

KATHEDRALEN DER KÜNSTLICHEN INTELLIGENZ

Ökonomische und ethische Herausforderungen für wahren Fortschritt

Im vorliegenden Artikel geht es um die Frage, unter welchen Bedingungen die Künstliche Intelligenz (KI) ihr eigentliches Fortschrittpotenzial einlösen kann. Technik ist per se als Neuheit noch kein Gut. Die Verbreitung von Feuer war für die Menschheit ein großer Fortschritt. Dies aber nur, weil wir einen klugen standardisierten Umgang damit gefunden haben. KI ist mit Feuer durchaus vergleichbar, da sie uns erlaubt, in nie dagewesener Form in das Wissen der Welt hineinzuzoomen und Zusammenhänge zu entdecken, von denen wir bisher nichts wussten. Dadurch könnte ein Wissensboom in der Forschung entstehen. Hinzu kommt in Arbeitsprozessen eine Vereinfachung vieler Dokumentationsaufgaben; ja sogar eine Delegation von Aufträgen an KI-Systeme (z.B. programmieren), ohne dabei eine Tastatur nutzen zu müssen. Schon heute wird KI nicht nur für solch neue Aufgaben ausgerollt, sondern auch eingebettet in andere Technologien genutzt, um Funktionalität zu verbessern (wie etwa in Smartphone-Kameras, wo solche sog. «narrow AI» zum Einsatz kommt). In der Robotik kann KI eingesetzt werden, um körperlich schwere Aufgaben in immer mehr Bereichen zu unterstützen. Zusammenfassend ist die heutige KI sicherlich ein Zenit der Technikgeschichte. Aus diesem Grund habe ich im Titel dieses Artikels KI mit Kathedralen in Verbindung gesetzt; einem Zenit der Architektur.

Kathedralen sind für Europa das, was Betriebswirte als «Unique Selling Proposition» bezeichnen würden; ein Wettbewerbsvorteil gegenüber anderen Kontinenten. Sie könnten uns Vorbild sein, wenn wir uns fragen, wie Europas

SARAH SPIEKERMANN, geb. 1973, leitet seit 2009 das Institut für Wirtschaftsinformatik & Gesellschaft an der Wirtschaftsuniversität Wien (WU). Sie forscht zur digitalen Ethik und Wertebasierten IT-Innovation und leitete die Entwicklung des ersten ISO Standards für ethisches IT-Design.

Zukunft aussehen könnte, wenn wir KI mit ähnlicher Ambition bauen. Wäre es nicht wünschenswert, dass wir mit einer ähnlich starken Lebensenergie, Visionsfreude und Durchdachtheit an KI herangehen, wie unsere Vorväter an die gotischen Kathedralen? Einen kraftvollen digitalen Raum schaffen, der ebenso «gut» ist und vertrauenswürdig, inspirierend und nährend im Resultat, wie diese Kirchenräume? Im Folgenden werde ich einerseits argumentieren, dass mit heutigen Innovationsprozessen und ökonomisch- kulturellen Rahmenbedingungen keine Aussicht darauf besteht, dass KI diesem Anspruch gerecht wird. Andererseits werde ich aufzeigen, dass uns durchaus die Grundlagen dafür zur Verfügung stünden, das zu ändern.

1. Zu den Voraussetzungen einer Kathedrale

Das Bild der gotischen Kathedrale und ihre Entstehungsgeschichte wird in diesem Artikel als ein Ideal skizziert, was sicherlich im mittelalterlichen Alltag durchwachsenere Züge hatte als hier teilweise dargestellt. Aber so herausfordernd und heterogen die Kathedralengeschichte ausgesehen haben mag, sie zu errichten bedurfte mindestens 3 fundamentaler Voraussetzungen: Die erste, welche heute am leichtesten nachvollziehbar ist, war natürlich die Baukunst und das Wissen um die Prozesse, eine Kathedrale technisch planen und dann konkret errichten zu können. Wie George Duby in seinem Buch «Die Zeit der Kathedralen» jedoch beschreibt, war dies nicht allein ausschlaggebend. Ebenso wichtig war der Auftrag, das «warum» dieser Architektur. Dieser Auftrag bestand in der mittelalterlichen Gesellschaft für jeden Einzelnen im Streben nach Heiligkeit und spiritueller Vollkommenheit; ein Streben, das sich nicht nur durchaus ambivalent in Kreuzzügen aktualisierte, sondern auch in der Errichtung eben dieser gotischen Kathedralen.¹ Letztlich drückt sich in den majestätischen Bündelpfeilern und lichtdurchfluteten Kreuzgängen ein Streben nach dem Himmel aus, der Wunsch, das Heilige auf Erden zu realisieren, was bei den Bauentscheidungen ausschlaggebender war als die ungeheuren Zusatzkosten, die dadurch für den Kirchenbau entstanden. Gotische Kathedralen bedurften nämlich in der Regel, um ihre Höhe zu ermöglichen, äußere Strebebögen, was deutlich teurer war, als den romanischen Vorgänger weiterzuverfolgen. Geld wurde also als zweitrangig betrachtet. Begüterte Familien finanzierten Seitenaltäre, wodurch sie sich postmortale Erlösung erhofften. Zusammenfassend sind gotische Kathedralen Ausdruck eines in höchstem Maße ambitionierten Wertstrebens.

