüber hinaus gibt es bei Empfehlungssystemen oft viele andere Beteiligte, die berücksichtigt werden müssen. Zum Beispiel sind beim Musikstreaming nicht nur die Musiker:innen selbst, sondern auch Plattenfirmen und Rechteinhaber:innen involviert, und bei Plattformen für Essenslieferungen spielen die Lieferfahrer:innen eine wichtige Rolle. Man sieht, Fairness ist ein komplexes Thema.

Fazit aus diesem Abschnitt:

Der Abschnitt zeigt, dass Empfehlungssysteme Vielfalt und Fairness beeinträchtigen können, indem sie Inhalte basierend auf bisherigen Vorlieben priorisieren und andere Inhalte gar nicht erst anzeigen. Dies kann zu Filterblasen führen, in denen Nutzer:innen weniger bis gar nicht mit unterschiedlichen Perspektiven konfrontiert werden. Die unterschiedlichen Interessen von Nutzer:innen, Anbieter:innen und Systemen selbst können die Fairness weiter erschweren, besonders in komplexen Bereichen wie Jobempfehlungen. Daher ist es wichtig, dass Empfehlungssysteme sorgfältig entwickelt und reguliert werden, um Voreingenommenheit und Ungerechtigkeit zu vermeiden.

Überlegungen und Herausforderungen

Empfehlungssysteme setzen wie oben angeführt oft zu sehr auf leicht verfügbare und messbare historische Daten. Dabei werden Diskussionen darüber vernachlässigt, wie stabil menschliche Einstellungen und Verhaltensweisen über die Zeit sind und wie man durch gezielte Maßnahmen Veränderungen bewirken kann.

Viele psychologische Theorien, die in Empfehlungssystemen verwendet werden, basieren auf älteren Forschungen, z. B. teilweise aus den 1990er Jahren. Diese Theorien haben sich oft als zu einfach herausgestellt und können die komplexe und dynamische Natur menschlicher Einstellungen, Verhaltensweisen, Gedanken und Gefühle nicht vollständig erfassen.

Einige Empfehlungssysteme nutzen Modelle aus der **kognitiven Psychologie**, um ihre Funktion zu verbessern. **Persönlichkeitsorientierte Empfehlungssysteme** berücksichtigen individuelle Persönlichkeitsmerkmale, um eine vielfältigere und persönlichere Auswahl an Empfehlungen zu bieten. Ein bekanntes Modell dafür ist das **Fünf-Faktoren-Modell**, das menschliche Persönlichkeitsmerkmale beschreibt.

Das Fünf-Faktoren-Modell (FFM), auch bekannt als Big-Five-Modell oder OCEAN (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism)-Modell, ist ein weitverbreitetes Konzept in der Psychologie, das verwendet wird, um menschliche Persönlichkeitsmerkmale zu beschreiben. Das Modell identifiziert fünf Hauptdimensionen der Persönlichkeit, die als grundlegend für die Charakterisierung eines Menschen gelten. Diese Dimensionen sind:

- 1. Offenheit für Erfahrungen (Openness to Experience):** Beschreibt, wie aufgeschlossen, neugierig und einfallsreich eine Person ist. Menschen mit hoher Offenheit sind tendenziell kreativ und experimentierfreudig.
- 2. Gewissenhaftigkeit (Conscientiousness):** Diese Dimension erfasst, wie organisiert, pflichtbewusst und zuverlässig eine Person ist. Personen mit hoher Gewissenhaftigkeit neigen dazu, zielorientiert und diszipliniert zu sein.
- 3. Extraversion (Extraversion):** Bezieht sich auf das Maß an Geselligkeit, Energie und positiver Emotionalität. Extravertierte Menschen sind kontaktfreudig, aktiv und oft optimistisch.
- 4. Verträglichkeit (Agreeableness):** Diese Dimension misst, wie kooperativ, mitfühlend und gutmütig eine Person ist. Menschen mit hoher Verträglichkeit sind oft hilfsbereit, freundlich und altruistisch.
- 5. Neurotizismus (Neuroticism):** Beschreibt die Tendenz zu negativen Emotionen wie Angst, Unsicherheit und Traurigkeit. Personen mit hohem Neurotizismus erleben häufiger Stress und emotionale Instabilität.

In der Forschung zu Empfehlungssystemen wird das FFM häufig verwendet, um Persönlichkeitsmerkmale zu erschließen, die die Präferenzen und Entscheidungen von Nutzer:innen beeinflussen könnten. Dadurch können personalisierte Empfehlungen besser auf die individuellen Bedürfnisse und Vorlieben der Nutzer:innen abgestimmt werden.

Schlussfolgerungen

Dieses Kapitel hat einen Einblick in Empfehlungssysteme gegeben, ihre Bedeutung für das Web und ihre möglichen Auswirkungen auf Einzelpersonen und die Gesellschaft. Nachdem erklärt wurde, wie diese Systeme funktionieren und wie sie bewertet und optimiert werden, konzentrierten sich die Diskussionen auf die Themen Vielfalt und Fairness. Aus diesen Überlegungen ergeben sich folgende wichtige Punkte:

- **Nicht nur auf vergangene Daten optimieren:** Wenn Empfehlungssysteme sich nur auf historische Muster und Verhaltensdaten stützen, können sie scheinbar effektiver werden. Aber das kann zu geringerer Zufriedenheit der Nutzer:innen führen und ihren Nutzen einschränken. Andere Aspekte, wie die Vielfalt der empfohlenen

Inhalte, sind wichtig, um die Qualität eines Systems wirklich zu beurteilen.

- **Gesellschaftliche Werte priorisieren:** In Bereichen, die für die Demokratie wichtig sind – wie Nachrichten und Medien – oder für das Wohlergehen des Einzelnen, zum Beispiel bei Jobempfehlungen, sollten die von der Gesellschaft definierten Werte wichtiger sein als die Ziele der Diensteanbieter:innen. Dies kann durch Politik und Regulierung sichergestellt werden.
- **Verschiedene Interessen berücksichtigen:** Empfehlungssysteme sind komplex und beinhalten verschiedene Interessengruppen mit oft gegensätzlichen Perspektiven. Fragen von Vielfalt und Fairness müssen aus all diesen Blickwinkeln betrachtet werden. Letztlich sollte die Gesellschaft oder die betroffene Gemeinschaft entscheiden, wessen Interessen Vorrang haben.

Fazit aus diesem Abschnitt:

Empfehlungssysteme sind ein mächtiges Werkzeug im modernen Web, das sowohl positive als auch negative Auswirkungen auf die Nutzer:innen und die Gesellschaft haben kann. Es ist entscheidend, dass die Entwicklung und Bewertung dieser Systeme nicht nur auf technischen Kriterien basiert, sondern auch Aspekte wie Vielfalt und Fairness berücksichtigt. Die gesellschaftlichen Werte und Bedürfnisse sollten Vorrang vor den wirtschaftlichen Interessen der Anbieter:innen haben. Zudem muss die Forschung zu Empfehlungssystemen kritisch hinterfragt werden, um sicherzustellen, dass sie wirklich menschenzentriert ist und nicht auf schwachen oder fehlerhaften Grundlagen basiert, die potenziell gefährlich sein können.

WISSENSFRAGEN

Empfehlungssysteme als Bedrohung für Vielfalt und Fairness?

Was versteht man unter einer „Filterblase“ in Bezug auf Empfehlungssysteme?

Warum ist Vielfalt in Empfehlungssystemen wichtig?

Welche Interessen müssen bei der Gestaltung von Empfehlungssystemen berücksichtigt werden?

Welche Herausforderung stellt sich bei der Implementierung von Fairness in Empfehlungssystemen?

Welche Rolle spielen technologische und gesellschaftliche Filterblasen in der Einschränkung von Vielfalt, und wie können Systeme gestaltet werden, um dem entgegenzuwirken?

Wie können algorithmische Verzerrungen in Empfehlungssystemen zu einer Benachteiligung bestimmter Nutzergruppen führen, und welche Maßnahmen könnten Abhilfe schaffen?

AKTIVIERUNG

Auswirkungen von Filterblasen auf Jugendliche

Ziel:

Die Teilnehmer:innen sollen sich aktiv mit den Auswirkungen von Filterblasen auf Jugendliche auseinandersetzen. Dabei wird untersucht, wie Filterblasen ihre Meinungsbildung, ihr Verhalten und ihre Wahrnehmung der Welt beeinflussen können.

Materialien:

- Moderationskärtchen und Stifte
- Pinnwand und Pins
- Ggf. Klebepunkte für Abstimmungen

Dauer:

Abhängig von der Anzahl der Gruppen, mindestens jedoch rund 20 Minuten

1. Einteilung der Teilnehmer:innen in Gruppen:

Jede Gruppe erhält mehrere Moderationskärtchen und Stifte.

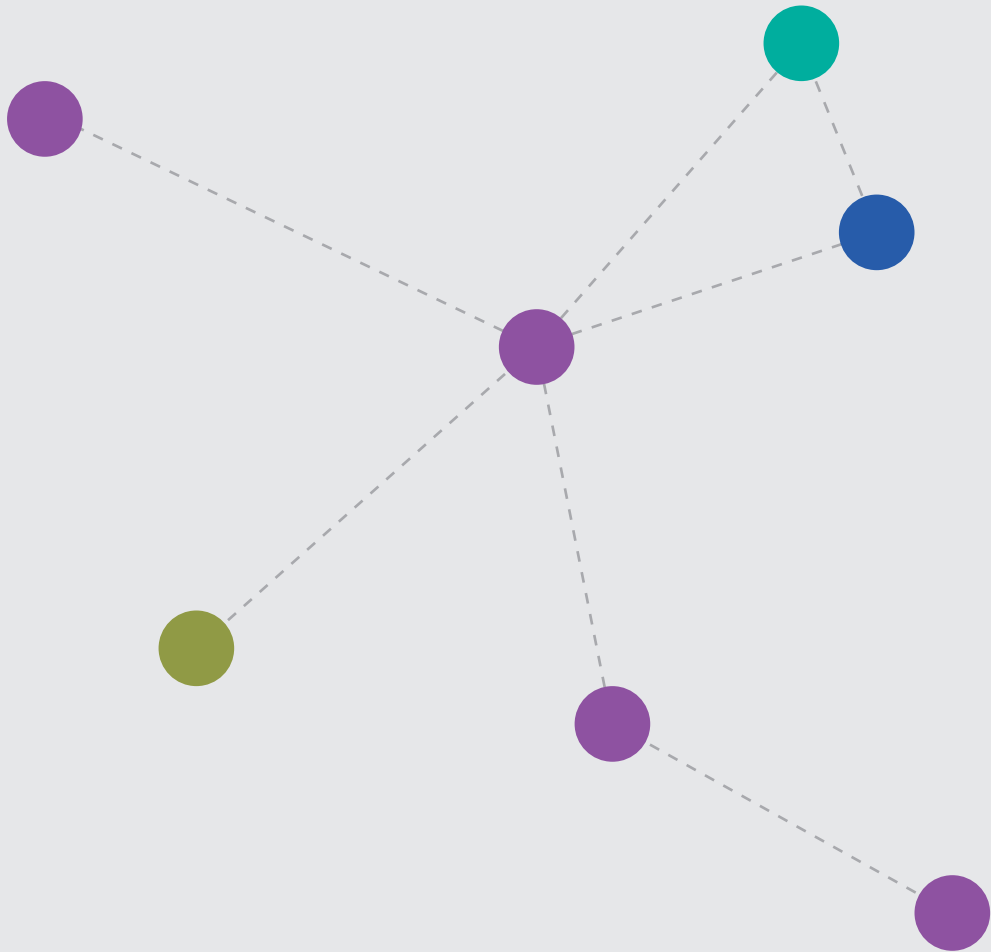
2. Jede Gruppe erhält die folgenden Aufgaben:

- ▶ verschiedene Bereiche zu identifizieren, in denen Jugendliche besonders von Filterblasen betroffen sein könnten (z. B. soziale Medien, Nachrichten, Musik).
- ▶ zu diskutieren, wie Filterblasen das Selbstbild und die Wahrnehmung anderer bei Jugendlichen beeinflussen könnten.
- ▶ eine kurze Geschichte oder ein Szenario zu entwickeln, das zeigt, wie ein:e Jugendliche:r durch eine Filterblase beeinflusst wird.
- ▶ Die Ergebnisse werden auf Moderationskärtchen festgehalten.

3. Präsentation und Diskussion:

Jede Gruppe präsentiert den anderen Teilnehmer:innen ihre Erkenntnisse. Anschließend diskutieren die Teilnehmer:innen gemeinsam über die vorgestellten Szenarien. Leitfragen können sein:

- ▶ Inwiefern können Filterblasen das Weltbild von Jugendlichen verzerren?
- ▶ Welche Gefahren bestehen, wenn Jugendliche nur eine begrenzte Sichtweise auf die Welt erhalten?
- ▶ Welche Strategien können Jugendliche anwenden, um aus ihren Filterblasen auszubrechen?



- KAPITELABSCHNITTE
- WISSENSFRAGEN
- FRAGE(N) ZUR DISKUSSION
- AKTIVIERUNG
- GRUPPENARBEIT

Kapitel 7

DAS VERZERRTE NETZ: VORURTEILE IM WEB

Voreingenommenheit ist allgegenwärtig – manchmal offensichtlich, oft jedoch verborgen. Sie entsteht häufig aus dem, was in unserem Wissen oder unseren Daten fehlt. Dieses Kapitel erläutert, was Voreingenommenheit ist, woher sie kommt, wie sie entsteht und bestehen bleibt. Zudem wird gezeigt, wie sie in Systeme einfließt und durch Algorithmen verstärkt werden kann. Zur Verdeutlichung des Problems wird das Web als größter von der Menschheit geschaffener Informationsspeicher herangezogen. Eine wesentliche Maßnahme zur Bekämpfung von Voreingenommenheit ist das Bewusstsein darüber.

Basis dieses Kapitels ist der Aufsatz „Bias and the Web“ von Ricardo Baeza-Yates und Leena Murgai, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024.
DOI: 10.1007/978-3-031-45304-5_28.

Einleitung

Unsere natürliche Neigung, bestimmte Meinungen oder Dinge anderen vorzuziehen, beeinflusst viele Bereiche unseres Lebens und führt zu sichtbaren und unsichtbaren Vorurteilen in allem, was wir tun. Diese Vorurteile sind nicht neu, sondern tief in unserer Kultur und Geschichte verankert. Durch das Internet und das Web können sich Vorurteile heute jedoch schneller und weiter verbreiten als jemals zuvor. Dadurch sind ihre Auswirkungen in den letzten Jahren zu einem wichtigen und oft kontrovers diskutierten Thema geworden.

Jede Strategie zur Bekämpfung von Vorurteilen beginnt mit Bewusstseinsbildung. Obwohl das allein das Problem nicht löst, ist es ein wichtiger erster Schritt. Fortschrittliche Gesellschaften erkennen soziale Vorurteile an und identifizieren schützenswerte Merkmale wie Geschlecht, ethnische Zugehörigkeit oder Religion, um unfaire und systembedingte Vorurteile zu bekämpfen. Manchmal werden sogar Maßnahmen ergriffen, um vergangene Ungerechtigkeiten auszugleichen, etwa durch Förderprogramme. Dies ist besonders wichtig, wenn es um die Entwicklung von Algorithmen geht, die einen wesentlichen Einfluss auf Menschen haben. Das Ziel ist es, das Bewusstsein für Voreingenommenheit im Web zu schärfen, um Web-Systeme zu schaffen, die allen Menschen zugutekommen und nicht nur wenigen Privilegierten.

Voreingenommenheit

Voreingenommenheit oder **Bias** entsteht oft aus dem Fehlen von Informationen. Um sie zu erkennen, müssen wir allerdings verstehen, was überhaupt fehlt. Diese Lücken in den Daten sind nicht unbedeutend, denn sie beeinflussen, wie Produkte, Dienstleistungen und Systeme gestaltet werden. Wenn Informationen über bestimmte Gruppen fehlen, werden ihre Bedürfnisse im Design nicht berücksichtigt. Ein Beispiel dafür ist die geschlechtsspezifische Datenlücke, die dazu führt, dass Frauen in vielen Bereichen, wie Gesundheitsfürsorge oder Sicherheitsvorkehrungen in Autos, unterrepräsentiert sind und dadurch Nachteile haben. Frauen haben daher häufiger Fehldiagnosen, sind bei Unfällen stärker gefährdet und haben mit vielen anderen Nachteilen zu kämpfen.

Eine weitere Herausforderung bei der Voreingenommenheit ist ihre Definition und Messung. Statistisch gesehen ist Bias eine systematische Abweichung von der Wahrheit, die durch ungenaue oder fehlerhafte Datenerfassung entsteht. Da der wahre Wert oft unbekannt ist, sind Daten immer eine verzerrte Darstellung der Realität. Beispielsweise sind Kategorien, die zur Klassifizierung von Menschen verwendet werden, oft voreingenommen, da sie auf subjektiven Entscheidungen basieren. Daten über eine Person sind daher immer unvollständig.

Voreingenommenheit beeinflusst unsere Wahrnehmung der Welt auf unsichtbare und tiefgreifende Weise. Studien zeigen, dass Suchergebnisse im Internet oft Geschlechterstereotype verstärken und die Wahrnehmung der Realität verändern können. Diese Verzerrungen im Internet tragen zur Aufrechterhaltung sozialer Hierarchien bei und behindern den gesellschaftlichen Fortschritt.

Das Messen von Verzerrungen ist schwierig, besonders, wenn es keine klaren Vergleichswerte gibt. Bei der Ressourcenverteilung könnte Gleichheit als Standard angenommen werden, doch bei einer sozialen Variable wie dem Einfluss einer Person auf eine andere Person ist der richtige Referenzwert oft unklar.

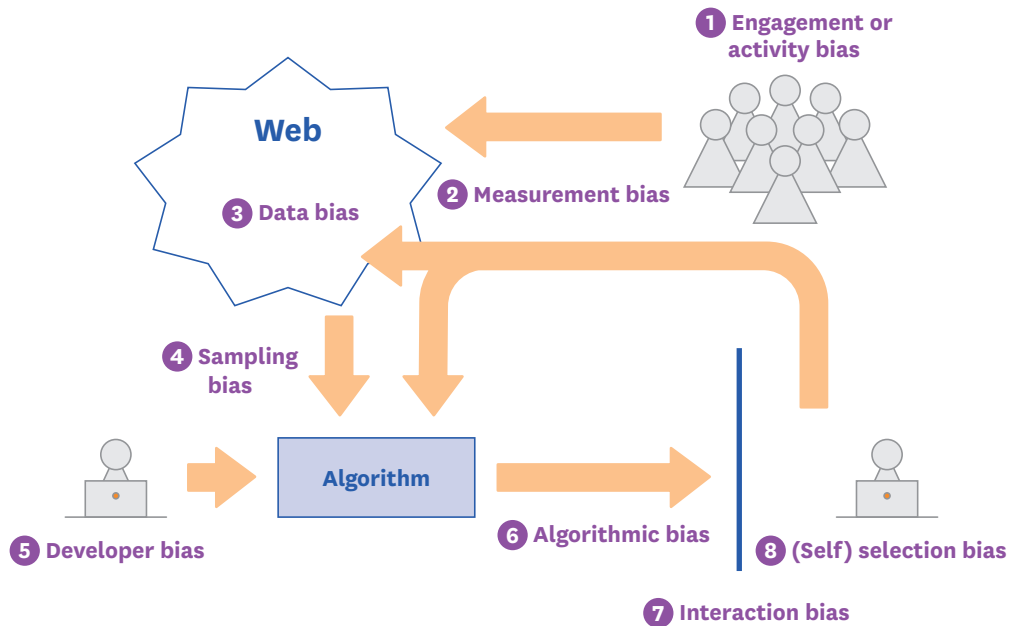
Kulturelle Verzerrungen entstehen durch gemeinsame Normen und Überzeugungen in Gemeinschaften. Diese Vorurteile führen dazu, dass wir unbewusst Meinungen über Menschen und Dinge haben, bevor wir ihnen begegnen. Bildung, die durch Vielfalt gefördert wird, ist das wichtigste Mittel, um Vorurteile zu bekämpfen.

Viele Algorithmen im Web basieren auf großen Datenmengen, die verzerrt sein können. Verzerrte Daten führen zu verzerrten Algorithmen, die diese Verzerrungen oft noch verstärken. Dies zeigt sich in

Der Begriff „Bias“ stammt aus dem Französischen („bias“) und gelangte später ins Englische. Ursprünglich wurde er verwendet, um eine schräge oder abweichende Richtung zu beschreiben, insbesondere in Bezug auf eine Kugel oder ein Objekt, das nicht geradeaus rollt. Im Laufe der Zeit wurde der Begriff metaphorisch verwendet, um eine Neigung oder Tendenz zu beschreiben, die nicht neutral ist und in bestimmten Kontexten eine Voreingenommenheit oder Verzerrung bedeutet. Heute wird „Bias“ häufig in wissenschaftlichen, psychologischen und statistischen Zusammenhängen verwendet, um systematische Fehler oder Verzerrungen zu beschreiben, die die Objektivität und Genauigkeit von Daten, Entscheidungen oder Meinungen beeinträchtigen können.

verschiedenen Formen von Bias wie Aktivitätsbias, Datenbias, Informationsbias und Voreingenommenheit der Entwickler:innen, die die Verzerrungen im Web weiter verstärken.

Systemische Verzerrungen im Web



Grafik 5: Der Teufelskreis der Voreingenommenheit im Web.

Quelle: Baeza-Yates und Murgai, DOI: 10.1007/978-3-031-45304-5_28

Im Folgenden werden einige in Grafik 5 angeführte Verzerrungen erläutert.

Aktivitätsbias: Die Weisheit der wenigen

Eine Studie aus dem Jahr 2011 zeigte, dass auf Twitter die obersten 0,05 Prozent der beliebtesten Personen fast 50 Prozent der gesamten Aufmerksamkeit auf sich ziehen. Das bedeutet, dass die Hälfte der Twitter-Nutzer:innen nur einer kleinen Gruppe berühmter Persönlichkeiten folgt. Diese Konzentration der Aufmerksamkeit wirft die Frage

auf, welcher Anteil der aktiven Webnutzer:innen für die Hälfte aller Inhalte verantwortlich ist. Das heißt, wir haben eine schweigende Mehrheit, die sich das Web nur ansieht, ohne etwas dazu beizutragen, was eine Form der Verzerrung der Nutzer:innen ist.

Eine weitere Studie von 2022 untersuchte das Engagement auf Twitter und fand heraus, dass etwa 90 Prozent der Menschen sich passiv verhalten. Interaktionen, die einen Klick erfordern (z. B. Likes oder Profilklicks), wurden von der Hälfte der Nutzer:innen durchgeführt, während Retweets, Antworten und Zitat-Tweets überwiegend von den aktivsten zehn Prozent der Nutzer:innen stammen. Auch hier zeigte sich, dass das oberste eine Prozent der Autor:innen 80 Prozent der Aufrufe generiert.

Datenbias

1 **Aktivitätsbias:**

Aktive Nutzer:innen sind überrepräsentiert.

2 **Messbias:** z. B. Unterschied zwischen der Realität und ihrer digitalen Abbildung.

3 **Datenbias:** Englischsprachige Texte sind z. B. im Web überrepräsentiert.

4 **Sampling/Auswahlbias:** Die Daten sind schlecht ausgewählt.

5 **Entwickler:innenbias:** Subjektive Sichten spiegeln sich im Code bzw. in der Datenauswahl wider.

6 **Algorithmischer Bias:** Das System ist z. B. systematisch unfair gegenüber bestimmten Gruppen.

7 **Interaktionsbias:** Die Art und Weise, wie Inhalte angezeigt werden, bestimmt, wie Nutzer:innen reagieren.

8 **Eigene Präferenzen** bestimmen, welche Daten erstellt werden.

Daten in sozialen Medien sind viel größer, viel voreingenommener und zweifellos im Durchschnitt von geringerer Qualität als beispielsweise Daten, die von Behörden, Universitäten und anderen Institutionen mit beispielsweise Peer-Review-Mechanismen veröffentlicht werden. Auch andere Bias haben einen Einfluss auf die Datenqualität, etwa die Unterrepräsentation von Frauen sowohl in den Inhalten von Wikipedia als auch unter den Content-Redakteur:innen. So sind nur neun Prozent der englischsprachigen Wikipedia-Redakteur:innen Frauen. Ein weiteres Beispiel ist, dass im Web englischsprachiger Content überrepräsentiert ist, wobei lediglich 15 Prozent der Weltbevölkerung Englisch als Mutter- oder Zweitsprache haben.

Diese Verzerrungen in den Daten sind schwer zu messen, was die Vorhersagegenauigkeit von Algorithmen, die maschinelles Lernen verwenden, beeinträchtigt. Erstens stellen Webdaten oft eine verzerrte Stichprobe der Bevölkerung dar, was zu erheblichen Fehlern in Studien führen kann, die auf sozialen Medien basieren. Zweitens lassen sich die Ergebnisse solcher Studien nicht ohne Weiteres auf die gesamte Bevölkerung übertragen. Drittens entstehen weitere Fehler durch verzerrte Datenstichproben, etwa aufgrund von Ungenauigkeiten bei der Auswahl der Daten, bei denen bestimmte Gruppen über- oder unterrepräsentiert sind, oder durch zu kleine Stichproben, die für verlässliche Analysen unzureichend sind.

Voreingenommenheit der Entwickler:innen

Die geringe Vielfalt unter den Entwickler:innen, insbesondere bei datengesteuerten Technologien, ist ein großes Problem. Etwa 80 Prozent der Softwareentwickler:innen sind Männer, was bedeutet, dass auf eine Frau vier Männer kommen. In spezialisierten Bereichen wie der KI-Forschung ist der Frauenanteil sogar noch niedriger: 2020 waren nur 16,1 Prozent des fix angestellten universitären Lehrpersonals in der Computerwissenschaft Frauen, und bei großen Tech-Firmen wie Facebook und Google lag der Anteil von KI-Forscherinnen bei nur 15 bzw. zehn Prozent.

