

Interview geführt von Claudia Schmid, [Kurzversion](#) veröffentlichte im HSG Focus, am 24.11.23

«Die digitale Gesellschaft braucht Leitlinien der Ethik»

Was hat digitale Ethik zum Stand der KI-Technologien zu sagen? «Wir tendieren dazu, auf die Künstliche Intelligenz den Heiligen Grahl oder die Büchse der Pandora zu projizieren. Eine realistische Einstellung müssen wir uns erarbeiten.», betont Philosophin Anita Horn.

Womit beschäftigt sich digitale Ethik und warum ist sie dringlich?

Der Begriff «digitale Ethik» ist missverständlich, weil es nicht primär darum geht, dass die Ethik «digital» wird. Es ist sinnvoll zwischen den spezifischen Bereichsethiken wie der Ethik der Künstlichen Intelligenz oder der Informationsverbreitung(ethik), und einer kritischen Ethik der digitalen Gesellschaft zu unterscheiden. In der Ethik der Künstlichen Intelligenz geht es darum, die Auswirkungen bestimmter Technologien wie generativer künstlicher Intelligenz (wie bspw. ChatGPT als sprach- und textbasierter Chatbot) oder der digitalen Besitzurkunden (NFTs) orientiert an bestehenden Werten zu beurteilen. Diese Analyse orientiert sich an moralischen Prinzipien und behandelt spezifische ethische Konflikte mit dem Ziel, Empfehlungen und Standards zu entwickeln. Gefragt wird beispielsweise, ob die technologischen Innovationen zu Verletzungen von Menschenrechten führen, ob Grundwerte und Rechte wie die Privatheit oder Urheberrechte gewahrt werden. Die digitalen Technologien der neusten Generation haben ein breites Feld an ethischen Fragen und Konflikten produziert, bspw. zum legitimen Einsatz der sozialen Robotik oder zur Gefahr der Desinformation.

Dringlich scheint darüber hinaus eine kritische Ethik der digitalen Gesellschaft zu sein, die sich einer erweiterten Reflexion widmet. Die normative Analyse von gesellschaftlichen, politischen, ökonomischen und ökologischen Auswirkungen digitaler Technologien und Geschäftsmodelle stützt sich hier auf grundsätzliche Überlegungen zum Selbst- und Weltverständnis in der digitalen Epoche. - Eine Ethik, die den Menschen auf transhumanistische Weise auf Information reduziert, versteht ihn als *homo optimus*, als ein Mängelwesen, das durch Technologie verbessert werden muss. Diese orientiert sich z.B. an Werten wie der Effizienz- oder Funktionalitätssteigerung. Eine humanistische Ethik anerkennt den Menschen als immer auch unergründliches, mitunter irrationales Wesen, das verwundbar und individuell ist. Sie zielt auf den Schutz der Menschenwürde und gewichtet Aspekte wie Bezogenheit, Sinnsuche oder Kreativität höher als Perfektion und Selbstoptimierung. – Das vorherrschende Selbstverständnis und die Vorstellungen von einem guten Leben entsprechen einem oft unbewussten Kompass im ethischen Urteilsprozess.

Welche spezifischen ethischen Herausforderungen stellen sich in der digitalen Gesellschaft?

Da die technologische Evolution meist schneller ist als unser Verstehen und die normative Beurteilung davon, hinken die ethische Beurteilung wie auch die rechtliche Regulierung oft

der Implementierung von Technologien hinterher. ChatGPT ist ein gutes Beispiel für das Phänomen der Nachregulierung. Die Technologie wird fortwährend weiterentwickelt, während wir versuchen abzuwägen, welche normativen Konsequenzen absehbar und welche Forderungen – etwa nach Moratorien – aus ethischer Sicht angezeigt wären. Das menschliche Denken und Erkenntnisvermögen ist nicht gleichermassen schnell, wie dasjenige der generativen Künstlichen Intelligenz. Die mathematische Rationalität der selbstlernenden Künstlichen Intelligenz steht einer vielschichtigen menschlichen Intelligenz und Entscheidungsfindung gegenüber, die immer auch durch irrationale, affektive und zwischenleibliche Aspekte mitbeeinflusst wird. In diesem Spannungsfeld zwischen technologischer Machbarkeit und dem Anspruch, Entwicklungen im Interesse Mensch und Gesellschaft zu steuern, bewegt sich die Ethik der digitalen Gesellschaft.