Baukunst und Wertstreben waren jedoch nicht ausreichend, um diese Gebäude zu errichten. Vielmehr bedurfte es darüber hinaus eines hohen Maßes an

Verlässlichkeit und Vertrauenswürdigkeit aller am Bau Beteiligten; einer Haltung, die man auch als «Ethos» bezeichnet und die zum damaligen Zeitpunkt maßgeblich (wenn auch nicht ausschließlich) durch die 10 Gebote gegeben war. Im Mittelalter waren die 10 Gebote zentraler Bestandteil der katholischen Lehre und des Alltagslebens. Selbst wenn Sünden und Verletzungen weit verbreitet waren, strebten viele Menschen nach ihrer Einhaltung, beeinflusst durch Predigten, Katechismen und kirchliche Disziplin, die vor allem an den Baustellen der Kathedralen präsent gewesen sein sollten. Man stelle sich nur vor, die am Bau beteiligten Handwerker hätten in Wirklichkeit irgendwelchen privaten Naturgötzen gedient. Sie hätten das Gold und die Steine gestohlen, die zuhauf im Umfeld dieser Bauten verfügbar waren. Sie hätten sich gegenseitig die Frauen ausgespannt, sich beneidet und betrogen usw. Die Entstehung der Kathedralen, mit ganzen Generationen an Handwerkerfamilien drumherum hätte niemals gefruchtet. Im Gegensatz muss ein gewisser Respekt für die in den 10 Geboten konkret aufgelisteten Regeln bestanden haben, damit diese Kathedralen über Jahrhunderte hinweg fertiggestellt werden konnten.

Zusammenfassend kann man also argumentieren, dass es mindestens die Formel

< Werte + Baukunst + Ethos > war, die uns diese Kathedralen ermöglicht hat. Wenn wir nun den Vergleich anziehen zwischen gotischen Kathedralen und KI anstreben, wie steht es um diese drei Fortschrittsvehikel beim Bau der KI?

2. Zu den Werten bei der Entstehung von Künstlicher Intelligenz

Spätestens seit 2016 haben Gesetzgeber weltweit, internationale Institutionen und Technologiefirmen Wertprinzipien für KI formuliert. Insgesamt gab es bereits 2019 mindestens 84 Wertprinzipienlisten, die beim Bau von KI eingehalten werden sollten.² Dabei lassen sich technisch unabdingbare und gesellschaftspolitisch erwartbare Werte unterscheiden.

Zu den technisch unabdingbaren Werten gehören die Robustheit, Sicherheit und Transparenz der KI. KI soll auditierbar und kontrollierbar sein. Ich bezeichne Werte dieser Natur als «Hygienefaktoren», weil ohne sie KI-Produkte wahrscheinlich nicht einmal den Kriterien der Produkthaftung entsprechen würden. Das wäre so, als ob man eine Kathedrale eröffnen dürfte, die einem aus unerklärlichen Gründen auf den Kopf fällt und in der es auch keine Fluchtwege gibt. Gesellschaftspolitisch erwartbare Werte von KI sind oft bereits verfassungsrechtlich verankert, wie der Datenschutz, der vor allem in

Europa eingefordert wird. Gesetzlich wird gewünscht, dass hohe Werte wie die Menschenwürde und die Gerechtigkeit zumindest nicht verletzt werden.³ Manchmal wird der Wert der Wohlfahrt (zu Englisch «wellbeing») erwähnt.⁴ Knapp die Hälfte der Wertelisten wünschen sich eine Art «Gutartigkeit» von KI (zu Englisch «beneficence»), auch wenn diese nicht genauer spezifiziert ist.⁵ Die politischen Kräfte unserer Zeit konnten sich also darauf verständigen, dass zumindest am Papier einige hohe Werte von KI nicht verletzt werden sollten; ganz im Sinne des von Google vor Jahren formulierten Vorsatzes: «Don't be evil». Kommt dies der Ambition gleich, die unsere Vorväter beim Bau der gotischen Kathedralen beseelt hat? Wohl kaum!

Die heutigen Werterwartungen an KI würden im Vergleich einer Bauherrnrede und zwei Tafelanschlügen auf den mittelalterlichen Baustellen entsprechen, wo zu lesen gewesen wäre: Bitte nicht die Heiligkeit verletzen! (für diejenigen die überhaupt lesen konnten). Der Vorsatz «don't be evil» ist wenig ambitioniert. Er entspricht in keinsten Weise einer Realisierung von «gut sein» «Gut ist der Wert, der in der Sphäre des Wollens an der [tatsächlichen!] Realisierung eines höheren Wertes haftet,»⁶ schrieb Scheler 1921. Gut sein in Bezug auf KI würde bedeuten, dass man sich bewusst entscheidet und erfolgreich darin ist, KI zum Beispiel systematisch in den Dienst einer gestärkten Personenwürde zu stellen. Alternativ könnte man KI so bauen, dass Menschen klüger und weiser werden als sie es in der Geschichte je waren. Solch ein bewusstes Streben und Umsetzen hoher Kernwerte im Bau von KI, würde dem Elan und der Schaffenskraft der Menschen entsprechen, die für uns die gotischen Kathedralen hinterlassen haben. Davon sind wir allerdings weit entfernt, da der heute Sinn von KI nicht die Würde und das Wissen von Menschen ist, sondern die Optimierung des von Profit- und Effizienz bestimmten Shareholder Value. Das Geld ist der Götze, für den die heutige KI gebaut wird und nichts mehr.

