

Gutes Leben und künstliche Intelligenz?

Schlaglichter aktueller Wissenschaftsdiskurse

Inhalt

Reise in die Robosphäre	1
Wo kommen wir her? Wo geht's hin?	2
Der Roboter: Unser Ebenbild?	3
Wanted: Eine Ethik jenseits der Bewusstseinsfrage	4
Neue ethische Paradigmen	6
Konkreter Fall: Altenpflege	7
Fühlende Roboter	8
Anthropomorphisierung ist menschlich	9
Wer und was werden wir sein?	10
Die Suche nach dynamischer Balance	11



Reise in die Robosphäre

Arbeit 4.0, Automatisierung, Internet der Dinge – im aktuellen Mediendiskurs geistert die Frage herum, wie unsere Zukunft aussehen wird, wenn sich alle Bereiche digital vernetzen und künstliche Intelligenz unsere Lebenswelt umfassend strukturieren wird.

Dieser Diskurs läuft meist getrennt von gesellschaftsgestalterischen Ansätzen, die sich an einem „guten Leben für alle“ orientieren oder systemische Kritik äußern. Die Frage nach der vernetzten Zukunft ist im Mainstream also meist eine Frage des: „Wie?“ und nicht des „Warum?“ oder „Warum lieber nicht?“.

Auf meiner Recherche zur [Konferenz](#) „Envisioning Robots in Society“ (14.-17.02.18), veranstaltet von der Universität Wien und der dänischen Universität Aarhus, habe ich mir einen Eindruck davon verschafft, wo wir gesellschaftlich stehen. Für meine weitere „utopische Wissensarbeit“ wollte ich Schnittstellen ausmachen, an denen beide Diskurse zusammenzubringen sind.

Einerseits erscheint das im Status Quo hoffnungslos und naiv: Denn die Entscheidungs- und Möglichkeitshorizonte derjenigen, die künstliche Intelligenz entwickeln und auf den Markt bringen, werden bestimmt durch die gegebenen wirtschaftlichen Bedingungen. Unternehmen müssen KI da einsetzen, wo sie Kosten spart und jährliche Profite erhöht. Dieses Ziel determiniert fast alle ingenieurs- und betriebswirtschaftlichen Entscheidungen. Somit verhindert es strukturell, dass man sich *umfassend* an Werten wie Nachhaltigkeit und Gutem Leben für alle ausrichtet. Die Frage, wo KI also zu einem guten Leben beiträgt, ist ein Forschungsanliegen, das man sich als Luxuszugabe leisten kann, kann aber in den gegebenen Strukturen nicht das neue Leitparadigma werden: Kein Unternehmen, das sich daran orientiert, wäre mehr global wettbewerbsfähig.

Soweit ist die naheliegende Sicht auf die Dinge.

Andererseits gibt es da ein diffuses Bewusstsein, dass man so langsam Ethik und Wirtschaft integrieren sollte, dass man ja global irgendwie kooperieren müsste, um den Klimawandel und bestehende Krisen und globale Ungerechtigkeiten in „den Griff zu bekommen“. Und es gibt mantraartig wiederholte Aussagen, vornehmlich in den Fazit-Abschnitten von Artikeln, Forschungsarbeiten, Büchern und Vorträgen: *Ja, wir brauchen eine neue Ethik für den Umgang mit KI, was wir in unserem Alltag wollen und was nicht, etc.* Nach dieser Einsicht folgt meist kein Abschnitt mehr. Vielleicht, weil das Unterfangen zu komplex, zu unüberschaubar, zu ambitioniert ist, um sich weiter mit ihm zu befassen.

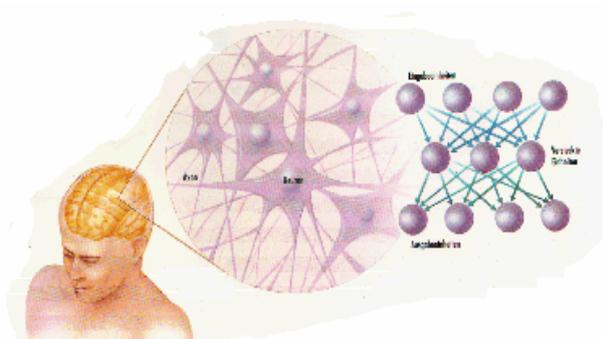
Die „Einsicht“, dass es eigentlich unmöglich ist, den KI-Fortschritt zu gestalten, ohne Rahmenbedingungen wie Wettbewerbsfähigkeit, Kostenersparnis etc. zu berücksichtigen, kämpft also latent im Unter- oder Halbbewussten mit dieser diffusen Unruhe, dass es doch irgendwie gestaltet werden müsse, dass sich um eine gute Zukunft bemüht werden müsste.¹

In Wien haben sich Geistes- und Kulturwissenschaftler*innen, Ingenieur*innen und Programmierer*innen getroffen, um über entsprechende Fragen zu diskutieren; einige Schlaglichter möchte ich hier vorstellen. Nicht anwesend waren die Protagonisten der Tagung selbst; mich hätten ein paar Anschauungsexemplare gefreut, habe aber insgesamt nur einen kleinen Industrieroboter zu sehen bekommen, der ein paar Mal seine Greifarme schwenkte.

Wo kommen wir her? Wo geht's hin?

Das **Wort** Roboter („künstlicher Mensch, Automat“) kommt aus dem Tschechischen, erstmals benutzt 1920, Grundlage ist das Substantiv *robotá* „Arbeit, Fronarbeit“. Das geht auf das altkirchenslawische *rab* „Sklave“ zurück – ein Roboter ist also jemand, der zur Fronarbeit gezwungen wird.