Wie bringt man der Technologie Ethik bei? Oder anders gefragt: Kann KI auch ethisch denken lernen?

Das maschinelle Lernen der Künstlichen Intelligenz ist fortgeschritten und es speist sich aus allen Informationen, die datafizzierbar und zugänglich sind. Sofern die Künstliche Intelligenz (am Beispiel von ChatGPT) Zugang hat zu verifizierten und qualitativ hochstehenden Grundlagenstudien, kann sie theoretisch auf deren Grundlage allgemeine ethische Fragestellungen stilischer beantworten.. – In der Realität ist die Quellengrundlage jedoch selten eindeutig. Die Validität von Informationen, die ChatGPT nutzt, kann nicht vorausgesetzt werden. Die generative Künstliche Intelligenz ist aktuell die effizienteste Recyclingmaschine für Informationen, die es gibt. Diese bunt gemischten Informationen müssen nach einem bestimmten Schlüssel algorithmisch vorsortiert werden müssten, um Verzerrungen und Desinformation vorzubeugen. Damit eine KI eine ethische Entscheidung treffen kann, muss der Mensch als Teil des Algorithmus u.a. Wertehierarchien vorgeben. Im Rahmen dieser Vorgaben sind Künstliche Intelligenzen der neusten Generation fähig, selber ein Wertesystem zu erweitern und zu entwickeln. Hier, v.a. aber bei ethischen Einzelfallbeurteilungen und gegeben des fehlenden Bewusstseins und der fehlenden Empathie der Technologie, stösst die «Urteilsfähigkeit» der KI an Grenzen.

Weshalb?

Ethisches Urteilen und verantwortliches Handeln sind in der angewandten Praxis keine, rein rationalen, logischen Ableitungen. Sie erfordern die sorgfältige Abwägung im Einzelfall und eine möglichst umfassende, vorausschauende, wertorientierte und empathische Vorgehensweise. In der Realität sind wir in der Ethik immer wieder mit unlösbaren Dilemmata konfrontiert, für die es keine eindeutig richtige Lösung gibt. Ein bewährtes Beispiel ist das moralphilosophische Trolley-Experiment, das seit der Erstfassung von Philippa Foot (1967) in beliebigen Varianten durchgespielt wird: Ein führerloser Zug fährt auf eine Weiche zu. Sie können die Weichenstellung beeinflussen. Auf der einen Spur rollt eine junge Mutter ihren Kinderwagen, auf der anderen sitzen vier ältere Professor*innen und diskutieren über den neusten Durchbruch bei der Entwicklung eines Sanitätsroboters, der im Krieg vielen Menschen das Leben retten könnte. – Wie stellen sie die Weiche? Und wie können Sie das ethisch begründen? – Bereits vor der Komplizierung des Experiments im Kontext der selbstfahrenden Autos, bei dem Fahrer*innen meist keine direkte Kontrolle über solche Entscheidungen mehr hat, unterschieden sich die Antworten je nach der zugrunde gelegten moralischen Referenztheorie. – Auch auf evidenzbasiertem Weg lassen sich keine

moralisch universellen Antworten finden. Zwar wünschen sich global alle Befragten einer MIT Studie, dass ihr Leben als Insassen selbstfahrender Autos prioritär geschützt wird, vor demjenigen der Verkehrsteilnehmer*innen. Hinsichtlich der Entscheidung, welche Passanten zuerst geopfert werden sollten, unterschieden sich die ethischen Präferenzen jedoch nach Alter, Geschlecht, Amt/Beruf, etc. – Wenn wir als Menschen oder in Ethik-Kommissionen Entscheidungen treffen, dann tun wir dies in der Regel deliberativ, unter Beachtung der geltenden Gesetzgebungen und in Anlehnung an moralphilosophische Werte und verschiedene Stellungnahmen, v.a. aber hinsichtlich der Frage nach dem Wohl der Betroffenen. Subjektive und zwischenmenschliche Erwägungen gehören mit dazu.