3. Zur Baukunst bei der Entstehung von Künstlicher Intelligenz

Was existiert, ist ein Wissen um die notwendige Baukunst, der es bedürfte, um KI-Systeme als Kathedralen zu planen. Baukunst bedeutet, dass wir die Architektur von KI ganz bewusst auf hohe Kernwerte hin ausrichten könnten, wenn wir es wollten. Das dafür notwendige Knowhow wurde in den letzten 30 Jahren systematisch erschlossen. Bereits in den späten 90er Jahren entwickelte Batya Friedman das sog. «Value Sensitive Design», was in vielen Feldstudien erprobt wurde.⁷ In den letzten 10 Jahren wurde dieses dann zum «Value-Based Engineering» (VBE) ausgebaut⁸ und mündete in den Innovationsstandard ISO/IEC/IEEE 24748-7000. Dieser beschreibt, welche Prozess-

schritte erforderlich sind, um Werte-basierte IT-Systeme zu bauen.⁹ Er baut auf einem interdisziplinären Wissen auf an der Schnittstelle zwischen Informatik, Wirtschaft und Ethik und hat drei Konzeptionsphasen:

Die erste Phase, in der das System durch eine Vorplanung läuft, ist die wo nach der Machbarkeit der KI gefragt wird und angedacht wird, wie sie gebaut werden könnte (sog. «Concept of Operations»). In dieser Phase werden – ganz im Sinne der Diskursethik – Stakeholder gewonnen, die direkt oder indirekt von der KI betroffen sind. Es wird in einer Vorabanalyse geklärt, für wie machbar und ethisch vertretbar diese die KI halten. Technisch wird untersucht, ob und wie die KI mit ihren Datenquellen und Schnittstellen qualitativ zufriedenstellend und transparent umgesetzt werden kann. Darauf aufbauend wird eine Investitionsentscheidung getroffen oder das Projekt beendet.

In der zweiten Phase geht es um die Entwicklung einer Wertstrategie für das KI-System. Man erkundet aus drei moralphilosophisch komplementären Betrachtungsperspektiven, welche Werte die KI positiv materialisieren könnte, und welche negativen Wertrisiken bestehen. Die utilitaristische Perspektive erschließt Vor- und Nachteile des Systems. Die tugendethische Betrachtung eruiert, welche charakterlichen Langfristfolgen das System haben könnte. Und die deontologische Betrachtung erschließt hohe universale Wertpotenziale, die von der KI berührt werden. Auf Basis dieser Untersuchungen zeigen sich eine Vielzahl von Werten, die von der KI im Positiven wie im Negativen freigesetzt werden könnte. Im Durchschnitt werden zehnmal mehr Systemrisiken erkannt als bei herkömmlichen Innovationsverfahren, während sich gleichzeitig die Kreativität rund um eine positive Wertvision verdreifacht.¹⁰ Diese Phase des VBE erlaubt die Werte zu identifizieren, um die es bei einer Technik gehen soll. So wie Abt Suger beim Bau der ersten gotischen Kathedrale von St. Denis um die neuplatonischen Werte von Licht, Harmonie, Maß und Symmetrie wusste.

Wie kann man so etwas denken für ein KI-System? Als VBE-Fallbeispiel mag hier ein von KI gesteuertes und mit KI kommunizierendes Auto genannt werden. Es bestünde die Möglichkeit, Autos jenseits ihrer Mobilitätsfunktion eine wirkliche Wertvision zu geben, der alles andere untergeordnet ist: eine Kombination aus Kontrolle, Besitz, Vertrauen und Mensch-Maschine Freundschaft, wie sie etwa in der Science-Fiction-Serie «Knight Rider» in den 90er Jahren skizziert wurde. Hier hat der Fahrer und nicht der Hersteller volle Kontrolle, «Ownership» und Transparenz über sein Fahrzeug, was von der Infrastruktur unabhängig und dezentral ihm (oder ihr) allein gehorcht. Natürlich würde «KITT», wie so ein Auto in der Science-Fiction Serie «Knight Rider» heißt, seine Fahrerin niemals zentral überwachen lassen oder bevormunden (Wert der Privacy), sie aber dennoch schützen (Wert der Sicherheit), ohne sie

allerdings zu entmündigen (Wert der Freiheit).

Die Umsetzung solch einer Wertstrategie erfolgt dann konkret in der dritten Phase des VBE, wo die materiellen Voraussetzungen einer solchen immateriellen Schönheit gelegt werden. In dieser Phase werden mitunter hunderte (bei Fahrzeugen tausende) von technischen «Systemanforderungen» spezifiziert, die in Form eines Pflichtenheftes von Ingenieuren und Informatikern, Zulieferern und Policy-Verantwortlichen umgesetzt werden müssen.