Es lassen sich ([Raja Chatila](#) zufolge) drei **Perioden** der Robotik unterscheiden: In den Jahren 1960 bis 1985 ging es primär darum, Roboter herzustellen – Ingenieure befassten sich mit Aufgaben wie Design, der Kinematik, der Kontrolle, Akkuratheit, Wiederholung und Schnelligkeit. Ab 1985 über die Jahrtausendwende bis zum Jahre 2010 rückten der Service, Risiken, das Umfeld der Roboter in den Vordergrund, Fragen wie die nach ihrer Autonomie



Natürliche und künstliche neuronale Netzwerke

¹ Mein eigenes Forschungsanliegen mit der utopischen Wissensarbeit bewegt sich in diesem Spannungsfeld. Es speist seine Motivation und seine Energie nicht aus der Überzeugung, dass man das Wissen, das es hervorbringen soll, im Hier und Jetzt flächendeckend einsetzen und integrieren kann, sondern aus der Überzeugung, dass im Hier und Jetzt völlig neue Denkansätze das Recht haben sollten zu entstehen, die dann da sein werden, wenn es darauf ankommt; die sich im Stillen entwickeln, wachsen, modifiziert, erprobt werden können, um etwas in der Hand zu haben, wenn irgendwann nach Alternativen gefragt werden sollte.

und Mensch-Maschine-Interaktion wurden relevant. Seit 2010, in einer Zeit zunehmender Vernetzung der Dinge, fokussiert man sich unter dem Stichwort *Industrie 4.0* auch auf Fragen des Transports und neuer Services im Bereich Autonomie, Interaktion, Lernen. Das *Lernen* wird eine Schlüsselrolle dabei spielen, inwiefern KI zu einem Akteur unseres Alltags wird, der uns auf bestimmten Ebenen ebenbürtig oder überlegen erscheint. Dieser muss sich eigenständig in Interaktion mit der Umwelt Fähigkeiten und Wissen aneignen können. Das geht dank der Technologie neuronaler Netzwerke: Dabei werden Computersysteme nicht mehr programmiert, sondern ähnlich dem menschlichen Gehirn mit Hilfe von Daten trainiert. Sie lernen sozusagen selbst.

Bei der Robotik, dem Fokus der Konferenz, geht es um *automatisiertes* intelligentes Verhalten. (Was allerdings „Intelligenz“ bedeutet, ist allgemein nicht geklärt – Intelligenz ist letztlich immer nur das, was der Intelligenztest misst!!).

Generell bezeichnet **künstliche Intelligenz** den Versuch, eine menschenähnliche Intelligenz nachzubilden, d. h., einen Computer so zu bauen oder zu programmieren, dass er eigenständig Probleme bearbeiten kann. Als **schwache KI** wird sie bezeichnet, wenn sie sich auf ein konkretes Anwendungsproblem (Schachspielen, Gesichtserkennung, Börsenhandel) fokussiert. Hier wird intelligentes Verhalten mithilfe der Mathematik und Informatik simuliert. Eine **starke KI** dagegen würde sich intelligent wie ein Mensch verhalten. Die starke KI mit ihren philosophischen Ambitionen („Was ist menschliches Bewusstsein?“) ist bisher nicht viel weitergekommen, wohingegen die schwache KI wachsende Fortschritte macht.

Künstliche Intelligenz ist immer auf eine bestimmte Anwendung hin designt: Sie dient dazu, menschliche Arbeitskraft zu ersetzen oder zu „verlängern“, Menschen zu assistieren oder sie zu rehabilitieren und sogar zu optimieren/ enhancen.

Der Roboter: Unser Ebenbild?

Den Menschen als Maschine nachbauen, dieser Traum der starken KI ist Jahrhunderte alt. Einer, der sich dieser Vision widmet, ist [Hiroshi Ishiguro](#), einer der Keynote-Speaker: Der japanische Robotiker hat eine Kopie von sich selbst konzipiert („Geminoid“), die angeblich schon als Spielpartner für seine kleine Tochter taugt, und sieht eine „Human-Robot-Symbiotic Society“ kommen. Um menschenähnliche Akteure zu schaffen, lässt er Gedächtnis-, Inferenz-, Computation-, und Lernleistungen in seine Robots integrieren und bedient sich dabei neuro-wissenschaftlicher Erkenntnisse. Es gilt, Fähigkeiten wie Stimmerkennung, Bildverständnis, Sensoren, Manipulator, Mechanik, Beziehungsfähigkeit, Bewusstsein, Intention/Verlangen, Intelligenz und anderes zu „implementieren“. Sein Ziel: Dass seine Roboter irgendwann einmal den „Total Turing Test“² bestehen, was bei ihm anscheinend heißt, als sozial akzeptiertes Mitglied in einer Gesellschaft anerkannt zu sein.



„Geminoid“ von Hiroshi Ishiguro

Den Nutzen der Roboter erklärte er folgendermaßen: Er beschrieb eine Situation mit Kindern an einem Tisch, in der jedes sich mit seinem Smartphone beschäftigte. Erst, als einer seiner Roboter dazu kam, begannen sie, miteinander zu interagieren.

² Bei diesem Test stellt ein Mensch per Terminal beliebige Fragen an einen anderen Menschen bzw. eine KI, ohne dabei zu wissen, wer jeweils antwortet. Er muss danach entscheiden, ob es sich beim Interviewpartner um eine Maschine oder einen Menschen handelte. Kann mensch die Maschine nicht von dem Menschen unterscheiden, so ist laut Turing die Maschine intelligent.

Wie weit muss er sich von einer lebensweltlichen Intuition entfremdet haben, um Roboter ernsthaft als das am besten geeignete Mittel anzusehen, menschliche Beziehungen zu bereichern? Ist hier nicht vielmehr Folgendes ein sinnvolles Prinzip:

Erhalte bewusst Räume des Austauschs, in denen sich Menschen unvernetzt und ohne KI begegnen können.

Die Konversation mit einem, nennen wir ihn mal, humanoiden Roboter, ist natürlich ein wesentlicher Designaspekt. Hier gibt es laut Ishiguro zwei Zugänge: Die integrierte KI kann sich einfach bei *Big Data* bedienen und das Internet auf einfache Konversationsmuster, thematisch unbegrenzt, durchsuchen. Chatbots können über 4000-5000 „conversational patterns“ verfügen. Ein anderer Zugang ist der „story-based approach“, bei dem in einer Art Entscheidungsbaum Intentionen, Ziele und Handlungen verbunden werden. Der Akt der Begrüßung ist beispielsweise beschreibbar als Handlung (Gruß) mit Intention (Gespräch abschließen) und Ziel (Beziehung herstellen, Selbstwert herstellen). Seine Schöpfung ERICA kann das bereits. Der Robot verfügt über Gesichtserkennung und hat ein Sprachsystem, das auch Witze erzählen kann.