Auf welche Weise kann digitale Ethik dazu beitragen, dass Technologien entwickelt werden, die dem Wohle aller dienen?

Die Auseinandersetzung mit der Ethik der digitalen Gesellschaft konfrontiert uns mit grundsätzlichen Fragen – wird sie ernst genommen, hat sie ein aufklärerisches Potential, dass nicht brach liegen gelassen werden sollte. – Wir tendieren dazu, auf die Künstliche Intelligenz den Heiligen Grahl oder die Büchse der Pandora zu projizieren. Eine realistische Einstellung müssen wir uns hartnäckig erarbeiten. Tatsächlich stehen wir mächtigen Technologien gegenüber, die wir zu Recht als schwer steuerbar erleben. Sinnvolle, normativ begründete Regulationen setzen eine besonnene Abwägung von Chancen und Risiken im Einzelfall voraus. Wenn die Risiken zu fundamental oder unabsehbar sind, müssen Entwicklungsprozesse trotz z.B. drohender finanzieller Verluste gestoppt werden. Eine Gefahr ist hier der Kontrollverlust durch die Eigendynamik der KI, welche wir nicht nachvollziehen können. Ansatzweise wurde dieses Risiko am Beispiel des Einsatzes von KI auf dem Finanzmarkt am AI Safety Summit im November diskutiert.

Ethik der digitalen Gesellschaft zielt auf eine möglichst eigenständige und informierte Reflexion. Sie setzt nicht nur ethische, sondern auch digitale Kompetenz voraus. Ethische Kompetenz erfordert ein moralisches Empfinden und die Motivation, reflektiert und moralisch verantwortlich zu argumentieren und handeln, um größeres Wohl der Gesellschaften anzustreben. Wir müssen festlegen, welche Werte als Leitlinien der Ethik in der digitalen Gesellschaft erhalten oder aufgegeben werden. – Ein vieldiskutiertes Beispiel ist Privatheit – lässt sie sich als Menschenrecht noch einfordern oder hat sie an normativer Bedeutung verloren? - Digitale Kompetenz erfordert ausreichende Kenntnisse über die Technologien, ihre Anwendung und Funktionsweise. Aufgrund der immensen Komplexität digitaler Technologien ist das Streben nach digitaler Kompetenz für Ethiker*innen eine grosse Herausforderung. Ein Mangel an digitaler Kompetenz kann zu unqualifizierten und zu allgemeinen Aussagen führen. – Wenn diese Überforderung jedoch dazu führt, dass wir uns nur noch an Einzelfallbeurteilungen trauen und die grossen Entwicklungslinien ausklammern, dann wird dies bildungspolitische folgenreich. Denn Aufklärung und Schulung in ethischer Reflexionskompetenz sollte bereits sehr früh, im Schulalter vermittelt werden. Die Tatsache, dass uns noch viele Antworten fehlen, sollte uns nicht davon abhalten, unbequeme Fragen zu stellen und auszuhalten. Dies betrifft aber nicht nur die Ethik, sondern die Wissenschaften an sich. Eindimensionale Paradigmen des Wachstums, der Datafizierung und Quantifizierung werden aus menschenrechtlicher Sicht in Frage gestellt.

Welche Rolle kommt dem Staat zu? Welche Rahmenbedingungen muss dieser schaffen, damit technologische Entwicklungen auch ethisch vertretbar sind?

