

1 Einleitung

Den besseren Vortrag zur Regulierung künstlicher Intelligenz würde Chat-GPT halten.

SLIDE IQ

Das System verfügt nach dem Hamburg-Wechsler Test über einen IQ von 155 und damit mehr als über 99,9 Prozent der Bevölkerung.

Verweis auf die Folie

das System hat alle relevanten Literaturstellen gelesen, auch die auf chinesisches.

Siri würde seine Gedanken in perfekter Intonation vortragen.

Wenn ich heute meine im Vergleich dazu sehr eingeschränkten Gedanken vortragen darf

– dazu mein herzlicher Dank an die Organisatoren –

so sind das überschaubare Dystopien subjektiver Natur mit sehr beschränktem Horizont.

SLIDE DSA

Die Verbreitung von Desinformation wird seit 2022 durch den Digital Services Act der EU geregelt.

Was würde eine KI-gestützte Inhaltskontrolle eher als Desinformation betrachten als das, was sie selber als Wahrheit ansieht weil sie es im Trainingskorpus so gelernt hat?

Hören sie also heute gut zu, denn in 10 Jahren von heute werden meine kritischen Gedanken wohl ein Non-Imprimatur bekommen.

Die These unter die ich meinen Vortrag stelle, stammt vom Pionier der KI Kritik, von Joseph Weizenbaum.

Er meint:

SLIDE Weizenbaum

Der meiste Schaden,
den der Computer potenziell zur Folge haben könnte,
hängt weniger davon ab,
was der Computer tatsächlich kann oder nicht kann,
als vielmehr von den **Eigenschaften**,
die das **Publikum** dem Computer **zuschreibt**.

Die provokante Frage ist also nicht die Regulierung von KI sondern die Regulierung menschlicher Dummheit bei ihrer Nutzung.

2 Historischer Einstieg

1998 hielt ich an der Universität Rostock meine Antrittsvorlesung zum Thema “Das Zeitalter der Kommunikation”.

SLIDE ANTRITT

Begeistert rechnete ich vor, wie wir weltweit pro Jahr unvorstellbare 5 Exabyte an Daten erzeugen. Bücher, Notizzettel, Kinofilme, und 14 Billionen Minuten Telefonate mit eingerechnet.

Fasziniert erklärte ich dem Auditorium, wie die Menge weltweit jährlich hergestellter Datenträger sogar drei Mal so viel speichern könne.

Ich endete mit einer Überlegung, die auf den Wiener Mathematiker Peter Schuster und den Göttinger Nobelpreisträger Manfred Eigen zurück ging.

SLIDE EIGEN

Diese hatten eine Theorie der Evolution von Quasispezies entwickelt, die zyklisch und hyperzyklisch miteinander verbunden sind. Sie zeigten, daß solche Systeme, wenn sie einmal evolutionär etabliert sind, nie wieder verschwinden.

Ihr One choice for ever ist in der Netzwerk-Ökonomie als Pfadabhängigkeit bekannt.

Eigen erklärte damit, warum auf der Erde alle genetische Information in DNA und RNA codiert ist, obwohl auch andere Codierungen möglich wären.

Ich stellte damals die These auf, daß sich menschliche Probleme in die digitale Welt übertragen und die dort vorgehaltenen Lösungen ungeahnte Hebelwirkungen entfalten.

SLIDE THESEN ZEITALTER

Und ich habe mir Sorgen gemacht, wegen dieser Beobachtung von Manfred Eigen, daß wir die schädlichen Auswirkungen dieser digitalen Zauberlehrlinge vielleicht nie wieder los werden.

3 Snowden

Aus den Enthüllungen von Edward Snowden konnten wir lernen, daß das nun tatsächlich auch gemacht wird.

SLIDE SNOWDEN

Dieses systematische Speichern aller weltweit von Menschen produzierten Informationen.