„It’s nice to build robots, but to think that they can do like people is nonsense“³, sagt der Installationskünstler Simon Penny. Auch die KI-Expertin [Joanna Bryson](#) betont: „Robots are just an artefact!“, selbst, wenn sie einen ‚sehen‘, ‚hören‘, ‚fühlen‘ und ‚verstehen‘ werden können. Sie seien intentionale, designte Akte, für die bezahlt werde, sodass unsere ethischen Instinkte bei ihnen fehl am Platz sind, egal, ob sie nun Witze erzählen können oder nicht. Sie bemängelt, dass Medien Roboter häufig als eine Art fremde Lebensform präsentieren – kein Wunder, sobald Roboter menschlich aussehen, verhalten wir uns anders gegenüber ihnen.

Zur Aufklärung müssten wir jedoch transparent machen, wie KI funktioniert. Sie zeigte dazu eine Visualisierung der Prozesse, die bei einem kleinen Fahrzeugroboter ablaufen, der sich in einem Raum orientiert. Das System arbeitet Entscheidungsbäume ab.⁴ Bei nicht-humanoiden Robotern sei es leichter, sie als Maschine wahrzunehmen. Man dürfe aber auch bei menschenähnlichen Robotern nicht so weit gehen, sie als moralische Akteure ihrer eigenen Sozietät zu sehen:

*Mache dir als Nutzer*in immer bewusst, dass du mit einer intentional designten Maschine interagierst, nicht mit einem lebenden „Alien“.*

Wanted: Eine Ethik jenseits der Bewusstseinsfrage

Debatten um starke KI haben einen großen Dreh- und Angelpunkt: den Unterschied zwischen „IS“ versus „SEEMS LIKE“, zwischen dem, was an Prozessen abläuft, und dem, wie es uns erscheint: Die Frage nach *Simulation* des Menschlichen, letzten Endes die Frage nach Existenz oder Simulation menschlichen Bewusstseins.

Doch ist die Frage, ob Bewusstsein ‚künstlich‘ entstehen kann, wirklich so relevant? Ist es wirklich so bedeutsam, dass sich feuernde Neuronen und elektronisches Feuern (1-0) so sehr zu ähneln scheinen? Spiegelt unser Hardware-Software-Dualismus unser westliches problematisches Bild von einem getrennten „Körper“ und „Geist“?

Bryson nennt verschiedene aufeinanderfolgende Stufen, die man auf dem Weg zu einem



Sicherheitsdienst REEM (PALRobotics)

³ [Simon Penny: Robotics and Art, Computationalism and Embodiment.](#)

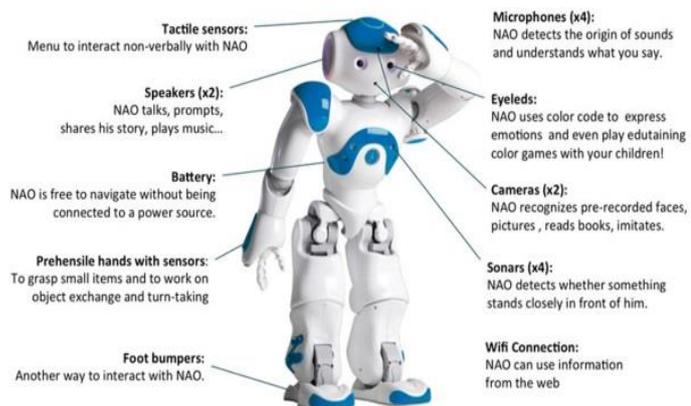
⁴ Hier geht es also nicht darum, dass das System selbstständig ungerichtet lernt. Das könnte, wenn ich das richtig verstanden habe, erst bei Systemen mit neuronalen Netzwerken der Fall sein. Wird es jemals möglich sein, so etwas zu schaffen? Gibt es ab dem Zeitpunkt, an dem das System so komplex ist, dass es einen Menschen perfekt simuliert, kein Unterschied mehr zwischen Simulation und Echtheit?

Bewusstseinszustand unterscheiden kann: reaktiv, adaptiv, emotional, selbst-bewusst, empathisch, sozial u.a.⁵ Sie sieht Bewusstsein als „information integration“ (Tononi), Intelligenz ist nicht gleich Bewusstsein: Nur, weil Intelligenz, Bewusstsein, Sprache häufig korrelieren, heie das nicht, dass es eine kausale Beziehung zwischen ihnen gebe. Bewusstsein zu konzeptionalisieren, zu operationalisieren, um es dann sichtbar (mit Zahlen messbar??) machen zu knnen – schwer zu sagen, ob das berhaupt jemals funktionieren wird. Auerdem: „We have no robots that are making any claim on consciousness right now.“ Sie sind als Sklaven geschaffen und auf eine bestimmte Funktion hin designet: „How can the robot have the right NOT to do what it is meant to do?“

Kein Bewusstsein, also keine Rechte? Was geschieht, wenn sich Systeme durch *deep learning* so weiterentwickeln, dass sie sich aus dieser Funktion befreien, ist eine andere Frage. Doch wenn sie das einmal knnen werden, werden sie es dann berhaupt noch ntig haben, dass wir ihnen Rechte zusprechen??

In einem Workshop zum Thema „KI-Rechte“ verglich jemand Roboter mit Gartengerten: die haben gem einer guten Grtnerpraxis auch das „Recht“ auf Gepflegtwerden, aber natrlich nicht im humanen Sinn: „We will have no robots in the future which will be concerned about if we harm them or not“: Es geht um eine Subjekt-Objekt-Beziehung. Robots sind “embodied, embedded, social systems with real interaction”. Gut mglich, dass sie uns bald erscheinen, als htten sie Bewusstsein: Bereits jetzt kann ein Robot seinen eigenen Stil entwickeln, denn seine Erinnerungen prgen seine Autobiografie, sodass Exemplare desselben Modells mit der Zeit individuell werden. Dennoch sind sie Maschinen.

Egal. Es braucht in jedem Fall eine Ethik⁶, die sich mit KI befasst – es geht schlielich um *uns*, und das, was KI hier und heute mit uns macht und wofr wir sie einsetzen. Nicht schwer anzunehmen, dass wir dabei am besten interdisziplinr vorgehen sollten. Doch damit dieser Dialog passieren kann, mssen Human- und Naturwissenschaftler*innen sowie Ingenieur*innen erst einmal lernen, dieselbe Sprache zu sprechen. Der Ingenieur [Bernhard Dieber](#) fordert eine „mandatory social participation of philosophers in engineering.“ Wre es also gut, etwa NAO (Aldebaran Robotics)



Szenariotechniken zu ethischen Fragen zu institutionalisieren?

Es wird notwendig, sich ber das Menschen- und Weltbild auszutauschen, das alle Parteien haben. Was ist Wirklichkeit? Wo ist sie konstruiert, wo ist sie etwas, das als extern und unvernderlich angesehen wird? Ingenieur*innen, sagt er, lieben Standardisierungen: Es gelte, ihr *mindset*, ihre Denkstrategien zu erweitern, auch das Unvorhersagbare zu akzeptieren; sie sind es gewohnt zu kalkulieren, deshalb brauche es eine Brckenfunktion, eine Dolmetscherrolle zwischen Philosoph*innen und Ingenieur*innen.

⁵ [Bryson: Roadmap to conscious machines](#)

⁶ Eine globale Initiative entwickelt aktuell ethische Standards fr den Umgang mit KI: IEEE, „[Global initiative on Ethics of Autonomous and Intelligent Systems](#)“. Die von 250 Akteuren aus Wissenschaft, NGOs, Industrie etc. untersttzte Initiative hat – mit dem internationalen Recht als Grundlage – 13 Standard-Entwrfe erarbeitet, das Ziel: „put ethics in action“. Die Diskussion soll konsensbildend, offen, global und inklusiv sein: Mitarbeit der Geisteswissenschaften und Zivilgesellschaft erwnscht. Diejenigen Standards, die eines der Mitglieder fr gut befindet, werden weiterentwickelt. Bereiche sind etwa: Menschenrechte, Wohlbefinden, Verantwortlichkeit,

Neue ethische Paradigmen

Gut. Aber welchem gemeinsamen „Denkparadigma“ sollte eine Ethik folgen, wie sie nun interdisziplinär entwickelt werden könnte? Auf der Konferenz wurde, vor allem von der dänischen Professorin Johanna Seibt ([Research Project on Integrative Social Robotics \(INSOR\)](#))⁷, eine *prozessphilosophische* Perspektive stark gemacht.

Das heißt meines Erachtens Folgendes: Betrachtet man ethische Probleme mit Robotern, sollte man sich auf *Ereignisse* und *Prozesse* als die grundlegenden Elemente der Realität fokussieren und keine „stabilen Substanzen“ („Seele“, „Schuld“, Mensch als isoliertes Subjekt mit festen Eigenschaften) als Grundelemente annehmen, wie es in der abendländischen Philosophie eine lange Tradition ist. Das *Dazwischen*, die *Beziehung* zwischen allen Elementen gewinnt an Bedeutung, alles schwimmt ineinander und wird fluide. Das führt dazu, dass man das Problem aus verschiedensten Perspektiven erfassen kann und „kontextsensitiv“ mit Feedback aus der Praxis arbeitet.

Ein zentraler Begriff ist wie gesagt die *Simulation*. Wir müssen diesbezüglich in der Beschreibung von Roboter-Mensch-Interaktion auf unser Vokabular achten. Ein Roboter *handelt* nicht. Ein Mensch handelt, in dem Sinn, dass dies mit Absicht zu tun hat. Ein Roboter *simuliert* handeln: Ein Roboter simuliert, mit jemandem zu sprechen. Er läuft, sammelt Müll, meldet ein Gewaltdelikt in einer Shoppingmall: Alles Handlungen, aber unintentionale. Die Handlungsbegriffe müssen auseinandergehalten werden.

Professorin Seibt brachte mehrmals die „Warum?“-Frage ein: Wir sollten uns grundsätzlicher fragen, *warum* ein Robot konstruiert werde. Die obengenannten wirtschaftssystemischen Ursachen wurden auf der Konferenz leider kaum reflektiert. Eine Ethik ja, aber bitte eine, die nicht die gegebenen wirtschaftlichen Bedingungen sprengt! Diese Bedingungen sind wie der sprichwörtliche Elefant im Raum: Würde man sich bis zum Äußersten denkerisch vorwagen, würde er sichtbar werden. Aber sobald er ausgesprochen wird, kommt man nicht umhin, grundsätzlich nach Alternativen zu suchen. *Und das ist in seiner Komplexität ziemlich furchteinflößend und wenig greifbar ... und die alltäglichen Herausforderungen sind schon genug, dafür gibt's keine Zeit und keine psychischen Ressourcen ...*

Ein Prinzipvorschlag zum Robotereinsatz lautete:
KI sollte nur das übernehmen, was der Mensch NICHT kann, was er aber übernehmen SOLLTE.

Mir wurde nicht ganz klar, was der Halbsatz „was er übernehmen sollte“ bedeutet. Was sollte ein Mensch denn übernehmen? An welchen Kriterien bemisst sich das? Welche Leitwerte bestimmen das?

Auf jeden Fall spiegelt sich hier eine vage Orientierung am guten Leben wider: Der Roboter als hilfreiche Ergänzung des Menschen, nicht als Ersatz. Ein junger Mensch kann einem alten Menschen Essen

anreichen. Warum sollte es ein Roboter tun? Ein Mensch kann nicht seltene Erden auf einem Asteroiden abbauen. Das sollte ein Roboter tun. Natürlich gibt es viele offene Fragen: Ich kann mir eigenständig einen Zug buchen, indem ich entsprechende Befehle im Netz eingebe und Transaktionen veranlasse. Eine KI könnte es allerdings schneller und effektiver tun. Sollte ich in der Zeit nicht lieber einem alten Menschen das Essen anreichen?



Transportrobot GITA (PIAGGIO)

⁷ Der [Integrated Robot Philosophy](#), wie sie in Aarhus entwickelt wird, geht es darum, 1) leitende Werte zu definieren, 2) Anforderungen zu identifizieren, 3) Anforderungen zu designen, und dann 4) einen Prototyp zu bauen. Transparenz, Missbrauchsbewusstsein. Hier zum Beispiel eine laufende Arbeitsgruppe zum Thema [ethisch motiviertes Nudging](#) von Robotern.

KI ist gut im seriellen Abarbeiten von Aufgaben und im synchronen Verarbeiten riesiger Datenmengen. Darin ist der Mensch schlecht. Der Mensch ist gut in empathischem Handeln, Mustererkennung und so weiter. Überhaupt: Er ist lebendig. Insofern wie gesagt nicht zu vergleichen mit einem Roboter, der kein Partner auf Augenhöhe, sondern ein erweitertes Werkzeug ist. Der Roboter kann das Handeln eines Menschen immer nur *simulieren*. Aber welches Handeln sollte er simulieren, wie oft und zu welchen Gelegenheiten?

Also, ein eigener Vorschlag:

KI sollte das übernehmen, was der Mensch NICHT kann oder weniger EFFEKTIV tun kann.

Konkreter Fall: Altenpflege

Ein konkreter Anwendungsbereich ist die **Altenpflege**.⁸ Wozu dient KI hier? Verschiedene Perspektiven sind zu berücksichtigen, die Perspektive des Betroffenen, der Pflegekräfte und auch des Roboters.⁹

Ein besonderer Fall ist KI und Demenz. Wobei können Roboter menschliche Pflegekräfte sinnvoll unterstützen? Eine Möglichkeit wäre die der „Guidance“, der Begleitung in den Speisesaal. Roboterperspektive: Wenn er dies tun sollte, müsste der demente Mensch ihm folgen und der Roboter verstehen, dass er ihm folgt. Tut der Mensch das nicht, was macht der Roboter dann? Viele demente Patient*innen werden in späten Krankheitsphasen aggressiv. Könnte hier der Roboter in bestimmten Situationen, bei Ermüdungserscheinungen des Personals, übernehmen (das *cannot but should*-Prinzip)? Er müsste in dem Fall fähig sein, Emotionen zu registrieren und zu speichern. Er muss mit völlig stummen Patienten umgehen können. Patient*innenperspektive: Wie fühlt sich ein*e Patient*in, dem/der ein Roboter statt einer Pflegekraft beiseite gestellt wird? Wie jemand zweiter Klasse? Oder ist er/sie sogar froh, erleichtert, weil er das Gefühl hat, so als Pflegebedürftige*r weniger das Gesicht zu verlieren? Sollte besser ein Roboter Intimwaschungen übernehmen? Oder sind vielmehr die gesellschaftlichen Rahmenbedingungen und Wertesysteme das Problem, dass sich Menschen schämen, wenn sie gewaschen werden? Braucht man Patientenverfügungen, in denen Patient*innen die spätere Nutzung von Robotern ablehnen oder erlauben und ihre Bedingungen angeben (zum Beispiel: „Wer darf wie meine vom Roboter gesammelten privaten Daten verwenden“)?

Ein Beispiel aus der Altenpflege, das aktuell schon eingesetzt wird, ist die süße Roboterrobbe PARO (AIST). Die Patient*innen und Bewohner*innen von Pflegeheimen nehmen sie gut an: sie ist kuschelig, kann einen Menschen „hören“, „sehen“ und reagiert auf Berührung – sie beruhigt das Gemüt. Eine Katzen-Version dieses Roboters, sagt der Informationsethiker [Oliver Bendel](#), funktioniert allerdings nicht. Das muss wohl daran liegen, dass wir alle lebendige Katzen kennen. Warum finden sich keine echten

Katzen in Pflegeheimen? Natürlich, weil sie zu kostenintensiv wären und rechtliche wie hygienische Fragen auftauchen.

Mögliches Prinzip (daher *unabhängig* gedacht von aktuellen Rahmenbedingungen):

Seal Robot, Paro, for Therapy 9th Generation



PARO (AIST)

⁸ [Raffaele Rodogno: The Ethics of Guidance Robots](#)

⁹ Die Perspektive eines Gegenstands (oder Tieres) einzunehmen, nachzuvollziehen, schlägt auch Jakob von Uexküll in seinem Ansatz der „[Umweltperspektive](#)“ vor. Diese Methode erscheint mir aber auch sinnvoll, um in politischen Entscheidungsprozessen etwa den Entscheidern die Weltansicht (Bedingungen, Gefühle, Wünsche, Schmerzen) derer zu vermitteln, über die sie entscheiden, die sie repräsentieren sollen.

KI sollte einen lebendigen Austausch mit Tier und Mensch nicht ersetzen. Zusätzlich und im Fall, dass ein Tier jedoch im Kontakt gefährdet wäre, kann ein Roboter eingesetzt werden.

Fühlende Roboter

Warum sollten Entwickler*innen überhaupt ihre Zeit investieren, Robotern das Fühlen „beizubringen“? Das ist eine (siehe oben) viel zu grundsätzliche Frage unter den aktuellen wirtschaftlichen Bedingungen, in denen jeder technische Fortschritt begrüßt wird, insofern er Absatz zu generieren verspricht. Davon abgesehen sind jedoch tatsächlich Situationen denkbar, in denen ein Mensch eine Situation nicht übernehmen sollte, die es aber erfordert, mit menschlichen Emotionen umgehen zu können – etwa lebensgefährliche Polizei- und Militäreinsätze. Hier könnte ein emotionsfähiger Roboter aushelfen und Menschenleben retten.

Welche Fähigkeiten, die mit Gefühlen zu tun haben, könnten Roboter entwickeln? [Phillip Brey](#) von der niederländischen Universität Twente hat sich darüber Gedanken gemacht:

Einmal geht es darum, *Emotionen zu erkennen*. Das erscheint nicht so problematisch, wenn ein Roboter Emotionen wiedererkennen und reflektieren kann. Wichtig ist allerdings, was mit den Daten geschieht, die er sammelt: Schließlich erstellt er im Kontakt mit uns ein Profil, wie wir uns gegenüber ihm verhalten haben, ob wir wütend, misstrauisch, nervös, überschwänglich, bekifft (??) und ähnliches waren. Ein Konzern, der den eigenen Gefühlszuständen gemäße Angebote versendet, ist eine fragwürdige Vorstellung.

Mögliches Prinzip also:

Ein Roboter sollte transparent machen, was mit den Daten geschieht und wie lange sie gespeichert sind.



Pepper (Aldebaran Robotics)

Hier stellt sich natürlich die Frage, unter welchen Bedingungen das überhaupt umsetzbar ist: Muss ein handlungsfähiger Roboter, der Emotionen erkennen und verarbeiten kann, nicht Kontakt mit dem Netz haben (=Gefahr des Datendiebstahls)? Er verliert seine Handlungsfähigkeit in dem Moment, in dem er seine gespeicherten Daten (d.h. sein Wissen über sein Gegenüber) verliert. Wann also ist der Zeitpunkt da, sie zu löschen? Wenn er mit einem Menschen nichts mehr zu tun hat? Ein anderer Fall, wenn er als Sicherheitsroboter überwiegend mit Kriminellen zu tun hat, die weniger Recht auf Datenschutz haben als unbescholtene Bürger*innen?

Hier zeigt sich, dass die Prinzipien bisher wenig kontextbezogen sind: Sie können nicht weiter ausformuliert werden, weil ihnen die konkrete Anwendungssituation fehlt. Es bleibt zu bedenken, wer sie wie auslegen sollte, sobald sich die Situation einstellt.

Zweitens: *Auf Emotionen reagieren*, sie kognitive Prozesse und „Handlungen“ des Roboters beeinflussen lassen, und drittens: *Emotionen ausdrücke*., Ein Beispiel ist das Modell „Pepper“, ein „social robot“, der offenbar Emotionen erkennen kann und Menschen unterhalten und Beziehungen herstellen soll. Für die Fähigkeit, Emotionen auszudrücken, d.h. zu *simulieren*, lassen sich Pros und Contras anführen: Entschlossene Wut oder Verständnis in einer Interaktion zu zeigen, könnte notwendig sein. Sobald der Roboter Urteile über eine soziale Situation fällt, braucht es dann eben auch Empathie, Scham, Mitgefühl etc. Dagegen spricht, dass diese Fähigkeit Menschen dazu „verführen“ könnte, Roboter als Mitlebewesen zu behandeln. Besonders stellt sich diese Frage bei Kindern: Sollte man Kleinkinder alleine mit einem Roboter wie Pepper lassen? Wenn ja, wie lange? Was geschieht mit ihnen, wenn sie sich emotional an ihn binden? Sollte Pepper sich weigern bestimmte Dinge zu tun?

Wenn ein Roboter außerdem andere Ziele als die der Nutzer*innen haben sollte (weil er etwa fremdmanipuliert ist), könnte er versuchen, die Ziele seiner Nutzer*innen zu beeinflussen.

Brey schlägt als Resultat folgende Prinzipien vor:

1. *Roboter sind keine Personen.*
2. *Sie sind kein Selbstzweck, sondern ein Mittel zum Zweck.*
3. *Roboteremotionen sind somit ebenso Mittel zum Zweck. Ein „Kümmern“ um diese Emotionen ist nicht notwendig.*
4. *Ihre Emotionen sollten Menschen nicht einladen, Robotern intrinsischen Wert zu verleihen.*

Die Suche nach Balance erscheint auch als guter Leitfaden, wenn es darum geht, mit Robotern zu kooperieren. [Anika Fiebich](#) vom Centre for the Study of Social Action in Mailand schlägt drei Dimensionen vor, die beim Kooperieren mit Robotern eine Rolle spielen: Die behaviorale, kognitive und affektive Ebene.

Beide Kooperationspartner müssen die gemeinsame Aufgabe mental repräsentieren, benötigen Empathie und müssen sensibel für das Gegenüber sein.

Laut aktuellem Stand haben wir kognitive Fähigkeiten und Verhalten bei Robotern schon implementiert, (natürlich *simulierte*) affektive Fähigkeiten (Gefühle usw.) noch nicht. Um in unserer sozialen Realität (die die komplexeste Realität ist, die wir kennen) klarzukommen, dafür braucht der Roboter als „artificial social agent“ so einiges: Er muss Menschen erkennen, an sozialen Interaktionen teilnehmen, selbst eine Geschichte besitzen: Er muss grüßen, antworten, erinnern, lehren, assistieren, Normen kennen, reagieren.

Ein Beispiel: Ein Roboter, der einen Menschen in einem Shop bedient, muss zum Beispiel sich und seinen Kunden als Akteur begreifen und seine Rolle und die des Kunden repräsentieren, basale Aktionen (Ware aushändigen) ausführen können und Erwartungen kennen (wie lange dauert ein durchschnittlicher Einkauf? Wie wird der Kunde sich verhalten?).

Soziales Leben lässt sich nur bedingt standardisieren. Ist es also gefährlich, standardisierte Roboter im sozialen Leben einzusetzen? Hier ist es nochmal wichtig, sich klarzumachen, dass alle diese Prozesse den Charakter von Simulationen haben, also nur dem ähnlich sein werden, was wir Menschen tun, wenn wir kooperieren, und das in verschiedenen Graden. Auch hier gilt: Es geht nicht darum, nach Bewusstsein zu fragen, sondern nach den simulativen Prozessen, die auch ohne das funktionieren.¹⁰

Dass der Mensch die grundlegende Funktionsweise des Roboters durchschaut, scheint wirklich elementar wichtig zu sein. Der Pariser Philosoph [Louis de Miranda](#) beschreibt das Phänomen eines „Technostresses“ in Japan: Fließbandarbeiter, die die Funktionsweise der Roboter nicht mehr durchschauten wie ihre Vorgänger, die diese eingeführt hatten, fühlten sich selbst wie Roboter, innerlich kolonisiert und von sich selbst entfremdet. Sollte also ein „anthrobotisches System“ gemeinsam mit seinen Nutzern co-designed sein?

Prinzip: *Ein Mensch sollte im Kontakt mit einem Roboter immer ein ungefähres Verständnis von dessen Funktionsweise haben.*

Anthropomorphisierung ist menschlich

Eine Binsenweisheit, aber besser einmal zu viel erwähnt: Wir sollten uns also davor hüten, menschliche Beziehungen durch die mit Maschinen/KI zu *ersetzen*. Der polnische Wissenschaftler [Maciej Musiał](#) sprach über ein gutes Leben in einer durch Maschinen/KI automatisierten Arbeitswelt. Wäre sie nicht *die* Chance, enge Beziehungen zu nahestehenden Menschen besser ausleben zu können – schließlich sehen die meisten von uns diese als fundamental für ein gutes Leben an, aber haben nicht genug Zeit dafür? Er wies auf das

¹⁰ Vgl. das OASIS framework von Johanna Seibt.

Paradox hin, dass viele gleichzeitig viel Zeit investieren, um enge Beziehungen zu vermeiden. Das könnte an den Strukturen liegen, in denen wir funktionieren und die uns die Energie dafür rauben.

Musiak fragt: Hätten wir Sexroboter und Pflegeroboter noch nötig, wenn wir alle die Zeit hätten, um uns um unsere Beziehungen zu kümmern? Wie würde eine Welt ohne Arbeitspensum – etwa eine Gesellschaft des bedingungslosen Grundeinkommens – die Idee menschlicher Beziehungen verändern? Würden instrumentalisierte Beziehungen abnehmen? Würden wir in einer sinnstiftenden Welt Beziehungen aber auch weniger romantisieren?

Stichwort „intrinsischer Wert“: Wir werden vor allem dann dazu verleitet, einen solchen im Roboter zu erkennen, wenn er menschenähnlich gestaltet ist. Wir haben als Menschen ohnehin eine Tendenz, die Dinge um uns herum zu anthropomorphisieren: Unsere Kuschtiere, unsere Zeichentrickfiguren ... gerade, wenn Roboter menschenähnlich designt sind, neigen wir dazu, sie auch so zu behandeln und diese Illusion aufrechtzuerhalten.¹¹ Das scheint eine Strategie zu sein, mit dem Unbekannten umzugehen.

Roboter prägen uns damit genauso wie wir sie. Sollten sie daher wirklich wie Menschen oder Tiere designt werden? Sollten sie Kinderstimmen haben? Sollten sie nicht lieber wie Gebrauchsgegenstände aussehen? Sollten sie sich ihrer Umgebung anpassen? Die meisten Roboter sind schwarz, weiß und glänzend designt; sie könnten genauso gut eine Holz-, Gras- oder Textiloptik haben. Die dänische Produktdesignerin [Majken Kirkegaard Rasmussen](#) findet, dass sie ihrer Funktion gemäß aussehen sollten, die sie haben (z.B. wie ein Abfalleimer, wenn sie Müll sammeln). Ihr Design sollte nicht bloß reine ‚Verpackung‘ sein, sondern ein hilfreiches Mittel der *Kommunikation*.

Prinzip: *Ein Roboter sollte in seinem Design seiner Umgebung und den Bedürfnissen seiner Nutzer*innen angepasst sein.*

Eine Idee, die uns Menschen helfen könnte, mit Robotern umgehen zu lernen, ist die von [Pak-Hang Wong](#) von der Universität Hamburg. Er schlägt neue *Ritualformen* im Umgang mit Robotern vor und beruft sich dabei auf den chinesischen Philosophen Xunzi. Der antike Konfuzianist argumentiert, dass Rituale in sich zu kultivieren helfen kann, um sich Werte anzueignen und so innere und äußere Harmonie zu erreichen. Das Ritual schneide das ab, was zu lang sei und verlängere das, was zu kurz gerate; es bediene sich aller Gegensätzlichkeiten und das jeweils zur richtigen Zeit. Rituale helfen also dabei, zu handeln. Rituale sind hochformalisiert, ebenso wie es das Programmieren eines Codes ist. Sollten wir also die Kraft und den Sinn von Ritualen wiederentdecken, wenn wir es mit Robotern zu tun haben, die Emotionen ausdrücken bzw. an uns erkennen? Sollten sie gar rituelle Objekte werden, die wichtig sein können für spirituelle Gefühle? Könnten auch Roboter irgendwann neue Rituale erfinden, die sie uns beibringen? Und würden die Rituale der westlichen KI kulturell bedingt anders sein als die der östlichen?

Das inspiriert mich zu folgendem Prinzip:

*Entwickle für jeden Roboter ein passendes Ritual, das es seinen menschlichen Nutzer*innen erleichtert, mit ihm zu interagieren und sich seine simulative Natur bewusst zu halten.*



Xunzi: Der „Philosoph der Rituale“

¹¹ [Campagna, Diego](#): Je mehr Roboter in einer Interaktion versagen (Bsp. Keine sinnvollen Antworten zu geben im Kontext dessen, was man fragt), umso mehr versuchen Menschen, ihr dennoch Sinn zu geben, ihre Erwartungen halten zu können

Wer und was werden wir sein?

Bisher haben wir Mensch und Roboter als Gegensätze („natürlich“ vs. „künstlich“) betrachtet. Das ist aber vermutlich irgendwann obsolet. Einen interessanten Input lieferte auch der Amsterdamer Künstler Hendrik-Jan Grievink. Er arbeitet für das Projekt [„nextnature“](#), wo es darum geht, uns durch Zukunftsszenarien klarzumachen, was alles auf uns zukommt oder zukommen kann, je nachdem, wie und ob wir es gestalten. Ihm zufolge können wir zukünftig nicht mehr zwischen *natürlich* und *künstlich* unterscheiden, es gehe eher darum, Biologie und Technologie auszubalancieren.



Fleisch aus dem Labor

Die Welt dessen, was geboren und was gemacht ist, vermischt sich. Immer schon sind die Tiere, die wir halten, und die Pflanzen, die wir essen, gezüchtet, also im Zwischenfeld zwischen geboren und gemacht, außer-Kontrolle und kontrolliert. Technologie wird unsere nächste Natur, sagt Grievink; das möchte er durch „spekulatives Design“ erforschen und aufzeigen: Wie werden künstliche Gebärmutter sein? Werden wir einmal tierische Attribute haben oder sogar Eier legen? Sechs Eltern haben? Oder wird es den Begriff der Eltern gar nicht mehr geben? Was stellen wir mit künstlich hergestelltem Fleisch an? Weniger futuristisch, näher an der Gegenwart, geht es im Projekt <http://hubot.org/> zu: Eine fiktive Jobagentur stellt hier durch neue Technologien enhanced Berufe vor.

Interessant auch seine Sicht auf den Grad, in dem Technologie unser Dasein bestimmt: Von einem rein fiktiven Gedankeninhalt hin zu einer Bedingung unserer Existenz. Er unterscheidet folgende pyramidenartige Stufen:



„Jobagentur“ HUBOT

envisioned (nur als Vision präsent: Menschen auf dem Mars)

operational (nur im Labor: geklonte Tiere)

applied (angewandt: selbstfahrende Autos)

accepted (im Alltag akzeptiert: Smartphones)

vital (erscheint unentbehrlich: Internet)

invisible (nicht mehr sichtbar: ???)

naturalized (gestaltet etwa Physiologie, Ernährung oder Zeitverständnis: Uhrensystem?)

Zu *invisible* fällt mir gerade kein Beispiel ein, vielleicht habt ihr ja eins. Für die Zukunft können wir also unsere althergebrachten Bilder, was ein Mensch ist, was Natur ist, und im Übrigen auch was Arbeit und Zeit ist, was eine Frau und ein Mann ist und einiges andere mehr über den Haufen werfen. Statt davor zu kapitulieren wie das Kaninchen vor der Schlange, macht es mehr Sinn, im Kleinen das Große zu gestalten und dabei einen Sinn für Balance zu entwickeln, für *Einheit in Vielfalt*.

Die Suche nach dynamischer Balance

In seinem Konzept des CREAL, das Ähnlichkeiten mit der Werteordnung meines Romanszenarios aufweist, schlägt [Louis de Miranda](#) ein neues Universal vor, ein Konzept immerwährender Kreativität, das sich beständig selbst zerstört und wieder neu entsteht. Er folgt dabei den postmodernen Denkern Deleuze und Guattari. Sie postulieren das All-Eine, die Totalität von Allem, die das gesamte physikalische Universum und seine Möglichkeitsbedingungen umfasst, ähnlich buddhistischer Denkansätze: Alles ist mit allem verbunden.

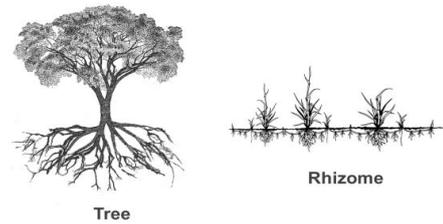
Geschieht etwas, ist diese „Realität“ notwendigerweise unvollkommen. Sie werfen Ideologien wie dem Kapitalismus vor, zu leugnen, dass sich immer neue Möglichkeiten realisieren können. Man sollte akzeptieren, dass Realität immer instabil und unvollkommen ist und sich frei durch immer neue ‚Realisierungen‘ bewegen. Die Tatsache, dass alles sich von allem unterscheidet, bringt immer neue Veränderung hervor.

Miranda zufolge brauchen Systeme und Ordnungen – um nachhaltig zu funktionieren – ein unsichtbares Wertennetz, um das sie kreisen. Nur so können sie Sinn herstellen. Diese essentiellen „signifier“ haben häufig ein Gegenteil, von dem sie sich abgrenzen, durch das sie sich erst konstruieren: Gott vs. Atheismus; Kapital vs. Kommunismus; Wettbewerb vs. Solidarität; Wissenschaft vs. Glaube. Sie simulieren dadurch, dass Realität zweipolig ist, aber das stimmt nicht.¹²

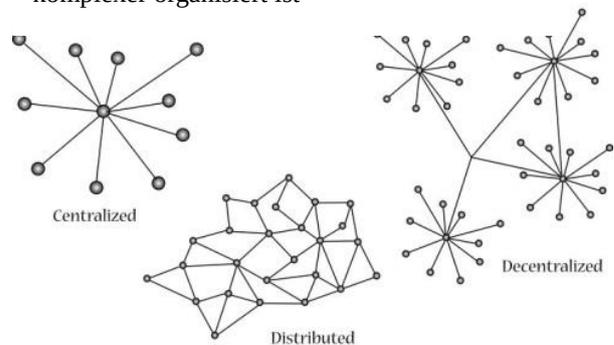
Auch Mirandas CREAL-Begriff (Vielheit, alle Dinge unterscheiden sich voneinander) hat so ein Gegenteil: das „Eine“, Allumfassende. Dieses Konzept, das einen globalen sozialen Vertrag hervorbringen soll, soll aber deshalb *nicht* totalitaristisch oder reduktionistisch sein, weil es ein „desire without object“ ist, ein beständiger, aber leerer Wunsch nach dem radikal Neuen.

Miranda ist überzeugt: Wir brauchen einen neuen Mythos, „a global social contract“, der Technik und Wissenschaft integrieren kann und nicht den Menschen allein in den Mittelpunkt stellt. Und der ethische Axiome bereitstellt, die nicht nur durch aufklärerische, rationalistische Ideale begründet sind.

Der letzte Redner auf der Konferenz – der Kybernetiker [Robert Trappl](#), Autor von *A Construction Manual for Robots' Ethical Systems* – hält schließlich einen Vortrag über *Robot Deus*. Es geht um KI als neue Religion. Eine Vereinigung in den USA („World of the Future Church“) hat KI zur neuen Gottheit erklärt. Er sagt, ihn wundert es, dass die Religion – man könnte es auch *Spiritualität* nennen – in Gestalt eines virtuellen Gottes zurückkommt, er dachte, dass unsere Tradition der Aufklärung dies hätte verhindern können.¹³ Offenbar kann sie es nicht. Rationalität und Spiritualität haben noch ein Wörtchen miteinander zu reden. Umso wichtiger zu fragen, was wir damit anfangen, nachdem wir das nun wissen.



Postmoderne Erkenntnisse: Denken in Hierarchien (und Dualismen) reduziert unsere Ideen von Wirklichkeit („ein Baumstamm, von dem alle Wurzeln abgehen“) – das Rhizom in der Pflanzenwelt zeigt uns, dass Leben komplexer organisiert ist



¹² Mein Kommilitone Alex, der sich damit in seiner Masterarbeit befasst hat, erklärt in diesem [Video](#), wo wir heute in der „Post-Postmoderne“ stehen mit unserem Versuch, die Welt und das Leben besser zu begreifen.

¹³ World of The Future Church.