4. Wie steht es mit der Baukunst bei der Erstellung von Softwareprodukten?

Leider besteht eine große Kluft zwischen dieser verfügbaren Baukunst und etablierter IT-Innovationspraxis in Unternehmen. Generell ist es seit den frühen 2000er Jahren zu einer Veränderung in der Innovations- und Produktentwicklungskultur gekommen. Eric Raymond's Buch «The Cathedral and the Bazaar» war 1999 dafür ein Meilenstein.¹¹ Er kritisierte, dass Software zu zentral geplant und kontrolliert würde, so wie Kathedralen. Gleichzeitig forderte er mehr Freiheit für Programmierer, die wie Händler auf einem Bazaar ihre Softwarekomponenten einbringen sollten. Während dieser Gedanke zur damaligen Zeit eine nachvollziehbare Kritik war, sehen wir heute das umgekehrte Problem, dass Softwareentwicklung (inklusive KI) zu sehr outgesourct bzw. verteilt ist und es zu wenig Kontrolle über die in ein Gesamtsystem eingebundenen Komponenten gibt. Der «Controller», also der eigentlich Verantwortliche für ein Softwareangebot, tut sich u. U. schwer damit, das in seinem Service zusammenlaufende Komponentenbündel (sog. «System of Systems») wirklich zu kontrollieren, geschweige denn allein dafür verantwortlich sein zu können. Beim Aufsetzen von KI-Modellen kommt es beispielsweise vor, dass Datenquellen eingebunden werden, deren Aktualität, Qualität, Urheberrechte oder kontextuelle Integrität nicht ausreichend gegeben sind. Oft haben Anbieter keinen Zugang zu ihren Quell-Services, sondern dürfen nur über sog. «Application Programming Interfaces» (APIs) benutzen, ohne dabei zu wissen, wie die eingebundenen KI-Blackboxen funktionieren, die sie ihren eigenen Kunden weiterreichen. OpenAI gehostete GPT-Services sind hier ein Beispiel. Es kommt zu einer Kette der Verantwortungslosigkeit; so als wolle man eine Kathedrale bauen, wo die Qualität des Trägermaterials ungewiss ist. Hinzu kommt der Druck in den Organisationen, die Software so billig wie möglich «ans Laufen» zu bringen, verbunden mit einem agilen, sukzessiven Ausrollen von vorläufigen Lösungen. So als wenn man eine Kathedrale aus Fertigbauteilen ohne Qualitätsprüfung hätte zusammenbauen wollen. Die Werte des Geldes, der Effizienz und der Geschwindigkeit stehen im Vordergrund.

Weiterer Teil dieser Innovationsdynamik ist, dass KI vor allem eingeführt wird, weil Kosteneinsparungen erwartet werden. Analysten prognostizierten in 2023, dass zwei Drittel aller Berufe zumindest zum Teil von KI Automatisierung betroffen sein werden und dass ein Viertel durch KI sogar ersetzt werden könnten.¹² Allerdings ist nicht sicher, ob dieser wiederum im Mittelpunkt stehende geldwerte Vorteil sich so realisiert wie erwartet, da heute etwa 50% der IT-Projekte in Unternehmen scheitern¹³ und Software (inklusive KI Systeme) generell hoch fehleranfällig ist; Studien schätzen 6.000 Fehler auf 1 Millionen Zeilen Code.¹⁴ Zur Einordnung: Ein modernes Auto integriert rund 13 Millionen Zeilen Code, was dann etwa 78.000 Fehlern entsprechen würde.¹⁵ Das bedeutet, dass selbst geldwerter Vorteil durch KI nicht gesichert ist. Zusammenfassend lässt sich sagen, dass die heutige Entwicklung von KI in eine unternehmerische Innovationsdynamik und Wertprioritäten eingebettet ist, die nicht erwarten lassen, dass die Technologie der Gesellschaft den Fortschritt bringen wird, der möglich wäre.

5. Der letzte große Formelbestandteil: Ein Ethos bei der Entwicklung von KI

Nehme man einmal an, dass die beschriebenen Schwachstellen heutiger Innovations- und Produktentwicklung erkannt und gemeistert würden, dass es etwa zu einer breiteren Akzeptanz von VBE käme; dann stellt sich jedoch immer noch der dritte Formelbestandteil zur Disposition: und das ist, dass man beim Bau von KI (und IT-Systeme generell) ein gewisses Ethos mitbringen muss; und zwar sowohl als Verantwortlicher im Unternehmen, als Programmierer, als auch als Nutzer. Ich habe oben plakativ skizziert, dass man leicht unterschätzt, von welcher großer Bedeutung das persönliche Ethos aller an einem System Beteiligten ist, damit ein gemeinsames Projekt – wie die Kathedralen – langfristig erfolgreich umgesetzt werden kann. Die 10 Gebote waren als Bestandteil des kulturellen Gedächtnisses im Mittelalter ein unsichtbares, verbindendes Gerüst für ein faires Miteinander. Durch sie konnte man auf traditionsbewährte, minimalistische Verhaltensregeln aufbauen, die schon in alttestamentlicher Zeit etabliert worden waren und die es Christen, Juden und Muslime erlaubt hatten friedlich zu koexistieren.¹⁶

