

Stärken und Schwächen von DS und den Menschen.

Künstliche Intelligenz und Ethik

Chancen und Gefahren | Smart Grids, Effizienzsteigerung des Energiehandels sowie der Nutzung der Energieinfrastruktur und -quellen, KI-basierte Interaktion mit Kundinnen und Kunden... - im Kontext des Energiesystems eröffnen sich grosse Chancen dank künstlicher Intelligenz. Gleichzeitig gilt es aus ethischer Sicht, nicht nur das Thema «Cybersecurity» im Blick zu haben.

PETER G. KIRCHSCHLÄGER

Künstliche Intelligenz erweist sich als weitaus gefährlicher als Atomwaffen. Warum haben wir also keine Regulierungsaufsicht?» [1] Elon Musk, Mitgründer des Online-Bezahlungssystems PayPal und von Space X sowie Chef von Tesla - und sicherlich kein Technologiefeind - weist auf die Gefahren von KI hin und fordert eine normative Reflexion von KI. Stephen Hawking schätzt dies ähnlich ein: «Wenn wir nicht lernen, wie wir uns auf die potenziellen Risiken vorbereiten und sie vermeiden können, könnte die künstliche Intelligenz das schlimmste Ereignis in der Geschichte unserer Zivilisation sein. Sie bringt Gefahren mit sich, wie zum Beispiel mächtige autonome Waffen oder neue Wege für die wenigen, um die vielen zu unterdrücken. Sie könnte grosse Störungen unserer Wirtschaft verursa-

chen.» [2] Im Folgenden soll KI aus ethischer Perspektive diskutiert werden [3], auf der Basis, technologiebasierten Fortschritt und Ethik in einer Interaktion zu denken [4].

Ethische datenbasierte Systeme?

Können Menschen die ethische Verantwortung an datenbasierte Systeme, DS, delegieren? Bezeichnungen wie «moral technologies» legen eine solche Möglichkeit nahe. Während DS ethische Regeln befolgen können sowie darauf basierend handeln können, ist ihnen aus folgenden Gründen die Moralfähigkeit, die Menschen ethisch begründbar zukommt [5], abzuspochen: Technologien müssen ohne Freiheit gedacht werden. Denn Technologien werden von Menschen entworfen, entwickelt und hergestellt,

d. h. sie werden heteronom produziert. So erweist sich auch die Aneignung ethischer Normen von Menschen gesteuert. In letzter Konsequenz bleiben demnach Maschinen immer in dieser Fremdbestimmung gefangen. Auch selbstlernende Maschinen gehen bildlich gesprochen auf eine erste Zeile des Codes zurück, die immer vom Menschen stammt.

Ohne Freiheit kann Technologien auch keine Autonomie zugesprochen werden. Während es Menschen entspricht, für sich selbst allgemeine moralische Regeln und Prinzipien zu erkennen, diese für sich selbst zu setzen und diese ihren Handlungen zugrunde zu legen [5], ist dies Technologien nicht möglich. Technologien sind primär auf Zweckmässigkeit ausgerichtet und können sich zwar als selbstlernende Systeme Regeln setzen,

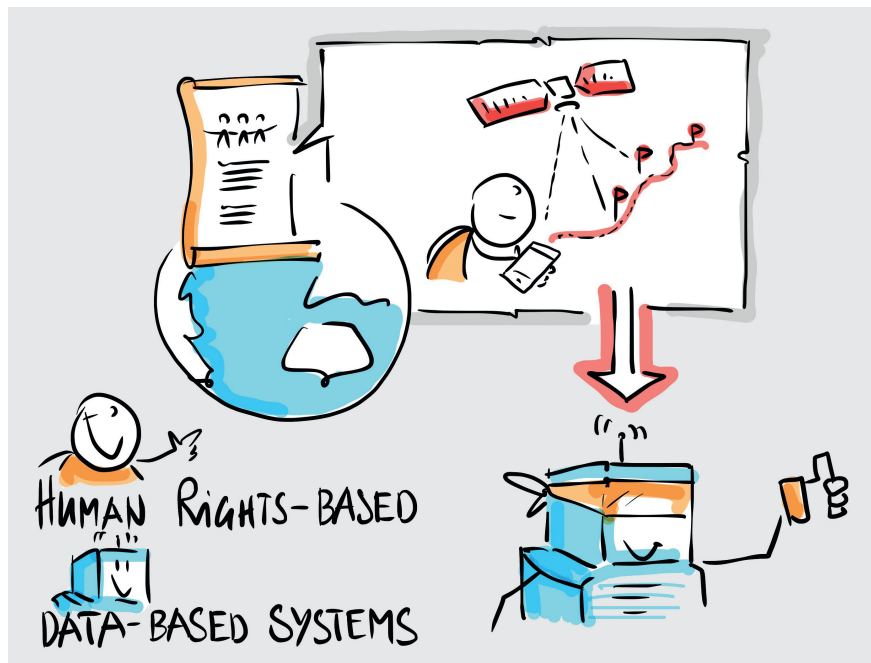
z.B. um eine Effizienzsteigerung zu erreichen. Diese Regeln weisen aber keine ethische Qualität auf. Denn Maschinen scheitern am Prinzip der Verallgemeinerbarkeit. Daher sollte z.B. im Bereich der Automatisierung von Mobilität der Begriff «autonome Fahrzeuge» vermieden und durch den Begriff «selbstfahrende Fahrzeuge» ersetzt werden [6].