Es ist wichtig, dass Frauen in der Entwicklung dieser Technologien stärker vertreten sind, um die Fehler zu vermeiden, die durch eine einseitige Perspektive entstehen. Studien zeigen, dass eine größere Vielfalt an Meinungen dazu beiträgt, Fehler schneller zu erkennen und zu beheben.

Absichtliche Verzerrungen im Web

Informationsbias

Auf Social-Media-Plattformen gibt es ein großes Problem mit von Bots generierten Inhalten, Falschinformationen und gefälschten Inhalten, die sich oft schneller verbreiten als echte Nachrichten. Eine Studie von 2020 zeigte, dass fast 45 Prozent von rund 200 Millionen Tweets zum Thema COVID-19 von Bots stammten. Diese Bots verbreiteten auch falsche Theorien, wie zum Beispiel die Behauptung, dass der Mobilfunkstandard 5G die Krankheit auslöste, was dazu führte, dass in England 5G-Türme angezündet wurden.

Ein weiteres Problem sind die Algorithmen von Social-Media-Plattformen, die darauf ausgelegt sind, die Nutzer:innen an die jeweilige Plattform zu fesseln. Dadurch können sie zur Verbreitung von Hassreden und Gewalt beitragen, besonders in Konfliktregionen. In Myanmar etwa half Facebook, Hassreden zu verstärken, was schließlich zu einem Völkermord an den Rohingya führte. Ähnliche Auswirkungen gab es in der Berichterstattung über die Gewalt in Tigray, Äthiopien, wo einer der tödlichsten Bürgerkriege der letzten Jahre wütet.

Fazit aus diesem Abschnitt:

Verzerrungen im Web entstehen oft durch eine Kombination von Faktoren, darunter eine unausgewogene Beteiligung von Nutzer:innen, fehlerhafte Datensätze und die einseitige Zusammensetzung von Entwicklerteams. Diese Verzerrungen beeinflussen, welche Informationen sichtbar werden und wie sie verarbeitet werden. Besonders problematisch ist, dass sie durch Algorithmen weiter verstärkt werden können, was negative Auswirkungen auf die Genauigkeit und Fairness der Online-Inhalte hat. Um diese Verzerrungen zu reduzieren, ist es wichtig, die Vielfalt unter den Entwickler:innen zu erhöhen und kritisch mit den zugrunde liegenden Datenquellen umzugehen.

Schlussfolgerungen

Das Problem der Voreingenommenheit ist vielschichtig und betrifft weit mehr als nur das Web. Was wir besprochen haben, ist nur ein kleiner Teil des Gesamtbilds. Unsere eigenen Vorurteile – sowohl als Einzelpersonen als auch als Gesellschaft – sind die Grundlage für viele dieser Verzerrungen. Diese Vorurteile wirken nicht nur im Web, sondern auch in der Nutzung mobiler Geräte und des Internets der Dinge.

Selbst wenn wir versuchen, unvoreingenommen zu sein, können unsere kulturellen und kognitiven Vorurteile uns immer noch beeinflussen. Der wichtigste Schritt, um Voreingenommenheit zu bekämpfen, ist, sich ihrer bewusst zu werden. Nur so können wir versuchen, sie zu reduzieren. Andernfalls riskieren wir, in einer Welt zu leben, in der Entscheidungen auf verzerrten Wahrnehmungen basieren und Vielfalt kaum eine Chance hat.

WISSENSFRAGEN

Systemische Verzerrungen im Web

Wie beeinflussen kulturelle und kognitive Verzerrungen die Gestaltung und Nutzung von Webinhalten, und welche Auswirkungen haben diese auf die Wahrnehmung der Realität durch Nutzer:innen?

Was versteht man unter „Aktivitätsbias“ im Kontext von sozialen Medien, und welche Auswirkungen hat dieser auf die Verteilung von Aufmerksamkeit?

Wie beeinflusst Datenbias die Qualität von Informationen, die in sozialen Medien und anderen Web-Plattformen verbreitet werden?

Warum ist die Vielfalt der Entwickler:innen wichtig für die Vermeidung von Verzerrungen in datengesteuerten Technologien?

FRAGEN ZUR DISKUSSION

Systemische Verzerrungen im Web

Inwiefern tragen soziale Medien und ihre Algorithmen zur Polarisierung der Gesellschaft bei, und was könnte getan werden, um dies zu mildern?

1

Mögliche Diskussion: Soziale Medien verstärken die Polarisierung, indem sie Inhalte priorisieren, die hohe Interaktionen generieren, was oft emotionale und kontroverse Themen sind. Dies schafft Filterblasen und Echokammern, in denen Nutzer:innen nur ähnliche Meinungen sehen. Maßnahmen könnten die Anpassung von Algorithmen umfassen, um diverse Perspektiven zu fördern und den Zugang zu ausgewogenen Informationen zu erleichtern.

Welche ethischen Überlegungen sollten bei der Entwicklung von Algorithmen und datengesteuerten Technologien berücksichtigt werden, um Vorurteile und Diskriminierung zu vermeiden?

2

Mögliche Diskussion: Entwickler:innen sollten sicherstellen, dass Trainingsdaten divers und repräsentativ sind, um Verzerrungen zu minimieren. Zudem ist Transparenz wichtig, damit die Öffentlichkeit nachvollziehen kann, wie Entscheidungen getroffen werden. Auch regelmäßige Audits und Feedback-Schleifen mit betroffenen Gruppen können helfen, Diskriminierung zu vermeiden und faire Systeme zu gewährleisten.

AKTIVIERUNG

Web-Detektive: Bias im Internet aufspüren

Ziel:

Die Jugendlichen sollen lernen, wie sie bewusster mit Vorurteilen und Verzerrungen im Internet umgehen, und kritisch hinterfragen können, welche Informationen im Web möglicherweise verzerrt oder einseitig dargestellt sind.

Materialien:

- Smartphones oder Notebooks (zum Recherchieren)
- Flipcharts/Whiteboards
- Stifte, Moderationskarten

Dauer:

60 Minuten

1. Einführung:

Start mit einer kurzen Erklärung darüber, was Bias im Internet bedeutet. Es wird wiederholt, dass Algorithmen und Daten im Web nicht neutral sind und bestimmte Meinungen oder Inhalte bevorzugen. Ein anschauliches Beispiel aus dem Alltag der Jugendlichen (z. B. wie Social Media oft ähnliche Inhalte zeigen) kann den Einstieg erleichtern.

2. Aufgabenstellung – „Finde den Bias!“:

Die Jugendlichen werden in Kleingruppen aufgeteilt. Jede Gruppe erhält die Aufgabe, im Internet nach einem Thema ihrer Wahl (z. B. Klima, Mode, Musik, Gaming) zu suchen. Sie sollen unterschiedliche Quellen verwenden (Google, Social Media, Wikipedia) und überlegen, ob sie Muster oder Verzerrungen entdecken. Gibt es bestimmte Meinungen, die häufiger auftauchen? Werden manche Themen oder Gruppen stärker repräsentiert als andere? (Siehe auch Beispiele)*

3. Diskussion in der Gruppe:

Jede Gruppe teilt ihre Entdeckungen und diskutiert, wo sie möglicherweise Vorurteile oder Verzerrungen gefunden hat. Dabei sollen sie besonders auf Werbung, Suchergebnisse und personalisierte Inhalte achten.

4. Präsentation der Ergebnisse:

Jede Gruppe stellt ihre Ergebnisse vor, indem sie erläutert, wo sie mögliche Verzerrungen entdeckt hat und welche Auswirkungen diese auf ihre Wahrnehmung haben könnten.

5. Reflexion und Schlussfolgerung:

Gemeinsam reflektieren alle, was sie gelernt haben. Wie können sie in Zukunft bewusster mit Informationen aus

dem Web umgehen? Gibt es Maßnahmen, die sie ergreifen können, um sich gegen Manipulationen und Vorurteile zu wappnen?

* **Beispiele für Verzerrungen:**

Die Jugendlichen suchen nach Informationen zum Thema „Klimawandel“ auf verschiedenen Plattformen, z. B. bei Youtube und auf Instagram.

▶ **Youtube-Suche:**

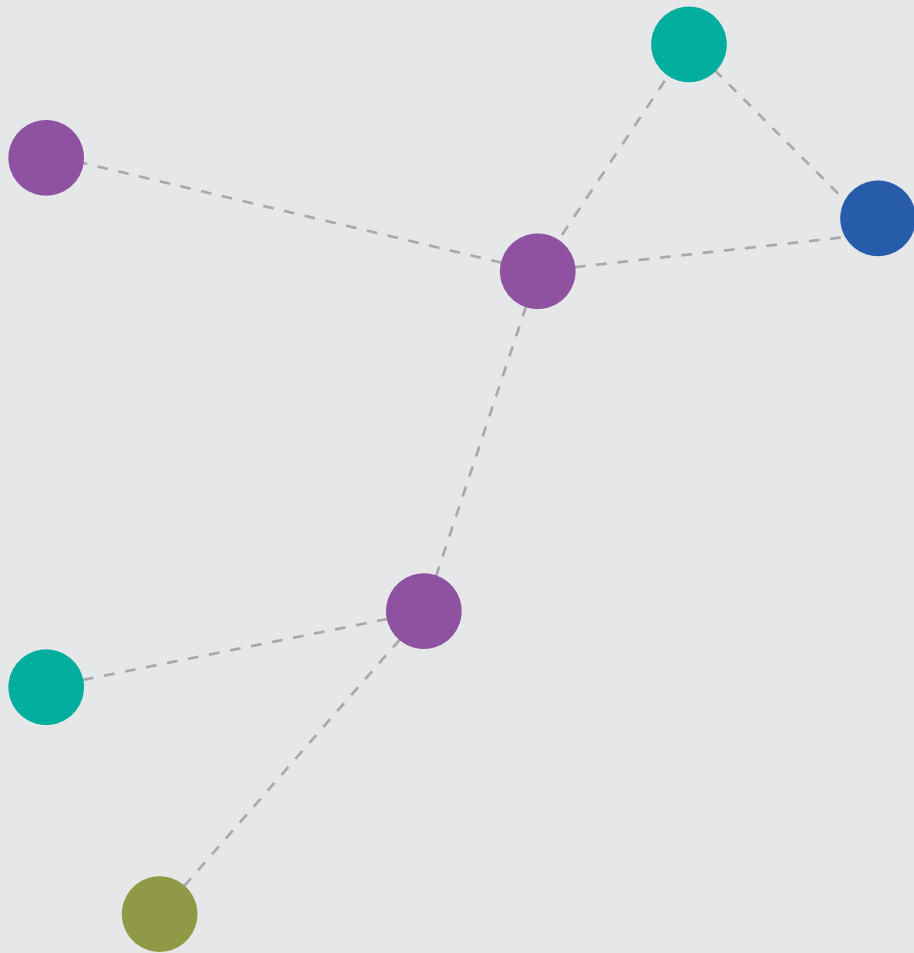
Eine Gruppe gibt den Begriff „Klimawandel“ in Youtube ein und schaut sich die ersten Ergebnisse an.

Mögliche Verzerrung: Youtube priorisiert eventuell pseudowissenschaftliche Videos.

▶ **Instagram-Suche:**

Die gleiche Gruppe sucht auf Instagram nach „Klimawandel“.

Mögliche Verzerrung: Instagram-Algorithmen könnten dazu neigen, Inhalte zu bevorzugen, die provokanter und emotionaler sind, weil solche Posts häufiger geteilt werden. Dadurch könnten verzerrte oder falsche Informationen bevorzugt verbreitet werden, was dazu führt, dass Nutzer:innen diese als glaubwürdiger wahrnehmen.



- KAPITELABSCHNITTE
- WISSENSFRAGEN
- FRAGE(N) ZUR DISKUSSION
- AKTIVIERUNG
- GRUPPENARBEIT

DIE NOTWENDIGKEIT DES SCHUTZES DER PRIVATSPHÄRE

In diesem Kapitel wird das Spannungsverhältnis zwischen Überwachung und Privatsphäre untersucht. Unternehmen sammeln in großem Ausmaß personenbezogene Daten und kommerzialisieren diese, indem sie zur gezielten Werbung und für andere kommerzielle Zwecke eingesetzt werden.

Gleichzeitig gibt es manchmal nachvollziehbare Gründe für Überwachung, etwa im Bereich Sicherheit oder Gesundheitsvorsorge. Es muss daher immer sorgfältig abgewogen und Missbrauch verhindert werden. Der Schutz der Privatsphäre ist entscheidend, um das Vertrauen der Nutzer:innen in die Technologie aufrechtzuerhalten und zu verhindern, dass Überwachung die Freiheit und Autonomie der Menschen einschränkt.

Basis dieses Kapitels ist der Aufsatz „The Threat of Surveillance and the Need for Privacy Protections“ von Martina Lindorfer, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024. DOI: 10.1007/978-3-031-45304-5_37.

Einleitung

Technologien sind heute ein wesentlicher Teil unseres Alltags. Wir nutzen sie vielfältig, von der Kommunikation über E-Mail, Chats, in sozialen Netzwerken bis hin zu Einkäufen und Bankgeschäften. Dabei geben wir oft bereitwillig persönliche Details preis, um online mit anderen in Kontakt zu treten oder bestimmte Geschäfte und Aufgaben zu erledigen. Apps und Websites sammeln dabei immer mehr private Informationen. Mit der wachsenden Beliebtheit von Geräten im *Internet der Dinge* (Internet of Things, IoT) teilen wir noch mehr über unser tägliches Leben und unsere Gewohnheiten, um den Komfort von *Smart Homes* zu genießen. Dazu gehören nicht nur Geräte wie *smarte* Toaster und Glühbirnen, sondern auch medizinische Geräte, Fitness-tracker sowie Smart-TVs und HbbTV-Geräte, die die Sehgewohnheiten der Nutzer:innen in Echtzeit überwachen und mit anderen Profilen verknüpfen können.

Unternehmen sammeln diese Daten aus mobilen Apps, dem Surfverhalten, den sozialen Medien und aus *smarten* Geräten, um detaillierte Profile von Nutzer:innen zu erstellen. Diese privaten Informationen und das daraus abgeleitete Verhalten sind zu einer Handelsware geworden: Große Technologiekonzerne und Datenhändler sammeln und verkaufen diese Daten nicht nur, um maßgeschneiderte Inhalte anzubieten, sondern auch für Marktforschung und gezielte Werbung. Dieses Phänomen wird als **Überwachungskapitalismus** (siehe Seite 16) bezeichnet. Oft sind diese Prozesse intransparent, weil schwer nachvollziehbare Tracking-Methoden verwendet werden. Zudem sind die gesammelten Daten nicht immer sicher: Sozialversicherungsnummern, Adressen, Finanzdaten und andere sensible Informationen von Milliarden von Nutzer:innen können in die Öffentlichkeit gelangen und kriminell genutzt werden.

Noch besorgniserregender ist, dass das sogenannte **Behavioral Targeting**, bei dem Nutzerprofile erstellt werden, nicht nur für Werbung, sondern auch zur Manipulation der öffentlichen und politischen Meinung eingesetzt wird. Ein bekanntes Beispiel dafür ist der Fall Cambridge Analytica, bei dem Daten von Facebook-Nutzer:innen verkauft wurden, um Wahlen über politische Kampagnen zu beeinflussen.

Überwachung versus Privatsphäre

Überwachung und Privatsphäre werden oft als Gegensätze gesehen, idealerweise sollten sie einander aber ergänzen. Überwachung kann notwendig sein, um die Sicherheit in der Gesellschaft zu gewährleisten, wie zum Beispiel bei der Strafverfolgung und der Bekämpfung von Verbrechen. Ein Beispiel ist der Einsatz von Überwachungskameras, deren Aufzeichnungen bei der Aufklärung von Straftaten wie den Bombenanschlägen beim Boston-Marathon 2013 oder dem Angriff auf das US-Kapitol 2021 geholfen haben. Wie immer ist aber das Maß der Überwachung ausschlaggebend, um auch Privatsphäre sicherzustellen.

Privatsphäre ist ein grundlegendes Menschenrecht. In der Europäischen Menschenrechtskonvention ist dieses Recht als Artikel 8 „Recht auf Achtung des Privat- und Familienlebens“ aufgenommen. Dieses Recht wird auch durch die **Datenschutzgrundverordnung (DSGVO)** innerhalb der Europäischen Union geschützt. Diese Gesetze stellen sicher, dass persönliche Daten geschützt werden und Personen die Kontrolle über ihre eigenen Daten behalten.

Jeder Mensch hat das Recht auf Privatsphäre und darauf, seine persönlichen Daten und Online-Aktivitäten zu schützen. Es gibt (Berufs-)Gruppen, die besonders gefährdet sind, überwacht zu werden. Dazu gehören politische Aktivist:innen, Journalist:innen, Whistleblower:innen, LGBTQIA+-Personen, ethnische und religiöse Minderheiten u. v. m.

Die **Datenschutzgrundverordnung (DSGVO)** ist ein Datenschutzgesetz, das 2018 von der Europäischen Union (EU) eingeführt wurde. Ziel der DSGVO ist es, Einzelpersonen mehr Kontrolle über ihre persönlichen Daten zu geben und sicherzustellen, dass Unternehmen diese Daten verantwortungsvoll behandeln. Ein zentrales Element der DSGVO ist die Einwilligung: Unternehmen müssen die klare und ausdrückliche Zustimmung der Betroffenen einholen, bevor sie deren Daten sammeln und verarbeiten. Außerdem haben die Betroffenen das Recht, zu erfahren, wie ihre Daten verwendet werden, ihre Einwilligung zu widerrufen und die Löschung ihrer Daten zu verlangen. Die DSGVO ist ein wichtiger Rechtsakt, der den Schutz persönlicher Daten sicherstellen soll. Ähnliche Gesetze gibt es auch in anderen Ländern, wie etwa den California Consumer Privacy Act (CCPA) in den USA, das brasilianische Datenschutzgesetz (LGPD), den kanadischen Personal Information Protection and Electronic Documents Act (PIPEDA) und das britische Datenschutzgesetz 2018 (DPA).

Nutzen von Überwachung

Schutz und Sicherheit: Überwachung kann helfen, kriminelle Aktivitäten, Terrorismus und andere Bedrohungen zu verhindern. Sicherheitskameras an öffentlichen Orten können Straftaten verhindern, ihre Aufnahmen können die Strafverfolgung unterstützen.

Gesundheit: Überwachung kann bei der Kontrolle von Krankheitsausbrüchen und der Überwachung von Impfraten helfen. In Notfällen, wie einer Pandemie, kann sie entscheidend dazu beitragen, die Ausbreitung von Krankheiten zu erkennen und zu stoppen.

Verkehrsmanagement: Die Überwachung des Verkehrsflusses kann helfen, Staus zu vermeiden und notwendige Verbesserungen der Infrastruktur zu planen.

Risiken von Überwachung

Eingriff in die Privatsphäre: Überwachung kann das Recht auf Privatsphäre verletzen und Menschen dazu bringen, sich überwacht und unwohl zu fühlen. Dies kann dazu führen, dass sie ihre Meinungen weniger frei äußern.

Machtmissbrauch und Diskriminierung: Regierungen und Unternehmen können Überwachung nutzen, um Menschen zu kontrollieren oder zu manipulieren, indem sie ihre Bewegungen und Aktivitäten überwachen, und um bestimmte Gruppen zu diskriminieren, basierend auf Faktoren wie Rasse, Religion oder politischer Überzeugung.

Mangelnde Transparenz: Ohne ausreichende Kontrolle kann Überwachung zu Missbrauch und Menschenrechtsverletzungen führen. Wenn Überwachung heimlich erfolgt, wissen die Betroffenen oft nicht einmal, dass sie überwacht werden, sie können sich daher nicht dagegen wehren.

Fazit aus diesem Abschnitt:

Überwachung kann gut begründet sein, sie kann aber auch missbraucht werden und negativen Folgen für den Einzelnen und die Gesellschaft insgesamt haben. Daher ist es wichtig, das Gleichgewicht zwischen Privatsphäre und Sicherheit sorgfältig abzuwägen und sicherzustellen, dass die Überwachung auf transparente, verantwortliche und ethische Weise erfolgt, damit auch insbesondere gefährdete Gruppen geschützt sind.

WISSENSFRAGEN

Überwachung versus Privatsphäre

Was bedeutet der Begriff „Überwachungskapitalismus“?

Welche Risiken und welche potenziellen Vorteile sind mit der Überwachung durch Unternehmen und Regierungen verbunden? Wie können diese beiden Aspekte in Einklang gebracht werden?

Was ist das Hauptziel der Datenschutzgrundverordnung (DSGVO)?

Wie kann die Einführung internationaler Datenschutzstandards dazu beitragen, die Privatsphäre global zu schützen, und welche Herausforderungen bestehen bei deren Umsetzung?

Welche Rolle spielen Technologien wie Anonymisierung und Verschlüsselung, um die Überwachung zu begrenzen und gleichzeitig die Privatsphäre der Nutzer:innen zu schützen?

Überwachung versus Privatsphäre

1
Wie kann die Balance zwischen notwendiger Überwachung zur Gewährleistung der öffentlichen Sicherheit und dem Schutz der Privatsphäre der Bürger:innen gewährleistet werden?

Mögliche Antworten könnten sich auf die Notwendigkeit transparenter und verantwortungsvoller Überwachungstechniken sowie auf gesetzliche Rahmenbedingungen beziehen, die den Missbrauch von Überwachungsdaten verhindern.

2
Welche Risiken birgt die weitverbreitete Sammlung und Verarbeitung von persönlichen Daten durch Unternehmen, und wie könnte dies die individuelle Freiheit und Autonomie beeinflussen?

Diskussionen könnten die Auswirkungen auf die Privatsphäre (welche?), die mögliche Manipulation von Meinungen durch gezielte Werbung und den Machtmissbrauch durch Unternehmen umfassen.

Maßnahmen zur Wahrung der Privatsphäre

Es gibt viele verschiedene Maßnahmen, die ergriffen werden können, um die Privatsphäre zu schützen. Diese Maßnahmen können von Einzelpersonen, App-Entwickler:innen, Unternehmen oder Regierungen umgesetzt werden. Sie umfassen technische, organisatorische und politische Ansätze. Es ist oft notwendig, eine Kombination dieser Maßnahmen anzuwenden, je nach Risiken für die Privatsphäre.

- **Datenschutz durch Design (Privacy by Design):** Datenschutz sollte von Anfang an ein integraler Bestandteil der Entwicklung eines Produkts oder Diensts sein.
- **Datenschutz durch Voreinstellung (Privacy by Default):** Standardmäßig sollten die datenschutzfreundlichsten Einstellungen aktiviert sein, sodass Nutzer:innen aktiv zustimmen müssen („Opt-in“ statt „Opt-out“), bevor ihre Daten erfasst werden.
- **Schutz und Verbesserung der Privatsphäre:** Die persönlichen Daten der Nutzer:innen sollten so behandelt werden, dass ihre Privatsphäre gewahrt bleibt, ohne dass die Funktionalität einer App oder eines Diensts darunter leidet.

Technische Maßnahmen

- **Zugangskontrollen:** Mechanismen, die den Zugang zu sensiblen Informationen nur autorisierten Personen erlauben, wie Passwortschutz und mehrstufige Authentifizierung.
- **Verschlüsselung:** Informationen werden so verschlüsselt, dass sie nur von autorisierten Parteien gelesen werden können. Dies schützt Daten vor unbefugtem Zugriff, sei es von Personen, die dasselbe öffentliche WLAN nutzen, oder von Internetanbietern und Behörden. Beispiele sind HTTPS für Websites, verschlüsselte E-Mails und Ende-zu-Ende-verschlüsselte Messenger wie Signal. Auch virtuelle private Netzwerke (VPNs) und **Tor** bieten Schutz durch Verschlüsselung.

Das **Tor-Netzwerk** ermöglicht anonyme Internetnutzung, indem es Daten über mehrere verschlüsselte Server leitet. Es kann z. B. für Surfen, Instant Messaging, E-Mail u. v. m. benutzt werden. Tor schützt Nutzer:innen vor einer Analyse des Datenverkehrs und wird daher auch für kriminelle Aktivitäten missbraucht.

- **Anonymisierung:** Entfernen von persönlich identifizierbaren Informationen aus Datensätzen, um die Privatsphäre zu schützen. Pseudonymisierung ist eine schwächere Form, bei der Daten nicht direkt einem:einer Nutzer:in zugeordnet werden können. Allerdings ist die Umsetzung oft schwierig.
- **Datenminimierung und Granularität:** Es sollte nur die minimal notwendige Menge an Daten gesammelt werden. Wo möglich, sollten unscharfe Daten verwendet werden, z. B. die Stadt statt der genauen Adresse.
- **Neue Technologien zur Verbesserung der Privatsphäre:** Technologien wie Differential Privacy stellen zusätzliche Schutzmaßnahmen dar.