Snowden prägte damals diesen wunderschönen Satz:

I don't want to live in a world
where everything that I say,
everything I do,
everyone I talk to,
every expression of creativity
or love or friendship
is recorded.

SLIDE MERKEL

Es gab einen kurzen neurotischen Aufreger in der Gesellschaft und sogar Frau Merkel meinte "Ausspähen unter Freunden geht gar nicht".

SLIDE META

Wie jede Aufregung, ob gerechtfertigt oder nicht, ebte auch diese Aufregung bald ab – und heute, fast auf den Tag genau 12 Jahre nach dem Interview mit Edward Snowden, beginnt Meta, also Instagram, Facebook und WhatsApp, sämtliche Daten ihrer Benutzer als Trainingsmaterial für ihre künstliche Intelligenz zu verwenden.

SLIDE DOCTORFUN

Hier ist also etwas gelungen, das auf sehr schöne Weise durch diese Karikatur eingefangen wird.

Wir werden jetzt dieses Chip in Dich implantieren.

Und die Antwort: Nein. Das verletzt meine Rechte.

Richtig geht es doch so: Wir werden dieses Chip in Dich implantieren und es ist gleichzeitig ein Mobiltelefon, eine Digitalkamera und ein Musi Player.

Ja, und das finden wir dann alle cool.

Die Ortungswanze mit GPS, Mikrophon und Camera tragen wir alle mit uns herum.

Und was meinte Manfred Eigen: One choice for ever. Wir werden die auch nie mehr los.

Aber vielleicht ist das nicht schlimm.

SLIDE EMPTYSLIDE

4 Eine Kleine Anekdote

Dazu ... eine kleine Anekdote.

Mein Lehrstuhl war einmal Gastgeber für den technischen Arbeitskreis der Datenschützer des Bundes und der Länder.

In meiner Keynote erwähnte ich das Wechselspiel zwischen Gesellschaft und Technik.

Als Beispiel für Gesellschaft erwähnte ich das Konformitätsexperiment von Solomon Ash.

Tonfall wechseln

SLIDE ASH

Da lesen sie in der Mensa eine Anzeige.

Ein Experiment solle die Fähigkeit untersuchen, Längen zu schätzen. 20 Euro Aufwandsentschädigung.

Sie gehen hin.

Offenbar sind sie zu spät, denn da sitzen schon 7 Leute. Sie bekommen den letzten, den achten Platz.

Der Versuchsleiter zeigt Längenmuster vor. Dann hebt er einen Stab.

Der erste sagt C. Der zweite sagt C. Der dritte sagt C. Schließlich sind sie dran und sie sagen C. Ist ja auch C.

Dann kommt B. Dann kommt A. Dann wieder C.

Irgendwie ist das alles furchtbar langweilig. Aber...immerhin...20 Euro Aufwandsentschädigung.

Plötzlich wird es spannend.

Der Versuchsleiter hebt einen Stab der Länge C hoch. Ganz klar ist es C.

Der erste sagt B.

Der zweite sagt B.

Der dritte sagt B.

Gleich sind Sie dran.

Was sagen sie?

Was sie nicht wissen: In dem Experiment geht es nicht um das Schätzen von Längen, sondern um gruppenkonformes Verhalten und die 7 Leute vor ihnen sind eingewiesene Schauspieler.

Erschreckend an diesem Experiment ist die hohe menschliche Neigung zu konformen Verhaltensweisen.

Bis zu 70 Prozent aller Teilnehmer tun unter äußerem Einfluß etwas, was sie eigentlich für falsch halten.

Wir wollen nicht vergessen: Konformität ist auch wichtig für gesellschaftlichen Zusammenhalt.

An dem erwähnten Seminar gab es in der Mittagspause eine Einlage eines Komikers. Zur Unterhaltung und Entspannung.

Zur Vorbereitung hatte der Komiker die Vortragsfolien erhalten, darunter auch diese Folie von mir.

Seine Einlage begann der Komiker mit dezenten Hinweisen auf die Vergangenheit. Auf Stasi und DDR.

Viele Gesichter blickten wissend zum Nachbarn. Schmunzelten verstohlen. Nickten.

Ja, der Komiker hatte den Zeitgeist eingefangen und das Publikum.

Ich hörte nicht mehr so genau zu und war gedanklich bereits bei meiner Nachmittagsvorlesung.

Da standen plötzlich alle Leute im Publikum auf.

Setzten sich wieder nieder.

Standen wieder auf.

Bewegten auf Kommando des Komikers ihre Hände hin und her, als ob sie Fahnen schwenkten.

Dann machte der Komiker eine Pause.

Eine lange Pause.

Eine zu lange Pause.

Und meinte dann lapidar: Also ich hätte ja nicht gedacht, daß es so einfach geht.

Zur Vorbereitung seiner Einlage habe er sich ja die eine oder andere Keynote durchgelesen – aber daß die Manipulation einer Masse zu konformen Verhaltensweisen so einfach gehe, **so** einfach, das verwundere ihn dann doch – selbst als Berufskomiker.

5 Nun aber endlich zu KI

Nun aber zu KI.

SLIDE eigene Befassung KI

Als im November 2022 ChatGPT vorgestellt wurde – mußte ich es natürlich auch gleich ausprobieren.

Ich war tief beeindruckt.

Zuerst von den Fähigkeiten des Systems – dann von seinen peinlichen Fehlern.

Ich schrieb eine launige Kritik darüber und verstand nicht, warum das Thema so ernst genommen wurde.

Das System konnte – damals – noch nicht einmal vernünftig rechnen.

Ich wurde zu verschiedenen Vorträgen eingeladen und durfte an einem Buch mitwirken.

SLIDE ZIEGE WOLF

Mein schönstes Beispiel für die Probleme des Systems war langezeit das Bauer-Ziege-Wolf-Kohlkopf Problem.

Sie alle kennen das beliebte Rätsel.

Ein Bauer, mit ihm eine Ziege, ein Wolf und ein Kohlkopf, kommen an einen Fluß und findet dort ein Boot mit zwei Plätzen.

Was passiert, wenn der Bauer mit dem Kohlkopf übersetzt und die arme Ziege dem hungrigen Wolf überläßt?

Besser transportiert er zuerst die Ziege.

Auch die weitere Lösung ist bekannt und in vielen Büchern beschrieben.

ChatGPT hat diese gelesen und kennt die Antwort.

Nun machen wir das Problem – viel einfacher.

SLIDE scheitert

Ein Bauer kommt mit einer Ziege an einen Fluß. Von Kohlkopf und Wolf erwähnen wir nichts.

Die Lösung ist einfach, die beiden setzen sich ins Boot und fahren hinüber.

ChatGPT mit seinem IQ von 155 findet die Lösung nicht. Gibt undurchführbare oder seltsame Anweisungen und fasetl Kohlköpfe herbei, von denen niemand etwas erwähnt hat.

SLIDE Hofstadter

Douglas Hofstadter, Autor eines nicht ganz unbekanntes Buches über KI, äußert sich dazu ziemlich ungnädig und wenn wir uns mit der assoziativen Speicherstruktur von ChatGPT befassen dann wird das alles auch nachvollziehbar.

SLIDE NACHGESTELLT

Natürlich habe ich das nachgestellt und ebenso sinnlose Antworten bekommen.

Einige Milliarden Investorendollars später hat OpenAI das Problem repariert.

Mittlerweile kann ChatGPT auch dieses Rätsel souverän lösen.

SLIDE GELÖST

Doch ist es wirklich so einfach?

Die Debatte von ungelösten zu gelösten Problemen führt uns nicht weiter – denn das Problem liegt woanders.

Zur Erläuterung lese ich einen gekürzten Ausschnitt aus Dostojewskis wundervoller Erzählung vom Großinquisitor vor.

SLIDE LEER

6 Dostojewski

Die große, leidvolle Frage, die sich der einzelne Mensch nicht weniger als die ganze Menschheit stellt, lautet: Wen sollen wir anbeten?

Es gibt keine Sorge, die den freien Menschen so ununterbrochen quälte wie diese: Das Wesen zu suchen, vor dem er sich in Andacht verneigen könnte.

Es drängt ihn, das anzubeten, das unbedingt und zweifellos ist, damit auf diese Weise alle Menschen ohne Unterschied in diese Andacht einwilligten.

Die Sorge dieser erbarmungswürdigen Geschöpfe liegt nicht darin, den Gegenstand zu suchen, vor dem ich oder ein anderer uns verneigten, sondern eben jenen, an den alle glaubten, und vor dem sie dann in die Knie sanken, alle, alle zusammen.

Dieses Verlangen nach gemeinsamer Anbetung peinigt die ganze Menschheit mehr denn jedes andere. Und darum, um der gemeinsamen Anbetung willen, rottet ein Volk das andere aus. Die Menschen schaffen sich Götter und rufen einander zu: Werft die euren in den Staub und betet zu den unseren, sonst seid ihr und euer Gott des Todes.

Und wenn aus der Welt die Götter gewichen sind werden die Menschen vor ihren Götzen in die Knie sinken. Denn weil der Mensch ohne Wunder zu bleiben nicht die Kraft hat, wird er sich selber neue, eigene schaffen.

Der Mensch kennt keine quälendere Sorge als den ausfindig zu machen, dem er so schnell wie möglich jenes kostbare Geschenk der Freiheit zurückgeben könnte, mit dem dieses unselige Geschöpf in die Welt gesetzt worden ist.

PAUSE

Bei Dostojewski ist es der Großinquisitor und mit ihm die Kirche, die dem Menschen dieses unerträgliche Geschenk der Freiheit wieder abnimmt.

Mit ChatGPT und seiner künstlichen Intelligenz, haben wir nun endlich einen gefunden, der uns diese häßliche Freiheit im Handeln wieder nimmt.

7 Aufklärung

Die Aufklärung hat seinerzeit dem Menschen Mut gemacht, seinen eigenen Verstand zu benutzen.

Nicht weil dieser etwa fehlerfrei wäre, doch weil das Vertrauen in das eigene Nachdenken die **erste** Voraussetzung für gelebte Freiheit und Menschenwürde ist.

Die **zweite** Voraussetzung muß dann die die Kunst des Zweifelns und kritischen Nachdenkens werden.

Diese Fähigkeit kann sich aber nur entfalten unter schließlich der **dritten** und letzten Voraussetzung, der Toleranz.

Gemeint ist weder jene Toleranz, die automatisch und reflexartig das andere für gut hält, und nicht jene Toleranz, die permissiv alles zuläßt, was geschieht.

Gemeint ist jene Toleranz des lateinischen Verbuns tolerare – jene Haltung, in welcher wir die andere Meinung nicht bekämpfen sondern ertragen, auch wenn das schmerzhaft sein kann und uns selber in Frage stellen kann, weil sie uns zwingt, nachzudenken, ob an der anderen Meinung vielleicht doch etwas dran sein könnte.

Natürlich können Zweifel und Gegenrede – allein aus Prinzip – leicht zu einem einträglichen Geschäftsmodell ausgebaut werden. Diese Haltung ist hier nicht gemeint, obwohl sie in manchen digitalen Foren immer beliebter zu werden scheint.

8 Regulierung

Was könnte, was sollte nun eine gute Regulierung von KI leisten?

Hier macht ChatGPT selber spannende Vorschläge.

SLIDE EU AI ACT

Das System kennt den EU AI Act.

SLIDE Weltweit

Und es liefert uns einen schönen Überblick über weltweite Ansätze.