Künstliche Intelligenz ist inzwischen in nahezu allen Lebensbereichen hintergründig mit dabei. Gesundheitswesen, Verkehr, Bildung, Rechtsprechung und Anstellungsverhältnisse sind nur einige Beispiele. Das Potential von KI-Technologien ist immens, sofern sie dafür eingesetzt werden können, nachhaltige Entwicklung oder inklusiveres Wirtschaftswachstum zu unterstützen. Auch im Bereich der Humantechnologie sind KI basierte Diagnosen und weitere bahnbrechende Erkenntnisse nennenswert, wie in der Paraplegieforschung. Um das phänomenale Potential der KI so zu nutzen, dass sie dem Wohl des Einzelnen und der Gesellschaft dient, braucht es bewusste und wertebasierte Regulationsmassnahmen. Leider zeigt die Realität gegenwärtig, dass viele Risiken zu lange unterschätzt wurden. Dies liegt auch an der Geschwindigkeit technologischer Innovation und des maschinellen Lernens. Die virtuelle Welt ist dadurch auch als Tummelplatz der Kriminalität gewachsen und zu einem Ort der teilweise mutwilligen Manipulation und Propaganda geworden. Otilie Normalbürgerin kann den unliebsamen Nachbarn mit minimaler technischer Fertigkeit als Killer in ein Jihad-Video einbetten und auf youtube präsentieren. Bilder von Kindesmissbrauch und Folter können durch KI generiert werden und zur gezielten Flutung einschlägiger Portale eingesetzt werden. Die kognitive Erfassung der Komplexität der Informationskommunikationstechnologien (IKT) und der Wirkmechanismen in der virtuellen Sphäre, die anhand der Künstlichen Intelligenz bereits geschaffen und implementiert wurde, ist nicht nur für den Einzelmenschen eine Überforderung. Nicht nur die ethische Beurteilung, auch die Gesetzgebung etwa im Bereich des Datenschutzes und die Regulierung des Zugangs zu KI-Technologien muss nachträglich erfolgen, weil sich ethisch problematische Folgen oft erst im Laufe der Implementation zeigen, geschieht diese vielfach als Reaktion auf Menschenrechtsverletzungen und nicht als rechtzeitige Prävention. - Im Bericht einer OHCHR-Arbeitsgruppe der UN zum Menschenrecht auf Privatheit im Kontext der Digitalisierung wurde 2021 ein sofortiges Moratorium für den Verkauf und die Verwendung von KI-Systemen gefordert, die wie CHATGPT oder DALL-E zu Menschenrechtsverletzungen führen können. Die politischen Reaktionen schlugen keine grossen Wellen. Gerade die ökonomische Machtballung der Tech-Unternehmen schien die Staaten immer wieder auf kleinere oder grössere Ohnmachtsposition zurückzuwerfen. Es gibt aber auch Gegenmassnahmen. Die Bletchley Declaration, die am AI Safety Summit Anfang November von Grossmächten wie Grossbritannien, USA und China unterzeichnet wurde, setzt einen Gegenakzent.

Die meisten Staaten erhoffen sich von KI viel. Gleichzeitig wächst die Sorge, dass die digitale Revolution Diktaturen effizienter macht als Demokratien. Gibt es so etwas wie eine digitale Planwirtschaft, die an die Stelle von Marktmechanismen treten wird?

Es fällt schwer die Träger*innen der wirtschaftlichen und politischen Macht in der virtuellen Realität zu identifizieren. Entsprechend abhängig werden die Nutzenden von staatlichen Deutungen und Regulationen. Polarisierung und Radikalisierung scheinen zuzunehmen – die Unsicherheit über das faktisch Gegebene und Verlässliche wird mit starken Meinungen und reduktiven Erklärungen gegengehalten. Der demokratische Staat gerät in eine ambivalente Lage – einerseits muss er die Freiheitsrechte und Sicherheit der Bürgerinnen und Bürger schützen, andererseits zwingen die Missstände zu einer zunehmenden nationalstaatlichen Regulation. Die Möglichkeiten der Zäsur im Internet werden von

diktatorischen Regierungen in der Regel vorbehaltlos und auf die eigenen politischen Interessen bedacht ausgeschöpft. Ein vieldiskutiertes Beispiel für die digitale Machtkummulation des Staates, die eine umfassende Kontrolle und Sanktionierung der Zivilbevölkerung erlaubt, ist das Social Credit System in China.