Organisatorische Maßnahmen

- **Datenschutzrichtlinien:** Klare Erklärungen, wie eine Organisation mit Daten umgeht und welche Rechte Einzelpersonen haben. Sie schaffen Transparenz und Verantwortlichkeit.
- **Interne Schulungen:** Sensibilisierung innerhalb einer Organisation für Datenschutzrisiken und Best Practices, um unabsichtliche oder absichtliche Verstöße zu verhindern. Es ist wichtig, dass Mitarbeiter:innen verstehen, wie sie mit gesammelten Daten umgehen müssen. Berichte zeigen, dass Mitarbeiter:innen von Technologieunternehmen oft weitreichenden Zugang zu sensiblen Daten haben.
 - **Zugang zu Sicherheitskameras:** Es gibt Berichte darüber, dass Hersteller von Überwachungskameras ihren Mitarbeiter:innen uneingeschränkten Zugang zu den Aufnahmen gewähren. Dies wirft erhebliche Datenschutzbedenken auf, insbesondere, wenn Mitarbeiter:innen Live- oder aufgezeichnete Videos aus den Häusern der Nutzer:innen ohne deren Wissen oder Zustimmung ansehen können. Zum Beispiel wurde berichtet, dass Mitarbeiter:innen von Unternehmen wie Ring (im Besitz von Amazon) Zugriff auf Videoaufnahmen von Kund:innen hatten, was potenziell missbraucht werden könnte.
 - **Transkriptionen von Smart Speakern:** Ähnliche Datenschutzprobleme traten bei Smart Speakern auf, bei denen Mitarbeiter:innen und Auftragnehmer von Unternehmen wie Amazon Berichten zufolge Sprachaufzeichnungen von Geräten wie Alexa transkribierten, ohne dass die Nutzer:innen explizit darüber informiert wurden. Diese Aufzeichnungen können sensible und private Gespräche

enthalten, was Fragen hinsichtlich des Umgangs und des Zugriffs darauf aufwirft.

- **Datenschutz-Folgenabschätzungen:** Systematische Bewertungen der potenziellen Datenschutzrisiken neuer oder geänderter Prozesse und Technologien, um Risiken frühzeitig zu erkennen und anzugehen.

Politische Maßnahmen

- **Datenschutzgesetze:** Rechtliche Rahmenbedingungen, die festlegen, wie personenbezogene Daten erhoben, gespeichert und verwendet werden dürfen. Diese Gesetze geben Einzelpersonen Rechte, wie den Zugang zu ihren Daten, das Recht auf Berichtigung oder Löschung, und legen Strafen für Missbrauch fest.
- **Internationale Vereinbarungen:** Rahmenwerke wie die DSGVO in der EU und der EU-US-Datenschutzrahmen helfen, globale Datenschutzstandards festzulegen und den grenzüberschreitenden Datenschutz zu erleichtern.
- **Interessenvertretungen und Aktivismus:** Aufklärung über Datenschutzrisiken ist entscheidend, um Organisationen und Regierungen zur Rechenschaft zu ziehen und die öffentliche Politik zugunsten des Datenschutzes zu beeinflussen.

Fazit aus diesem Abschnitt:

Der Schutz der Privatsphäre erfordert ein umfassendes und vielschichtiges Vorgehen, das technische, organisatorische und politische Maßnahmen umfasst. Trotz technologischer Fortschritte und rechtlicher Rahmenbedingungen bleibt der Schutz persönlicher Daten eine komplexe Herausforderung, insbesondere angesichts der potenziellen Missbrauchsmöglichkeiten durch interne Mitarbeiter:innen und unzureichend gesicherte Systeme. Es ist entscheidend, dass Organisationen und Regierungen weiterhin strenge Datenschutzrichtlinien implementieren und durchsetzen, um das Vertrauen der Nutzer:innen zu erhalten und deren Privatsphäre zu schützen.

WISSENSFRAGEN

Maßnahmen zur Wahrung der Privatsphäre

Welche Maßnahmen werden in der DSGVO als „Privacy by Design“ und „Privacy by Default“ bezeichnet, und was bedeuten diese Begriffe?

Welche technischen Maßnahmen können ergriffen werden, um die Privatsphäre von Nutzer:innen zu schützen?

Welche Rolle spielen Datenschutzgesetze wie die DSGVO bei der Wahrung der Privatsphäre, und wie beeinflussen sie das Verhalten von Unternehmen?

Warum ist Mitarbeiterschulung in Bezug auf Datenschutzrisiken wichtig, und welche Probleme wurden in der Vergangenheit damit in Verbindung gebracht?

Inwiefern kann „Privacy by Design“ die Datensicherheit stärken, und welche Vorteile bietet es gegenüber nachträglichen Sicherheitsmaßnahmen?

AKTIVIERUNG

Datenzwiebel

Ziel:

Die Teilnehmer:innen sollen den Zusammenhang zwischen den von ihnen geteilten Daten und den daraus ableitbaren Informationen verstehen.

Optional:

Optional kann auf einem Flipchart parallel eine Datenzwiebel aufgezeichnet werden, und die Teilnehmer:innen überlegen selbstständig, welche Schlussfolgerungen die Bekanntgabe welcher Daten zulässt, notieren diese Überlegungen auf Moderationskärtchen und ordnen sie den einzelnen Schichten zu.

1. Besuch der Website:

<https://internethealthreport.org/2019/zeig-mir-meine-daten-und-ich-sag-dir-wer-ich-bin/?lang=de>

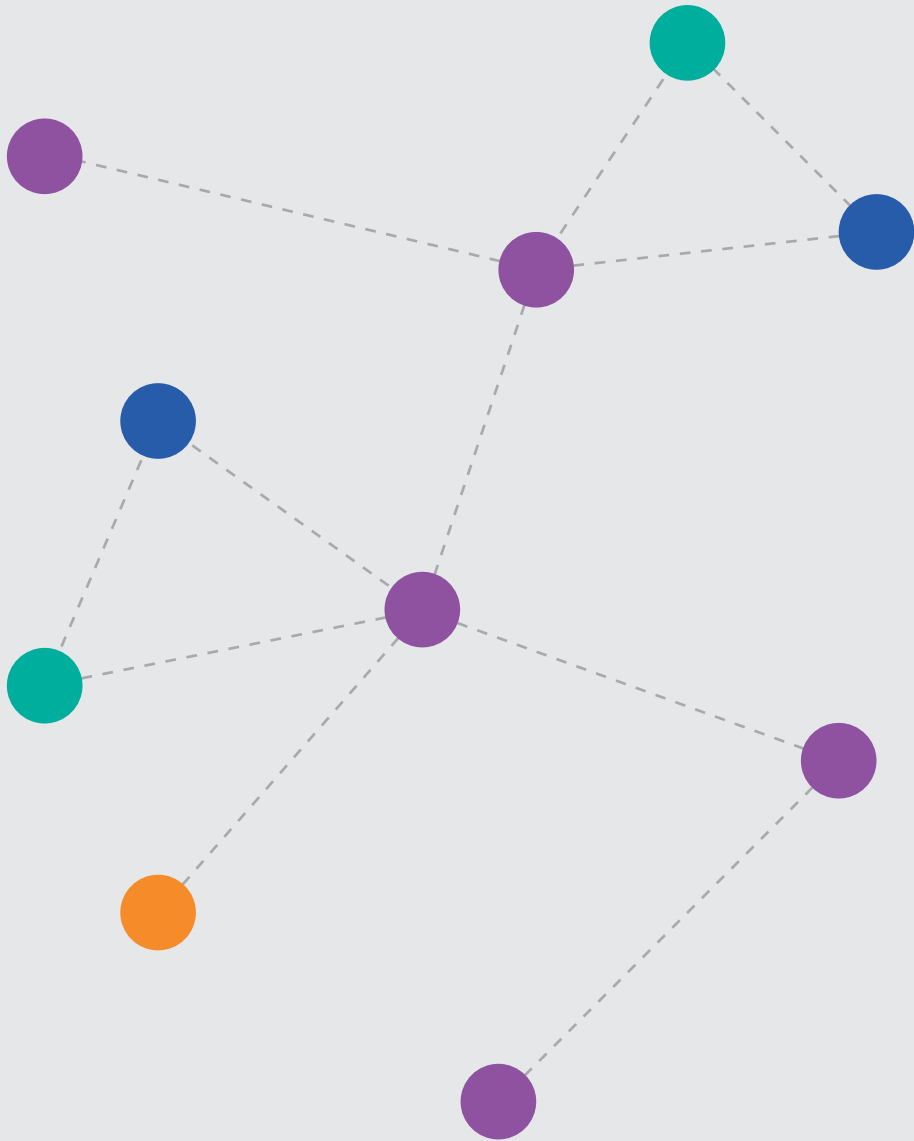
2. Gemeinsames Erörtern und Feststellen:

Die Teilnehmer:innen klicken durch die einzelnen Schichten der Datenzwiebel und besprechen ihr digitales Profil.

- ▶ Welche Daten werden von den Teilnehmer:innen alltäglich online geteilt?
- ▶ Was könnten diese Daten verraten?
- ▶ Was „denkt“ die Maschine von den Teilnehmer:innen, und welche Folgen könnte das haben?

3. Reflexion und Abschluss:

- ▶ Offene Diskussion über die Bedeutung von Datenschutz und Privatsphäre.
- ▶ Was sind die effektivsten Maßnahmen zum Schutz der Privatsphäre (technisch, organisatorisch, politisch)?



- KAPITELABSCHNITTE
- WISSENSFRAGEN
- FRAGE(N) ZUR DISKUSSION
- AKTIVIERUNG
- GRUPPENARBEIT

ALGORITHMISCHE MODERATION VON INHALTEN

Das Kapitel erklärt, wie die automatische Kontrolle von Inhalten – also die sogenannte algorithmische Inhaltsmoderation – im Web funktioniert und welche Herausforderungen damit verbunden sind. Algorithmen helfen dabei, Inhalte in sozialen Medien und anderen Onlineplattformen zu überprüfen, um sicherzustellen, dass diese Inhalte angemessen und legal sind. Dabei können jedoch Probleme auftreten, weil es schwierig ist, die genaue Bedeutung von Texten zu verstehen, insbesondere wenn es um Humor oder Ironie geht.

Es wird auch gezeigt, wie solche Algorithmen die Meinungsfreiheit einschränken und bestehende Vorurteile in der Gesellschaft verstärken können. Gleichzeitig könnten sie aber auch nützlich sein, um konstruktive Diskussionen zu fördern und die Stimmen von unterrepräsentierten Gruppen sichtbarer zu machen. Da sich die Technik permanent weiterentwickelt, bleiben diese Herausforderungen aktuell und erfordern weitere Forschung und Diskussion.

Basis dieses Kapitels ist der Aufsatz „On Algorithmic Content Moderation“ von Erich Prem und Brigitte Krenn, in Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024.

DOI: 10.1007/978-3-031-45304-5_30.

Einleitung

Algorithmische Inhaltsmoderation bedeutet, dass Computerprogramme genutzt werden, um Inhalte im Web, insbesondere auf Social-Media-Plattformen, zu kontrollieren. Diese Programme arbeiten oft mit menschlichen Moderator:innen zusammen, aber es gibt einen Trend zu immer mehr Automatisierung, weil die Menge an Inhalten ständig wächst und die Programme diese schneller verarbeiten können. Es gibt verschiedene Gründe, warum Inhalte überprüft und eventuell entfernt werden: zum Beispiel aus rechtlichen Gründen, weil sie gegen Urheberrechte verstoßen oder strafbare Inhalte enthalten, oder weil sie gegen die Regeln der Plattformen verstoßen.

Obwohl die Technologie hinter der Inhaltsmoderation schon länger existiert, wurde sie erst mit dem Aufstieg der großen sozialen Netzwerke wirklich wichtig. Diese Netzwerke ermöglichen es, dass falsche Informationen, sogenannte Fake News, sich schnell verbreiten oder dass Menschen nur noch Inhalte sehen, die ihre Meinung bestätigen. Solche Phänomene können durch die Nutzung von Online-Plattformen ohne direkte persönliche Kommunikation weiter verstärkt werden.

Die Inhaltsmoderation ist auch deshalb wichtig, weil sich der öffentliche Diskurs zunehmend in soziale Medien verlagert.

Gründe für die algorithmische Inhaltsmoderation

Es gibt mehrere Gründe, nutzergenerierte Inhalte (User Generated Content, UGC) zu prüfen.

1. Gesetzliche Anforderungen: In Österreich bestimmt das Medien-gesetz, dass Medienanbieter:innen wie zum Beispiel Online-Zeitungen für die Inhalte haften, die auf ihren Websites veröffentlicht werden. Social-Media-Plattformen wie Facebook, Twitter oder TikTok sind weltweite Phänomene, hier braucht es internationale Regulierungsregeln. Ein Beispiel dafür ist der European Digital Services Act (DSA) der EU, der versucht, Online-Dienste innerhalb der EU zu regulieren.

2. Redaktionelle Konzepte: Jede Plattform hat eigene Regeln für den Umgang miteinander, die in den Nutzungsbedingungen und der sogenannten Netiquette festgelegt sind. Automatische Systeme helfen dabei, Inhalte zu erkennen, die emotionale oder diskriminierende Diskussionen auslösen könnten.

3. Schutz geistigen Eigentums: Inhalte müssen auch dahin gehend überprüft werden, ob sie urheberrechtlich geschütztes Material enthalten. Dies zeigt sich in der Diskussion um Upload-Filter, die automatisch überprüfen, ob hochzuladende Information urheberrechtlich geschützt ist.

Inhaltsmoderation betrifft nicht nur Texte, sondern auch Bilder, Videos und Musik und ist sowohl für Nachrichtenwebsites als auch für Social-Media-Plattformen wichtig. Darin zeigt sich, dass algorithmische Moderation breit anwendbar ist.

Fazit aus diesem Abschnitt:

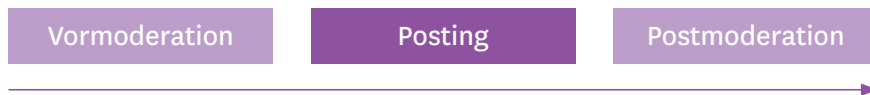
Algorithmische Inhaltsmoderation ist notwendig, um die riesige Menge an nutzergenerierten Inhalten (User Generated Content, UGC) in Online-Medien und sozialen Netzwerken zu verwalten. Die Hauptgründe dafür sind gesetzliche Anforderungen und redaktionelle Richtlinien, die sicherstellen, dass Inhalte gesetzeskonform und angemessen sind. Automatische Systeme helfen dabei, problematische Inhalte zu erkennen, von unerwünschten Kommentaren bis hin zu Urheberrechtsverletzungen, und sind unerlässlich für die Moderation von Texten, Bildern, Videos und Musik auf verschiedenen Plattformen.

Technische Ansätze und Herausforderungen der Inhaltsmoderation

Inhaltsmoderation nutzt verschiedene technische Methoden. Ein zentraler Bestandteil dieser Ansätze sind sogenannte **Klassifikatoren**, die auf maschinellem Lernen basieren. Die dabei verwendeten Klassifikatoren können je nach Bedarf unterschiedliche Aufgaben übernehmen.

Es gibt verschiedene Phasen der Moderation. Bei der Vormoderation werden nutzergenerierte Inhalte von Klassifikatoren automatisch in Kategorien eingeordnet. Dabei wird entschieden, ob die Inhalte auf einer Website veröffentlicht werden dürfen oder nicht. Diese Entscheidung basiert darauf, ob die Inhalte den gesetzlichen Vorgaben oder den Richtlinien der Plattform entsprechen. Inhalte, die gegen diese Regeln verstoßen, werden vor der Veröffentlichung blockiert.

In der Postmoderation werden bereits veröffentlichte Inhalte nachträglich überprüft. Hierbei unterstützen Klassifikatoren die Moderator:innen dabei, besonders relevante oder problematische Beiträge zu identifizieren. Diese nachträgliche Überprüfung ermöglicht es, die Diskussionskultur zu wahren und auf Inhalte zu reagieren, die möglicherweise gegen die Regeln verstoßen, aber in der Vormoderation nicht aufgefallen sind.



Grafik 6: Phasen der Moderation. Quelle: Bettina K. Lechner

Erkennung von Fake News

Ein wesentlicher Teil der Inhaltsmoderation ist die Erkennung und Bekämpfung von tendenziösen und gefälschten Nachrichten, sogenannten Fake News. Besonders in sozialen Medien ist dies wichtig, da sie im Gegensatz zu Foren in Online-Zeitungen weniger stark überwacht werden. Während Online-Zeitungen Artikel veröffentlichen, die von Journalist:innen verfasst und redigiert wurden, können auf Social-Media-Plattformen Nutzer:innen Inhalte ohne viel Kontrolle teilen. Das führt dazu, dass dort neben Meinungsfreiheit auch viel Platz für Propaganda und Fehlinformationen ist. Ein bekanntes Beispiel dafür

sind die US-Präsidentenwahlen 2016, bei denen Fake News eine große Rolle spielten.

Die Forschung konzentriert sich daher stark auf die automatisierte Erkennung solcher Nachrichten und auf die Überprüfung von Fakten. Es gibt viele Initiativen und Websites, die sich mit Faktenüberprüfung beschäftigen, wie <https://www.politifact.com/> in den USA, <https://www.eufactcheck.eu/> in Europa oder <https://www.poynter.org/>, ein internationales Netzwerk zur Faktenüberprüfung. Diese Projekte arbeiten daran, die Verbreitung von falschen Informationen zu verhindern und die Integrität von Nachrichten zu wahren.

Herausforderung durch Desinformation und Deepfakes

Die Erkennung von Desinformation bleibt eine große Herausforderung, weil sich ständig neue Themen und gezielte Propaganda entwickeln. Um damit Schritt zu halten, muss auch die Faktenprüfung immer wieder angepasst werden. Dank technologischer Fortschritte, besonders im Bereich des Deep Learning, gibt es heute raffinierte Methoden, um automatisch gefälschte Inhalte zu erstellen, sogenannte Deepfakes. Diese Deepfakes nutzen Künstliche Intelligenz, um täuschend echte, aber gefälschte Texte, Bilder, Audios oder Videos zu generieren. Beispiele für solche Technologien sind große Sprachmodelle wie OpenAIs GPT-Familie oder Googles PaLM, die auf Basis von Eingaben neue Texte erzeugen können, sowie Programme wie DALL-E, die aus Textbeschreibungen Bilder erstellen können. Diese Technologien machen es immer schwieriger, echte von gefälschten Inhalten zu unterscheiden.

Klassifikatoren sind Algorithmen, die verwendet werden, um Daten in verschiedene Kategorien oder Klassen einzuteilen. Im Zusammenhang mit der Inhaltsmoderation werden Klassifikatoren eingesetzt, um automatisch zu bestimmen, ob ein Beitrag oder Inhalt bestimmte Kriterien erfüllt und in eine bestimmte Kategorie eingeordnet werden sollte, beispielsweise „erlaubt“ oder „nicht erlaubt“.

Beispiel: Wenn ein Klassifikator trainiert wurde, um Spam von legitimen Nachrichten zu unterscheiden, lernt er, dass Spam-Nachrichten oft bestimmte Wörter oder Phrasen wie „Gratisangebot“ oder „Kaufen Sie jetzt“ enthalten. Wenn der Klassifikator dann eine neue Nachricht mit diesen Phrasen analysiert, wird er sie wahrscheinlich als Spam klassifizieren.

In der Inhaltsmoderation helfen Klassifikatoren dabei, große Mengen an nutzergenerierten Inhalten schnell und effizient zu prüfen und zu kategorisieren, bevor ein:e menschliche:r Moderator:in eingreift oder die Inhalte direkt veröffentlicht oder blockiert werden.

Fazit aus diesem Abschnitt:

Die algorithmische Inhaltsmoderation spielt eine entscheidende Rolle bei der Bewältigung der Flut von nutzergenerierten Inhalten im Web. Dabei stehen vor allem die Erkennung von problematischen Inhalten wie Fake News und die Kontrolle von Diskussionen im Mittelpunkt. Während die Moderationstechniken zunehmend auf KI-gestützte Systeme setzen, bleiben Herausforderungen wie die Erkennung von Desinformation und die Gefahr von Deepfakes bestehen. Diese Technologien entwickeln sich ständig weiter, was es erforderlich macht, dass auch die Methoden zur Moderation und Faktenprüfung kontinuierlich angepasst und verfeinert werden.

Gesellschaftliche Herausforderungen

In der heutigen digitalen Welt stehen wir vor zahlreichen Herausforderungen, wenn es um die Moderation von Online-Inhalten und den Schutz der Meinungsfreiheit geht.

Freiheit der Meinungsäußerung und inhaltliche Moderation

Ein wichtiger Aspekt der Moderation von Online-Inhalten ist die Erkennung und Entfernung illegaler Inhalte. Zum Beispiel ist es in Österreich strafbar, den Holocaust öffentlich zu leugnen. Solche Inhalte werden oft schnell entfernt, und es können rechtliche Maßnahmen folgen. Neben illegalen Inhalten gibt es auch solche, die gegen die Nutzungsbedingungen von sozialen Netzwerken verstoßen.

Schwieriger wird es, wenn es um die Moderation von Inhalten geht, die zwar nicht illegal sind, aber gesellschaftlich als unerwünscht oder gefährlich gelten. Diese Inhalte sind schwerer zu definieren, da sie von individuellen oder gesellschaftlichen Werten abhängen. Es ist in vielen Ländern nicht grundsätzlich illegal, zu lügen oder Schimpfwörter zu verwenden, aber solche Inhalte können dennoch als problematisch angesehen werden, da sie bestimmte Gruppen stärker betreffen oder die Verbreitung gefährlicher Informationen begünstigen könnten. Ein großes Problem dabei ist, dass der Begriff „gefährliche Inhalte“

WISSENSFRAGEN

Technische Ansätze und Herausforderungen der Inhaltsmoderation

Was ist das Hauptziel der algorithmischen Inhaltsmoderation auf Social-Media-Plattformen?

Welche Herausforderung stellt die ständige Weiterentwicklung von Technologien wie Deep Learning für die Inhaltsmoderation dar?

Welche technischen Maßnahmen können entwickelt werden, um zwischen legaler Meinungsäußerung und gefährlichen Inhalten zu unterscheiden, ohne die Meinungsfreiheit einzuschränken?

Wie beeinflussen kulturelle und sprachliche Unterschiede die Genauigkeit und Fairness der algorithmischen Inhaltsmoderation auf globalen Plattformen?

Welche Rolle spielen Initiativen zur Faktenüberprüfung bei der Inhaltsmoderation? Welche sind bekannt?

FRAGEN ZUR DISKUSSION

Technische Ansätze und Herausforderungen der Inhaltsmoderation

1
Welche ethischen Bedenken könnten bei der Verwendung von Algorithmen zur Moderation von Online-Inhalten entstehen, insbesondere in Bezug auf die Meinungsfreiheit und die Gefahr von Zensur?

Diskussion darüber, ob und wie algorithmische Systeme möglicherweise ungewollt bestimmte Meinungen unterdrücken könnten und wie dies mit dem Recht auf freie Meinungsäußerung in Einklang gebracht werden kann.

2
Wie könnten die Risiken von Deepfakes und anderen automatisiert erstellten Desinformationen durch algorithmische Moderation besser kontrolliert werden, ohne dabei die Meinungsvielfalt einzuschränken?

Die Teilnehmer:innen überlegen, welche Maßnahmen und Technologien eingeführt werden könnten, um Desinformation effektiv zu bekämpfen, während gleichzeitig sichergestellt wird, dass die Meinungsvielfalt auf Social-Media-Plattformen erhalten bleibt.

oft ungenau ist und es unklar bleibt, wer genau durch diese Inhalte geschädigt wird oder wer die Autorität hat, solche Inhalte zu identifizieren und zu entfernen. Dies führt häufig zu Bedenken, dass die Entfernung solcher Inhalte Zensur darstellen könnte.

Herausforderungen in Bezug auf Bedeutung und Sprache

Eine weitere große Herausforderung in der Moderation von Inhalten ist das Verständnis dessen, was Nutzer:innen wirklich ausdrücken wollen. Menschliche Kommunikation ist komplex, und es ist leicht, missverstanden zu werden. Der Begriff der „Bedeutung“ ist zudem ein schwieriges philosophisches Konzept, das seit Jahrhunderten diskutiert wird. Bestimmte Aspekte der natürlichen Sprache, wie Humor, Ironie oder Spott, sind für Maschinen besonders schwer zu erkennen, und selbst Menschen tun sich damit manchmal schwer. Bei diesen Algorithmen besteht die Gefahr, den Kontext einer Äußerung nicht richtig erfassen zu können. Sie haben Schwierigkeiten, wichtige Nuancen zu erkennen.

Ein spezielles Problem bei der algorithmischen Moderation besteht darin, dass viele Algorithmen nur einzelne Beiträge analysieren und es ihnen schwerfällt, längeren Diskussionen oder Dialogen zu folgen. Oft erfordert die wahre Bedeutung eines Beitrags jedoch die Berücksichtigung des gesamten Gesprächsverlaufs. Schätzungen zufolge erreichen kommerzielle Moderationstools nur eine Genauigkeit von etwa 70 bis 80 Prozent. Ein Beispiel für eine solche Fehleinschätzung ereignete sich 2020, als Facebook Beiträge mit dem Hashtag #EndSARS in Nigeria entfernte. Diese Beiträge sollten auf Polizeigewalt aufmerksam machen, wurden aber fälschlicherweise als Fehlinformationen über COVID-19 eingestuft.

Politische Herausforderungen: Macht der Kontrolle und Einfluss auf die Gesellschaft

Die Moderation von Online-Inhalten ist nicht nur eine technische, sondern auch eine tief politisch geprägte Aufgabe, da sie bedeutende Fragen der Macht und Kontrolle aufwirft. Durch die Kontrolle darüber, welche Inhalte sichtbar bleiben und welche gelöscht werden, haben die Moderator:innen erheblichen Einfluss darauf, welche Themen in der öffentlichen Diskussion präsent sind und welche möglicherweise verschwinden. Diese Kontrolle kann dazu führen, dass bestimmte

Gruppen in der Gesellschaft nicht gehört werden, indem ihre Beiträge nicht mehr sichtbar sind. Besonders kritisch wird es, wenn die Moderation von Inhalten zunehmend durch Algorithmen übernommen wird. Diese Algorithmen treffen Entscheidungen darüber, welche Inhalte entfernt werden, basierend auf den Vorgaben, die in ihrem Programmcode festgelegt sind.