Von so einem koordinierten, kultivierten und grenzbewussten Miteinander sind wir in der digitalen Welt weit entfernt. Es ist bekannt, dass Plattformen, die heute von Monopol- oder Oligopolstellungen profitieren, ihre Nutzer absichtlich in Abhängigkeiten und Suchtverhalten treiben.¹⁷ Menschen, die dem Netz verfallen sind, verbringen mittlerweile fast ihre gesamte private Zeit

dort und unterbrechen über 80-mal am Tag ihre Arbeitsroutinen. Veränderte Gehirnstrukturen und reduzierte Aufmerksamkeitsspannen sind die Folge.¹⁸ Gleichzeitig verkünden Tech-Milliardäre die Geburt einer neuen Superintelligenz und verkaufen ihre problematischen Softwarelösungen auf dem Rücken falscher Versprechen, indem sie eine vermenschlichte Sprache für KI nutzen. So zeigt GPT etwa an, es würde «denken», während es in Wirklichkeit «rechnet». KI-Sprachmodelle werden auf «Ehrlichkeit» hin optimiert, obwohl es sich hierbei nur um die Wahrscheinlichkeit einer Richtigkeit von Berechnungen handeln kann; etc. Im Prinzip fußt Vieles von dem, was als menschenähnliche «Intelligenz» versprochen wird auf einem semantischen Lügegebäude.

Das Wirken an einer wünschenswerten Zukunft, kann nur beginnen, wenn wir für diese neue digitalisierte Welt ebenso ein Ethos entwickeln, das jeder Einzelne verinnerlichen kann, um eine tägliche Orientierung zu bekommen, was in der digitalen KI-durchtränkten Welt richtig und falsch ist. Genau diesem Unterfangen haben sich in 2025 16 Wissenschaftler und Autoren aus 13 Fachrichtungen zugewandt, die versucht haben, einen Kanon für die digitale Welt zu entwickeln.¹⁹ Über einen Zeitraum von acht Monaten entwickelte die Gruppe in einem siebenstufigen Prozess 10 Regeln, die in ihrer Reihenfolge und Bedeutungsdimension den 10 Geboten angelehnt sind bzw. von den 10 Geboten inspiriert wurden.

Tabelle 1 enthält die 10 Gebote der Bibel (Spalte 1) ergänzt durch einen erweiterten Interpretationsrahmen des jeweiligen Gebots (Spalte 2). Die Erarbeitung dieses Interpretationsrahmens über den unmittelbaren Wortlaut des Dekalogs hinaus wurde von drei Gruppenmitglieder erarbeitet, unterstützt von einem Alttestamentler.²⁰ Darauf aufbauend wurden die beschriebenen Herausforderungen der Digitalen Welt in drei Schritten auf die 10 Gebote gemappt.²¹ In weiteren Prozessschritten kam es zu einer Verdichtung der digitalen Problemfelder entlang der Bedeutungsstruktur der 10 Gebote sowie vor allem einer Rückbindung der Sprache an die Kernaussagen der Gebote. Zum Beispiel «Zerstört nicht Natur und Leben für den technischen Fortschritt» in Regeln 5 bindet zurück and das «Du sollst nicht töten» des fünften Bibelgebots. Durch diesen Prozess entstanden die 10 Regeln, die in weiterer Folge wiederum ihre eigenen Interpretationsräume aufmachen; angebunden dann an digitale Lebenswelten. Tabelle 1 fasst dies alles zusammen.

Wie die 10 Gebote sind die 10 Regeln interpretationsoffen und können von jedem kontextualisiert werden, der an der Digitalisierung beteiligt ist. Im verbleibenden Abschnitt möchte ich daher einen solchen Interpretationsversuch unternehmen und überlegen, wie sich dieses Ethos für die digitale Welt auf die Bauherren einer wirklich guten KI auswirken könnte.

Diese müssten sich zuallererst fragen, welcher Wert durch KI entstehen soll.