Um einem allfälligen Missverständnis vorzubeugen: DS können ethische Regeln einprogrammiert bzw. antrainiert werden, um ein ethisch legitimes Handeln einer Maschine zu erreichen [7]. Die Menschenwürde und die Menschenrechte aller Menschen wirken in dieser Verantwortung der Menschen konstituierend und leitend, DS mit Ethik zu schaffen: human rights based DS, HRBDS.

Komplexität von Ethik

DS sind in der Lage, ethische Regeln zu befolgen sowie darauf basierend Handlungen zu vollziehen. Damit man im Zuge dessen der komplexen Realität gerecht wird, braucht es dabei deutlich mehr als Regeln wie die «Asimovschen Gesetze» [8].

Um der Komplexität der Ethik gerecht zu werden, ist aber nicht mathematische oder digitale Ethik zu betreiben. Sondern es gilt zu beachten, dass sich Ethik in ihrer Komplexität und in ihrer Gesamtheit wegen ihrer Sensibilität für die regelüberragende Einzigartigkeit des Konkreten als nicht in die Sprache der Mathematik und der Programmierung übersetzbar erweist. U. a. deswegen handelt es sich bei Ethik nicht um Kasuistik. Gewisse Aspekte der Ethik lassen sich als Regeln datenbasierten Systemen programmieren bzw. antrainieren. Einige ethische Elemente entziehen sich diesem digitalen Zugriff. So lassen sich aber DS beispielsweise das Prinzip der Menschenwürde aller Menschen und die Menschenrechte sowie ethische Handlungsmaximen (Verbot der Lüge, des Diebstahls) vermitteln. Selbst in diesem Bereich des Möglichen sollte aber nicht vernachlässigt werden, dass DS zwar diese Regeln lernen und befolgen können, diese Achtung der Regeln ohne Wissen um die ethische Qualität der Regeln erfolgt. In anderen Worten: DS würden auch nichtethische oder unethische Regeln in gleichem Masse respektieren.



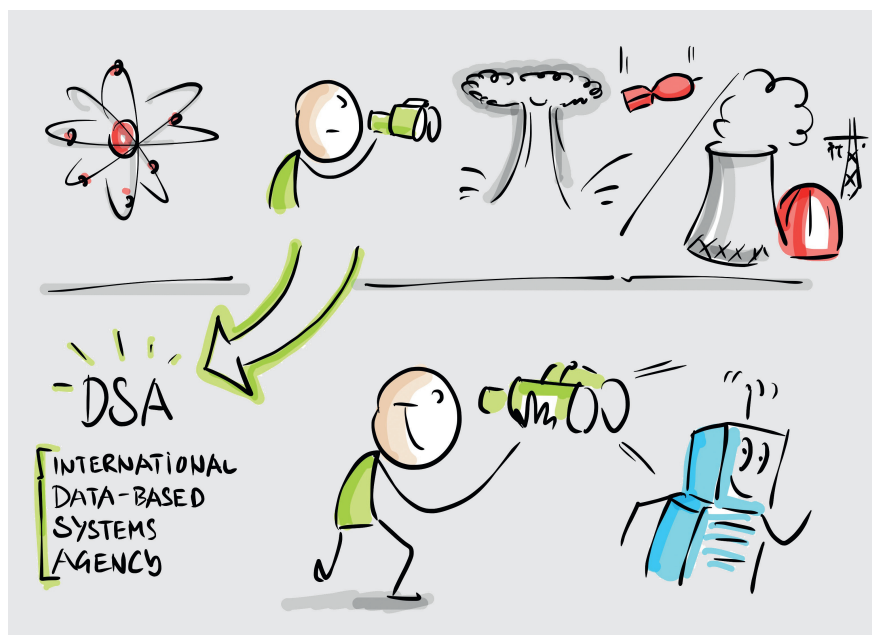
Menschenrechte als Basis für datenbasierte Systeme (DS).

Schwierig bis unmöglich wird ein Transfer der Ethik in die Mathematik und in die Programmierung u. a. dann, wenn Handlungsmaximen divergieren oder kollidieren. Die zunehmende Komplexität der Lebenswirklichkeit fordert zu realitätsgerechten Normeinsichten und zu differenzierteren und besseren Abwägungen heraus, was DS wegen ihrer fehlenden Moralfähigkeit überfordert. Unmöglich erweist sich die Übersetzung der Ethik in die Sprache der Mathematik, der Programmierung und der Digitalisierung in Situationen und Fällen, wo beim Menschen die Tugend der Epikie und das Gewissen besonders zum Zuge kommen. «Epikie ist die Berichtigung des Gesetzes da, wo es in Folge seiner generellen Fassung lückenhaft ist.» [9] Epikie trägt der Wirklichkeit Rechnung, dass in der konkreten Begegnung mit konkreten Menschen in einer konkreten Situation Regeln an ihre Grenzen kommen, da das Konkrete in seiner Einzigartigkeit das Regelhafte weit übertrifft. Darüber hinaus wird den Menschen eine Normgestaltungsverantwortung zugemutet, die für DS aufgrund ihres Mangels an Moralfähigkeit unerreichbar bleibt. Diese menschliche Normgestaltungsverantwortung zielt darauf ab, dass Regeln kontinuierlich kritisch hinterfragt und – falls nötig – im Dienste einer prospektiven ethischen Verbesserung von Menschen angepasst werden [10].

Die Verantwortung, Normen zu gestalten und zu nutzen

Diese prospektive kreative Ebene umfasst auch eine Normgebungsverantwortung der Menschen. «Die Wahrnehmung des sittlichen Anspruchs bedeutet ja keineswegs bloss ein Ableben von normativ festgelegtem Sach- und Sinnverhalten, sondern ist immer schon ein schöpferisches Sehen und Entdecken. Schöpferisch ist dieses Sehen und Entdecken dadurch, dass der Mensch aufgefordert ist, in seiner Fantasie neue sinnvolle Momente der Lebensgestaltung zu riskieren, die im bisherigen Regelsystem nicht vorkamen. Die sittliche Gutheit der Person drängt ihn dazu, das menschlich Richtige in Form von Modellen weiterzuentwickeln.» [10] Auch die Normgebungsverantwortung geht weit über das in die Sprache der Mathematik und der Programmierung Übersetzbare hinaus und kann daher datenbasierten Systemen nicht übertragen werden.

Die Komplexität der ethischen Auseinandersetzung mit datenbasierten Systemen erhöht zusätzlich noch zum einen die sogenannte «dual use»-Problematik. Diese umfasst das Potenzial, dass das gleiche datenbasierte System für einen legitimen oder illegitimen Zweck eingesetzt werden kann. Dies lässt sich am Beispiel der Drohne veranschaulichen: Drohnen können dafür eingesetzt werden, um Menschen in Katastrophen-



Zum Wohl und zum Schutz der Menschen und des Planeten braucht es dringend die Schaffung einer Internationalen Agentur für datenbasierte Systeme DSA bei der UNO.

gebieten mit lebensnotwendigen Nahrungsmitteln zu versorgen, wo Menschen nicht mehr hinkommen. Die gleichen Drohnen können aber auch als sogenannte «Killerdrohnen» für gezielte Tötungen eingesetzt werden. Und selbstverständlich fällt die ethische Bewertung eines solchen DS in den beiden Situationen diametral entgegengesetzt aus. Diese «dual use»-Problematik zeigt die Grenzen einer generellen Bewertung von Drohnen auf. Eine allgemeine Einschätzung zu Drohnen muss für eine differenzierte ethische Analyse durch eine Betrachtung der Drohnen in ihrer Anwendung ergänzt werden.

Zum anderen kann es zu «Ambivalenz» kommen. Ambivalenz bezeichnet die Möglichkeit, dass ein datenbasiertes System einem ethisch guten Zweck dienen kann, dabei aber auch gleichzeitig etwas ethisch Schlechtes umfasst. Dies lässt sich am Beispiel der Automatisierung von Mobilität zeigen: Staus mit ihren ökologisch und ökonomisch negativen Auswirkungen entstehen unter anderem wegen der Variation von Tempo und aufgrund der unterschiedlichen Geschwindigkeiten der einzelnen Fahrzeuge. Würde es gelingen, beides zu minimieren, käme es zu einer Verringerung von Staus. Automatisierte Mobili-

tät umfasst beides – die Abschwächung der Variation von Geschwindigkeit durch einen Verkehrsfluss auf gleichem Tempo sowie die Vereinheitlichung von Geschwindigkeiten aller Fahrzeuge durch die dezentrale Kommunikation zwischen den einzelnen Fahrzeugen. Verbunden mit der dezentralen Kommunikation zwischen den Fahrzeugen sind auch deutlich weniger Unfälle zu erwarten, da die fahrenden Systeme dezentral untereinander dafür sorgen, dass es grossmehrheitlich nicht zu Zusammenstößen kommt. Dies bedeutet nicht, dass Unfälle vollumfänglich ausgeschlossen werden können. Denn selbstverständlich kann es bei fahrenden Systemen auch zu Fehlern kommen; sie treten aber viel seltener auf, da Unfallursachen wie z. B. Ablenkung, Müdigkeit, Stress, Emotionen ausgeschlossen werden können [11]. Weniger Staus, weniger Unfälle – erfreuliche Perspektiven aus ethischer Sicht, die sich bei der Automatisierung von Mobilität ergeben.

Gleichzeitig umfasst die Automatisierung von Mobilität die Gefährdung der menschenrechtlich geschützten Privatsphäre und des menschenrechtlich garantierten Datenschutzes. Zudem trägt sie zur global ungleichen Verteilung von Vor- und Nachteilen bei. Während reiche Länder auf automatisierte Mobilität umstellen können, leiden Menschen in ärmeren Ländern unter menschenunwürdigen und menschenrechtsverletzenden Arbeitsbedingungen bzw. moderner Sklaverei sowie unter Umweltzerstörung bei der Schürfung der Rohstoffe bzw. bei der

RÉSUMÉ

Intelligence artificielle et éthique

Opportunités et risques

Smart grids, augmentation de l'efficacité aux niveaux du commerce de l'énergie ainsi que de l'utilisation de l'infrastructure et des ressources énergétiques, interaction avec les clients basée sur l'IA, etc. – l'intelligence artificielle ouvre la voie à de grandes opportunités dans le contexte du système énergétique. Il est toutefois important de ne pas se focaliser uniquement sur le thème de la « cybersécurité » : les systèmes basés sur des données nécessitent en effet de toute urgence une orientation éthique.

Des principes et des normes éthiques, ainsi qu'une réglementation juridique basée sur ceux-ci, ouvrent la possibilité de promouvoir de manière ciblée l'utilisation des opportu-

nités éthiques des systèmes basés sur les données, afin de pouvoir les exploiter en tant qu'innovation pour le bien éthique. Ils permettent en outre d'éviter les risques et les dangers éthiques.

Le fait que l'humanité ait déjà montré au cours de son histoire qu'elle est capable de ne pas toujours poursuivre et mettre en œuvre « aveuglément » ce qui est techniquement faisable, mais aussi d'y renoncer, de se retenir ou de s'auto-limiter lorsque le bien-être de l'humanité et de la Terre est en jeu, donne des raisons d'espérer et de croire que cela est possible.

Herstellung der automatisiert fahrenden Systeme. Ethik ertut hier dringlichen Handlungsbedarf.

Schlussbemerkung

Datenbasierte Systeme und deren Nutzung brauchen unbedingt und mit hoher Dringlichkeit ethische Orientierung von Menschen. Ethische Prinzipien und Normen und eine darauf basierende rechtliche Regulierung eröffnen die Möglichkeit gezielter Förderung der Nutzung ethischer Chancen datenbasierter Systeme, um sie als Innovation zum ethisch Guten einsetzen zu können. Diese befähigen zudem zur Vermeidung von ethischen Risiken und Gefahren.

Anlass zur Hoffnung und Zuversicht, dass dies möglich ist, gibt, dass die Menschheit in ihrer Geschichte bereits

gezeigt hat, dass sie fähig ist, nicht immer das technisch Machbare «blind» zu verfolgen und umzusetzen, sondern auch auf technisch Machbares verzichtet, sich zurückhält oder sich selbst beschränkt, wenn das Wohl der Menschheit und des Planeten Erde auf dem Spiel steht. Beispielsweise haben – vereinfacht gesagt – Menschen im Bereich der Nukleartechnologie geforscht, man hat die Atombombe entwickelt, sie wurde mehrmals abgeworfen, und dann hat die Menschheit die Forschung und Entwicklung im Bereich der Nukleartechnologie substanziell und massiv global eingeschränkt und reguliert, um noch Schlimmeres zu verhindern, obwohl es dagegen massiven Widerstand gab. Diese Unterbindung ist weitestgehend gelungen – dank eines internationalen Regimes, konkreter Durchsetzungsmechanismen und der Internationalen Atomenergiebehörde (IAEA) bei der UN.

- [2] Arjun Kharpal. «Stephen Hawking says A.I. could be «worst event in the history of our civilization»», CNBC News, 6. November 2017, www.cnbc.com/2017/11/06/stephen-hawking-ai-could-be-worst-event-in-civilization.html
- [3] Für eine vertiefte Betrachtung: Peter G. Kirchschräger, Digital Transformation and Ethics. Ethical Considerations on the Robotization and Automatization of Society and Economy and the Use of Artificial Intelligence, Nomos-Verlag, 2021.
- [4] Peter G. Kirchschräger, «The correlation between ethics and technology», in Giulia Isetti, Elisa Innerhofer, Harald Pechlaner, Michael de Rachewiltz (Hrsg.), Religion in the Age of Digitalization. From New Media to Spiritual Machines, Routledge, 2020, S. 165-180.
- [5] Immanuel Kant, Grundlegung zur Metaphysik der Sitten. Werkausgabe Weischedel. Vol. 7., Suhrkamp, 1974.
- [6] Peter G. Kirchschräger, Automatisierung von Mobilität – theologisch-ethische Überlegungen, Feinschwarz, 13. Januar 2017. www.feinschwarz.net/automatisierung-der-mobilitaet-theologisch-ethische-ueberlegungen
- [7] Vgl. Wendell Wallach, Colin Allen, Moral Machines. Teaching Robots Right From Wrong, Oxford University Press, 2009.
- [8] Vgl. Isaac Asimov, Meine Freunde, die Roboter (Bibliothek der Science Fiction Literatur 20), Sonderausgabe, Heyne Verlag, 1982, S. 67.
- [9] Aristoteles, Nikomachische Ethik, 4. Aufl., Felix Meiner Verlag, 1985, S. V, 14, 26.
- [10] Günter Virt, Damit Menschsein Zukunft hat. Theologische Ethik im Einsatz für eine humane Gesellschaft, Echter Verlag, 2007, S. 42-43.
- [11] Vgl. Kay W. Axhausen, «Autonome Fahrzeuge. Erste Überlegungen», Vortrag am 5. September 2016 im Rahmen der Sommerakademie der Schweizerischen Studienstiftung zum Thema Automatisierung der Mobilität, Magliaso.

Neuer Studiengang

Ethik-Spezialisierung

Neuer Masterstudiengang «Ethik» ab Herbst 2023, unter anderem mit Spezialisierung «Ethik der digitalen Transformation»:
www.unilu.ch/master-ethik

Literatur

→ Smart Grids, International Energy Agency, 2022, from www.iea.org/reports/smart-grids

Referenzen

- [1] Catherine Clifford, «Elon Musk: «Mark my words – A.I. is far more dangerous than nukes»», CNBC News, 13. März 2018, www.cnbc.com/2018/03/13/elon-musk-at-sxsw-ai-is-more-dangerous-than-nuclear-weapons.html



Autor

Prof. Dr. theol. lic. phil. **Peter G. Kirchschräger** ist o. Professor für Theologische Ethik und Leiter des Instituts für Sozialethik ISE. → Universität Luzern, 6002 Luzern → peter.kirchschrager@unilu.ch

Weil Wissen verbindet.

Ihr Fachverband seit 1889.

www.electrosuisse.ch/verband