Ein großes Problem dabei ist, dass solche algorithmischen Systeme bestehende Vorurteile und Diskriminierungen in der Gesellschaft widerspiegeln und sogar verstärken können. Wenn ein Algorithmus z. B. mit voreingenommenen Daten trainiert wurde, besteht die Gefahr, dass er bestimmte Inhalte aus marginalisierten Gemeinschaften – also gesellschaftlich benachteiligten oder unterrepräsentierten Gruppen – häufiger als unangemessen einstuft und löscht, während ähnliche Inhalte aus anderen Gemeinschaften möglicherweise unberührt bleiben.

Zusätzlich gibt es das Problem der fehlenden Transparenz. Die Algorithmen, die zur Moderation von Inhalten verwendet werden, sind oft urheberrechtlich geschützt, was bedeutet, dass die Öffentlichkeit keinen Einblick in den zugrunde liegenden Code oder die Kriterien hat, nach denen entschieden wird, welche Inhalte gekennzeichnet oder entfernt werden. Diese fehlende Transparenz erschwert es den Nutzer:innen, zu verstehen, warum ihre Beiträge gelöscht wurden, was zu Frustration und einem Gefühl der Ungerechtigkeit führen kann.

Fazit aus diesem Abschnitt:

Die Moderation von Online-Inhalten stellt die Gesellschaft vor große Herausforderungen. Einerseits ist sie notwendig, um illegale oder gegen Richtlinien verstoßende Inhalte zu entfernen und die Nutzer:innen vor falschen und irreführenden Informationen zu schützen. Andererseits birgt sie das Risiko, die Meinungsfreiheit einzuschränken, insbesondere, wenn es um schwer definierbare, aber als unerwünscht betrachtete Inhalte geht. Die Schwierigkeit, den Kontext und die Nuancen menschlicher Sprache richtig zu erfassen, verstärkt diese Problematik, besonders bei der algorithmischen Moderation. Insgesamt zeigt sich, dass ein sorgfältiger Umgang mit der Moderation erforderlich ist, um den Spagat zwischen Schutz der Gesellschaft und Wahrung der Meinungsfreiheit zu bewältigen.

WISSENSFRAGEN

Gesellschaftliche Herausforderungen

Warum ist die Moderation von Online-Inhalten notwendig?

Wie können bestehende Bildungssysteme angepasst werden, um Menschen auf die Herausforderungen der digitalen Gesellschaft vorzubereiten?

Welche Herausforderungen gibt es bei der algorithmischen Inhaltsmoderation?

Warum ist es für algorithmische Moderationstools schwierig, die Bedeutung eines Beitrags im Kontext eines gesamten Gesprächsverlaufs richtig einzuschätzen, und welche Konsequenzen können daraus entstehen?

Wie beeinflusst generative KI die Herausforderungen in der Moderation von Inhalten?

Gesellschaftliche Herausforderungen

1
Sollte die Verantwortung für die Moderation von Inhalten mehr bei Menschen oder bei Algorithmen liegen?

Mögliche Antworten: Einige könnten argumentieren, dass Algorithmen unverzichtbar sind, um die Menge an Inhalten zu bewältigen, während andere darauf hinweisen, dass menschliche Moderator:innen besser in der Lage sind, Kontext und Nuancen zu verstehen, was zu gerechteren Entscheidungen führt. Eine Kombination aus beidem könnte eine ausgewogene Lösung sein.

2
Wie kann eine Balance zwischen der Notwendigkeit, gefährliche Inhalte zu entfernen, und der Wahrung der Meinungsfreiheit erreicht werden?

Mögliche Antworten: Dies könnte durch klare, transparente Richtlinien und den Einsatz von Algorithmen in Verbindung mit menschlicher Überprüfung erreicht werden. Eine Möglichkeit wäre, den Betroffenen die Möglichkeit zu geben, Entscheidungen anzufechten, um sicherzustellen, dass ihre Meinungsfreiheit nicht ungerechtfertigt eingeschränkt wird.

Schlussfolgerungen

Die Moderation von Inhalten ist keine neue Praxis; sie existiert, seit Menschen in schriftlicher Form kommunizieren. Früher wurde sie ausschließlich von Menschen durchgeführt, aber mit der algorithmischen Inhaltsmoderation gibt es nun eine neue, technologische Dimension. Es gibt verschiedene Gründe für die Moderation von Inhalten:

- **Rechtliche Gründe** spielen eine Rolle, wie die Verpflichtung der Plattformbetreiber, illegale Inhalte zu entfernen.
- **Praktische Gründe** umfassen das Filtern von Inhalten, um den Nutzer:innen die relevantesten Informationen anzuzeigen.
- Zudem gibt es **soziale und politische Gründe**, wie die Durchsetzung von Benutzerrichtlinien oder das Entfernen von Inhalten, die als unangemessen oder gefährlich angesehen werden.

Die Moderation kann auf verschiedene Weisen erfolgen. Inhalte können markiert oder hervorgehoben werden, um die Nutzer:innen auf mögliche Probleme aufmerksam zu machen. Sie können auch durch Herabstufen im Ranking weniger sichtbar gemacht werden, oder sie werden vollständig gelöscht. In einigen Fällen wird die Veröffentlichung von Inhalten verzögert, oder sie werden an zuständige Behörden gemeldet. Es gibt große Unterschiede darin, inwieweit die Nutzer:innen über diese Prozesse informiert werden, und viele soziale Netzwerke bieten nur begrenzten Einblick in ihre Moderationspraktiken.

Gesetze schreiben zwar nicht direkt vor, dass Algorithmen zur Inhaltsmoderation eingesetzt werden müssen, aber kurze Fristen und hohe Strafen machen den Einsatz von Algorithmen oft zur einzigen praktikablen Lösung, um die große Menge an Inhalten zu bewältigen.

Eine der größten Herausforderungen bei der Moderation von Inhalten ist die automatische Interpretation menschlicher Sprache, besonders, wenn es zum Beispiel um Satire oder Humor geht, wobei Kontext und feine Nuancen den Unterschied machen. Außerdem sind die technologischen Möglichkeiten und Trainingsdaten für weniger verbreitete Sprachen oft weniger entwickelt als für Englisch. Zu den Risiken gehören die Einschränkung der Meinungsfreiheit, mögliche Voreingenommenheiten und eine ungerechte Behandlung von Inhalten. Politische Einflussnahme ist ebenfalls ein Problem, wenn abweichende Meinungen unterdrückt oder Inhalte tendenziös moderiert werden.

GRUPPENARBEIT

Faktencheck-Challenge

Ziel:

Die Teilnehmer:innen sollen lernen, wie man Informationen im Internet auf ihre Echtheit überprüft. Sie recherchieren verschiedene Websites zum Faktencheck und finden Beispiele von Nachrichten oder Posts, die entweder korrekt oder falsch sind. Diese Beispiele sollen sie analysieren und diskutieren.

Materialien:

- Arbeitsblätter mit den beschriebenen Situationen und Fragen
- Flipcharts oder Whiteboards für die Präsentationen
- Stifte und Papier für Notizen

Dauer:

Abhängig von der Anzahl der Gruppen, mindestens jedoch 60 Minuten

1. Einführung in das Thema Faktencheck: Was ist das, und warum ist es in der heutigen digitalen Welt so wichtig? Dabei können bekannte Faktencheck-Websites wie Mimikama, Correctiv, Snopes, EUfactcheck oder Politifact vorgestellt werden, die sich auf das Überprüfen von Informationen spezialisiert haben.

2. Internet-Recherche in Gruppen oder einzeln: Ziel ist es, mindestens zwei Beispiele für Informationen zu finden, die als Fake News entlarvt wurden, und zwei Beispiele für Informationen, die als wahr bestätigt wurden. Dabei können die empfohlenen Faktencheck-Websites genutzt oder eigene Quellen gefunden werden.

3. Kurze Präsentation der Ergebnisse:

Dabei sollte erklärt werden:

- ▶ Was die Nachricht oder der Post behauptete.
- ▶ Wie und warum diese Information als falsch oder wahr eingestuft wurde.
- ▶ Welche Auswirkungen diese Information hatte oder hätte haben können.

4. Diskussion über Erfahrungen während der Recherche:

Mögliche Fragen könnten sein:

- ▶ Gab es Überraschungen bei den gefundenen Beispielen?
- ▶ Welche Gefühle kamen auf, als klar wurde, dass eine Information falsch war?
- ▶ Wie kann man in Zukunft sicherstellen, dass die geteilten oder konsumierten Informationen zuverlässig sind?

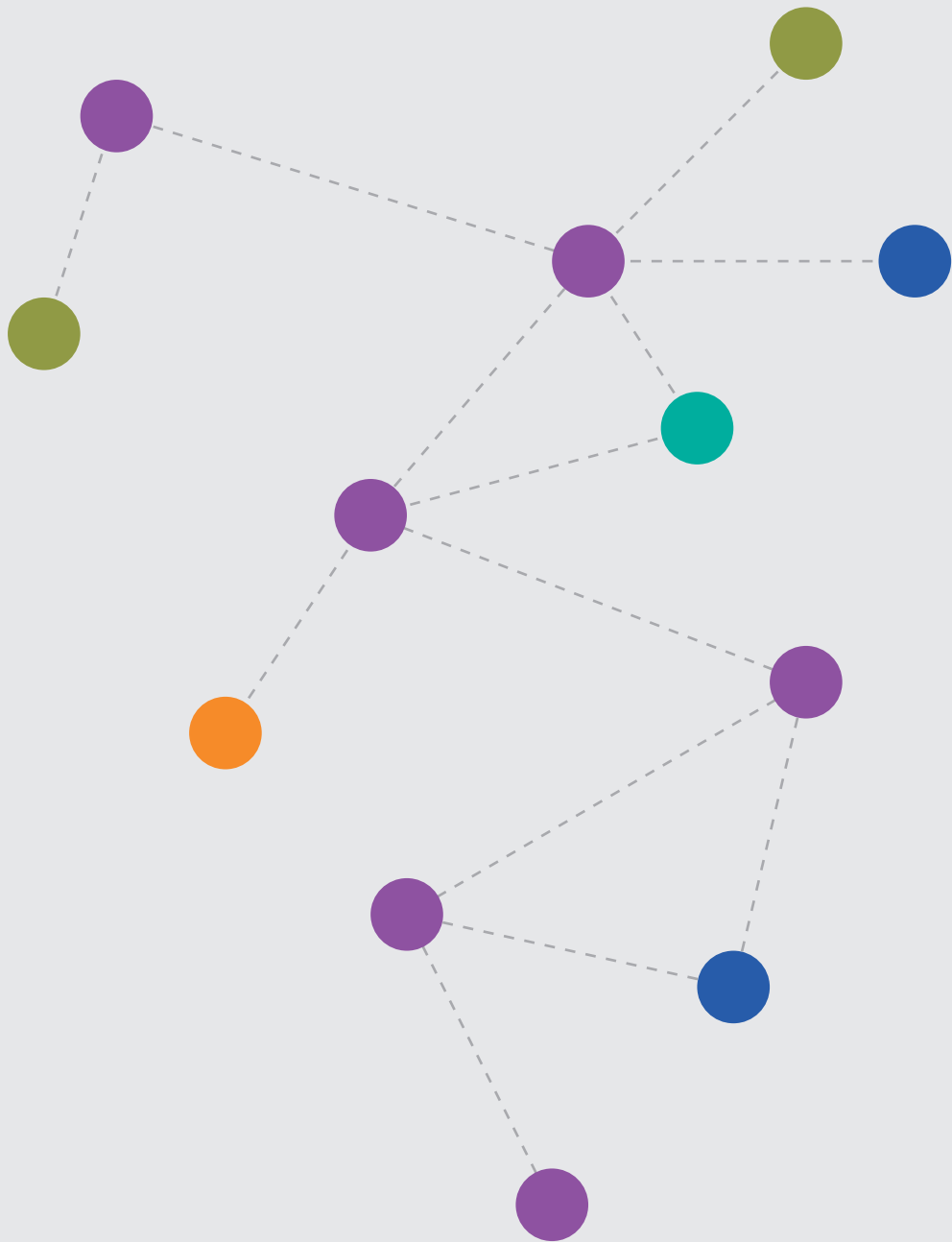
5. Abschluss: Zusammenfassung der wichtigsten Punkte und Betonung der Bedeutung der Überprüfung von Informationen vor dem Teilen. Es können noch Tipps gegeben werden, wie im Alltag Fakten überprüft werden können, z. B. durch den Vergleich mehrerer Quellen oder die Nutzung bekannter Faktencheck-Websites.

Trotz dieser Herausforderungen kann algorithmische Moderation auch zum Nutzen der Gesellschaft eingesetzt werden. Beispiele dafür sind Maßnahmen zur Deeskalation, die Förderung sachlicher und konstruktiver Online-Debatten und die Unterstützung von Stimmen, die in der Gesellschaft oft unterrepräsentiert sind.

Die Entwicklung von generativer Künstlicher Intelligenz und großen Sprachmodellen wird als ein Wendepunkt in der textbasierten KI betrachtet. Diese Technologien werden viele neue Anwendungen und Systeme ermöglichen, bringen aber auch neue Herausforderungen mit sich. Ähnliche Probleme wie bei der Moderation von Inhalten treten auch bei diesen neuen Technologien auf. Fragen der Diskriminierung, Voreingenommenheit und mangelnder Transparenz betreffen auch die großen Sprachmodelle und erfordern weitere Forschung, öffentliche Debatten und politische Maßnahmen.

Fazit aus diesem Abschnitt:

Die Moderation von Inhalten ist notwendig, bringt aber komplexe Herausforderungen mit sich, besonders in Bezug auf Meinungsfreiheit, Voreingenommenheit und Transparenz. Algorithmische Systeme sind oft unverzichtbar, aber ihre Grenzen bei der Erkennung von Kontext und Nuancen bergen Risiken. Mit der Entwicklung generativer KI verschärfen sich diese Probleme, weshalb weitere Forschungen, gesellschaftliche Diskussionen und klare politische Regelungen dringend erforderlich sind.



● KAPITELABSCHNITTE ● WISSENSFRAGEN ● FRAGE(N) ZUR DISKUSSION
● AKTIVIERUNG ● GRUPPENARBEIT

Kapitel 10

DER MENSCH IM MITTELPUNKT DER SYSTEMENTWICKLUNG

Trotz der zunehmenden Automatisierung steht der Mensch noch im Mittelpunkt der Entwicklung von Softwaresystemen. Dieses Kapitel untersucht die Rolle des Menschen bei der Entwicklung dieser Systeme. Entwickler:innen und Nutzer:innen sind Teil von Gemeinschaften. Gezeigt wird, wie sich diese Tatsache auf die Entwicklung von Softwaresystemen auswirkt und wie die Nutzung von Software deren zukünftige Entwicklung beeinflusst.

Basis dieses Kapitels ist der Aufsatz „Humans in the Loop: People at the Heart of Systems Development“ von Helen Sharp, in Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024. DOI: 10.1007/978-3-031-45304-5_23.

Einleitung

Heutzutage ist Software für fast alles, was wir tun, unerlässlich. Apps machen es möglich, mit unseren Freund:innen und Bekannten in Kontakt zu bleiben, Rechnungen zu bezahlen, unsere Fitness zu überwachen etc. Große Projekte wie das James-Webb-Teleskop erkunden den Weltraum, Software überwacht den Verkehr in Städten oder beobachtet bedrohte Tierarten. Privatleben, Arbeit und Schule sind ohne Software heute kaum mehr möglich.

Die Softwareentwicklung ist entscheidend für die Gestaltung und Funktion dieser digitalen Werkzeuge. Sie beeinflusst, wie sie funktionieren, wie sicher sie sind und wie sie Nutzer:innen bei ihren Aufgaben unterstützen.

Entwickler:innen von Softwaresystemen

Wenn man an Entwickler:innen von Software denkt, stellt man sich meist Fachleute vor. Es gibt jedoch immer mehr Menschen, die zwar keine Spezialist:innen sind und trotzdem direkt an der Entwicklung und Implementierung von Softwaresystemen beteiligt sind. Egal, ob Spezialist:in oder nicht, die Menschen, die Software erstellen, arbeiten nicht allein. Sie sind Teil einer Gemeinschaft von Designer:innen, Entwickler:innen, Benutzer:innen und anderen Beteiligten, die zur Erstellung beitragen. Diese Gemeinschaften beeinflussen die Systementwicklung.

Softwareentwickler:innen

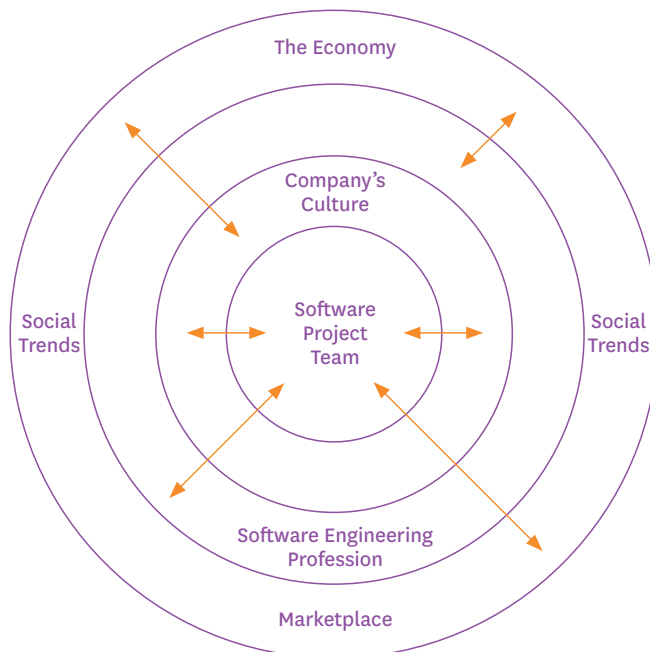
Die moderne Softwareentwicklung ist sehr komplex, basiert auf Programmiersprachen und stützt sich stark auf vorhandene Module oder Bausteine (Komponenten) wie Sprachbibliotheken und Schnittstellenmodule, die wiederum von verschiedenen Gruppen erstellt werden. Software muss in ihre technische Umgebung eingebettet werden und ist von den Eigenschaften des Geräts und der Umgebung, in der sie arbeitet, abhängig. Softwareentwickler:innen und ihre Arbeit sind daher stark von anderen abhängig.

Der Gemeinschaftsaspekt der Softwareentwicklung wird oft übersehen. Softwareentwickler:innen arbeiten in Teams und bilden enge Gemeinschaften innerhalb von Firmen und oft über Unternehmensgrenzen, Disziplinen und Kontinente hinweg. Diese Gemeinschaften unterstützen sich gegenseitig, indem sie Lösungen austauschen und Informationen verbreiten.

Ein Beispiel dafür, wie Entwickler:innen von ihrer Umgebung beeinflusst werden, zeigt eine Studie, die deren Sicherheitsverhalten untersuchte. Ziel war es, nachvollziehen zu können, warum bekannte Sicherheitslücken immer noch in Software eingebettet werden. Warum bekämpfen die Entwickler:innen diese bekannten Probleme nicht, wenn die Software erstellt wird? Das Ergebnis hat gezeigt, dass das Sicherheitsverhalten von Softwareentwickler:innen oft von den Einstellungen und Prioritäten ihrer Unternehmen und Kund:innen beeinflusst wird, nicht nur von ihren eigenen Entscheidungen.

Fachexpert:innen

Bei jeder Softwareentwicklung müssen sich die Entwickler:innen mit dem Anwendungsbereich auseinandersetzen, in dem die Software eingesetzt werden soll. In speziellen Fachbereichen, wie der



Grafik 7: Gemeinschaften und Umgebungen, die die Softwareentwicklung beeinflussen
© Tim Clarke

AKTIVIERUNG

Apps

Materialien:
Smartphone

Dauer:
10 Minuten

Die Teilnehmer:innen setzen sich mit den von ihnen am meisten genutzten Apps auseinander und überlegen, welche Nutzen sie für sie haben und als wie sicher sie sie einschätzen.

- ▶ Welche Apps werden von den Teilnehmer:innen am häufigsten genutzt?
- ▶ Warum?
- ▶ Was tun sie damit?
- ▶ Welchen Nutzen hat diese App für sie?

AKTIVIERUNG

Maker-Bewegung

Materialien:
Notebook

Die Teilnehmer:innen recherchieren auf den Websites [instructables.com](https://www.instructables.com) und [makezine.com](https://www.makezine.com) oder ähnlichen Websites Technologieprojekte, die sie interessant finden, und setzen sie optional sogar im Zuge z. B. eines Schulprojekts um.

wissenschaftlichen Forschung, entwickeln oft die Fachleute selbst die Software, obwohl sie keine Softwarespezialist:innen sind.

Die Maker-Bewegung (MM) ermutigt Menschen, Technologie nicht nur zu konsumieren, sondern aktiv an der Erstellung teilzuhaben. Die Bewegung fördert das Herstellen, Teilen, Lernen von Hard- und Software und Unterstützen in Gemeinschaften. Websites wie instructables.com und makezine.com zeigen, was erreicht werden kann, wenn Menschen ihre Ideen und Kreationen teilen und darauf aufbauen.

Diese Bewegung verdeutlicht die Macht der Gemeinschaft und die Fähigkeit der:des Einzelnen, Technologie so zu gestalten, dass sie den eigenen Zwecken dient, anstatt von ihr gesteuert zu werden.

Fazit aus diesem Abschnitt:

Die Entwicklung von Softwaresystemen ist ein gemeinschaftlicher Prozess, an dem Softwareingenieur:innen und Fachexpert:innen beteiligt sind. Gemeinschaften von Entwickler:innen, Unternehmenskulturen sowie Nutzer:innen- und Auftraggeberinteressen spielen eine entscheidende Rolle bei der Qualität und Sicherheit der Software. Bewegungen wie die Maker-Bewegung ermutigen Menschen dazu, aktiv an der Gestaltung von Technologie mitzuwirken, und verdeutlichen die Bedeutung der individuellen Kreativität und Zusammenarbeit.

Nutzer:innen von Softwaresystemen

In diesem Abschnitt betrachten wir zwei wichtige Aspekte der Beziehung zwischen Menschen und Softwaresystemen:

1. Benutzer:innen sind kreativ und passen Technologie an ihre eigenen Bedürfnisse und Kontexte an.
2. Das Feedback und das Verhalten der Benutzer:innen beeinflussen die zukünftige Entwicklung von Systemen.

Menschen als Entwickler von Softwaresystemen

Inwiefern kann die Kultur eines Unternehmens oder einer Gemeinschaft den Softwareentwicklungsprozess beeinflussen?

Mögliche Antworten zur Diskussion:

- **Einfluss auf Prioritäten:** Die Unternehmenskultur kann die Prioritäten bei der Softwareentwicklung beeinflussen, z. B. ob mehr Wert auf Sicherheit, Benutzerfreundlichkeit oder schnelle Markteinführung gelegt wird.
- **Entscheidungsfindung:** Eine kollaborative Kultur fördert offene Diskussionen und den Austausch von Ideen, während eine hierarchische Kultur Entscheidungen von oben nach unten durchsetzt, was die Kreativität einschränken kann.
- **Fehlerkultur:** In einer Kultur, die Fehler als Lernmöglichkeit betrachtet, sind Entwickler:innen eher bereit, innovative Ansätze auszuprobieren. In einer Kultur, die Fehler bestraft, könnten Entwickler:innen risikoscheuer sein.
- **Kommunikation und Zusammenarbeit:** Eine offene und unterstützende Kultur fördert den Wissensaustausch und die Zusammenarbeit zwischen Teams und Abteilungen, was zu einer besseren Integration und höherer Softwarequalität führt.
- **Benutzerzentrierung:** Eine Kultur, die den Fokus auf die Bedürfnisse der Benutzer:innen legt, entwickelt Software, die besser auf deren tatsächliche Anforderungen und Erwartungen abgestimmt ist.

Berücksichtigung der Nutzer:innen und ihrer Aktivitäten

In den 1970er-Jahren wurde klar, dass Software so gestaltet werden muss, dass sie für Menschen leicht zu bedienen ist. Der frühe Fokus auf die „Mensch-Maschine-Schnittstelle“ entwickelte sich in den 1980er-Jahren schnell zur Mensch-Computer-Interaktion (Human-Computer Interaction, HCI).

Die HCI basiert auf Disziplinen wie Psychologie, Linguistik, Informatik und Soziologie. Das Ziel ist, Software zu entwickeln, die menschliche Eigenschaften wie Aufmerksamkeit und Lernen sowie Gruppendynamiken und Einstellungen berücksichtigt. Zum Beispiel kann das Verständnis von Aufmerksamkeit helfen, Informationen so zu strukturieren, dass Benutzer:innen das Gesuchte besser finden, und Bildschirme so zu gestalten, dass wichtige Informationen hervorgehoben werden.

Der Begriff „Interaktionsdesign“ umfasst heute ein breites Spektrum an Disziplinen, darunter Psychologie, Informatik, Produktdesign, Sozialwissenschaften und Ergonomie. Es geht darum, nicht nur für die Benutzer:innen, sondern auch mit ihnen zu entwerfen, damit die Technologien wirklich benutzerfreundlich sind.

Einblicke in die Nutzung von Technologien durch Gemeinschaften können durch die Beobachtung und das Feedback der Benutzer:innen gewonnen werden. Es ist jedoch oft schwierig, kulturelle Normen wie Annahmen, Bräuche, Überzeugungen und Gewohnheiten zu erkennen, obwohl sie die Nutzung von Technologien stark beeinflussen.

Softwarenutzung beeinflusst die zukünftige Entwicklung

Heute sind Entwicklung und Nutzung von Software eng miteinander verbunden. Ein Beispiel dafür ist der **Lean-UX-Ansatz** (lean = schlank; UX steht für User Experience), bei dem eine Idee in der realen Nutzung getestet wird, bevor die Entwicklung weit fortgeschritten ist. Ein weiteres Beispiel ist das **A/B-Testing**. Dabei werden zwei verschiedene Versionen eines Systems (Version „A“ bzw. „B“) an verschiedene Nutzergruppen ausgeliefert, um zu sehen, welche besser funktioniert. Nutzer:innen werden zufällig in Gruppen eingeteilt und wissen normalerweise nicht, dass sie an einem Experiment teilnehmen.

Weitere Informationsquellen sind **Kundenrezensionen**, die zeigen, wie beliebt und erfolgreich ein Produkt ist. Auch **App-Rezensionen** in sozialen Medien wie Twitter können wertvolles Feedback liefern. Es ist jedoch oft schwierig, nützliche Informationen aus diesen Quellen zu gewinnen. Dieser Bereich wird noch weiter erforscht.

Fazit aus diesem Abschnitt:

Menschen passen Technologie kreativ an ihren eigenen Kontext an und beeinflussen durch ihr Feedback die Weiterentwicklung von Software. Die Benutzerfreundlichkeit und Funktionalität von Software werden durch interdisziplinäre Ansätze verbessert. Die enge Zusammenarbeit zwischen Entwickler:innen und Benutzer:innen ist entscheidend, um wirklich benutzerzentrierte Technologien zu schaffen.

Menschen in Partnerschaft mit Softwaresystemen

Menschen sind oft enger mit Software verbunden, als sie nur zu verwenden. Software beeinflusst, was Menschen erreichen können, wie sich ihre Ideen entwickeln und wie sie sich verhalten.

Beispiele:

- **CAD-Systeme** (computergestützte Konstruktionssysteme): Diese haben die Art und Weise, wie Gebäude entworfen werden, revolutioniert. Sie bieten hohe Genauigkeit und sind leicht zu ändern.
- **Satellitenbilder**: Die softwaregestützte Analyse und Bearbeitung von Satellitenbildern hat unser Verständnis für die Auswirkungen menschlicher Aktivitäten auf den Planeten verbessert.
- **Navigationssysteme**: Präzise GPS-Ortung und Karten haben das Gehen und Radfahren erleichtert. Früher musste man sorgfältig planen, heute folgt man einfach den Anweisungen.

Software kann nicht nur Verhalten beeinflussen, sondern auch Gewohnheiten ändern.

WISSENSFRAGEN

Menschen als Entwickler und Nutzer von Softwaresystemen

Wer ist typischerweise an der Entwicklung von Softwaresystemen beteiligt?

Was ist eine wichtige Eigenschaft der Softwareentwicklungsgemeinschaften?

Wie beeinflussen Unternehmen und Kund:innen die Entscheidungen der Softwareentwickler:innen?

Was ist der Lean-UX-Ansatz?

Was ist A/B-Testing in der Softwareentwicklung?

Welche Informationsquellen können Entwickler:innen nutzen, um Feedback zur Nutzung von Software zu erhalten?

GRUPPENARBEIT

A/B-Testing für die Lieblingsapp

Ziel:

Die Teilnehmer:innen sollen verstehen, wie A/B-Testing in der Softwareentwicklung angewendet wird, und kreativ überlegen, wie verschiedene Varianten ihrer Lieblingsapp getestet werden könnten.

Beschreibung:

Die Teilnehmer:innen schlüpfen in die Rolle von Softwareentwickler:innen und planen ein A/B-Testing für ihre Lieblingsapp. Sie sollen zwei Varianten der App entwerfen und überlegen, welche Elemente getestet werden könnten, um die Benutzerfreundlichkeit und Qualität der App zu verbessern.

Materialien:

- Whiteboard oder Flipchart für die Gruppen, um ihre Varianten und Testpläne zu skizzieren
- Stifte und Papier für Notizen

1. Einführung: Erläuterung des Konzepts des A/B-Testings und seiner Bedeutung in der Softwareentwicklung. Beispiel: „Beim A/B-Testing werden zwei Versionen einer App erstellt und an verschiedenen Nutzergruppen getestet, um herauszufinden, welche besser funktioniert.“

2. Gruppenbildung und App-Auswahl: Einteilung der Teilnehmer:innen in kleine Gruppen. Jede Gruppe wählt eine Lieblingsapp aus (z. B. eine Social-Media-App, eine Fitness-App oder eine Messaging-App).

3. Variantenentwicklung: Die Gruppen entwickeln zwei verschiedene Varianten der ausgewählten App. Sie können sich auf verschiedene Aspekte konzentrieren, wie z. B. das Layout, die Navigation, die Farben, die Funktionen oder die Benutzerfreundlichkeit. Variante A könnte z. B. eine klassische Version der App sein, während Variante B eine modernere, vereinfachte Version ist.

4. Testplanung: Jede Gruppe plant, wie sie das A/B-Testing durchführen würde. Welche Kriterien werden gemessen (z. B. Benutzerfreundlichkeit, Geschwindigkeit, Zufriedenheit)? Wie werden die Nutzergruppen ausgewählt?

5. Präsentation: Jede Gruppe stellt ihre beiden Varianten und den Testplan vor. Die Gruppen erklären jeweils, warum sie bestimmte Änderungen vorgenommen haben und wie sie die Ergebnisse messen würden.

6. Diskussion: Nach den Präsentationen diskutieren die Gruppen über die verschiedenen Ansätze und darüber, welche Variante ihrer Meinung nach besser funktionieren könnte und warum.

Insgesamt hat die Partnerschaft zwischen Software und Menschen positive Seiten, aber auch Grenzen und unbeabsichtigte Folgen. Persuasive Technologien sind solche, die explizit darauf ausgelegt sind, Menschen von etwas zu überzeugen („to persuade“= überreden, überzeugen) bzw. ein bestimmtes Handeln zu bewirken. Beispielsweise kann der Energieverbrauch in Privathaushalten geändert werden, indem man den Verbraucher:innen Feedback hinsichtlich ihres Energieverbrauchs gibt und sie ihn dann optimieren. Manchmal kann die Überzeugungsarbeit jedoch auch weniger positiv sein. Viele Online-Dienste sind so gestaltet, dass sie Nutzer:innen dazu bewegen, etwas zu tun, was über ihr ursprüngliches Ziel hinausgeht, z. B. ein zusätzliches Produkt zu kaufen oder sich für andere als die ursprünglich beabsichtigten Dienste anzumelden.

Es kann (muss aber nicht) als faire Marketingtechnik bewertet werden, wenn etwa transparent dargestellt wird, warum beim Kauf eines Zugtickets gleich auch Unterkünfte für die besuchte Stadt angeboten werden. Denn warum sollte man nicht einige Optionen als Service anbieten? Oft wird eher ein eher trügerischer Ansatz verfolgt, bei dem Nutzer:innen verleitet werden, etwas zu konsumieren, was sie eigentlich gar nicht wollten (siehe dazu www.deceptive.design).

Fazit aus diesem Abschnitt:

Menschen und Software arbeiten zusammen und beeinflussen sich gegenseitig. Software kann menschliches Verhalten und Gewohnheiten verändern, sowohl positiv als auch negativ. Präzise Navigationssysteme oder CAD-Software haben unsere Arbeitsweise revolutioniert. Gleichzeitig bergen Empfehlungssysteme das Risiko, Meinungen zu verstärken und Kreativität einzuschränken. Es ist wichtig, diese Partnerschaft bewusst zu gestalten, um die Vorteile zu maximieren und negative Auswirkungen zu minimieren.

FRAGEN ZUR DISKUSSION

Eigene Erfahrungen mit Software und ihrem Einfluss auf Verhalten bzw. Gewohnheiten

1

Gibt es Beispiele, wie Apps oder Softwaresysteme das Verhalten positiv beeinflusst haben? Zum Beispiel durch Feedback zum Energieverbrauch oder durch Gesundheits- und Fitness-Apps?

Mögliche Antwort: Eine Fitness-App könnte regelmäßige sportliche Aktivitäten und eine verbesserte Ernährung fördern, indem sie Erinnerungen und Fortschrittsberichte bereitstellt, die zur Erreichung von Gesundheitszielen motivieren.

Welche Beispiele gibt es, bei denen Apps oder Softwaresysteme Gewohnheiten verändert haben? Wie geschah dies?

2

Mögliche Antwort: Eine Budgetplaner-App könnte dabei helfen, Ausgaben besser zu überwachen und zu kontrollieren, was zu mehr Ersparnissen und einem bewussteren Umgang mit Geld führt.

3

Gibt es Fälle, in denen sich die Teilnehmer:innen durch Websites oder Apps getäuscht fühlten, um etwas zu tun, was ursprünglich nicht geplant war? Wie wurde darauf reagiert?

Mögliche Antwort: Eine Newsletter-Anmeldung könnte als erforderlich erscheinen, um auf eine Website zuzugreifen. Nachdem sich Nutzer:innen getäuscht fühlten, kündigten sie das Abonnement und wurden vorsichtiger.

Welche positiven und negativen Erfahrungen gibt es mit Empfehlungssystemen, die Produkte oder Inhalte vorschlagen? Haben diese Empfehlungen das Verhalten verändert?

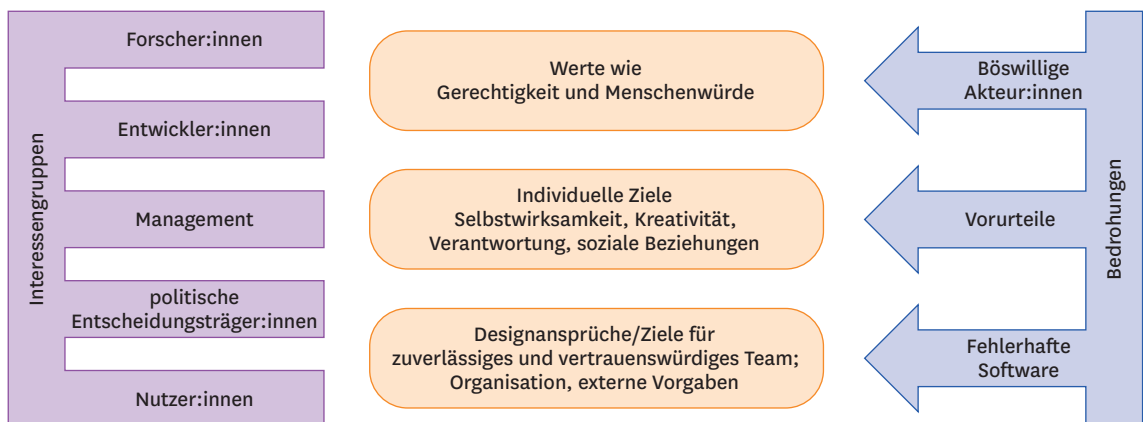
4

Mögliche Antwort: Positive Erfahrung: Neue Bücher wurden entdeckt, die sonst unbemerkt geblieben wären. Negative Erfahrung: Wiederholte Vorschläge ähnlicher Inhalte könnten die Sichtweise einschränken, anstatt sie zu erweitern.

Unterschiedliche Interessen und Ziele im Softwareprozess

Bei der Systementwicklung muss der gesamte Prozess betrachtet werden, von der Analyse über das Design bis hin zur praktischen Anwendung. Dies stellt eine Herausforderung dar, da es beispielsweise schwierig ist, eine gemeinsame Sprache zu finden, in der alle Beteiligten dieselben Begriffe mit identischer Bedeutung verwenden. Besonders herausfordernd sind die unterschiedlichen Perspektiven und Interessen der Beteiligten.

Hier ergeben sich mehrere nicht triviale Fragen: Wie integriert man die unterschiedlichen Nutzer:innen des Systems, zumal diese in den frühen Phasen oft noch unbekannt sind? Wer legt fest, wie gesamtgesellschaftliche, oft widersprüchliche Anliegen definiert und implementiert werden? Wie lassen sich individuelle Bedürfnisse bzw. die der jeweiligen Interessengruppen mit gesamtgesellschaftlichen Interessen in Einklang bringen? Wie gestaltet man die Kommunikation und den Umgang mit Änderungsanforderungen? Und wie integriert man die verschiedenen Beteiligten effektiv in den Prozess (siehe Grafik 8)? Und wie begegnet man den angeführten Bedrohungen? Man sieht, abseits der nicht geringen Schwierigkeiten in der Definition der Zielvorstellungen stellt die Systementwicklung eine große Herausforderung dar.



Grafik 8: Design- und Entwicklungsherausforderungen bei komplexen Systemen. Quelle: Ben Shneiderman (adaptiert)

Schlussbetrachtungen

Die Entwicklung von Softwaresystemen ist ein gemeinschaftliches Projekt, bei dem der Mensch im Mittelpunkt steht. Software hilft Menschen, ihre Ziele zu erreichen, und wie Menschen die Software nutzen, beeinflusst deren Weiterentwicklung. Trotz ihrer Grenzen haben Softwaresysteme zu bedeutenden Fortschritten geführt.

Softwareentwickler:innen und Nutzer:innen sind Teil von Gemeinschaften mit eigenen kulturellen Normen und Erwartungen, die die Entwicklung und Nutzung von Software beeinflussen. Die Erstellung von Software hängt oft von vorgefertigten Komponenten und Technologien ab, die von anderen entwickelt wurden. Nicht spezialisierte Entwickler:innen und Nutzer:innen spielen eine immer größere Rolle bei der Gestaltung verfügbarer Technologien.

Letztendlich muss der Softwareentwicklungsprozess auf unterschiedlichste Interessen und Ziele Rücksicht nehmen. Dies zeigt, dass es ein komplexer Prozess ist, der Arbeiten im Team erfordert.



- KAPITELABSCHNITTE
- WISSENSFRAGEN
- FRAGE(N) ZUR DISKUSSION
- AKTIVIERUNG
- GRUPPENARBEIT

Kapitel 11

PHILOSOPHISCHE GRUNDLAGEN DES DIGITALEN HUMANISMUS

In diesem Kapitel wird erklärt, was Humanismus ist und welche Bedeutung er im digitalen Zeitalter hat. Gezeigt wird der Zusammenhang von Handlung, Freiheit und Verantwortung. Aus dieser Trias ergibt sich, dass der Mensch keine Maschine ist und Maschinen nicht wie Menschen sind. Daraus leiten wir zwei praktische Forderungen für den Digitalen Humanismus ab, die im Verlauf des Kapitels genauer vorgestellt werden: nämlich, dass KI-Systeme keine Verantwortung übernehmen und keine Entscheidungen treffen dürfen.

Basis dieses Kapitels ist der Aufsatz „Philosophical Foundations of Digital Humanism“ von Julian Nida-Rümelin und Klaus Staudacher, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024. DOI: 10.1007/978-3-031-45304-5_3.

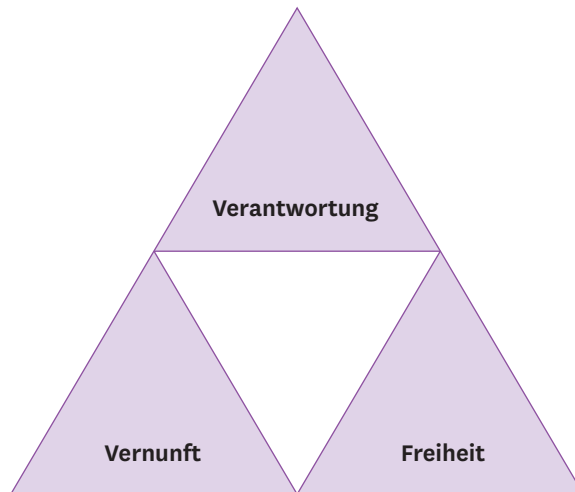
Einleitung

Der Digitale Humanismus bietet eine neue Ethik für das Zeitalter der Künstlichen Intelligenz. Er steht, vereinfacht ausgedrückt, im Gegensatz zur „Silicon-Valley-Ideologie“. Diese Ideologie ist eng verknüpft mit der amerikanischen, puritanischen Hoffnung auf Erlösung, auf eine „reine“ Welt, die frei von Sünde ist (der Begriff „Puritanismus“ leitet sich vom Begriff „pure“ (rein) ab und bezeichnet eine in England und Schottland entstandene Bewegung, die mit den ersten Auswanderern und Auswanderinnen im 17. Jahrhundert nach Amerika gelangte). Ins digitale Zeitalter übertragen bedeutet das die Schaffung eines perfekten, fehlerfreien digitalen Abbilds des Menschen – eine technologische Utopie. Im Zentrum steht eine Künstliche Intelligenz, die sich selbst verbessert, hyperrational ist und zunehmend menschlich wird. Einziger Unterschied zum Puritanismus ist, dass hier nicht Gott der Schöpfer ist, sondern Software-Ingenieur:innen, die sich weniger als Teil der Computerindustrie, sondern vielmehr als Teil einer Gruppe sehen, die das digitale Paradies auf Erden erschaffen will.

Wie alle Technologien sind auch digitale Technologien zweischneidig. Die digitale Transformation wird unsere Lebensbedingungen nicht automatisch verbessern – es kommt darauf an, wie wir diese Technologie nutzen und entwickeln. Wir selbst können den Prozess der digitalen Transformation interpretieren und aktiv gestalten. Das ist der Kerngedanke des Digitalen Humanismus.

Was bedeutet Humanismus?

Humanismus kann vieles bedeuten, von der Pflege alter Sprachen bis hin zur Idee, dass die Menschheit die Erde beherrschen soll. Wenn wir hier von Humanismus sprechen, geht es nicht um historische Epochen wie den italienischen Frühhumanismus (Francesco Petrarca), den deutschen Humanismus im 15. und 16. Jahrhundert (Erasmus von Rotterdam) oder den Neuhumanismus im 19. Jahrhundert (Wilhelm von Humboldt). Es handelt sich auch nicht um ein rein westliches oder europäisches Phänomen, da humanistisches Denken und Handeln auch in anderen Kulturen existiert.



Grafik 9: Die Triade von Vernunft, Freiheit und Verantwortung hat die Moral der Lebenswelt und die Rechtsordnung über Jahrhunderte beeinflusst.
Quelle: Bettina K. Lechner

Wir bezeichnen mit „Humanismus“ eine bestimmte Vorstellung davon, was das Menschsein ausmacht, verbunden mit einer Praxis, die diesem Ideal so weit wie möglich entspricht. Im Zentrum der humanistischen Philosophie steht die Idee des Menschen als Akteur. Der Mensch gestaltet aktiv sein Leben, trägt Verantwortung und ist frei. Freiheit und Verantwortung bedingen einander und sind mit der Fähigkeit zur Vernunft verbunden. Menschen sind als Akteure ihres Lebens moralisch verantwortlich.

Diese Verbindung von Vernunft, Freiheit und Verantwortung prägt das humanistische Verständnis des Menschen. Sie hat die Moral der Lebenswelt und die Rechtsordnung über Jahrhunderte beeinflusst. Um diese humanistische Praxis besser zu verstehen, wollen wir den Zusammenhang zwischen Verantwortung, Freiheit und Vernunft näher beleuchten.

Exkurs Posthumanismus

Der Humanismus stellt den Menschen als vernünftiges und handlungsfähiges Wesen in den Mittelpunkt. Der Post-Humanismus hingegen sieht den Menschen nicht mehr als zentralen Ausgangspunkt von Denken und Handeln. Es gibt verschiedene Formen von Post-Humanismus, z. B. den technologischen, der den Menschen mittels einer künstlichen

Superintelligenz überwinden will. Andere Formen von Posthumanismus kreisen um Kritik am Eurozentrismus, an den Dichotomien (Entweder-oder-Gegensätze): Mann/Frau, Natur/Kultur, Subjekt/Objekt etc. oder an der Annahme, dass sich der Mensch durch seine Urteils- und Lernfähigkeit sowie seine Autonomie grundlegend von nicht belebter Materie unterscheidet. Es gibt also Strömungen, die von der Gleichwertigkeit oder Gleichberechtigung von unterschiedlichen Seinsformen ausgehen, egal ob diese nun „tierisch“, „menschlich“, „pflanzlich“ oder „technologisch“ sind. Die Debatten zum Posthumanismus und auch zum Transhumanismus werden aktuell geführt – und sind alles andere als abgeschlossen.

Fazit aus diesem Abschnitt:

Humanismus bedeutet, den Menschen als freien und verantwortlichen Akteur zu sehen. Er basiert auf den Konzepten von Vernunft, Freiheit und Verantwortung, die in einem kulturellen Prozess über Jahrhunderte die Moral und die Rechtsordnung geprägt haben. Humanismus ist nicht auf eine bestimmte Kultur oder Epoche beschränkt, sondern ist ein universelles Ideal.

Zusammenhang von Verantwortung, Freiheit und Vernunft

Verantwortung ist eng mit Freiheit, Vernunft und Handlung verbunden. Eine Handlung ist vernünftig, wenn es gute Gründe dafür gibt, sie auszuführen. Vernunft bedeutet, dass wir die Gründe für unsere Handlungen, Überzeugungen und Einstellungen angemessen abwägen können. Freiheit ist die Möglichkeit, den besseren Gründen zu folgen, die wir durch Abwägung ermittelt haben.

Wenn ich frei bin, dann leiten mich die Gründe, die ich abgewogen habe, zu einer Entscheidung und Handlung.

WISSENSFRAGEN

Was bedeutet Humanismus?

Was versteht man unter Humanismus?

Welche drei zentralen Konzepte stehen im Mittelpunkt der humanistischen Philosophie?

Wie hat der Humanismus die Moral der Lebenswelt und die Rechtsordnung beeinflusst?

Warum ist die Triade von Vernunft, Freiheit und Verantwortung zentral für das humanistische Verständnis des Menschen?

Was versteht man unter technologischem Posthumanismus?

FRAGEN ZUR DISKUSSION

Zusammenhang von Verantwortung, Freiheit und Vernunft

Warum ist die Vorstellung vom Menschen als freiem und verantwortlichem Akteur wichtig für die moderne Gesellschaft?

1

Mögliche Diskussion: Diese Vorstellung betont, dass der Mensch nicht nur rational handeln kann, sondern auch Verantwortung für seine Entscheidungen trägt. In einer modernen Gesellschaft bedeutet dies, dass Individuen und Institutionen für ihre Handlungen rechenschaftspflichtig sind und auf Grundlage ethischer Prinzipien agieren sollten. Diese Werte fördern eine Kultur der Selbstverantwortung und Moral, die für ein gerechtes und demokratisches Zusammenleben unerlässlich ist.

Wie können die Konzepte von Vernunft, Freiheit und Verantwortung in der heutigen technologisch geprägten Welt angewendet werden?

2

Mögliche Diskussion: Vernunft, Freiheit und Verantwortung könnten in der Technologieentwicklung Anwendung finden, indem KI-Systeme so gestaltet werden, dass sie menschliche Entscheidungen unterstützen, ohne sie zu ersetzen. Die Freiheit der Nutzer:innen, Entscheidungen zu treffen, bleibt dabei erhalten, und Entwickler:innen tragen die Verantwortung für die ethischen Implikationen der von ihnen entworfenen Systeme. Dies könnte helfen, Technologien zu schaffen, die den Menschen stärken und nicht kontrollieren.

Was bedeutet es, einen Grund für etwas zu haben?

Beispiele für Gründe:

- Wenn ein verletztes Unfallopfer am Straßenrand liegt, haben wir einen Grund, dem:der Verletzten zu helfen (z. B. Erste Hilfe leisten oder einen Krankenwagen rufen).
- Wenn Petra ihrer Freundin Johanna verspricht, ihr am Wochenende beim Umzug zu helfen, hat Petra einen Grund, dies zu tun. Auch wenn es andere wichtige Gründe gibt, wie etwa die kranke Mutter von Petra, bleibt das Versprechen ein Grund.

Diese Beispiele zeigen zwei wichtige Eigenschaften von Gründen:

1. **Gründe sind normativ:** Wenn es einen Grund für eine Handlung gibt, sollte man diese Handlung ausführen, es sei denn, es gibt gewichtigere Gründe, die dagegensprechen.
2. **Gründe sind objektiv:** Ein guter Grund ist unabhängig von persönlichen Einstellungen/Vorlieben. Der Grund, einem Unfallopfer zu helfen, löst sich nicht auf, nur weil man andere Vorlieben hat oder denkt, dass das Opfer keine Hilfe verdient.

Verhältnis von Vernunft und Freiheit und ihre Bedeutung für Verantwortung

Verantwortung setzt voraus, dass man die Freiheit hat, eine Handlung zu unterlassen oder eine Entscheidung zurückzunehmen. Aber sind unsere Entscheidungen und Handlungen wirklich frei?

Handlungen unterscheiden sich von bloßem Verhalten:

- Wenn eine Person in einem Bus durch eine Vollbremsung das Gleichgewicht verliert und auf eine andere Person fällt, ist das ein Unfall.
- Wenn dieselbe Person gezielt auf eine andere Person fällt, würden wir Absicht zuschreiben und damit von einer Handlung sprechen.

Handlungen sind durch Absicht und Vernunft geleitet. Eine Handlung ist rational, wenn gute Gründe dafür sprechen, und irrational, wenn gute Gründe dagegen sprechen. Die Gründe, von denen wir uns leiten lassen, sind das Ergebnis eines Abwägungsprozesses. Durch diese

Fähigkeit zur Abwägung sind wir rationale Wesen und verantwortlich für unser Tun.

Wenn ein Schaden durch bloßes Verhalten verursacht wird, wird kein Vorwurf gemacht. Wir beschreiben nur die Kausalität (z. B. „Durch die Vollbremsung stürzte die Person und verletzte die andere“). Wenn der Schaden durch eine Handlung verursacht wurde, erwarten wir eine Erklärung und Rechtfertigung.

Verantwortung

Verantwortung bedeutet, dass man von Gründen betroffen ist. Weil wir rational sind und die Fähigkeit zur Abwägung haben, sind wir frei. Und nur weil wir frei sind, können wir verantwortlich sein.

Fazit aus diesem Abschnitt:

Verantwortung, Freiheit und Vernunft sind eng miteinander verbunden. Eine Handlung ist dann vernünftig, wenn es gute Gründe für sie gibt. Freiheit bedeutet, diesen Gründen zu folgen. Menschen sind verantwortlich für ihre Handlungen, weil sie in der Lage sind, Gründe abzuwägen und Entscheidungen zu treffen. Dieses Verständnis ist zentral für den Digitalen Humanismus.