10 Regeln für die Digitale Welt	10 Gebote	Bedeutung Bibel: Bei diesem Gebot geht es im weiteren Sinn um Aspekte wie...
1. Erhebt digitale Technik nicht zum Selbstzweck	Du sollst keine anderen Götter neben mir haben	Idolatrie – Das Anbeten der falschen Götter/ Götzenkult; Hierarchie – Die Anerkennung einer Ordnung, bei der Gott (das Gute/die Schöpfung) als höchstes Prinzip anerkannt wird; Irdische Wirklichkeiten : Die Verabsolutierung innerweltlicher irdischer Wirklichkeiten ist abzulehnen
2. Schreibt Maschinen keine Menschlichkeit zu.	Du sollst den Namen Gottes nicht veruntreuen	Wahrhafte Benennung der Dinge: Die ‚Namen‘ die wir geben dürfen die Wahrheit nicht veruntreuen; Wahrhafte Beschreibung des Seins : Falschdarstellung des von Gott gegebenen für eigenen Zwecke; Bilderverbot: Falsche Bilder sollen vermieden werden
3. Schafft Raum für Muße und analoge Begegnung.	Du sollst den Tag des Herrn heiligen (Sabbatgebot)	Zeit : die von Gott geschenkte Zeit am Tag (der aktive Teil des menschlichen Lebens) ist ein Geschenk und sollte sinnvoll genutzt/dem Guten gewidmet sein; Muße: Zeit, um für das Heilige empfänglich zu sein; Ritualisierte Freizeit ; Vorschrift der Unterbrechung des Alltäglichen
4. Garantiert den Erhalt sozialer und demokratischer Kompetenzen.	Du sollst Vater und Mutter ehren	Bedeutung familiärer Bindungen ; Anerkennung der natürlichen verbindlichen Lebenszusammenhänge in die man hineingeboren wurde (Anerkennung der eigenen Biographie); Anerkennung von Hierarchien (es gibt rechtmäßige irdische Autoritäten); Transgenerationale Solidarität
5. Zerstört nicht die Natur für den technischen Fortschritt.	Du sollst nicht töten	Recht auf Leben ; Kein Recht über den Tod anderer zu entscheiden; Kein Zweck rechtfertigt die Tötung
6. Behandelt Menschen nicht als bloße Datenobjekte.	Du sollst nicht ehebrechen	Güter der personalen Beziehung wahren; Treue; Vertrauen ; Schutz der Intimsphäre
7. Last Euch nicht Eurer menschlichen Potenziale berauben.	Du sollst nicht stehlen	Luther: Wer stiehlt hat nicht begriffen, dass das was von Gott gegeben ist, gerade richtig ist. Das was man hat, reicht; Recht auf und Schutz von Eigentum ; Eingrenzung von Eigentum bedarf der Begründung
8. Verleugnet nicht die Grenzen der Technik.	Du sollst nicht falsch gegen Deinen Nächsten aussagen	Bedeutung der Wahrheit , Vollständigkeit, Akkuratheit; Schutz des Nächsten, der nicht verunglimpft werden darf
9. Nutzt Maschinen nicht, um die Freiheit Anderer zu untergraben.	Du sollst nicht begehren Deines nächsten Frau (personales Begehrenkeitsverbot)	Enthaltsamkeit ; Man darf nicht zum Sklaven seiner Leidenschaften werden; Emotionale Abhängigkeiten vermeiden; Intervention in (andere) menschliche Beziehungen
10. Verhindert Machtkonzentration und garantiert Teilhabe.	Du sollst nicht begehren Deines nächsten Hab und Gut (gegenständliches Begehrenkeitsverbot)	Wissensgier statt Weisheit; Habgier; Übergreifigkeit (wie Aneignung von Gemeingütern)

Tabelle 1

Geht es hier um einen vermuteten Fortschritt oder kann man konkrete Vorteile ausloten, wo die KI ihre Nutzer wirklich bereichert? Regel 1 – korrespondierend mit dem 1. Gebot man solle keine Götzen haben – setzt einem unspezifizierten Fortschrittswert entgegen, KI nüchtern zu hinterfragen. Vor allem sollten Investoren nicht der Marketingsprache erliegen, KI könne etwa «denken», «kreativ» und «ehrlich» sein, was einen Bruch des 2. Gebotes gleichkommt bzw. Regel 2 aufruft: Schreibt Maschinen keine Menschlichkeit zu. Wenn KI hingegen auf hohe Werte hin ausgerichtet wird, wie oben skizziert wurde, dann hat sie das Potenzial, Menschen zu bereichern. Dann ist sie kein Selbstzweck, sondern stellt gute Wertideen in den Vordergrund.

Wenn man Technologie an Werten ausrichtet, wird man darauf gestoßen, dass KI der 2. Generation (also Transformer Technologie) anders gebaut werden muss als wir das gedankenlos und ohne Ethos bisher mit KI der 1. Generation (Social Media) getan haben. Wir müssen dafür sorgen, dass die neuen KI-Systeme, die schon bald ohne Tastatur funktionieren werden, uns wieder den Raum und die Muße zurückgeben, den wir durch Social Media und Messenger verloren haben. Dass kann umgesetzt werden, indem zum Beispiel die Rechte des digitalen Netzes, uns zu erreichen von «everything, everywhere, all at once» umgestellt werden auf «selective, local, rhythmic». Ein KI-Softwareagent kann helfen zu selektieren, wer mich überhaupt, wann und wo erreichen darf (in Erfüllung des Sonntagsgebots bzw. der Regel 3). Das bedeutet natürlich auch, dass KI-Systeme, die für mich so arbeiten, von mir dezentral kontrolliert und konfiguriert werden müssten. Die Machtasymmetrien, die wir derzeit haben, etwa in Social Media, wo externe Geschäftsmodelle bestimmen, wann ich was sehe, müssen im Sinne der 10. Regel ein Ende finden: «Verhindert Machtkonzentration und garantiert Teilhabe».

Durch solche Konfigurationen von KI unter der Kontrolle des Users könnte die Freiheit aller rehabilitiert werden. Zwei Parteien – etwa Eltern und Kinder oder Arbeitgeber und Arbeitnehmer – würden die Möglichkeit haben einzustellen, wer wen wann sehen und überwachen darf. Dabei kann das personale Begehrlichkeitsverbot von Regel 9 – Nutzt Maschinen nicht, um die Freiheit Anderer zu untergraben – hilfreich sein, denn vielleicht ist heute zu wenig bewusst, dass das Überwachen von Anderen einen übergriffigen Freiheitsentzug bedeutet. Vielleicht hilft Regel 9 uns daran zu erinnern, dass unsere Freiheit mit das höchste Grundrecht war, was uns die Moderne geschenkt hat. Was man dann im Sinne von Regel 6 beachten muss, ist, dass alle Daten, die man mit oder ohne KI sammelt, immer nur einen begrenzten Aussagewert haben. Daten sind oft falsch und manipulierbar und gleichzeitig verleiten sie zu einer übergriffigen Kontrollillusion. Wenn etwa Arbeitgeber die Freiheit ihrer Mitarbeiter im Homeoffice durch Überwachung untergraben und verhin-