Partiell mag die Anwendung von Theorien wie der Planwirtschaft auf die Marktlogik der virtuellen Sphäre kreativ und strukturierend sein. Soziale Ungleichheiten und Fragen der distributiven Verteilungsgerechtigkeit stellen sich auch in der virtuellen Sphäre und verlangen nach einer Entgegnung. Geht man mit Eduard Kaeser davon aus, dass «das Zeitalter der Theorien vorbei ist», scheint es wichtig, es bei Problemlösungen nicht bei der Anwendung von Einzeltheorien bewenden zu lassen, die für die analoge Welt entwickelt wurden. Die technologische Komplexität und Undurchschaubarkeit im Bereich der IKT führen zu Abhängig- und zu Beeinflussungsmöglichkeiten, die uns als Nutzerinnen und Nutzern erst teilweise bewusst sind. – Um Veränderungsprozesse anzustossen, muss die einzelne Bürgerin zunächst die Vogelstrausshaltung aufgeben und aktiv zu einer ethisch vertretbaren Praxis beitragen. Dazu gehört es, sich bei validierten Quellen zu informieren und z.B. die Konsumverantwortung möglichst bewusst wahrzunehmen.

Und was für eine Auffassung von Staat führt uns erfolgreich in ein ressourcenschonendes und technologiereiches Zeitalter?

Dazu gibt es verschiedene Auffassungen. Tendenziell nimmt der Ruf nach staatlicher Regulierung zu, sobald die Gesellschaft mit den Folgen von technologischen Innovationen überfordert ist oder negative Konsequenzen deutlich werden.

Ausgehend von einem demokratischen Verständnis des Staates kann dieser nur dann funktionieren, wenn die einzelnen Bürgerinnen und Bürger sich aktiv an Meinungsbildung und Politik beteiligen. Der Einzelne kann sein Bewusstsein über die politischen Ungerechtigkeiten und Möglichkeiten im Kontext der digitalen Lebenswelt nicht per Knopfdruck erlangen. Aufklärung setzt auch Interesse und Beteiligung voraus. Zugleich stehen die Staaten in der Pflicht, den Schutz der Grundrechte zu garantieren. Kompliziert wird es hier, wenn Grundrechte gegeneinander abgewogen werden müssen. Die freie Meinungsäußerung darf zur Verhinderung von Diskriminierung aus ethischer Sicht einerseits eingeschränkt werden, andererseits ist die Gewichtung umstritten. Während das Abschalten von Hate-Speech Seiten wenig kontrovers ist, ist es bei Verschwörungstheorien oder Propaganda-Seiten schon weniger eindeutig. Zum Schutz der Grundrechte und der Sicherheit scheinen auch demokratische Staaten teilweise zu restriktiveren Mitteln und Gesetzgebungen greifen zu müssen.

Das Leitmotiv dieser Ausgabe ist «ANPASSUNG: Wie wir gewinnen, indem wir verzichten». – Was könnte dieses Motiv für die digitale Gesellschaft bedeuten?

Das Motiv lässt eine zaghafte Gegenbewegung zur in unserer Gesellschaft populären Selbstoptimierung, Wachstums- und Effizienzsteigerungslogiken anklingen. Die Aussicht darauf, dass uns individuell und gesellschaftlich Selbstbeschränkungen bevorstehen, scheint aus ethischer Sicht stimmig. – Gegeben der Ressourcenknappheit, der Herausforderungen von Energiepolitik und Klimaschutz kann beispielsweise der erschreckende ökologische Fussabdruck, der mit der Nutzung und Anwendung von KI-Technologien auch einher geht,

nicht langfristig ausgeblendet werden. Die Frage, in welchem Bereich der Einsatz von KI und die grenzenlose Digitalisierung überhaupt sinnvoll sind, wird enttabuisiert. Wir haben in den letzten Jahren allenfalls zu unhinterfragt angenommen, dass Digitalisierung in allen Lebensbereichen und Branchen die bestmögliche Option ist. Je länger und konkreter wir mit den Nebenwirkungen von Künstlicher Intelligenz und Digitalisierung konfrontiert sind, desto deutlicher wird der Bedarf nach einer normativen Ausrichtung und entsprechenden Massnahmen.

Dr. phil. Anita Sophia Horn ist politische Philosophin und Psychoanalytikerin. An der Universität St.Gallen arbeitet sie als Dozentin für Ethik der digitalen Gesellschaft