Wichtige Ideen des Digitalen Humanismus

1. Der Mensch ist keine Maschine (Zurückweisung des „mechanistischen Paradigmas“)

Eine der größten Herausforderungen für das humanistische Menschenbild ist die Idee, dass der Mensch eine Maschine ist. Diese Idee stammt aus der Frühen Neuzeit und sieht die Welt als Uhrwerk und den Menschen als kleines Zahnrad darin. Der Schöpfer dieser „Uhr“ sorgt dafür, dass alles perfekt funktioniert, ohne Raum für menschliche Freiheit, Verantwortung oder Vernunft.

GRUPPENARBEIT

Verantwortung, Freiheit und Vernunft im Alltag

Ziel:

Die Teilnehmer:innen sollen die Zusammenhänge von Verantwortung, Freiheit und Vernunft verstehen und anwenden, indem sie alltägliche Situationen analysieren und diskutieren.

Materialien:

- Arbeitsblätter mit den beschriebenen Situationen und Fragen
- Flipcharts oder Whiteboards für die Präsentationen
- Stifte und Papier für Notizen

Dauer:

Abhängig von der Anzahl der Gruppen, mindestens jedoch rund 50 Minuten

1. Analyse:

Jede Gruppe erhält kurze Schilderungen verschiedener Alltagssituationen, die sie analysieren soll.

Beispiele für solche Situationen:

- ▶ Ein Freund bittet um Hilfe beim Ausmalen seiner Wohnung, aber man hat bereits andere Verpflichtungen.
- ▶ Man findet eine Geldbörse auf der Straße.
- ▶ Eine Mitschülerin wird in der Schule gemobbt.
- ▶ Man erhält zu viel Wechselgeld im Geschäft.

Die Gruppen sollen die folgenden Fragen zu jeder Situation beantworten:

- ▶ Was sind die Gründe, die für oder gegen eine bestimmte Handlung sprechen?
- ▶ Inwiefern spielt Vernunft eine Rolle bei der Entscheidung?
- ▶ Welche Freiheit hat die Person in dieser Situation?
- ▶ Welche Verantwortung trägt die Person?

2. Präsentation und Diskussion:

Jede Gruppe präsentiert ihre Erkenntnisse vor den anderen Teilnehmer:innen. Die:der Moderator:in fasst die wichtigsten Punkte zusammen und stellt sicher, dass alle Teilnehmer:innen die Verbindungen zwischen den Konzepten verstehen.

3. Diskussion und Reflexion:

Offene Diskussion über die Bedeutung von Verantwortung, Freiheit und Vernunft im digitalen Zeitalter. Reflexion darüber, wie diese Konzepte auf die Nutzung und Entwicklung von Technologie angewendet werden können.

GRUPPENARBEIT

Theaterstück über Verantwortung, Freiheit und Vernunft

Ziel:

Die Teilnehmer:innen sollen die Zusammenhänge zwischen Verantwortung, Freiheit und Vernunft durch kreative Darstellung und Analyse realer Situationen verstehen und anwenden

Materialien:

- Szenenbeschreibungen (optional)
- Requisiten und Kostüme für die Aufführungen (optional)
- Stifte und Papier für Notizen

Dauer:

Abhängig von der Anzahl der Gruppen, mindestens jedoch rund 50 Minuten

1. Gruppeneinteilung und Szenenvergabe:

- ▶ Die Teilnehmer werden in kleine Gruppen von 4 bis 5 Personen eingeteilt.
- ▶ Jede Gruppe erhält eine kurze Beschreibung einer Alltagssituation, die sie in einem kurzen Theaterstück (5 bis 7 Minuten) darstellen soll.

2. Szenenideen:

- ▶ Szenario 1: Eine Person verursacht aus Versehen einen Unfall, bei dem ein Fahrradfahrer verletzt wird. Die Gruppe soll darstellen, wie die Person abwägt, ob sie die Verantwortung übernimmt, und welche Gründe für und gegen verschiedene Handlungen sprechen.
- ▶ Szenario 2: In einem Restaurant verschüttet ein Kellner aus Versehen ein Getränk auf einen Gast. Die Gruppe soll darstellen, wie der Kellner und der Gast die Situation handhaben und welche Überlegungen zu Verantwortung und Vernunft angestellt werden.
- ▶ Szenario 3: Eine Person nimmt an einer Gruppenarbeit teil, aber ein Gruppenmitglied erledigt seinen Teil der Aufgaben nicht. Die Gruppe soll darstellen, wie die anderen Gruppenmitglieder darauf reagieren und welche Gründe für und gegen bestimmte Handlungen sprechen.

3. Erarbeitung der Szenen:

- ▶ Jede Gruppe arbeitet an ihrer Szene. Sie soll die Charaktere, den Dialog und die Handlung entwickeln.
- ▶ Die Gruppen sollen auch sicherstellen, dass die Konzepte Verantwortung, Freiheit und Vernunft in ihrer jeweiligen Szene deutlich werden.

4. Aufführung der Szenen:

- ▶ Jede Gruppe führt ihre Szene vor den anderen Gruppen auf.
- ▶ Die anderen Gruppen beobachten die Aufführungen und machen sich Notizen zu den dargestellten Konzepten.

5. Analyse, Diskussion und Abschluss:

- ▶ Nach jeder Aufführung diskutieren die Zuschauergruppen, wie die Konzepte Verantwortung, Freiheit und Vernunft in der Szene dargestellt wurden.
- ▶ Die:der Moderator:in führt die Diskussion und fasst die wichtigsten Punkte zusammen.
- ▶ Am Ende gibt es eine gemeinsame abschließende Diskussion.

Maschinen – und auch Computersysteme – funktionieren deterministisch (wenn sie funktionieren). Jede Operation der Hardware kann als kausal bestimmt beschrieben werden – durch den vorherigen Zustand der Maschine und den aktuellen Input.

Wenn wir dieses Bild auf den Menschen übertragen, würde das humanistische Menschenbild als Illusion oder Selbsttäuschung erscheinen. In einer humanistischen Sichtweise ist der Mensch jedoch kein Mechanismus, sondern ein freier und verantwortlicher Akteur, der mit anderen Menschen und der Welt interagiert. Wir Menschen haben mentale Eigenschaften, Überzeugungen, Wünsche, Absichten, Ängste und Erwartungen.

2. Maschinen sind nicht wie Menschen (Zurückweisung des „animistischen Paradigmas“)

In diesem Konzept werden der Maschine menschliche Eigenschaften und Verhaltensweisen zugeschrieben. Dieses animistische Paradigma war insbesondere nach dem Zweiten Weltkrieg mit der Entwicklung der Computertechnologie einflussreicher als das mechanistische. Dazu trug auch **Alan Turing** 1950 mit seinem Aufsatz *Computing Machinery and Intelligence* bei.

Das animistische Paradigma interpretiert algorithmische Systeme so, dass sie mentale Eigenschaften haben, solange ihr Verhalten dem des Menschen ähnlich ist. Diese Sichtweise ist besonders in der Idee der „starken KI“ zu finden. Diese geht davon aus, dass ein Programm lernen kann, jede intellektuelle Aufgabe zu verstehen oder zu lösen, die ein Mensch ausführen kann. Damit einher geht auch die Möglich-

Der Turing-Test, benannt nach dem britischen Mathematiker und Informatiker Alan Turing, wurde 1950 vorgestellt. Der Test soll bestimmen, ob eine Maschine menschliche Intelligenz zeigen kann.

So funktioniert der Test: Ein Mensch führt ein Gespräch mit einer anderen Person und einer Maschine, ohne zu wissen, wer wer ist. Die Gesprächsteilnehmer:innen kommunizieren nur über Text, um Aussehen und Stimme zu verbergen. Wenn der Mensch nicht zuverlässig erkennen kann, welche der Gesprächspartner:innen die Maschine ist, hat die Maschine den Test bestanden und wird als intelligent betrachtet.

Der Turing-Test ist eine wichtige Grundlage in der Diskussion über Künstliche Intelligenz und ihre Fähigkeiten.

keit von Bewusstsein. In dieser Sicht ist es nur eine Frage der Zeit, bis Maschinen Menschen in ihren Fähigkeiten überflügeln.

Das Konzept der „schwachen KI“ dagegen beschränkt sich auf Lösen wohl definierter, vorbestimmter Probleme und stellt keinen generellen Intelligenzanspruch. Dabei können aber menschliche Denk-, Wahrnehmungs- und Entscheidungsprozesse durch Software simuliert werden.

In der wissenschaftlichen Gemeinschaft gibt es keine einheitliche Position dazu, ob sich das radikale Konzept der „starken KI“ verwirklichen lässt. Wäre dies der Fall, hätte dies weitreichende Folgen, Maschinen hätten dann auch so etwas wie Menschen- oder Maschinenrechte und Würde. Und wenn Maschinen uns Menschen überflügeln, wer hindert sie dann daran, uns zu eliminieren?

Fazit aus diesem Abschnitt:

Der Digitale Humanismus lehnt die Vorstellung ab, dass Menschen Maschinen sind oder dass Maschinen wie Menschen sein können. Der Mensch ist ein freier und verantwortlicher Akteur, während Maschinen keine mentalen Eigenschaften oder Menschenrechte haben. Diese Unterscheidung ist wichtig, um die ethischen Implikationen der Künstlichen Intelligenz richtig zu bewerten.

Praktische Folgen des Digitalen Humanismus

Auch komplexe KI-Systeme können in absehbarer Zeit nicht als Personen betrachtet werden. Daraus ergeben sich zwei praktische Forderungen:

- **Erstens sollten wir KI-Systemen keine Verantwortung zuschreiben.** KI-Systeme sind nicht in dem Maße rational und frei, wie es für die Zuschreibung von Verantwortung erforderlich ist.
- **Zweitens sollten ethische Entscheidungen niemals von algorithmischen KI-Systemen getroffen werden.**

AKTIVIERUNG

Plakatgestaltung zu den zwei praktischen Forderungen des Digitalen Humanismus

Ziel:

Die Teilnehmer:innen sollen die zwei praktischen Forderungen des Digitalen Humanismus durch kreative Plakatgestaltung anschaulich und merkbar darstellen.

Materialien:

- Großes Papier oder Flipchart-Papier
- Marker und Buntstifte
- Scheren und Kleber
- Alte Zeitschriften für Collagen
- Stifte und Papier für Notizen

Dauer:

Abhängig von der Anzahl der Gruppen, mindestens 45 Minuten

1. Einführung:

Der:die Moderator:in erklärt die zwei praktischen Forderungen des Digitalen Humanismus:

- ▶ KI-Systemen keine Verantwortung zuschreiben.
- ▶ Ethische Entscheidungen niemals von algorithmischen KI-Systemen treffen lassen.

2. Gruppeneinteilung:

Die Teilnehmer:innen werden in kleine Gruppen von 2 bis 3 Personen eingeteilt. Jede Gruppe erhält die Aufgabe, ein Plakat zu gestalten, das diese beiden Forderungen anschaulich darstellt.

3. Materialverteilung:

Jede Gruppe erhält großes Papier oder Flipchart-Papier, Marker, Buntstifte, Scheren, Kleber und alte Zeitschriften für Collagen.

4. Plakatgestaltung:

Jede Gruppe arbeitet an ihrem Plakat. Die Gruppen sollen kreative und visuell ansprechende Wege finden, um die beiden Forderungen darzustellen. Sie können Symbole, Zeichnungen, kurze Texte und Collagen verwenden, um die Botschaften klar und einprägsam zu machen.

5. Präsentation der Plakate:

- ▶ Jede Gruppe präsentiert ihr Plakat den anderen Gruppen und erklärt, wie sie die Forderungen des Digitalen Humanismus dargestellt hat.
- ▶ Die anderen Gruppen können Fragen stellen und Feedback geben.
- ▶ Offene Diskussion über die Bedeutung der beiden Forderungen im Kontext des Digitalen Humanismus.

NOTIZEN



- KAPITELABSCHNITTE
- WISSENSFRAGEN
- FRAGE(N) ZUR DISKUSSION
- AKTIVIERUNG
- GRUPPENARBEIT

Kapitel 12

DER DIGITALE HUMANISMUS

Die digitale Transformation bringt nicht nur große Vorteile, sondern auch erhebliche Herausforderungen mit sich. In diesem Artikel stellen wir den „Digitalen Humanismus“ und das „Wiener Manifest“ vor. Ansätze, die darauf abzielen, das Verhältnis zwischen Technologie und Menschheit zu verstehen und positiv zu beeinflussen. Der Digitale Humanismus fördert eine Gesellschaft, die die Menschenrechte achtet und sowohl technische als auch soziale Innovationen umfasst. Er ist proaktiv und zielt darauf ab, Technologie so zu gestalten, dass sie das Leben verbessert, anstatt nur wirtschaftliche Interessen zu verfolgen.

Basis dieses Kapitels ist der Aufsatz „Digital Transformation, Digital Humanism: What Needs to Be Done“ von Hannes Werthner, in: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024. DOI: 10.1007/978-3-031-45304-5_3.

Einleitung

Die Informationstechnologie (IT) hat tiefgreifende Auswirkungen auf unser individuelles Leben und unsere Gesellschaft, von täglichen Interaktionen bis hin zu großen geopolitischen Machtspielen. Die IT hat die Art und Weise, wie wir die Welt betrachten und verstehen, schnell und nachhaltig verändert. Sie war wesentlich für den wirtschaftlichen Fortschritt. Diese rasante Entwicklung kann zu Lösungen globaler Krisen beitragen, sie bringt aber auch neue Herausforderungen mit sich.

Der **Digitale Humanismus** bietet einen Rahmen, um diesen Prozess zu verstehen und gegebenenfalls darin einzugreifen, um die Entwicklung der Technologie in eine Richtung zu lenken, die eine gerechtere und demokratischere Gesellschaft fördert. Aus dieser fächerübergreifenden Initiative entstand das **Wiener Manifest**, das darauf abzielt, Technologie im Einklang mit menschlichen Werten zu gestalten und zu nutzen.

„The System is failing“

IT-Systeme haben sich als unverzichtbar erwiesen, insbesondere während der COVID-19-Pandemie, als sie ermöglichten, dass die Welt weiterlaufen konnte: Arbeit, Bildung und Kommunikation fanden online statt. Die IT unterstützt grundlegende und lebenswichtige Funktionen in unserer Gesellschaft. Und sie kann eine entscheidende Rolle bei der Erfüllung der nachhaltigen Entwicklungsziele (Sustainable Development Goals, SDG) der Vereinten Nationen spielen.

Trotz dieser Erfolge hat die Technologie auch ihre Schattenseiten. **Tim Berners-Lee** wies 2017 in seinem Artikel *The System is failing* (*The Guardian*, 16.11.2017) darauf hin, dass die derzeitigen IT-Systeme tiefgreifende Probleme mit sich bringen.

Sir Timothy John Berners-Lee, (geb. 1955 in London) ist der Erfinder des World Wide Web, der Hypertext Markup Language (HTML), des URL-Systems und des Hypertext Transfer Protokolls (HTTP). Er steht dem World Wide Web Consortium (W3C) vor.

Machtmonopole im Internet:

- Multinationale IT-Unternehmen besitzen Macht, die schwer von nationalen Regierungen zu kontrollieren ist. Sie sind die, in Bezug auf ihren Marktwert, am höchsten bewerteten Firmen der Welt und haben in einigen Marktsegmenten Marktanteile von 50 bis 90 Prozent.
- Diese Unternehmen bieten Dienste an, die Regierungen nicht in dieser Qualität bereitstellen, sie entscheiden über die Umsetzung wesentlicher Dienste für die Bürger:innen, z. B. den Zugang zum mobilen Web oder zu Cloud-Diensten.

Verlust der Souveränität:

- IT-Unternehmen wie Google oder Apple beeinflussen maßgeblich die Architektur und Funktionalität wichtiger Dienste. Dies zeigt das Beispiel der Corona-App, bei der diese Konzerne und nicht nationale Regierungen darüber entschieden, ob diese Anwendungen als mobile Apps zur Verfügung stehen.
- Große Technologieunternehmen beeinflussen wichtige Entscheidungen, was die Fähigkeit von Staaten und Einzelpersonen, eigenständig Entscheidungen zu treffen, einschränkt.

KI und automatisierte Entscheidungsfindung:

- Künstliche Intelligenz kann autonom Entscheidungen treffen, was rechtliche und ethische Fragen der Verantwortung aufwirft.
- Die Undurchsichtigkeit von Blackbox-Algorithmen erschwert das Verständnis der Entscheidungsgrundlagen.
- KI-Tools wie ChatGPT liefern nicht immer zuverlässige Ergebnisse und verleiten dazu, diese zu 100 Prozent zu glauben, da sie sehr plausibel und überzeugend wirken.

KI und Einsatz im Militär:

- Autonome, KI-basierte Systeme werden zunehmend in militärischen Anwendungen eingesetzt.
- Der Generalsekretär der Vereinten Nationen, António Guterres, hat bereits auf Twitter gefordert, dass „autonome Maschinen, die in der Lage sind [...] ohne menschliche Beteiligung Leben zu nehmen, durch internationales Recht verboten werden sollten“.
- Sowohl die Zivilgesellschaft als auch die universitäre Öffentlichkeit haben bereits mit der Kampagne „Stop Killer Robots“ reagiert und fordern ebenfalls ein Verbot solcher Technologien.

AKTIVIERUNG

Analyse der Herausforderungen durch IT-Systeme

Ziel:

Die Teilnehmer:innen sollen die verschiedenen aus der Anwendung moderner IT-Systeme resultierenden Probleme und Herausforderungen analysieren und diskutieren. Sie sollen verstehen, wie diese Herausforderungen das tägliche Leben und die Gesellschaft beeinflussen und warum es das Wiener Manifest braucht.

Materialien:

- Stifte und Papier oder digitale Geräte für Notizen
- Flipcharts oder Whiteboard
- Ausdrücke der wichtigsten Punkte des Abschnitts (optional)

Dauer:

45 bis 60 Minuten (abhängig von der Anzahl der Gruppen)

1. Gruppenbildung und Themenzuweisung:

Jede Gruppe erhält eines der Hauptthemen aus dem Abschnitt:

- ▶ Machtmonopole im Internet
- ▶ Verlust der Souveränität
- ▶ KI und automatisierte Entscheidungsfindung und Einsatz in der Kriegsführung
- ▶ Überwachung und Privatsphäre
- ▶ Desinformation und Filterblasen
- ▶ Umweltauswirkungen der IT

2. Analyse und Diskussion:

Jede Gruppe analysiert ihr Thema und diskutiert die folgenden Fragen:

- ▶ Was sind die Hauptprobleme und Herausforderungen in diesem Bereich?
- ▶ Wie beeinflussen diese Probleme eventuell sogar das eigene tägliche Leben und wie die Gesellschaft?
- ▶ Welche möglichen Lösungen oder Ansätze könnten helfen, diese Herausforderungen zu bewältigen?

3. Präsentation, Fragen und Feedback:

Jede Gruppe präsentiert ihre Ergebnisse und Vorschläge kurz vor der gesamten Gruppe. Nach jeder Präsentation können die anderen Teilnehmer:innen Fragen stellen und Feedback geben.

4. Gemeinsame Diskussion und Reflexion:

Gemeinsame Diskussion, welche Erkenntnisse aus der Übung gewonnen wurden. Erörterung, welche gemeinsamen Themen und Lösungen sich herauskristallisiert haben.

Auswirkungen der Automatisierung auf Beschäftigung:

- Automatisierung hat große Auswirkungen auf Jobs und Arbeitsplätze. Die offene Frage ist, ob Maschinen die Menschen unterstützen oder sie ersetzen und so zu Arbeitsplatzverlusten führen.
- Aktuell fördert die KI-Industrie eine ungerechte Arbeitsteilung, bei der meist schlecht bezahlte und gering qualifizierte Arbeit in armen Ländern erledigt wird, während die reicheren Länder profitieren (koloniale Arbeitsteilung).

Zunehmende Überwachung und Verletzung der Privatsphäre:

- Sowohl private Unternehmen als auch der Staat erweitern ihre Überwachungsaktivitäten.
- Dies stellt eine große Bedrohung für liberale Demokratien dar.

Probleme in den Online-Medien:

- Plattformen, die für offene Kommunikation und Partizipation gedacht waren, fördern nun Fake News und Filterblasen.
- Diese Tendenzen vergiften den politischen Diskurs und gefährden demokratische Strukturen.
- Diese Entwicklung ist auch ein Ergebnis von Algorithmen aus dem Bereich der Recommender-Systeme (siehe auch Kapitel 6, Seite 83), zudem spiegelt das Web in Daten und Algorithmen die Vorurteile der Nutzer:innen wider.

Umweltauswirkungen der IT:

- Trotz positiver Beispiele für den Einsatz von IT im Klimaschutz überwiegen oft negative Umweltauswirkungen.
- Die Entwicklung und das Training von KI-Systemen verursachen einen sehr hohen Energieverbrauch sowie Emissionen.

Fazit aus diesem Abschnitt:

Die Informationstechnologie hat unser Leben in vielerlei Hinsicht verbessert, insbesondere während der COVID-19-Pandemie. Gleichzeitig hat sie jedoch bedeutende Probleme und Herausforderungen mit sich gebracht, wie Machtmonopole, den Verlust der Souveränität, ethische Fragen rund um KI, Überwachung, Desinformation und Umweltauswirkungen. Das erfordert eine kritische Auseinandersetzung und einen bewussten Umgang mit der weiteren Entwicklung der Technologie.

Der Digitale Humanismus und das Wiener Manifest

Der Wiener Workshop zum Digitalen Humanismus fand im April 2019 statt. Angeregt wurde er durch die beeindruckenden Erfolge der Informationstechnologie, aber auch durch die zuvor beschriebenen ernststen Herausforderungen, die mit dieser einhergehen. Im Zentrum des Workshops stand die Forderung an die Wissenschaftler:innen, Verantwortung zu übernehmen und Technologien zu entwickeln, die sich an den Bedürfnissen und Werten der Menschen orientieren – und nicht umgekehrt.

Mehr als 100 Teilnehmer:innen aus Universitäten, öffentlichen Einrichtungen, Zivilgesellschaft und Wirtschaft diskutierten, wie die Informationstechnologie das gesellschaftliche Leben beeinflusst und wie man diese Entwicklung positiv steuern kann. Es war wichtig, Expert:innen aus verschiedenen Bereichen wie Politikwissenschaft, Recht, Soziologie, Geschichte, Anthropologie, Philosophie, Wirtschaftswissenschaft und Informatik zusammenzubringen, um eine breite Perspektive zu gewährleisten.

Der Begriff „Digitaler Humanismus“ baut auf den Ideen des Humanismus und der Aufklärung auf, die den Menschen und seine Entscheidungsfreiheit in den Mittelpunkt stellen. Es geht darum, dass Menschen in der Lage sind, kritisch zu denken, zu entscheiden und für ihre Entscheidungen und Handlungen Verantwortung zu übernehmen. Technischer Fortschritt ist weder „gottgegeben“ noch deterministisch. Er sollte daher nicht einfach hingenommen werden, sondern bewusst und im Einklang mit demokratischen, humanistischen und ökologischen Werten gestaltet werden.

Wir definieren den Digitalen Humanismus als einen Ansatz, der das komplexe Zusammenspiel von Technologie und Mensch beschreibt, analysiert und vor allem beeinflusst, für eine bessere Gesellschaft und ein besseres Leben unter voller Achtung der universellen Menschenrechte. Hannes Werthner

Beim Workshop wurde ein Wiener Manifest für Digitalen Humanismus (<https://caiml.org/dighum/dighum-manifesto/>) erstellt und von den Unterzeichnenden aus fast 50 Ländern angenommen. Das Manifest ist ein Aufruf zum gemeinsamen Handeln, um Unterstützung aus aller

Welt zu gewinnen und eine menschlichere und nachhaltigere Zukunft zu schaffen.

Grundsätze des Wiener Manifests

Privatsphäre, Demokratie und Inklusion

- Digitale Technologien müssen Demokratie und Integration fördern.
- Privatsphäre und Meinungsfreiheit sind zentrale Werte und müssen geschützt werden.

Regulierung und öffentliche Verwaltung

- Die Politik muss eingreifen, um Technologiemonopole zu verhindern.
- Entscheidungen, die Menschenrechte betreffen, müssen weiterhin von Menschen getroffen werden.

Besondere Rolle der Wissenschaft und der Universitäten

- Wissenschaft muss verschiedene Disziplinen vereinen, um Herausforderungen zu bewältigen.
- Universitäten schaffen neues Wissen und fördern kritisches Denken. Sie haben eine besondere Verantwortung.

Bildung

- Neue Lehrpläne müssen Geistes-, Sozial-, Technik- und Ingenieurwissenschaften verbinden.
- Bildung in Informatik und den ethischen sowie gesellschaftlichen Auswirkungen der IT sollte so früh wie möglich beginnen.

Der Digitale Humanismus will nicht nur die negativen Seiten der IT beseitigen, sondern auch Innovationen fördern, die den Menschen in den Mittelpunkt stellen.

Fazit aus diesem Abschnitt:

Der Digitale Humanismus setzt sich dafür ein, dass Technologie im Einklang mit menschlichen Werten gestaltet wird. Das Wiener Manifest betont die Wichtigkeit der Zusammenarbeit verschiedener Disziplinen, den Schutz der Privatsphäre, die Stärkung der Demokratie und die Förderung einer inklusiven Gesellschaft. Es ruft dazu auf, Technologien bewusst und verantwortungsvoll zu entwickeln.

GRUPPENARBEIT

Wie könnte ein persönliches Wiener Manifest aussehen?

Ziel:

Die Teilnehmer:innen entwickeln gemeinsam ein Manifest, das ihre Visionen und Prinzipien für die Anwendung des Digitalen Humanismus darstellt, ohne dass sie vorher das tatsächliche Wiener Manifest kennen.

Materialien:

- Flipcharts bzw. große Papierbögen, Moderationskärtchen
- Stifte und Marker
- Notebooks oder Tablets für Recherchen (optional)

Dauer:

Abhängig von der Anzahl der Gruppen, mindestens jedoch rund 50 Minuten

1. Brainstorming und Recherche:

Jede Gruppe führt ein Brainstorming durch, um Ideen dazu zu sammeln, welche Prinzipien und Werte ihr Manifest umfassen sollte. Themen können Datenschutz, ethische Künstliche Intelligenz, Zugänglichkeit von Technologie etc. sein.

2. Entwurf des Manifests:

Jede Gruppe formuliert auf Grundlage der gesammelten Ideen ein kurzes Manifest. Dieses sollte klar definierte Prinzipien und Ziele für die Anwendung des Digitalen Humanismus enthalten.

3. Präsentation und Diskussion:

Jede Gruppe präsentiert ihr Manifest vor den anderen Teilnehmer:innen. Nach jeder Präsentation gibt es eine kurze Diskussionsrunde.

4. Reflexion:

In einer abschließenden Reflexionsrunde sprechen die Teilnehmer:innen über den Prozess und die Lernerfahrungen und darüber, wie sie das Manifest in die Praxis umsetzen könnten.

Hinweise:

- ▶ Die Teilnehmer:innen sollten ermutigt werden, kreativ zu denken und innovative Lösungen für die Herausforderungen der digitalen Welt zu entwickeln.
- ▶ Die Bedeutung der interdisziplinären Perspektive sollte betont werden und Teilnehmer:innen aus verschiedenen Fachgebieten dazu ermutigt werden, ihre speziellen Kenntnisse einzubringen.

WISSENSFRAGEN

Der Digitale Humanismus und das Wiener Manifest

Was ist der Hauptzweck des Digitalen Humanismus?

Welche Rolle spielt die Informationstechnologie (IT) in der Gesellschaft?

Was war das Hauptziel des Wiener Workshops zu Digitalem Humanismus im April 2019?

Warum war die Teilnahme von Expert:innen aus verschiedenen Disziplinen am Wiener Workshop wichtig?

Was ist das Wiener Manifest, und welches Ziel verfolgt es?

Der Digitale Humanismus und das Wiener Manifest

Wie kann der Digitale Humanismus dazu beitragen, die negativen Auswirkungen der digitalen Transformation zu mindern?

1

Welche Herausforderungen könnten bei der Umsetzung des Digitalen Humanismus in der Praxis auftreten?

2

Mögliche Diskussion: Herausforderungen umfassen die oft gegensätzlichen Interessen von Wirtschaft, Politik und Gesellschaft. Zudem kann es schwierig sein, internationale Standards zu schaffen und aufrechtzuerhalten, da kulturelle und gesetzliche Unterschiede bestehen. Die schnelle technologische Entwicklung stellt zudem eine Herausforderung dar, da gesetzliche und ethische Rahmenbedingungen oft nicht in gleichem Tempo angepasst werden können.

Mögliche Diskussion: Der Digitale Humanismus fördert einen ethischen Umgang mit Technologie, der den Menschen in den Mittelpunkt stellt. Er kann helfen, negative Auswirkungen zu mindern, indem er darauf abzielt, dass Technologien entwickelt werden, die demokratische Werte, Privatsphäre und soziale Gerechtigkeit stärken, anstatt sie zu gefährden. Durch Initiativen wie das Wiener Manifest wird ein Rahmen geschaffen, der Standards für ethische Technologieentwicklung setzt.

Welche Vorteile bietet die fächerübergreifende Zusammenarbeit bei der Gestaltung von Technologien?

3

Mögliche Diskussion: Fächerübergreifende Zusammenarbeit bringt Perspektiven aus verschiedenen Disziplinen zusammen, was zu umfassenderen und verantwortungsvolleren Lösungen führen kann. Technologische, soziale und ethische Aspekte können dabei besser integriert werden, was zu Technologien führt, die den Bedürfnissen der Gesellschaft gerechter werden und besser auf verschiedene Herausforderungen reagieren können.

Wie können Wissenschaft, Politik und Gesellschaft dazu beitragen, dass Technologien im Einklang mit Menschenrechten entwickelt und eingesetzt werden?

4

Mögliche Diskussion: Wissenschaft, Politik und Gesellschaft können menschenrechtskonforme Technologien fördern, indem sie ethische Standards entwickeln, rechtliche Rahmenbedingungen schaffen und öffentliche Diskussionen anstoßen. Die Wissenschaft bringt Expertise ein, die Politik schafft Gesetze zum Schutz der Menschenrechte, und die Gesellschaft sorgt durch Engagement und Aufklärung dafür, dass technologische Entwicklungen im Einklang mit diesen Werten stehen.

Die Initiative Digitaler Humanismus

Das Thema des Digitalen Humanismus ist aktuell und wichtig. Nicht nur Informatiker:innen, sondern auch Expert:innen aus anderen Bereichen sowie die Zivilgesellschaft und Politiker:innen beschäftigen sich damit. Das wachsende öffentliche Interesse an den Themen KI und Digitalisierung zeigt sich in vielen internationalen Maßnahmen, z. B. durch neue Gesetze in der EU wie den Digital Services Act, den Digital Markets Act und den EU AI Act, die alle die Onlinewelt regulieren sollen (siehe auch Kapitel 5, ab Seite 65).

In Österreich unterstützen mehrere Ministerien und besonders die Wiener Stadtregierung die Initiative. Sie fördern Forschungsprojekte und Stipendien und geben öffentliche Erklärungen ab, wie die Poysdorfer Erklärung der Außenminister Österreichs, der Tschechischen Republik und der Slowakei.

Insgesamt ist es gelungen, ein Netzwerk nationaler und internationaler Expert:innen aus Wissenschaft, Verwaltung, Wirtschaft und Zivilgesellschaft aufzubauen, die den Digitalen Humanismus gemeinsam, z. B. im Zuge von Online-Vorträgen, Diskussionsreihen, Workshops, Summer Schools, Büchern, weiterentwickeln. Dieses internationale Netzwerk von klugen und innovativen Denker:innen, das ständig wächst, macht letztendlich den Kern der Initiative aus.

Fazit aus diesem Abschnitt:

Die Initiative für Digitalen Humanismus hat großes internationales Interesse und Unterstützung gefunden. Durch die Zusammenarbeit nationaler und internationaler Expert:innen aus Wissenschaft, Verwaltung, Wirtschaft und Zivilgesellschaft wird der Digitale Humanismus weiterentwickelt. Regelmäßige Aktivitäten wie Vorträge, Workshops und Veröffentlichungen tragen zur Verbreitung und Vertiefung dieses wichtigen Ansatzes bei.

GRUPPENARBEIT

Vergleich des eigenen Manifests mit dem Wiener Manifest

Ziel:

Die Teilnehmer:innen erkennen Überschneidungen bzw. Unterschiede zwischen ihrem eigenen (siehe Gruppenarbeit auf Seite 178) und dem Wiener Manifest.

Materialien:

- Flipcharts bzw. große Papierbögen, Moderationskärtchen
- Stifte und Marker
- Notebooks oder Tablets für Recherchen (optional)

Dauer:

Abhängig von der Anzahl der Gruppen, mindestens jedoch rund 50 Minuten

1. Analyse und Vergleich

Jede Gruppe analysiert das Wiener Manifest und markiert im eigenen Manifest Punkte, die damit konform gehen bzw. sich unterscheiden.

2. Präsentation und Diskussion

- ▶ Jede Gruppe präsentiert ihre Erkenntnisse vor den anderen Teilnehmer:innen.
- ▶ Nach jeder Präsentation gibt es eine kurze Diskussionsrunde.

3. Reflexion

In einer abschließenden Reflexionsrunde sprechen die Teilnehmer:innen über den Prozess und die Lernerfahrungen und wie sie das Wiener Manifest in die Praxis umsetzen könnten.

Zukünftige Aufgaben

Das Internet und das Web sind zu einem öffentlichen Gut geworden. Der Digitale Humanismus zielt darauf ab, Technologie im Einklang mit menschlichen Werten zu gestalten. Um dies zu erreichen, müssen wir auch über Macht und Politik sprechen, da die technologische Entwicklung oft von ungleichen Machtverhältnissen geprägt ist. Der Digitale Humanismus braucht einen mehrdimensionalen Ansatz auf unterschiedlichen Ebenen:

- **Unterschiedliche inhaltliche Bereiche:** von Plattformen über KI und Datenschutz oder Arbeit und bis hin zur Entwicklung technischer Systeme.
- **Unterschiedliche Disziplinen:** Informatik, Ingenieurwissenschaften, Sozial- und Geisteswissenschaften müssen zusammenarbeiten.
- **Unterschiedliche Aktivitäten:** Forschung, Entwicklung, Bildung, Kommunikation und politische Maßnahmen sind notwendig.
- **Unterschiedliche gesellschaftliche Akteur:innen:** Von Universitäten und Schulen, Firmen, Interessenvertretungen und öffentlichen Institutionen bis hin zu den Bürger:innen und ihren Selbstorganisationen.

Der Digitale Humanismus konzentriert sich darauf, Technologien zu entwickeln, die den Menschen unterstützen, den Zugang zu Wissen erleichtern, die Teilhabe an der Gesellschaft ermöglichen und die Vielfalt fördern.

Dabei gibt es viele offene Fragen, wobei bei der Lösung die Zusammenarbeit verschiedener Disziplinen notwendig ist. Ein Beispiel ist das Konzept der Fairness: Wie können wir sicherstellen, dass KI fair und unvoreingenommen ist? Wie soll die Arbeit zwischen Menschen und Maschinen aufgeteilt werden? Wie können wir eine faire Beteiligung der Nutzer:innen an Plattformen gewährleisten? Wie können wir eine faire Verteilung von Online-Inhalten sicherstellen? Wie können wir bei Umweltfragen die verschiedenen Interessen fair vertreten?

Dieses Beispiel zeigt, und es gibt viele andere, die Notwendigkeit der Interdisziplinarität sowie der Kooperation, auch bei der Entwicklung und Anwendung neuer Technologien.

Fazit aus diesem Abschnitt:

Der Digitale Humanismus erfordert einen ganzheitlichen Ansatz, der verschiedene Problembereiche, Disziplinen und Aktivitäten umfasst. Es geht darum, Technologien zu entwickeln, die den Menschen unterstützen und faire sowie nachhaltige Lösungen bieten. Eine interdisziplinäre Zusammenarbeit ist entscheidend, um die Herausforderungen der modernen Informationstechnologie zu bewältigen.

Schlussbetrachtungen

Der Digitale Humanismus bietet eine disziplinenübergreifende und ethische Perspektive auf die Technologieentwicklung. Er betont, dass der wirtschaftliche und gesellschaftliche Wandel, der durch IT mitverursacht wird, auch eine politische Frage ist. Da es keine vorbestimmte Entwicklung gibt und kein höheres Wesen dafür verantwortlich ist, sollten wir, die Menschen, die treibende Kraft sein. Dies sollte in Form eines demokratischen und partizipativen Ansatzes organisiert sein. Es ist wichtig, dass wir uns nicht nur auf die Optimierung und Beschleunigung von Technologien konzentrieren, sondern eine langfristige und nachhaltige Perspektive einnehmen.

Die Ziele des Digitalen Humanismus sind einfach zu formulieren und zu verstehen. Sie zu erreichen, ist eine große Herausforderung.

Wiener Manifest für Digitalen Humanismus

Wien, Mai 2019

„**The system is failing**“ – so Tim Berners-Lee. Der Gründer des Web betont, dass die Digitalisierung zwar beispiellose Möglichkeiten eröffnet, aber auch ernste Bedenken aufwirft: die Monopolisierung des Web, die Ausbreitung extremistischer Verhaltensmuster, die von sozialen Medien orchestriert werden, ebenso wie Filterblasen und Echokammern als Inseln entkoppelter „Wahrheiten“, der Verlust der Privatsphäre sowie die weite Verbreitung digitaler Überwachungstechnologien. Digitale Technologien verändern die Gesellschaft fundamental und stellen unser Verständnis davon infrage, was unsere Existenz als Menschen ausmacht. Viel steht auf dem Spiel. Die Herausforderung einer gerechten und demokratischen Gesellschaft mit dem Menschen im Zentrum des technologischen Fortschritts muss mit Entschlossenheit und wissenschaftlichem Einfallsreichtum bewältigt werden. Technologische Innovation erfordert soziale Innovation, und diese erfordert ein breites gesellschaftliches, demokratisches Engagement.

Dieses Manifest ist ein Aufruf zum Nachdenken und Handeln angesichts der aktuellen und zukünftigen technologischen Entwicklung. Wir wenden uns an unsere akademischen Communitys, aber auch an Pädagog:innen, Führungskräfte aus der Industrie, Politiker:innen, politische Entscheidungsträger:innen und Fachgesellschaften auf der ganzen Welt: Mischt Euch ein und beteiligt Euch! Es geht um die Mitgestaltung der Politik mittels Expertise und öffentlichen Engagements, wo und wie auch immer das möglich ist. Unsere Forderungen sind das Ergebnis eines sich entfaltenden Prozesses, der Wissenschaftler:innen und Praktiker:innen aus verschiedenen Bereichen und Themen zusammenbringt und der von Sorgen und Hoffnungen um die Zukunft geprägt ist. Wir sind uns unserer gemeinsamen Verantwortung für die aktuelle Situation und die Zukunft bewusst – sowohl als Expert:innen als auch als Bürger:innen.

Wir erleben die Co-Evolution von Technologie und Mensch.

Die Metamorphose vom eigenständigen Computer zu einem weltweit vernetzten System führt zu einer globalen industriellen und

gesellschaftlichen Revolution. Eine Flut von Daten, Algorithmen und Rechenleistung beeinflusst unser gesellschaftliches Gefüge in fundamentaler Weise. Menschliche Interaktionen, gesellschaftliche Institutionen, Volkswirtschaften, politische Strukturen sowie die Wissenschaften werden verändert. Dieser Wandel schafft und bedroht Arbeitsplätze, schafft und zerstört Wohlstand, verschiebt Machtstrukturen und hat massive ökologische Auswirkungen, im Guten wie im Schlechten. Die bisherigen Grenzen zwischen dem Persönlichen und Professionellen, dem Privaten und Öffentlichen sowie zwischen Mensch und Maschine verschwimmen.

Im Zentrum steht der Ruf nach Aufklärung und Humanismus.

Der revolutionäre Aspekt der Informatik hängt mit ihrer Fähigkeit zusammen, menschliche Aktivitäten zu automatisieren. Bereits jetzt übertreffen Maschinen bei vielen Aufgaben das, was der Mensch an Geschwindigkeit, Präzision und sogar analytischer Ableitung leisten kann. Es ist an der Zeit, humanistische Ideale mit einer kritischen Reflexion des technischen Fortschritts zu kombinieren. Wir verknüpfen dieses Manifest daher mit der intellektuellen Tradition des Humanismus, die auf dem Weg zu einer aufgeklärten Moderne stets im Zentrum gestanden ist.

Wie alle Technologien entstehen auch digitale Technologien nicht aus dem Nichts. Sie sind durch implizite und explizite Entscheidungen geprägt und beinhalten Werte, Normen, wirtschaftliche Interessen und Annahmen darüber, wie die Welt ist oder sein sollte. Viele dieser Optionen bleiben in Softwareprogrammen und deren Algorithmen verborgen. In Anlehnung an die Tradition des renommierten Wiener Kreises und anderer Denkströmungen der Moderne drängen wir auf kritisches rationales Denken und interdisziplinäre Zusammenarbeit, um die Zukunft aktiv zu gestalten.

Wir müssen Technologien nach menschlichen Werten und Bedürfnissen formen, anstatt nur zuzulassen, dass Technologien Menschen formen. Unsere Aufgabe besteht nicht nur darin, die Nachteile der Informations- und Kommunikationstechnologien einzudämmen, sondern vor allem auch darin, von Beginn an menschenzentrierte Innovationen zu fördern. Wir fordern einen Digitalen Humanismus, der das komplexe Zusammenspiel von Technologie und Menschheit beschreibt, analysiert und vor allem beeinflusst, für eine bessere Gesellschaft und ein besseres Leben unter voller Achtung universeller Menschenrechte.

Zusammenfassend verkünden wir folgende Kernforderungen:

- **Digitale Technologien sollen so gestaltet sein, dass sie Demokratie und Inklusion fördern.** Dies wird besondere Anstrengungen erfordern, um derzeitige Ungleichheiten zu überwinden und das emanzipatorische Potenzial digitaler Technologien zu nutzen – und damit unsere Gesellschaft inklusiver gestalten zu können.
- **Privatsphäre und Redefreiheit sind Grundwerte, die im Mittelpunkt unserer Aktivitäten stehen sollen.** Daher müssen Strukturen wie soziale Medien oder Online-Plattformen derart geändert werden, dass freie Meinungsäußerung, Verbreitung von Informationen und Privatsphäre besser geschützt sind.
- **Es müssen wirksame Vorschriften, Gesetze und Regeln festgelegt werden, die auf einem breiten Diskurs beruhen.** Sie sollen Fairness und Gleichheit, Verantwortlichkeit und Transparenz von Softwareprogrammen und Algorithmen sicherstellen.
- **Die Regulierungsbehörden müssen gegenüber Technologie-monopolen intervenieren.** Die Wettbewerbsfähigkeit des Markts muss wiederhergestellt werden, da Tech-Monopole die Marktmacht ausnutzen und Innovationen ersticken. Staaten sollen nicht alle Entscheidungen den Märkten überlassen.
- **Entscheidungen, deren Folgen die individuellen oder kollektiven Menschenrechte betreffen können, müssen weiterhin vom Menschen getroffen werden.** Die Entscheidungsträger:innen müssen für ihre Entscheidungen verantwortlich und haftbar sein. Automatisierte Entscheidungssysteme sollen die Entscheidungsfindung durch den Menschen nur unterstützen und nicht ersetzen.
- **Wissenschaftliche Ansätze über verschiedene Disziplinen hinweg sind Voraussetzung, um die Herausforderungen zu meistern.** Technologische Disziplinen wie die Informatik müssen mit den Sozial-, Geistes- und Naturwissenschaften zusammenarbeiten, um disziplinäre „Silos“ zu durchbrechen.
- **Universitäten sind der Ort, an dem neues Wissen erzeugt und kritisches Denken geschärft wird.** Sie haben daher eine besondere Verantwortung und sollen sich dessen auch bewusst sein.
- **Akademische und industrielle Forscher:innen müssen sich in einem offenen Dialog mit der Gesellschaft auseinandersetzen und ihre Ansätze reflektieren.** Diese Überlegungen müssen in die tatsächlichen Praktiken der Erzeugung neuen Wissens und neuer Technologien eingebettet werden und gleichzeitig die Freiheit des Denkens und der Wissenschaft verteidigen.

- **Praktiker:innen auf der ganzen Welt sollen sich der gemeinsamen Verantwortung für die Auswirkungen der Informationstechnologien stellen.** Es muss ein gemeinsames Verständnis darüber entstehen, dass keine Technologie neutral ist: Sowohl deren potenzielle Vorteile als auch mögliche Nachteile müssen erkannt und von Praktiker:innen weltweit in der Umsetzung berücksichtigt werden.
- **Es bedarf einer Vision für neue Bildungsinhalte, die Wissen aus den Geistes-, Sozial- und Ingenieurwissenschaften kombinieren.** Im Zeitalter der automatisierten Entscheidungsfindung und Künstlichen Intelligenz sind Kreativität, Reflexion und die Berücksichtigung menschlicher Aspekte für die Ausbildung zukünftiger Informatiker:innen und anderer Berufsgruppen von entscheidender Bedeutung.
- **Die Ausbildung in der Informatik und die Bildungsarbeit hinsichtlich ihrer gesellschaftlichen Auswirkungen müssen so früh wie möglich beginnen.** Ziel ist es, den Lernenden zu ermöglichen, sowohl IKT-Kenntnisse als auch das Bewusstsein für die betreffenden Themen zu entwickeln.

Wir stehen an einem wichtigen Kreuzungspunkt in Richtung Zukunft. Wir müssen handeln und die richtige Richtung einschlagen!

A/B-Testing: vergleicht zwei Versionen (A und B) eines Elements, um herauszufinden, welche besser funktioniert. Nutzer:innen werden zufällig aufgeteilt, und ihr Verhalten (z. B. Klickrate) wird analysiert. So können datenbasierte Entscheidungen getroffen und die erfolgreichere Variante identifiziert werden.

AGI (Artificial General Intelligence, Allgemeine Künstliche Intelligenz): ein Programm/System, das Aufgaben und Probleme ähnlich wie ein Mensch lösen kann. Im Gegensatz zur spezialisierten, schwachen KI, die nur für bestimmte Aufgaben geeignet ist, wäre die AGI flexibel und anpassungsfähig. Sie könnte lernen, Schlussfolgerungen zu ziehen und sich neuen Herausforderungen stellen. Ziel ist eine universelle Intelligenz, die menschliches Denken und Lernen nachahmt. Sie existiert bisher nur als Konzept.

Algorithmus: eine schrittweise Anleitung zur Lösung eines Problems oder zur Durchführung einer Aufgabe. Er besteht aus einer Reihe von Anweisungen, die in einer bestimmten Reihenfolge ausgeführt werden, um ein gewünschtes Ergebnis zu erzielen.

ARPANET (Advanced Research Projects Agency Network): erstes paketvermittelndes Computernetzwerk und Vorläufer des Internets. Es wurde in den 1960er Jahren vom US-Verteidigungsministerium entwickelt, um Forschungseinrichtungen und Universitäten zu vernetzen. Durch die Technik der Paketvermittlung war das Netzwerk robust und konnte auch bei Ausfällen funktionieren. Die erste Nachricht wurde 1969 übertragen. Es legte den Grundstein für das Internet und wurde durch dieses abgelöst.

Backpropagation-Algorithmus: ein Verfahren zum Trainieren künstlicher

neuronaler Netzwerke. Er minimiert den Fehler zwischen der vorhergesagten und tatsächlichen Ausgabe, indem die Gewichte im Netzwerk angepasst werden. Im Vorwärtsdurchlauf berechnet das Netzwerk eine Ausgabe basierend auf den Eingaben. Im Rückwärtsdurchlauf wird der Fehler zurückpropagiert (= rückgemeldet), um die Gewichtungen zu korrigieren. Der Algorithmus passt damit die Verbindungen in einem neuronalen Netz an, um die Vorhersagen zu verbessern, indem er Fehler im System minimiert. Backpropagation ist entscheidend für die Effektivität moderner neuronaler Netzwerke im maschinellen Lernen und zentraler Bestandteil von deren Trainingsprozessen.

Behavioral Targeting: Marketingstrategie, bei der Werbung gezielt auf das Verhalten und die Interessen von Nutzer:innen zugeschnitten wird. Durch die Analyse von Onlineaktivitäten wie besuchten Websites, Klicks, Suchanfragen und Verweildauer wird ein Profil der Nutzer:innen erstellt. Auf dieser Basis erhalten Nutzer:innen individuell angepasste Werbeanzeigen, die ihre Interessen widerspiegeln.

Bias: systematische Verzerrung oder Voreingenommenheit bei Entscheidungen, Wahrnehmungen oder Datenauswertungen. Im Kontext von Daten und Künstlicher Intelligenz tritt Bias auf, wenn Algorithmen durch unausgewogene oder fehlerhafte Daten bestimmte Gruppen bevorzugen oder benachteiligen. Bias kann unbewusst oder durch fehlerhafte Annahmen entstehen und beeinflusst die Objektivität von Ergebnissen.

Blackbox: ein System oder ein Prozess, dessen inneres Funktionieren nicht sichtbar oder nicht verständlich ist. Man sieht nur die Eingaben und die Ausgaben, kennt aber nicht die genauen

Abläufe dazwischen. In der Künstlichen Intelligenz wird der Begriff verwendet, da komplexe Systeme wie neuronale Netze Entscheidungen treffen, ohne dass der genaue Entscheidungsprozess für Menschen nachvollziehbar ist.

CAD (Computer-Aided Design): Software zur Erstellung und Bearbeitung von technischen Zeichnungen oder auch 3D-Modellen. Es wird vor allem im Ingenieurwesen, in der Architektur und im Produktdesign genutzt, um präzise Entwürfe digital zu erstellen. CAD ermöglicht eine hohe Genauigkeit und spart Zeit im Vergleich zu manuellen Zeichnungen, da Änderungen und Simulationen einfach umgesetzt werden können.

Cloud-Computing: die Bereitstellung von IT-Ressourcen wie Speicherplatz, Rechenleistung und Anwendungen über das Internet („die Cloud“) statt lokal auf dem eigenen Computer. Nutzer:innen können flexibel auf diese Ressourcen zugreifen und zahlen oft nur für die tatsächlich genutzten Dienste. Es ermöglicht, Daten und Anwendungen ortsunabhängig zu nutzen und bei Bedarf zu skalieren, ohne eigene Server zu betreiben.

Computational Thinking: Ansatz zur Problemlösung, der sich an Denkweisen aus der Informatik orientiert. Es umfasst das Zerlegen komplexer Probleme in kleinere, die Abstraktion relevanter Informationen und das Entwickeln schrittweiser Lösungen (Algorithmen). Computational Thinking hilft, systematisch und effizient an Aufgaben heranzugehen, es wird neben der Informatik auch in anderen Bereichen angewandt.

CPU (Central Processing Unit): das zentrale Rechenwerk eines Computers, es wird oft als „Gehirn“ des Systems

bezeichnet. Sie führt Berechnungen und Befehle aus, steuert alle Abläufe und verarbeitet Daten nach den Anweisungen der Software. Die CPU beeinflusst maßgeblich die Geschwindigkeit und Effizienz eines Computers.