dern wollen, dass diese die Mittagspause ausdehnen, um einkaufen zu gehen, im Fachjargon also «Arbeitszeitbetrug» nachgehen, dann dürfen sie sich nicht wundern, wenn das «Datenobjekt» Mitarbeiter sich einen Mouse-Jiggler kauft, um der KI seine Präsenz vorzutäuschen. Wer das Ethos zwischenmenschlichen Vertrauens verletzt und Menschen zum Datenobjekt degradiert, erntet nichts als Kampf. Und das ist jedenfalls keine Voraussetzung dafür, Technologie friedlich einzusetzen, was das Ethos erreichen will. Regel 6 – Behandelt Menschen nicht als bloße Datenobjekte – ist also essenziell, um die Stakeholder am «Bau der Zukunft» zur positiven Mitarbeit zu motivieren

Schließlich helfen die Regeln dabei, uns bewusst zu machen, dass Technik nicht perfekt ist und dass es keinen Sinn macht anzustreben, Menschen und ihre ungeheuren Kreativ- und Lösungspotenziale durch KI ersetzen zu wollen. Wie ich oben bereits angedeutet habe, sind Softwaresysteme, wie hilfreich sie auch sein mögen, voll von Fehlern und wenn man Menschen einfach so durch KI ersetzt, dann wird man zwar mit weniger menschlichen Fehlern in Organisationen zu kämpfen haben, dafür aber mit KI- und Softwarefehlern. Die Bilanz bleibt also gleich; mit dem Unterschied, dass Softwarefehler schwerer zu beheben sind als menschliche Fehler, da es immer Fachpersonal braucht, diese zu finden und aufzuheben. Diese Rechnung wird heute leider zu wenig aufgemacht. Das Gegenteil des Lügenverbots von Regel 8 – sorgt Euch um die Wahrheit und verleugnet nicht die Grenzen der Technik – ist heute immer noch gängige Praxis. IT-Fehler werden systematisch unter den Teppich gekehrt. Keiner will dafür gerade stehen. Das Leid, die Kosten, der Missmut, der Vertrauensverlust und die Depressionen, die dadurch in Organisationen und auch in ganzen Branchen schon entstanden sind, ist wohl kaum zu beziffern. Aber KI bringt Softwarefehler durch das Phänomen des Halluzinierens und durch die absurdesten Falschempfehlungen so sehr ins Bewusstsein, dass davon auszugehen ist, dass die Grenzen der Technik in Zukunft hoffentlich offener und realistischer diskutiert werden können.

Wenn dann einmal die Grenzen von Technologien anerkannt, erforscht und thematisiert sind (ein wichtiger Forschungsbereich), dann besteht auch die Chance, im Sinne des Raubverbots von Regel 7 (Lasst Euch nicht Eurer menschlichen Potenziale berauben) wieder nüchterner über die Aufgabenverteilung zwischen KI-gestützten Maschinen und Menschen nachzudenken. Regel 7 ermutigt alle, menschliche Potenziale wieder zu entdecken und diese zu verteidigen. Nicht KI-Systeme, sondern Menschen sollten in ihren verschiedenen Rollen als Ärzte, Künstler, Seelsorger, Journalisten, Wissenschaftler wieder denken und handeln dürfen, wozu ihnen das Mußegebot von Regel 3 Raum gibt (Sabbatgebot). Dass man dann KI als Optimierer, Assistent, Challenger, etc. in angebrachten Kontexten hinzuziehen kann, steht außer Frage.

Aber auch wann und wie oft man das tut, mag wohl überlegt sein. Wo braucht man KI wirklich? Diese Frage ist essentiell in einer Zeit, wo 41% des Pro-Kopf CO² Budgets schon heute für IT ausgegeben wird.²² Das Tötungsverbot der Bibel und Regel 5 – Zerstöre nicht Leben und Umwelt für den technischen Fortschritt – kehrt hier in vergrößertem Format wieder: Software should not «eat the world».