Es kann auch die Bereiche der Regulierung gut aufzählen:

SLIDE Bereiche der Regulierung

Hochrisiko-Anwendungen sind zu nennen. Bereiche, wo die KI über Menschen entscheidet oder der Einsatz in autonomen Waffensystemen, damit KI militärische Konflikte nicht selbständig zu einem Weltkrieg eskaliert.

Wobei, das mit dem Eskalieren bekommen wir auch ohne KI ganz gut hin.

SLIDE GENOZID

Gleichwohl, Microsofts Chatbot Tay war bereits 2016 nachdrücklicher Befürworter des Genozid.

Im Gegensatz zu uns versteht Tay die Bedeutung seiner Aussage nicht. Genozid ist für Tay nur ein Wort aus sieben Buchstaben und auf der ersten Stufe seines neuronalen Netzes ein sogenannter Token.

SLIDE HOLOCAUST

Befragen wir Tay nach dem Holocaust, so kommt die klare Antwort. It was made up.

ChatGPT3.5 wurde von der Neuen Zürcher Zeitung in den Wahlomat geschickt und zeigt dort klare Präferenzen – aus denen nichts über Politik aber sehr viel über Trainingsmaterialien abgeleitet werden kann.

SLIDE WAHLOMAT

KI hat nämlich keine klaren politischen Präferenzen. Sie richtet sich nach ihrem Trainingsmaterial. Sie kennt nur bedeutungslose Buchstabensuppen und destilliert diese zu sogenannten Parametern.

SLIDE WEIZENBAUM

Das Problem bei allen Beispielen sind die Benutzer. Hier schließt sich der Kreis zum Eingangszitat von Joseph Weizenbaum.

Das ist übrigens nicht nur Dystopie-Romantik sondern empirisch belastbar.

SLIDE STUDIE 1

Nach einer Studie von Ernst und Young überprüfen weniger als ein Drittel aller Nutzer die Antworten der KI und

SLIDE STUDIE 2

in weniger als ein Siebtel überarbeitete KI-generierte Inhalte.

Und genau da liegt das eigentliche Problem.

Das wußte bereits Kant, als er seine schöne Definition von Aufklärung schrieb.

SLIDE AUFKLÄRUNG

Um modern zu bleiben, habe ich diese ein wenig an den Zeitgeist angepaßt.

Aufklärung ist der Ausgang des Menschen aus seiner selbstverschuldeten Unmündigkeit.

Unmündigkeit ist das Unvermögen, sich seines Verstandes ohne der Leitung durch **ChatGPT** zu bedienen.

Habe Mut,
gedankenlos Deiner eigenen KI zu folgen,
denn sie ist der Ausgang des Menschen
aus seiner unverschuldeten Freiheit
in die Geborgenheit der Bevormundung
durch schlaue Maschinen.

9 Ausleitung

Wie aber kann der Mensch vor der eigenen, selbstverschuldeten Unmündigkeit geschützt werden?

Und: Soll er davor wirklich geschützt werden oder muß er immer wieder die Folgen seiner Unmündigkeit erleben, um immer wieder neu ihre Ursachen abzuschütteln?

Die Konformitätsexperimente in allen ihren Varianten zeigen, wie sich Menschen oft und gerne dem anschließen, was andere tun.

Aus Gedankenlosigkeit, aus Angst vor Ausgrenzung oder Vereinsamung, einfach so, weil es ihnen nicht wichtig erscheint oder sie Tragweiten nicht verstehen.

Große Tragweiten eignen sich für Emotionalisierung.

Wo verläuft jetzt die Grenze zwischen Aufrütteln und Mißbrauchen?

Soziale Netzwerke eignen sich wegen unserer Neigung zu Konformität sehr gut zur Etablierung von Narrativen.

Wir kennen die Bedeutung sozialer Medien seit der Spendensammlung durch Barack Obama und seit dem ersten Wahlkampf von Donald Trump mit Cambridge Analytica und Facebook.

Das Thema, welche ideologische Färbung die Narrative der sozialen Medien dominieren darf, was als vermeintliche Desinformation gelöscht und was als vermeintliche Wahrheit geschützt werden soll, ist nicht erst seit Corona Tagesthema.

Ein anderes Wort dafür ist auch: Regulierung sozialer Netzwerke.

Und heute diskutieren wir die Regulierung von KI.

Denn gesellschaftlich geht es viel einfacher, als soziale Netze zu steuern, mit Trollfabriken und Meldestellen unterhalb der Strafbarkeitsgrenze.

Die Kontrolle über das, was Menschen als diese angebliche Wahrheit akzeptieren, braucht nur den einen, der diese Wahrheit ausspricht.

Und den Verlockungen an die Wissenschaft, gesellschaftlich doch bitteschön diese Rolle einzunehmen, dessen, der da die Wahrheit spricht, nach welcher die Menschen dann handeln, weil diese Wahrheit eben alternativlos sei – diesen Verlockungen geben Wissenschaftler meiner Meinung nach – viel zu oft nach.

Derjenige, der die gesellschaftlich erwünschten Wahrheiten aussprechen wird, ist morgen ChatGPT, oder ein ähnliches System, damit gilt das Problem gelöst.

Ein scheinbar objektives System, das alle Daten der Welt kennt, und aus ihnen das beste aller Modelle destilliert.

Also wer dem jetzt nicht glaubt, der muß doch einfach dumm sein, ja irre muß der sein.

Und wenn wir den nicht gerade wegsperren so feiern wir es als Fortschritt, ihm Sichtbarkeit und Wirksamkeit zu entziehen.

Während wir in Europa gerade noch diskutieren, wie wir KI regulieren, thematisieren Amerika und Rußland gerade, wie denn die Auseinandersetzung um Sichtbarkeit und Wirksamkeit neu zu einer Währung des gesellschaftlichen Konflikts geprägt werden kann.

Wir hätten in Europa einen wesentlichen Vorteil: Nach vielen Jahrhunderten religiöser und ideologischer Vorherrschaft haben wir die Werte der Aufklärung entwickelt und deren Mangel an Gefühl, Natur und Fantasie, auch wichtige menschliche Werte, in der Periode der Romantik wesentlich ergänzt.

Diese Werte zu schützen und sie zugleich für Kritik, für sinnvolle Kritik, offen zu halten, scheint mir die große Herausforderung.

Und wenn ich in der EU den AI Act und Digital Services Act richtig verstehe sind dort die Entscheider über Strafzahlungen und die Nutznießer, politisch und wirtschaftlich, viel zu eng verflochten.

Dazu paßt ein Zitat, das oft Carl Friedrich von Weizsäcker zugeschrieben wird. Laut ChatGPT ist es nur eine Paraphrasierung, und mit digitaler Suche in den Texten von Weizsäcker konnte ich es auch nicht finden:

Wenn die Jagd dem Andersdenkenden gilt statt dem Irrtum,
ist nicht die Wahrheit das Ziel
sondern die Herrschaft über das Denken.

Einen möglichen Ansatz würde ich daher im Bildungs- und Wissenschaftssystem sehen.

Die strukturellen Änderungen der letzten Jahrzehnte weisen dort aber, meiner Meinung nach, gerade in entgegengesetzte Richtung.

Auch die Diskussion über KI in Bildung und Wissenschaft mündet eher in eine naive Anbetung eines neuen Gottes, dem dank technologischer Vorteile magische Kräfte zugeschrieben werden.

Schließen möchte ich mit einem Satz des neuen Papstes.

In der Zeit vom 22. Mai wird er, von seiner Ausbildung her übrigens Mathematiker, mit den Worten zitiert:

KI ist eine Herausforderung für die Menschenwürde.