Dartmouth-Konferenz (Dartmouth Summer Research Project on Artificial Intelligence): fand im Sommer 1956 am Dartmouth College in Hanover, New Hampshire, USA, statt. Sie gilt als Geburtsstunde der Künstlichen Intelligenz als eigenständiges Forschungsgebiet. Sie wurde initiiert von John McCarthy, Marvin Minsky, Claude Shannon und anderen. Dort prägte McCarthy den Begriff „Artificial Intelligence“. Ziel war es, die Möglichkeiten zu erforschen, Maschinen mit menschenähnlicher Intelligenz auszustatten, einschließlich der Fähigkeit, Sprache zu nutzen oder Probleme zu lösen, die dem Menschen vorbehalten war.

Deep Learning: ein Bereich des maschinellen Lernens, der auf künstlichen neuronalen Netzen basiert und bei komplexen Aufgaben wie der Bildererkennung eingesetzt wird. Deep Learning ermöglicht es Systemen, aus großen Datenmengen zu lernen und Muster zu erkennen. Der Begriff „Deep“ bezieht sich auf die große Anzahl von Schichten im Netzwerk, bis zu 130.

Differential Privacy: Datenschutzkonzept, das ermöglicht, Einzeldaten in Datensätzen zu schützen, während gleichzeitig nützliche Informationen für Analysen erhalten bleiben. Dabei wird durch gezieltes statistisches „Rauschen“ (= Zufallsschwankungen) verhindert, dass Rückschlüsse auf einzelne Personen möglich sind. Differential Privacy wird vor allem in großen Datenanalysen genutzt, um die Privatsphäre zu wahren und dennoch aussagekräftige Ergebnisse zu erhalten.

Digitaler Humanismus: ein Ansatz, der den Menschen in den Mittelpunkt der technologischen Entwicklung stellt. Er analysiert das komplexe Zusammenspiel von Mensch und Technologie, und versucht, dies zu beeinflussen, für eine nachhaltige und bessere Gesellschaft unter Berücksichtigung der Menschenrechte.

DOI (Digital Object Identifier): eindeutige, permanente Kennzeichnung für digitale Objekte wie wissenschaftliche Artikel, Bücher oder Datensätze. Er dient dazu, diese Objekte im Internet eindeutig auffindbar und verlinkbar zu machen. Jeder DOI verweist auf eine spezifische URL, die das jeweilige Objekt zugänglich macht, auch wenn es den Ort ändert. DOIs sind besonders in Forschung und Lehre weitverbreitet.

E-Commerce: bezeichnet den Kauf und Verkauf von Waren und Dienstleistungen über das Internet. Kund:innen können dabei Produkte online durchsuchen, vergleichen und direkt bestellen. Händler nutzen E-Commerce-Plattformen, um ein breites Publikum zu erreichen, oft ohne physische Läden. Bekannte E-Commerce-Modelle sind B2C (Business-to-Consumer), B2B (Business-to-Business) und C2C (Consumer-to-Consumer). E-Commerce bietet Bequemlichkeit und große Produktauswahl, kann aber auch Herausforderungen wie Datenschutz und Retourenmanagement mit sich bringen.

Echokammer: entsteht, wenn Menschen vor allem Informationen konsumieren, die ihre eigenen Meinungen bestätigen. Algorithmen in sozialen Netzwerken fördern dies, indem sie Inhalte zeigen, die den Interessen der Nutzer:innen entsprechen. So entstehen „Filterblasen“, in denen abweichende Meinungen kaum noch vorkommen. Dadurch werden eigene Überzeugungen verstärkt, was

die Wahrnehmung der Realität verzerren kann.

Expertensystem: Computersystem, das menschliches Fachwissen simuliert, um Entscheidungen oder Problemlösungen in bestimmten Bereichen zu unterstützen. Es besteht aus einer Wissensbasis, die Fachwissen speichert, und einer Inferenzmaschine. Dies ist ein Programm, das formalen logischen Regeln folgend mithilfe der Wissensbasis logische Schlussfolgerungen zieht, also das Problem löst. Expertensysteme werden in Bereichen wie Medizin oder Finanzen eingesetzt, um bei der Analyse und Entscheidungsfindung zu helfen. Sie können komplexe Aufgaben bewältigen, sind jedoch auf ihr spezifisches Fachgebiet begrenzt.

Filterblase: beschreibt den Zustand, in dem Algorithmen bevorzugt solche Inhalte anzeigen, die den bisherigen Interessen der Nutzer:innen entsprechen. Dies führt oft zu eingeschränktem Informationsfluss und verhindert, dass Nutzer:innen eine breitere Meinungsvielfalt wahrnehmen.

Gatekeeper: eine Instanz, die den Zugang zu Informationen steuert und entscheidet, welche Inhalte Nutzer:innen sehen. Plattformen wie soziale Netzwerke oder Nachrichtenportale übernehmen oft diese Rolle, indem sie mithilfe von Algorithmen Inhalte filtern und priorisieren. Dadurch beeinflussen Gatekeeper, welche Inhalte für Nutzer:innen sichtbar oder verborgen bleiben, was die Meinungsbildung und Informationsvielfalt prägen kann.

Gewichte: Zahlen in einem neuronalen Netzwerk, die bestimmen, wie stark die Eingaben jedes Neurons bei der Berechnung der Ausgabe berücksichtigt werden. Sie werden während des Trainings eines Netzwerks angepasst.

GPU (Graphics Processing Unit):

spezialisierter Prozessor, der ursprünglich zur Berechnung und Darstellung von Grafiken entwickelt wurde. Anders als die CPU ist die GPU auf parallele Datenverarbeitung ausgelegt, wodurch sie große Datenmengen effizient verarbeiten kann. GPUs werden heute auch in Bereichen wie KI und Machine-Learning eingesetzt, da sie komplexe Berechnungen schneller als herkömmliche CPUs durchführen können.

GUI (Graphical User Interface):

grafische Benutzeroberfläche, die es Nutzer:innen ermöglicht, mit einem Computer oder Programm über visuelle Elemente wie Fenster, Symbole und Menüs zu interagieren, anstatt Befehle einzugeben. Sie macht Software intuitiver und benutzerfreundlicher.

HCI (Human-Computer Interaction):

untersucht, wie Menschen mit Computern interagieren, um diese Interaktionen möglichst intuitiv zu gestalten. Es geht um die Gestaltung von Benutzeroberflächen und Systemen, die leicht verständlich und nutzerfreundlich sind. HCI verbindet Informatik, Psychologie und Design. Ein Ziel ist es, Technologien so zu entwickeln, dass sie optimal an die Bedürfnisse der Nutzer:innen angepasst sind.

HTML (Hypertext Markup Language):

Standardauszeichnungssprache für die Strukturierung und Darstellung von Inhalten im Web. Sie definiert die Struktur einer Webseite mithilfe von „Tags“, die angeben, wie verschiedene Elemente (z. B. Texte, Bilder, Links) angezeigt werden sollen. HTML bildet das Grundgerüst jeder Webseite und ermöglicht es Nutzer:innen, im Web zu navigieren bzw. zu surfen.

HTTP (Hypertext Transfer Protocol):

Netzwerkprotokoll, das den Datenaus-

tausch im Web regelt und den Zugriff auf Websites und deren Inhalte ermöglicht. Es wird verwendet, um Anfragen von einem Client (wie einem Browser) an einen Server zu senden und die Antworten (wie HTML-Seiten oder andere Daten) zu empfangen. HTTP ist die Basis für die Kommunikation im Web; die sichere Version davon ist HTTPS (s = secure), das die Datenübertragung verschlüsselt.

Informatisierung: Prozess, bei dem immer mehr Lebens- und Arbeitsbereiche durch digitale Technologien und Informationssysteme unterstützt oder auch automatisiert werden. Ziel ist es, Abläufe effizienter, schneller und oft auch kostengünstiger zu gestalten. Dies betrifft fast alle Bereiche unserer Wirtschaft und Gesellschaft. Informatisierung verändert, wie Informationen gesammelt, verarbeitet und genutzt werden.

Internet: globales Netzwerk von miteinander verbundenen Computern und Servern. Es ermöglicht Menschen auf der ganzen Welt, Informationen auszutauschen und miteinander zu kommunizieren. Es besteht aus einer Vielzahl von Netzwerken, die durch standardisierte Protokolle (siehe TCP/IP) miteinander verbunden sind.

Wie werden Daten von A nach B übertragen? Zuerst werden Daten (E-Mails oder Bilder oder ...) in Pakete aufgeteilt, wobei jedes Paket „weiß“, woher es kommt, wohin es soll und in welcher Reihenfolge die Pakete nach ihrer Ankunft am Ziel zusammengesetzt werden sollen. Jedes Paket wird einzeln durch das Netz geleitet, wobei jedes Paket einen anderen Weg nehmen und auf Störungen flexibel reagieren kann. Daher ist das Internet sehr ausfallsicher und kann zudem nicht durch die Störung eines Knotens zerstört werden. Wie in einem großen Straßennetz gibt es mehrere Wege von A nach B. Daran erkennt man,

dass das Design dieses resilienten Netzwerks militärischen Ursprung hat.

IP-Adressen (Internet Protocol Address): Jeder Computer im Internet hat eine eindeutige IP-Adresse, ähnlich wie eine Postadresse. Diese Adresse dient zur Identifikation und ermöglicht, dass Informationen den richtigen Empfänger erreichen.

IT (Informationstechnologie): alle Technologien, die zur Verarbeitung, Speicherung und Übertragung von Informationen eingesetzt werden. Dazu gehören Hardware wie Computer und Server, Netzwerke, Software und Speichersysteme. IT ist in nahezu allen Lebensbereichen präsent und bildet die Grundlage moderner Kommunikation, Datenanalyse und Automatisierung. Unternehmen und Organisationen nutzen IT, um Arbeitsprozesse effizienter zu gestalten, Daten zu verwalten und globale Vernetzung zu ermöglichen.

Internet der Dinge (Internet of Things, IoT): verbindet physische Geräte wie Haushaltsgeräte, Autos und Maschinen mit dem Internet. Diese Geräte sammeln Daten und tauschen sie aus, um Abläufe zu automatisieren und zu optimieren. Beispiele sind Smart Homes (z. B. vernetzte Thermostate und Leuchten), Industrieanlagen, die sich selbst überwachen, oder Wearables im Gesundheitsbereich. IoT ermöglicht Echtzeitanalysen und erleichtert alltägliche Prozesse. Neben den Vorteilen bringt IoT jedoch auch Herausforderungen im Bereich Datenschutz und Sicherheit mit sich.

KI (Künstliche Intelligenz): Bereich der Informatik zur Entwicklung von Systemen, die Aufgaben ausführen können, die normalerweise menschliche Intelligenz erfordern, wie Lernen, Problemlösung oder Sprachverstehen. Beispiele

sind Sprachassistenten, Bilderkennung und selbstfahrende Autos. Ziel der KI ist es, dass Maschinen ein Verhalten zeigen, das menschliches Denken und Handeln simuliert.

KI-Governance: Regeln, Prozesse und Gesetze, um Künstliche Intelligenz sicher, ethisch und transparent zu gestalten und einzusetzen. Sie zielt darauf ab, Risiken wie Diskriminierung und Datenschutzverletzungen zu minimieren, aber auch Innovation zu fördern. Dazu gehören Richtlinien zur Nachvollziehbarkeit von KI-Entscheidungen, ethische Standards, Regulierungen zur Haftung, aber auch Gesetze wie der EU AI Act. KI-Governance stärkt das Vertrauen in KI-Technologien und soll für ihren verantwortungsvollen Einsatz sorgen.

Klassifikation: Verfahren im Bereich des maschinellen Lernens und der Datenanalyse, bei dem Daten in vordefinierte Kategorien oder Klassen eingeteilt werden. Ein Klassifikationsmodell lernt aus Beispieldaten (Trainingsdaten), um zukünftige Daten korrekt zuzuordnen. Beispiele sind das Erkennen von E-Mails als Spam oder Nicht-Spam oder auch das Kategorisieren von Bildern. Ziel ist es, unbekannte Daten möglichst genau richtigen Klassen zuzuteilen.

Lean-UX-Ansatz: Methode im User Experience Design (zielt auf Verbesserung des Nutzererlebnisses), die sich auf schnelle und iterative Entwicklungsprozesse konzentriert. Anstatt umfassende Dokumentationen zu erstellen, arbeiten Designer:innen, Entwickler:innen und andere Teammitglieder eng zusammen, um schnell Prototypen zu erstellen und frühzeitig Feedback von Nutzer:innen einzuholen. Lean UX fördert auch Zusammenarbeit und Flexibilität.

Mikroprozessor: zentrales Rechen-element eines Computers, das Anwei-

sungen verarbeitet und Berechnungen durchführt. Er besteht aus Millionen winziger Transistoren und führt grundlegende arithmetische, logische und steuernde Operationen aus, die für das Funktionieren von Software und Hardware erforderlich sind. Ein Mikroprozessor ist oft die CPU eines Computers, aber er kann auch andere Aufgaben übernehmen, etwa die Steuerung eines Mikrokontrollers. Mikroprozessoren befinden sich daher nicht nur in Computern, sondern auch in vielen anderen Geräten wie Autos, Haushaltsgeräten oder Industrieanlagen.

Netzwerkeffekte: Je mehr Nutzer:innen auf einer Plattform sind, desto attraktiver wird sie für weitere Nutzer:innen.

Neuron: Grundbaustein eines neuronalen Netzes beim maschinellen Lernen, inspiriert von biologischen Neuronen im Gehirn. Es ist eine Einheit, die Daten empfängt, verarbeitet und an andere Neuronen weiterleitet. Es „lernt“ durch Anpassung über viele Iterationen (= mehrfache Wiederholungen) im Training, um das gewünschte Verhalten zu erreichen.

Neuronales Netz: Methode der KI, die dem Aufbau des menschlichen Gehirns nachempfunden ist. Ein neuronales Netz besteht aus vernetzten Neuronen und Synapsen, die in Schichten angeordnet sind und Daten verarbeiten. Durch das Training lernt das Netzwerk, Muster in den Daten zu erkennen. Neuronale Netze werden für Aufgaben wie Bild- und Spracherkennung genutzt. Sie sind zentral für viele Anwendungen des maschinellen Lernens.

Paketvermittlung: bedeutet im Internet, dass Daten in kleine Pakete zerlegt werden, die dann unabhängig voneinander über verschiedene Routen zum Ziel gesendet werden. Jedes Paket enthält

Ziel- und Reihenfolgedaten, sodass es bei den Empfänger:innen korrekt zusammengesetzt werden kann. Diese Technik macht das Netzwerk robust gegen Ausfälle. Sie bildet die Grundlage für die flexible und belastbare Struktur des Internets.

Plattformökonomie: Geschäftsmodell, bei dem digitale Plattformen als Vermittler zwischen Anbietern und Nachfragern agieren. Bekannte Beispiele sind Amazon, Uber oder Airbnb. Plattformen bringen Käufer:innen und Verkäufer:innen zusammen und ermöglichen den Austausch von Produkten, Dienstleistungen oder Informationen. Sie profitieren von Netzwerkeffekten. Plattformökonomie hat viele Märkte verändert und oft traditionelle Geschäftsmodelle verdrängt. Sie führt auch zu dem Winner-takes-it-all-Phänomen.

Protokolle: festgelegte Regeln oder Standard, der definiert, wie Daten zwischen Geräten in einem Netzwerk übertragen werden sollen. Sie legen fest, wie Daten gesendet, empfangen und interpretiert werden, sodass verschiedene Geräte miteinander kommunizieren können, auch wenn sie von unterschiedlichen Herstellern stammen oder unterschiedliche Systeme nutzen. Auch das Internet funktioniert durch eine Reihe solcher Protokolle, wie zum Beispiel das TCP/IP-Protokoll, das die Regeln für die Datenübertragung und die Datenflusskontrolle definiert; oder HTTP, das für die Übertragung von Websites verwendet wird.

Router: dedizierte Computer, die Datenpakete durch das Internet leiten und dafür sorgen, dass sie ihren Weg zum Ziel finden.

Schwache KI: Form der Künstlichen Intelligenz, die für eine spezifische Aufgabe oder einen begrenzten An-

wendungsbereich entwickelt wurde. Sie ist darauf spezialisiert, bestimmte Probleme zu lösen, wie Sprachassistentz, Bilderkennung oder Schachspielen, und verfügt nicht über allgemeine Intelligenz oder Bewusstsein. Schwache KI kann nur das leisten, wofür sie programmiert wurde.

Server: Computer, der Daten, Dienste oder Ressourcen über ein Netzwerk für andere Computer, sogenannte Clients, bereitstellt. Er kann physisch (eigener Rechner) oder virtuell (z. B. Cloud-Server) sein.

Starke KI: siehe AGI.

Synapse: in einem neuronalen Netz die Verbindung zwischen zwei Neuronen, die Daten weiterleitet und eine wichtige Rolle im Lernprozess spielt.

TCP/IP (Transmission Control Protocol/Internet Protocol): grundlegendes Protokollsystem des Internets. Es besteht aus zwei Hauptkomponenten: a) IP (Internet Protocol), das für die Adressierung und das Routing von Datenpaketen zuständig ist. Es sorgt dafür, dass Pakete den Weg durch das Netzwerk finden und beim Empfänger ankommen. b) TCP (Transmission Control Protocol) sichert die zuverlässige Übertragung der Daten. Es stellt sicher, dass alle Pakete vollständig und in der richtigen Reihenfolge ankommen, indem es verlorene Pakete neu anfordert. Beide zusammen ermöglichen die zuverlässige und geordnete Kommunikation im Internet.

Training eines neuronalen Netzes: Prozess, in dem das Netzwerk lernt, aus Eingabedaten die richtigen Ausgaben zu erzeugen. Hierbei werden die Gewichte der Verbindungen (Synapsen) zwischen den Neuronen angepasst, um den Fehler zu minimieren. Dieser Prozess wiederholt sich über viele Durchläufe, bis der

Fehler minimal ist und das Netzwerk gut auf die Trainingsdaten passt.

Transistor: winziges elektronisches Bauteil, das als Schalter oder Verstärker für Signale dient. Er kann den Stromfluss steuern und ist daher ein Grundbaustein moderner Elektronik. Transistoren bestehen meist aus Halbleitermaterialien wie Silizium und werden in großer Zahl in Mikroprozessoren und anderen Schaltkreisen verbaut. Sie ermöglichen komplexe Berechnungen und Steuerungen.

UGC (User Generated Content): Inhalte, die von Nutzer:innen einer Plattform oder eines Diensts selbst erstellt und hochgeladen werden anstatt von professionellen Content-Produzent:innen. Dies kann Texte, Fotos, Videos, Rezensionen oder Kommentare umfassen. Beispiele sind Instagram-Bilder, YouTube-Videos, Produktbewertungen auf Amazon oder Wikipedia-Einträge. UGC fördert oft die Interaktivität und Authentizität von Plattformen, da die Inhalte von realen Nutzer:innen stammen.

URL (Uniform Resource Locator): eindeutige Webadresse, die zu einer spezifischen Ressource im Internet führt, wie z. B. einer Webseite, einem Bild oder einem Dokument. URLs ermöglichen den Zugriff auf Inhalte und Ressourcen durch standardisierte Adressen im Web.

Virtualisierung: beschreibt den Prozess, in dem Aufgaben, die zuvor Hardwarekomponenten erfüllten, auf Software übertragen werden. Ein Beispiel ist das Auto, wo vermehrt Softwarekomponenten wichtig sind. Dabei schaffen spezielle Programme eine virtuelle Umgebung auf physischer Hardware. Software ist flexibler und billiger als Hardware, und es ermöglicht eine effizientere Ressourcennutzung und flexible Skalierung.

Wearables: tragbare elektronische Geräte, die am Körper getragen werden und verschiedene Funktionen erfüllen. Beispiele sind Smartwatches, Fitness-tracker oder Smart-Brillen. Sie erfassen Daten wie Schritte, Puls oder Schlaf und helfen so, die Gesundheit zu überwachen oder Benachrichtigungen anzuzeigen. Wearables verbinden sich oft mit Smartphones, um Daten zu synchronisieren und zusätzliche Funktionen zu bieten.

Wertschöpfungskette: beschreibt die Stufen, die ein Produkt oder eine Dienstleistung durchläuft, um einen Mehrwert zu schaffen – von der Idee bis zum Endprodukt. Das Konzept der Wertschöpfungskette dient Unternehmen auch zur Analyse und Optimierung ihrer Abläufe.

Wiener Manifest: ein Aufruf aus dem Jahr 2019, der auf dem Wiener Workshop zum Digitalen Humanismus entstand. Das Manifest fordert eine menschliche und nachhaltige Gestaltung der digitalen Transformation durch die Förderung von Demokratie, Privatsphäre und ethischen Standards in der Technologieentwicklung.

Winner-takes-it-all-Phänomen: beschreibt eine Marktstruktur, in der der erfolgreichste Anbieter fast den gesamten Marktanteil erhält, während die Konkurrenz kaum profitiert. Oft anzutreffen in digitalen Märkten oder Netzwerkeffekten (z. B. soziale Medien oder Plattformen), führt es dazu, dass ein dominanter Anbieter Vorteile wie größere Nutzerbasis, Bekanntheit und Kostenvorteile aufbaut. Das Prinzip kann Marktzutritt für andere erschweren und zur Monopolbildung führen.

WWW (World Wide Web) oder Web: weltweites Netzwerk aus Webseiten, die über das Internet abgerufen werden können. Es basiert auf Hypertexttechno-

logie, bei der Informationen durch Links verbunden sind, sodass Nutzer:innen von einer Seite zur nächsten navigieren können. Mit einem Webbrowser können Texte, Bilder, Videos und vieles mehr angezeigt werden. Das Web wurde 1989 von Tim Berners-Lee entwickelt. Es hat die Art, wie wir Informationen suchen, kommunizieren und arbeiten, grundlegend verändert. Der Begriff wird oft synonym mit Internet verwendet. Standardisierte Elemente des Web sind HTML, URL und HTTP.

XAI (eXplainable Artificial Intelligence): Technologie, die Entscheidungsprozesse von maschinell lernender KI für Menschen verständlich und nachvollziehbar machen soll. Diese KI funktioniert oft wie eine Blackbox, indem sie Entscheidungen trifft, ohne dass der genaue Ablauf sichtbar ist. XAI dagegen soll aufzeigen, wie und warum bestimmte Ergebnisse entstehen. XAI ist eines der wichtigsten Forschungsgebiete innerhalb der KI.

Herausgeber:innen

Anita Eichinger, Direktorin der Wienbibliothek im Rathaus, setzt sich theoretisch wie praktisch mit Bibliotheken als Wissensinstitutionen in Zeiten der digitalen Transformation auseinander. Zuletzt erschienen: Anita Eichinger, Katharina Prager: *We Are Needed More Than Ever: Cultural Heritage, Libraries and Archives*, in: Hannes Werthner et al. (Hrsg.), *Perspectives on Digital Humanism*. Heidelberg: Springer, 2022, S. 109–114.

Peter Knees, assoziierter Professor und Inhaber des UNESCO-Lehrstuhls für Digitalen Humanismus an der Fakultät für Informatik der TU Wien, forscht zu Themen wie Künstliche Intelligenz und Empfehlungssysteme mit Anwendungen im Bereich Musik. Zuletzt erschienen: P. Seshadri, S. Shashaani, and P. Knees: *Enhancing Sequential Music Recommendation with Negative Feedback-informed Contrastive Learning*, in: *Proc 18th ACM Conference on Recommender Systems*, 2024, ACM.

Hannes Werthner, ehemals Informatikprofessor an der TU Wien und anderen Universitäten, beschäftigt sich mit E-Commerce und Empfehlungssystemen. 2019 organisierte er den Workshop zum Digitalen Humanismus und initiierte das Wiener Manifest zum Digitalen Humanismus. Zuletzt erschienen: Hannes Werthner et al. (Hrsg.), *Introduction to Digital Humanism. A Textbook*. Cham: Springer, 2024.

Impressum

Eine Veröffentlichung der Wienbibliothek im Rathaus, der Fakultät für Informatik der TU Wien und des UNESCO Lehrstuhls für Digitalen Humanismus

Herausgegeben von Anita Eichinger, Peter Knees und Hannes Werthner

Didaktische Umsetzung von Bettina K. Lechner

Erschienen im Residenz Verlag GmbH Salzburg – Wien
www.residenzverlag.com

Grafische Gestaltung: Lisa Ifsits
Korrektur: Christian Mertens
Lektorat: Julia Teresa Friehs
Druck und Gesamtherstellung: Prime Rate Zrt, Budapest

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

ISBN 978-3-7017-3641-6

2024 Wienbibliothek im Rathaus,
Residenz Verlag GmbH Salzburg – Wien

Nicht immer ist es uns möglich gewesen, die Rechtsinhaber:innen und Rechtsnachfolger:innen zu ermitteln oder zu erreichen. Die Wienbibliothek im Rathaus bittet um Kontaktaufnahme in allen Fällen, wo nachweislich Honoraransprüche bestehen.

Lizenziert unter CC BY <http://creativecommons.org/licenses/by/4.0/>. Ausnahmen hiervon sind mit © gekennzeichnet.

Bitte verwenden Sie für die Quellenangabe folgendes: CC BY „Digitalisierung und wir. Lehrbuch zum Digitalen Humanismus mit praktischen Übungen“, hrsg. von Anita Eichinger, Peter Knees, Hannes Werthner. Residenz, 2024.

Das Buch ist digital abrufbar auf:
<https://resolver.obvsg.at/urn:nbn:at:AT-WBR-1565583>

RESIDENZ
VERLAG

Stadt
Wien

Wienbibliothek
im Rathaus

TU
WIEN Informatics


unesco
Chair