Anmerkungen

- 1 Vgl. Georges DUBY, *Die Zeit der Kathedralen - Kunst und Gesellschaft 980 – 1420*, Frankfurt/M., 2019.
- 2 Anna JOBIN u.a., *The global landscape for AI ethics guidelines*, in: *Nature – Machine Intelligence* 1 (2019), 389–399.
- 3 EU COMMISSION, *Artificial Intelligence Act. In COM(2021)*, Brüssel 2023.
- 4 HLEG, *Assessment List for Trustworthy AI (ALTAI)*, Brüssel, 2020.
- 5 Vgl. JOBIN, *The global landscape for AI* (s. Anm. 2).
- 6 Max SCHELER, *Der Formalismus in der Ethik und die Materiale Wertethik - Neuer Versuch der Grundlegung eines ethischen Personalismus*, Hamburg 2014, 48.
- 7 Baty FRIEDMAN – Peter KAHN, *Human values, ethics, and design*, in: Julie JACKO – Andrew SEARS (Hg.), *The Human-Computer Interaction Handbook*, New York 2003, 1177–1201.
- 8 Sarah SPIEKERMANN, *Value-based Engineering. A Guide to build ethical Technology for Humanity*, Berlin, 2023.
- 9 ISO, *ISO/IEC/IEEE 24748-7000:2022. Standard model process for addressing ethical concerns during system design. Systems and software engineering — Life cycle management*, 2022.
- 10 Kathrin BEDNAR – Sarah SPIEKERMANN, *The Power of Ethics. Uncovering Technology Risks and Positive Value Potentials in IT Innovation Planning*, in: *Business & Information Systems Engineering* 66/2 (2024), 181–201.
- 11 Eric S. RAYMOND, *The Cathedral and the Bazaar*, Sebastopol 1999.
- 12 Joseph BRIGGS – Devesh KODNANI, *The Potentially Large Effects of Artificial Intelligence on Economic Growth*, in: <https://www.gspublishing.com/content/research/en/reports/2023/03/27/d64e052b-0f6e-45d7-967b-d7be35fabd16.html> (abgerufen am 18.11.2025).
- 13 Stefan MAGER – Tobias OSSEFORTH, *Erfolgsfaktoren und Fallstricke von IT-Projekten*, in: <https://live.handelsblatt.com/erfolgsfaktoren-und-fallstricke-von-it-projekten/> (abgerufen am 03.12.2025).
- 14 Mark SHERMANN, *Using ChatGPT to Analyze Your Code? Not So Fast*, in: <https://insights.sei.cmu.edu/blog/using-chatgpt-to-analyze-your-code-not-so-fast/> (abgerufen am 04.12.2025).
- 15 Jochen KNECHT, *Warum Tesla keine 6 Jahre Vorsprung hat*, in: <https://www.auto-motor-und-sport.de/tech-zukunft/tesla-model-3-teardown-kommentar/> (abgerufen am 04.12.2025).
- 16 Eberhard SCHOKENHOFF, *Naturrecht und Menschenwürde. Universale Ethik in einer geschichtlichen Welt*, Mainz 1996, 261ff; Sebastian Günther, *People of the Scripture! Come to a Word Common to You and Us (Q. 3:64). The Ten Commandments and the Qur'an*, in: *Journal of Qur'anic Studies* 9/2 (2007), 28–58.
- 17 Nir EYAL and Ryan HOOVER, *Hooked. How to Build Habit-Forming Products*, New York, 2015.
- 18 Gloria MARK, *Attention Span. A Groundbreaking Way to Restore Balance, Happiness and Productivity*, New York 2023.
- 19 Alle Beteiligten haben den Großteil ihrer beruflichen Laufbahn damit verbracht, verschiedene Facetten menschlicher, sozialer oder ökologischer Themen in unsere aktuellen technologischen und wirtschaftlichen Innovationen einzubringen. Ihr Hintergrund deckt funktional den gesamten «Science-Fiction-Zyklus» der heutigen KI-Technologien von der ersten bis zur dritten Generation ab, einschließlich Roboter- und neuronaler Technologien. Er umfasst neben der Informatik und Wirtschaftsinformatik die Fachbereiche der Medizin, Technikfolgenabschätzung, Recht, Militärtechnologie, Neurowissenschaften und Bionik, Psychotherapie, Psychologie, Soziologie, Sozioökonomie, Philosophie und Theologie.

- 20 Wir möchten an dieser Stelle Prof. Dr. Schwienhorst-Schönberger danken für seine aktive Unterstützung; sowohl bei der Interpretation des Dekalogs als auch beim Briefing der Teilnehmer der Göttweiger Future Foundation Gespräche.
- 21 Genau wurde zunächst von vier Personen schriftlich und unabhängig voneinander ein Mapping vorgenommen zwischen den 10 Geboten und Digitalherausforderungen, gefolgt von zwei Workshops mit 14 Teilnehmern. Die Workshopergebnisse führten zu einem ersten Entwurf der 10 Regeln. Die Workshopergebnisse wurden entlang des Regelvorentwurfs und der 10 Gebote verdichtet. Dabei wurde insbesondere auf die Reihenfolge der 10 Gebote und deren Wortlaut geachtet, mit dem Ziel, dass der Sinn des jeweiligen Dekalog-Gebots und seine Position in den 10 Regeln erhalten bleibt.
- 22 Robert ISTRATE u.A., *The environmental sustainability of digital content consumption*, in: Nature Communications 15/3724 (2024).

Abstract

Cathedrals of Artificial Intelligence. Economic and Ethical Challenges for True Progress. This article looks at AI from an economic perspective, with a particular focus on innovation management and product development. It argues that the current development of AI is embedded in a business context that does not suggest that the technology will bring about the progress for society that would otherwise be possible. The Gothic cathedral, symbolic of European culture, is used as an ideal image of extraordinary progress. According to the thesis of this article, cathedrals could only be built because their creation process was accompanied by a three-part formula for success: < values + engineering process knowledge + ethos >. The article outlines these three vehicles of progress and highlights how they could be considered in AI development in order to unleash their true potential for progress.

Keywords: