

A.

Digitalisierung und Theologie: die Vermessung des Feldes
Digitization and Theology: Surveying the Field

Die Theologie im Prozess der Quantifizierung von Lebenssachverhalten. Beschreibungshoheit und Deutungshoheit

Carina Geldhauser / Hermann Diebel-Fischer / Martin Kutz

1. Ausgangslage

Beziehungen sind nicht immer einfach – vor allem dann nicht, wenn die an ihnen Beteiligten mit unterschiedlichen Interessen agieren. Das Beziehungsgeflecht, das uns im Rahmen der Digitalisierung begegnet, ist ein solches. Was Digitalisierung alles einschließt, ist dabei zunächst ebenso undurchsichtig wie ihre Konsequenzen. Als Prozess ist sie ein Geschehen, das ein sozio-technisches Großprojekt darstellt. Ihr transformierender Charakter scheint bisweilen gar einem Selbstzweck zu gleichen. Nichts, so scheint es, ist noch sicher davor, *digitalisiert* zu werden.

Die Welt des sogenannten *Digitalen* hat Auswirkungen, die Felix Stalder als „Kultur der Digitalität“¹ bezeichnet und die Sabine Müller-Mall grundlegend als politisch begreift.² Beide Deutungsangebote zeichnen sich dadurch aus, dass sie ein Hereinbrechen einer technischen Sphäre in die Gesellschaft als in hohem Maße transformativ verstehen.

Es ist unbestritten, dass die gegenwärtigen Entwicklungen auf dem Gebiet der Informatik Auswirkungen haben, die weit über ihr eigenes Feld hinausreichen. Maßgeblich vorangetrieben wird dies von den Entwicklungen auf dem Gebiet des Machine Learnings (ML) im Speziellen bzw. der Künstlichen Intelligenz (KI) im Allgemeinen. Findet eine Informatisierung unserer Lebenswelt statt – und wenn ja, wie ist sie zu verstehen und wie kann ihr angemessen begegnet werden?

2. Problembeschreibungen

Mit der Digitalisierung gehen Prozesse einher, die wir zunächst abstrakt darstellen und dann anhand des durchaus streitbaren Beispiels eines Chat-

1 Stalder, Felix (2016): Kultur der Digitalität, Berlin.

2 Vgl. Müller-Mall, Sabine (2020): Freiheit und Kalkül, Ditzingen.

bots im Seelsorgeeinsatz erläutern wollen. Die relevanten Prozesse sind (a) die Modellierung sowie (b) die Quantifizierung, um Sachverhalte berechenbar, also mit einem Algorithmus entscheidbar, zu machen.

2.1 Modellierung

Wann immer Prozesse unserer Lebenswelt mit Computern bearbeitbar gemacht werden sollen, müssen sie vorher modelliert werden. Diese Modellierung bedeutet, dass eine abstrakte und damit verallgemeinernde Beschreibung des Abzubildenden erstellt werden muss. Dies führt zwangsläufig dazu, dass bestimmte Aspekte keine Berücksichtigung finden, während andere in das Modell einfließen. Dazu müssen nicht nur Entscheidungen über die Relevanz der einzelnen Bestandteile und Faktoren eines Prozesses getroffen werden, sondern zunächst gilt es, diese überhaupt zu identifizieren. An dieser Stelle wird deutlich, dass *Beschreiben* und *Werten* bzw. *Werturteile* zusammenhängen. Vor dem Hintergrund der Frage nach der Modellierung von Lebenssachverhalten lassen sich folgende Punkte unterscheiden:

2.1.1 Was wird modelliert?

Nicht jeder Lebenssachverhalt lässt sich abstrakt abbilden, ohne gleichzeitig unsachgemäß vereinfacht zu werden. Die Möglichkeit einer Modellierung besteht vor allem dort, wo bereits klare Strukturen ein Geschehen gliedern und eindeutige Zuordnungen möglich sind – dies wird später noch relevant werden. Diese Modellierung kann aber auch automatisiert erfolgen, wie etwa in künstlichen neuronalen Netzen. Sprachmodelle, Large Language Models (LLMs), sind wahrscheinlickeitsbasierte Abstraktionen, die ein Modell von Sprache(n) darstellen. Geht es jedoch darum, bestimmte Aspekte der Sprache darzustellen, um bei der Verwendung solcher Modelle in Chatbots bestimmte *Antworten*, also Sprachausgaben, zu verhindern oder zu vermeiden, wird es komplizierter: Wenn man verhindern möchte, dass ein Chatbot als Sprachagent eine Nutzerin in ihrem Vorhaben, einen Suizid zu verüben, bestärkt, gleichzeitig aber die Möglichkeit wahren möchte, dass der Chatbot mit der Nutzerin über diese Themen sprechen kann, dann reicht es nicht aus, das Thema *Suizid* zu modellieren und aus den Gesprächen auszuschließen (und auf professionelle Hilfe durch einen Menschen zu verweisen). Vielmehr müssten zusätzlich die Situationen *über ein Thema*

sprechen, jemanden bezüglich eines Themas im themenbezogenen Vorhaben bestärken etc. modelliert werden.

2.1.2 Wer trifft die Entscheidung darüber, was in ein Modell überführt wird?

Eng verbunden mit der Frage danach, was modelliert wird, steht jene nach den Entscheidungsstrukturen. In der Regel sind dies keine öffentlichen Stellen, die diese Entscheidungen treffen, weswegen die Strukturen undurchsichtiger sind. Es sollte jedoch beachtet werden, wer die Entscheidung mit welchen Intentionen getroffen hat. Gerade für öffentliche Angelegenheiten mag es naheliegen, ein demokratisches Verfahren zu verlangen. Jedoch bleibt zu bedenken, wie Johannes Himmelreich ausgeführt hat, dass die Demokratisierung solcher Prozesse nicht notwendig dazu führt, dass die Ergebnisse besser oder gar in vorzugswürdiger Weise ausfallen.³ Davon nicht vollkommen losgelöst steht die Frage nach der gesellschaftlichen oder gruppenbezogenen Akzeptanz: Wie bereitwillig wird eine Technologie eingesetzt und lassen sich anfallende Entwicklungskosten über den Einsatz refinanzieren?

2.1.3 Welchen Stellenwert hat das Modell im Lebenssachverhalt?

Unsere Lebenswelt ist komplex und unsere Ressourcen, sie zu beherrschen und zu gestalten, sind limitiert. Man kann in Luhmann'scher Tradition der Komplexität begegnen, indem man in der (soziologischen) Beschreibung nicht mehr auf das Ganze der Gesellschaft verweist, sondern einzelne Funktionsbereiche sowie ihre Binnenlogiken (ihre *Kommunikationen*, ihre binären Codes), nach denen sie operieren, betrachtet.⁴ Auf diesem Wege ist es möglich, Teilaspekte zu beschreiben, ohne das Gesamtsystem und die in ihm handelnden Individuen betrachten zu müssen. Aus diesen Beschreibungen von Zusammenhängen können dann Erklärungen gewonnen werden.

Was für die Soziologie innerhalb ihrer Grenzen hilfreich ist, geht an den Erfordernissen des Lebens bisweilen vorbei: Stellen wir uns vor, wir wollten

3 Vgl. Himmelreich, Johannes (2022): Against "Democratizing AI", in: *AI & Society* 38 (2022/4), 1333–1346.

4 Vgl. Luhmann, Niklas (1998): *Die Gesellschaft der Gesellschaft*. Bd. 1 und 2. Frankfurt/M.

– losgelöst von allen rechtlichen Bedenken – das Thema *Seelsorge* so weit in (teil-)automatisierte Prozesse überführen, dass der Einsatz von menschlichen Seelsorgenden weitgehend reduziert werden kann. Wenn man davon ausgeht, dass Seelsorge einem professionellen Selbstverständnis unterliegt und dass die Ziele einer seelsorglichen (Aus-)Bildung hinreichend präzise beschrieben sind, sodass gemessen werden kann, dass und zu welchem Grad sie erreicht werden, dann spricht zunächst nichts dagegen, anhand der festgelegten Kriterien zu überprüfen, inwieweit diese Standards eingehalten werden. Eine solche Kriteriologie erlaubt es zu überprüfen, ob und zu welchem Grad das Modell dasjenige abbildet, was ihm in der Realität zugrunde liegt.

Aus den bis hierhin dargelegten Umständen wird deutlich, dass mit dem Modell und seinem Einsatz in der Welt verschiedene Kriterien, Maßzahlen und somit auch Schwellenwerte verbunden sind. Dies führt uns zum zweiten Punkt, dem Paradigma der Berechnung.

2.2 Quantifizierung und Berechenbarkeit

Auf das Thema *Modellierung* folgt unmittelbar das Paradigma der Berechnung bzw. der Berechenbarkeit. Die mit einem Modell einhergehende Abstraktion kann sich zunächst darauf beschränken, aufzuzeigen, wo Entscheidungen wie getroffen werden.

Geht es um den Vergleich von Möglichkeiten, so muss gerade bei an sich qualitativ erlebten Lebenssachverhalten wie *Gleichberechtigung*, *Fairness* oder eben *toxic language* eine Messbarkeit hergestellt werden. Diese ermöglicht dann, unscharfe Konzept miteinander zu vergleichen und eine Auswahl oder Entscheidung zu treffen. Das Problem der Operationalisierung ist jedoch nicht trivial, weil jede Operationalisierung mit einer Modellbildung verbunden ist, bei der auf dem Wege der Abstraktion Teile eines Konzeptes entweder entfallen oder stärker betont werden. Bei unscharfen Konzepten ist zudem die Subjektivität der konkreten Bestimmung, um was genau es sich handelt, ein Faktor, der nicht vernachlässigt werden darf.

Die Entscheidung für eine quantitative Abbildung nichtquantitativer Sachverhalte ist somit möglich, bedarf aber in der Regel einer Erklärung. Damit kann nachvollziehbar gemacht werden, welche konkrete Vorstellung wie quantifiziert worden ist. Wenn es um die automatisierte Erzeugung von Entscheidungen geht, dann bedarf es der Festlegung von Schwellenwerten – d.h. von Werten, ab deren Erreichen etwas von vornherein anderes gilt.

Dies wird insbesondere dort wichtig, wo sich dieser Schwellenwert auf ein Abstraktum bezieht.

Nimmt man das Beispiel eines *toxicity scores*, also eines Werts, der eine Maßzahl für die *Toxizität* einer Textausgabe darstellt, wird deutlich, worin das Problem besteht: Daran kann man zeigen, dass die Abstraktion zweiter Ordnung eine durchaus wichtige – weil entscheidende und auf die reale Welt zurückwirkende – ist. Das Anliegen, die Ausgaben eines LLM-basierten Chatbots nicht *toxisch* sein zu lassen, beruht darauf, dass eine bestimmte Anspruchshaltung mit einer Kommunikation verbunden steht. Diese Anspruchshaltung wird dann in ein Modell überführt, innerhalb dessen festgelegt ist, welche Sprachausgaben (Beleidigungen, Obszönitäten, rassistische Aussagen, sexistische Aussagen etc.) vermieden werden sollen.

Dies lässt sich nicht mit einer simplen Liste von unerwünschten Lemmata bewerkstelligen, da somit zu viele Aussagen verhindert würden, indem Verneinungen als *Entschärfungen* ignoriert oder die Thematisierung derartiger Begriffe auf der Metaebene verunmöglicht werden. Aus diesem Grund bedient man sich weiterer Modelle, deren Aufgabe es ist, einen Wert zu ermitteln, bei dessen Über- bzw. Unterschreiten eine Aussage ausgegeben oder vermieden wird. Die Festlegung dieser Grenze geschieht jedoch einigermaßen willkürlich – zwar schon so, dass die Erfahrung der Festlegenden mit einfließt, aber eben doch so, dass ein (fiktiver) *Score* von $> 0,05$ als Schwellenwert etwas möglich bzw. unmöglich werden lässt, was bei einem Wert von $0,04$ nicht geschehen würde. Hier bleibt einerseits die Frage nach der semantischen Bedeutung der Differenz von $0,01$ bestehen, andererseits wird die Frage danach aufgeworfen, welches Vertrauen man in diese *Scores* haben sollte.

Damit ist nicht gemeint, dass das Festlegen von Schwellenwerten an sich schädlich sei – dies ist gerade dann nicht der Fall, wenn es darum geht, bestimmte Aufmerksamkeit auf Sachverhalte zu lenken, um diese dann genauer anzuschauen. Im Falle der Automatisierung jedoch entfällt dieser Schritt des *genauen Hinschauens* und die Nutzenden bzw. vom System Betroffenen müssen sich darauf verlassen können, dass diese Arbeit vorher vollumfänglich geleistet worden ist.

3. Deutungsansprüche – Chatbots in der Seelsorge?

Das Angebot von Chatbots, mit denen Themen auf dem Gebiet der mentalen Gesundheit besprochen werden können, nimmt stetig zu.⁵ Gleichzeitig lassen die zunehmende Entfremdung von der Kirche und ihren Angeboten, aber auch die Hemmung, mit fremden Menschen über die eigene Endlichkeit zu sprechen, das Seelsorgeangebot durch Geistliche nicht für jeden Menschen als attraktiv erscheinen.⁶ Das weitestgehend unregulierte Angebot an Chatbots steht im Kontrast zu einem professionellen Angebot, wie es von den Kirchen gemacht wird.

Ohne das Gebiet des Therapeutischen betreten zu wollen, haben wir in einem ersten Schritt ausgelotet, inwieweit sich LLM-basierte Chatbots für Seelsorgekontexte eignen.⁷ Dieses Einsatzszenario illustriert die im vorangegangenen Abschnitt beschriebenen Probleme, wenn es darum geht, einen solchen Chatbot mit Sicherheitsmechanismen auszustatten, mit denen vermieden werden soll, dass ein Chatbot diejenigen gefährdet, die ihn nutzen.

Wenn diese Nutzung ohne *Human-in-the-Loop* geschehen soll, d.h. keine Überwachung stattfindet, dann muss sichergestellt werden, dass der Chatbot entsprechend vorher festgelegter Limitationen agiert und aus diesen auch nicht ausbrechen kann, ganz gleich, welche Eingaben – sei es durch Sprache oder durch Text – durch die Nutzenden getätigt werden. Um diesen Anspruch gerecht werden zu können, müssen sowohl die für die Nutzenden sichtbare Anwendungsoberfläche als auch der Chatbot selbst entsprechend modelliert werden. Ersteres ist wichtig, da zu Testzwecken Simulationen durchgeführt werden müssen, in denen jedes noch so abwegige Nutzungsverhalten getestet werden kann.

Dafür ist es unumgänglich, dass die gewünschten wie auch die unerwünschten Gesprächsverläufe identifiziert werden. Betrachtet man nur den Fall der Suizidprävention, so wird deutlich, dass dies eine äußerst komplexe

5 Vgl. Bernhardt, Klaus (2024): Der KI-Therapeut. Psychische Probleme mit künstlicher Intelligenz überwinden – KI-Tools als erste Hilfe für Betroffene, München; vgl. auch te Wildt, Bert/Lauer, Gerhard/Schmidt, Robin (Hg., 2024): Was machen Digitalisierung und Künstliche Intelligenz mit der Psychotherapie? Einwürfe und Provokationen, Berlin/Boston.

6 Vgl. Blackstein, Achim/Krabbes, Carsten (2024): Formen und Chancen Digitaler Seelsorge. Erfahrungen aus der Praxis, in: Praktische Theologie. Zeitschrift für Praxis in Kirche 59 (2024/1), 14–17.

7 Vgl. Diebel-Fischer, Hermann u.a. (in Vorbereitung): CareGPT? – Exploring the Requirements for and the Capabilities of LLM-based Counseling Bots, München.

Aufgabe ist. Während die Pipeline zur Vermeidung als schädlich (*harmful*) betrachteter Trainingsdaten im Rahmen der Erzeugung eines LLMs verhältnismäßig simpel ist,⁸ wird mit Blick auf die Dialogauswertung und -bewertung für einen unüberwachten Chatbot das Problem akut, dass der Kippunkt, ab dem eine *normale* Unterhaltung besorgniserregend wird, nicht eindeutig benannt werden kann. Dies macht es selbst für den Fall, dass sich dieses Kippen in eine unerwünschte Richtung operationalisieren lässt, zu einem risikoreichen Unternehmen, entsprechende Schwellwerte festzulegen. In diesem Zusammenhang eine vertretbare Entscheidung zu treffen, erfordert das Zusammenspiel mehrerer Disziplinen. Wer hat die Deutungshoheit darüber, was angebracht und zumutbar ist?

4. Interdisziplinarität

Die Mathematik gilt gemeinhin als Sprache der Quantifizierbarkeit, dennoch ist sie mehr als das: sie hilft mit ihrer abstrakten Sprache, Strukturen zu erkennen und zu beschreiben. Jedoch sind Mathematiker:innen nicht daran gebunden, aus einem gegebenen Sachverhalt Modellannahmen und -parameter zu extrahieren und mit diesen – und nur diesen – Aussagen zu treffen, im Gegenteil: Zwar lassen sich mathematische Modelle durch ihre Anlehnung an Phänomene der Natur motivieren, jedoch sind die konkreten Modellannahmen allein Sache der Mathematiker:innen. Oft entstehen Modellannahmen daraus, dass nur unter diesen Umständen mit bestimmten Werkzeugen eine rigorose Aussage getroffen werden kann.

Im akademischen, innermathematischen Umfeld ist die Anlehnung eines mathematischen Modells an ein physikalisches oder biologisches Modell nur eine von vielen Indikatoren zur Beurteilung des Erkenntnisinteresses. Es zählt auch, ob eine Fragestellung schwer zu lösen ist, ob neue Methoden entwickelt werden oder ob der aktuelle Wissensstand erweitert wird, beispielsweise durch eine Verallgemeinerung auf weitere Objekte, weniger restriktive Voraussetzungen und so weiter.

Über allem jedoch schwebt der Geist der universellen und zeitlosen Gültigkeit einer Aussage durch die Verwendung eines Beweises, also einer Verkettung logischer Aussagen, die uns von Axiomen und a priori festgelegten Modellannahmen hin zur Aussage führen. Nun lässt sich trefflich

8 Vgl. Brandizzi, Nicolo' u.a. (2024): Data Processing for the OpenGPT-X Model Family. arXiv:2410.08800v1.

darüber streiten, ob diese Beweise tatsächlich voraussetzungslos sind, d.h. alles aus Definitionen und Axiomen ableitbar ist, und folglich mathematische Erkenntnisse apodiktisch gewiss oder analytische Urteile im Sinne von Kant sind. Ebenso werden wir nie erfahren, ob es, wie Paul Erdős glaubte, tatsächlich ein Buch gibt, in dem Gott die perfekten Beweise gesammelt hat.⁹ Fest steht jedoch die „Angemessenheit und Genauigkeit der mathematischen Formulierung der Naturgesetze“¹⁰ und die fundamentale Bedeutung mathematischer Theorien für heutige technische Lösungen, vom Internet (Graphentheorie) über die Kryptographie zum Universal Approximation Theorem¹¹, dem theoretischen Fundament eines jeden Neuronalen Netzwerks.

Allerdings ist beim Ziehen von Schlussfolgerungen aus mathematischen Prinzipien über die Welt, in der wir leben, große Vorsicht angebracht: So hat schon Kurt Gödel sich „verheddert“, als er exakt bewiesene Theoreme mit umgangssprachlichen Begriffen verband, die keine scharfen Konturen haben. Ulrich Felgner konstatiert deswegen: „Aus einer solchen Verbindung können sich keine klaren Einsichten ergeben.“¹²

Es gilt also zu präzisieren, was die *Mathematisierung der Welt* leisten kann und was nicht. Da bei der Auswertung großer Datenmengen a priori nicht klar ist, welche Eigenschaften das Modell hat – diese zu derivieren, ist ja Teil der Fragestellung –, kann auch kein universell gültiger mathematischer Beweis einer Hypothese gegeben werden.

Hier ist die *data science* verantwortlich, aus ungeordneten *big data* interessante Fakten zu derivieren. Allerdings muss bei jeder Modellbildung und bei jeder Wahl der statistischen Methode zwischen verschiedenen Möglichkeiten gewählt werden, sodass hier interessanterweise auch traditionell geisteswissenschaftliche Modellbildungsmethoden zum Zuge kommen. So schreibt Mikhail Belkin, dass die Anwendung von Wilhelm von Ockhams

9 Vgl. Aigner, Martin/Ziegler, Günter Matthias (¹2002, 2018): Das BUCH der Beweise, Berlin/Heidelberg.

10 Wigner, Eugene (1960): The Unreasonable Effectiveness of Mathematics in the Natural Sciences, in: Mickens, Ronald Elbert (Hg., 1990): Mathematics and Science, Atlanta, 291–306.

11 Vgl. Cybenko, George (1989): Approximation by Superpositions of a Sigmoidal Function, in: Mathematics of Control, Signals, and Systems 4 (1989/2), 303–314; vgl. Hornik, Kurt (1991): Approximation Capabilities of Multilayer Feedforward Networks, in: Neural Networks 4 (1991/2), 251–257.

12 Felgner, Ulrich (2020): Philosophie der Mathematik in der Antike und in der Neuzeit, Berlin, 224.

Rasiermesser der Auswahl derjenigen Funktion unter den Funktionen, die sich perfekt an die Daten anpassen, gleiche, die nach einem bestimmten Begriff der Glattheit der Funktion am besten geeignet ist: „Select the smoothest function according to some notion of function smoothness, among those that fit the data perfectly.“¹³

Allerdings ist die Aussagekraft der Ergebnisse einer statistischen Auswertung ohne die Verwendung von in anderen Wissenschaften fundierten Hypothesen sehr beschränkt. So sind die bei Franz Messerli berichtete Korrelation von Schokoladenkonsum pro Kopf und der Anzahl der Nobelpreisträger pro zehn Millionen Einwohner pro Land wohl eher eine Kuriosität als bedingt durch einen kausalen Zusammenhang¹⁴ und die (auch mediale) Überbewertung statistischer Korrelationen eine Gefahr für die Wissenschaft¹⁵.

Zusammenfassend sehen wir: Im naturwissenschaftlichen Modellbildungszyklus wird zuerst eine Fragestellung außerhalb der Mathematik auf ein mathematisches Modell reduziert, dieses dann mit mathematikinhärenten Methoden bearbeitet und dann wiederum in das Originalproblem überführt. Dieser Modellbildungszyklus setzt allerdings voraus, dass die wichtigsten Wirkmechanismen (zum Beispiel Naturgesetze) vorausgesetzt und gegeben sind. Bei ML-Algorithmen verhält es sich aber anders: hier sollen nicht induktiv aus gegebenen Axiomen und Annahmen Aussagen getroffen, sondern aus Daten *deduktiv* Modellannahmen und Gesetzmäßigkeiten statistisch deriviert werden. Der ML-Algorithmus orientiert sich am statistischen Verhalten *im Mittel* – der gelegentliche statistische Ausreißer gehört also dazu –, was der mathematischen absoluten Sicherheit widerspricht.

Soll also ein Modell, welches der Lebenswirklichkeit nahekommt, aufgestellt und mit quantitativen Methoden untersucht werden, so entsteht eine Lücke zwischen mathematisch-rigorosen, eben *beweisbaren* Aussagen über ein klar definiertes Modell mit eindeutig gegebenen Annahmen und den mittels statistischer Methoden a posteriori vorgeschlagenen Hypothe-

13 Belkin, Mikhail (2021): Fit without Fear. Remarkable Mathematical Phenomena of Deep Learning through the Prism of Interpolation, in: Acta Numerica 30 (2021), 203–248, hier 218.

14 Vgl. Messerli, Franz (2012): Chocolate Consumption, Cognitive Function, and Nobel Laureates, in: New England Journal of Medicine 367 (2012/16), 1562–1564.

15 Vgl. Maurage, Pierre u.a. (2013): Does Chocolate Consumption Really Boost Nobel Award Chances? The Peril of Over-interpreting Correlations in Health Studies, in: The Journal of Nutrition 143 (2013/6), 931–933.

sen über das in unserem Datensatz vorliegende quantitative Abbild eines (möglicherweise verzerrten) Ausschnitts unserer Lebenswirklichkeit. Wenn letzterer dann wiederum in Wahrscheinlichkeiten (zum Beispiel die wahrscheinlich passendste Antwort auf einen Satz des mit dem Chatbot interagierenden Individuums) übersetzt wird, geht nochmals Schärfe und Aussagekraft verloren.

5. Deutungshoheit(en) – oder noch einmal: Wozu noch Theologie?

Friedrich Wilhelm Graf fragt in „Die Wiederkehr der Götter“ im letzten Abschnitt: „Wozu noch Theologie?“¹⁶ Seine Antworten auf diese damals rhetorische Frage legen Zeugnis ab von einem kulturprotestantisch-historisch geprägten Zugang zur Theologie und kulminieren in der Aussage: „[U]m in den Arenen von Wissenschaft, Zivilgesellschaft, Kirche und Politik der heilsamen Unterscheidung von Gott und Mensch Geltung zu verschaffen.“¹⁷ Stellt man diese Frage zwanzig Jahre später erneut, dann muss die Antwort auch dem Umstand Rechnung tragen, dass die zunehmende Entkirchlichung der Gesellschaft langfristig nicht spurlos an der Theologie als Wissenschaft vorbeigehen wird. Dass sie im breiten Bestand als Disziplin an Universitäten gesellschaftlich in Frage gestellt wird, ist nur noch eine Frage der Zeit. Umso wichtiger ist es daher, ihren Beitrag für die Gesellschaft über die–nach wie vor richtigen–Antworten Grafs hinaus deutlich zu artikulieren und dort, wo jene Unterscheidung als aufgehoben postuliert wird, kritisch hinzuschauen, was mit einer solchen Behauptung eigentlich bezweckt werden soll.

Die Theologie als Kulturwissenschaft mit originärer Weltdeutungsaufgabe aus christlicher Perspektive – oder genauer: der Aufgabe, diese Weltdeutung fundiert vorzubereiten – muss sich zu diesen Prozessen verhalten, wenn sie ihren Weltbezug erhalten möchte. Die Ethik hat in diesem Zusammenhang eine integrative Funktion: als hermeneutisches Instrument, um Probleme auf dem Gebiet der Diesseitgestaltung zu beschreiben und zu bearbeiten,¹⁸ ist sie Repräsentantin eines qualitativen Zugangs zur Welt, der

16 Graf, Friedrich Wilhelm (2004): Die Wiederkehr der Götter. Religion in der modernen Kultur, München, 249.

17 Ebd., 278.

18 Vgl. Fischer, Johannes (2012): Verstehen statt Begründen, Stuttgart; vgl. Rendtorff, Trutz (2011): Ethik, Tübingen.

einem quantitativen – mithin berechnenden – Zugang zunächst diametral entgegensteht: Das Fundament des KI-Modells ist in der universellen, formalisierten und dem Anschein nach objektiven Sprache der Definitionen, Theoreme und Beweise konstruiert. Im Kontrast dazu gibt es in der Ethik viele Akteure aus zahlreichen Wissenschaften; weder die Ziele noch die notwendigen Voraussetzungen oder gar Definitionen sind geklärt.¹⁹ Eine angewandte Ethik für ein Feld, das auf mathematischen Füßen steht und berechnet, während die Ethik bedenkt, erscheint nicht nur dann als große Herausforderung, wenn KI-Systeme mit einer Ethik ausgestattet werden sollen. Schon die notwendige Operationalisierung von Ethik erweist sich als schwierig. In Anlehnung an Charles Percy Snow kann man diesbezüglich durchaus von zwei Kulturen²⁰ eines gesellschaftlichen Transformationsprozesses sprechen.

Doing Theology Digitally, gelesen als Theologie Treiben für das Digitale, hat zur Aufgabe, verschiedenartige Zugänge zur Welt miteinander ins Gespräch zu bringen und Übersetzungsprozesse aufzudecken und zu moderieren. Ihre eigene Positionalität – gegeben durch ihre jeweilige Bekenntnisorientierung – ist dabei insofern hilfreich, als die Theologie durch diese um die grundsätzliche Bedeutung von Vorannahmen weiß und gleichzeitig die Möglichkeit kennt, dass ebenjene Vorannahmen auch andere sein könnten.

Vor dem Hintergrund des Aufeinandertreffens verschiedener Kulturen der Weltdeutung (qualitativ vs. formalisiert-quantitativ) in der Digitalisierung kann die Theologie auf der Metaebene moderierend und auf der materialetischen Ebene (mit-)gestaltend aktiv werden. Anthropologisch betrachtet steht nämlich die Fehlbarkeit in Form der Unzulänglichkeit des Menschen im Raum, solche Probleme angemessen zu bearbeiten, die hier in ihrer profanisierten Bedeutung zum Tragen kommt. Nicht die gestörte Gottesbeziehung des Menschen ist der relevante Makel, sondern seine gestörte Beziehung zu sich selbst und der Welt. Der Gedanke, der eigenen Imperfektion und den eigenen Unzulänglichkeiten über Umwege Abhilfe verschaffen zu können, also über das Erschaffen von Maschinen, die das, was fehlt, ausgleichen, tilgt die Imperfektion nicht, sondern verlagert diese nur. Christian Schwarke hat diese Verlagerung anhand der Robotik in Aufnahme des schöpfungstheologischen Paradigmas und des Bildes der *imago Dei* treffend beschrieben:

19 Vgl. Hagendorff, Thilo (2020): The Ethics of AI Ethics. An Evaluation of Guidelines, in: *Minds and Machines* 30 (2020/1), 99–120.

20 Vgl. Snow, Charles Percy (1959): *The Two Cultures*, Cambridge.

„Die Figur der Gottebenbildlichkeit hilft, solche Zusammenhänge zu erschließen, und darin ein Verständnis dessen zu erreichen, was geschieht, wenn wir Roboter machen, die uns ähnlich sind. Auch eine so weitgehende technische Entwicklung wäre keine Gegenschöpfung und keine Usurpation göttlichen Territoriums. Sie wäre vielmehr die Realisierung der Gottebenbildlichkeit. Ob man aber die Konsequenzen schätzen wird, steht auf einem anderen Blatt.“²¹

Bei der Frage nach einem Chatbot, der auch zur Seelsorge fähig ist, geht es nicht so sehr um Roboter als Artefakte, die *Hülle* und *Inhalt* des Dargestellten bzw. Simulierten sind – sie sind noch mal ein eigenes Thema der Mensch-Maschine-Interaktion. Sondern es geht um das, was als Intelligenz im weitesten Sinne des Wortes verstanden wird und mit menschlichen kognitiven Fähigkeiten assoziiert ist.²² Indem wir Maschinen bauen, die Aufgaben übernehmen, die für einen Menschen nicht zu bewältigen sind, schafft sich der imperfekte Mensch ein Gegenüber, dessen Imperfektion für ihn selbst zum Problem werden kann. Um im Bilde zu bleiben: um den Konsequenzen in der von Schwarke aufgeworfenen Analogie Herr zu werden, sieht jener die biblische Überlieferung als Anhaltspunkt für Orientierung.²³ Damit ist sein Ausblick wesentlich optimistischer als der jener Zeitgenossen, die darin eine Aufhebung der Grenze zwischen Mensch und Maschine sehen.²⁴ Doch was folgt, wenn man diese Analogie weiterdenkt? Die Schöpfungserzählung führt weiter in jene von der Erlösungsbedürftigkeit der Geschöpfe, da ihre Beziehung zum Schöpfer *gestört* ist. Das Erlösungsangebot in Form des Christus stellt den (vorläufigen) Höhepunkt der Heilsgeschichte dar – aber was ergibt sich daraus für die Welt?

Theologisch lässt sich hier lernen, dass die Beziehung von Schöpfer und Geschöpf dann keine einfache ist, wenn dem Geschöpf Autonomie zuteilwird und der Schöpfer eine Bindung zum Geschöpf erhalten will. Interpretiert man das Schaffen des Menschen vor diesem Hintergrund,

21 Schwarke, Christian (2017): Technik und Christentum. Anmerkungen zu einem verkanteten Verhältnis, in: Böhmer, Sebastian u.a. (Hg.): Technologien des Glaubens. Schubkräfte zwischen technologischen Entwicklungen und religiösen Diskursen (Acta Historica Leopoldina 71), Stuttgart, 131–142, hier 141.

22 Vgl. Seising, Rudolf (2024): Geist, Intelligenz, Information und Daten – Artificial Intelligence im Wandel der Wissenschaftskulturen. Eine ideengeschichtliche Begriffsverortung, in: Reder, Michael/Koska, Christopher (Hg.): Künstliche Intelligenz und ethische Verantwortung, Bielefeld, 23–47.

23 Vgl. ebd.

24 Vgl. Burckhardt, Martin (2018): Philosophie der Maschine, Berlin, 303.

dann werden die unserer Problembeschreibung artikulierten Fragen zwar nicht gelöst, sondern es wird deutlich, dass diese ungeklärten Fragen nicht erst heute auftreten und akut werden.

Nun bleiben mehrere Deutungsmöglichkeiten: Es ist zunächst möglich, mit theologischem Bestock die Beziehunghaftigkeit des Menschen herauszupräparieren und festzustellen, dass es sich bei den gegenwärtigen Entwicklungen um ein großes Missverständnis handelt. Diese Interpretation würde David Franklin Nobles *The Religion of Technology* folgen, der argumentiert, dass es eine religiös motivierte Entwicklung von Technik in der christlich dominierten westlichen Welt gegeben habe. Diese hätte einstmals dem Wohle des Menschen gedient, sei aber von diesem Wege abgekommen. Schon 1998 führt er KI neben Atomwaffen und Gentechnik als Beleg dafür auf: „Put simply, the technological pursuit of salvation has become a threat to our survival.“²⁵ Die andere Interpretation wäre sehr viel biblischer und folgt Schwarkes Zugang einer theologischen Technikdeutung: Im Wagnis der Weltgestaltung war der Mensch seit jeher mit Kontingenz und Ambivalenzen konfrontiert. Die Welt zu bebauen und zu bewahren, das *dominium terrae*, das dem Menschen nach Gen 1,28 zuteilgeworden ist, hat keine Gelingensgarantie. Wohl aber kann der Mensch auf einen *concursum divinum* hoffen und unter den ihm gegebenen Möglichkeiten versuchen, sein Bestes zu geben.

Literaturverzeichnis

- Aigner, Martin/Ziegler, Günter Matthias (¹2002, 2018): Das BUCH der Beweise, Berlin/Heidelberg.
- Belkin, Mikhail (2021): Fit without Fear. Remarkable Mathematical Phenomena of Deep Learning through the Prism of Interpolation, in: Acta Numerica 30 (2021), 203–248.
- Bernhardt, Klaus (2024): Der KI-Therapeut. Psychische Probleme mit künstlicher Intelligenz überwinden – KI-Tools als erste Hilfe für Betroffene, München.
- Blackstein, Achim/Krabbes, Carsten (2024): Formen und Chancen Digitaler Seelsorge. Erfahrungen aus der Praxis, in: Praktische Theologie. Zeitschrift für Praxis in Kirche 59 (2024/1), 14–17.
- Brandizzi, Nicolò u.a. (2024): Data Processing for the OpenGPT-X Model Family. arXiv:2410.08800v1.
- Burckhardt, Martin (2018): Philosophie der Maschine, Berlin.

25 Noble, David Franklin (1999): *The Religion of Technology*, New York, 208.

- Cybenko, George (1989): Approximation by Superpositions of a Sigmoidal Function, in: *Mathematics of Control, Signals, and Systems* 4 (1989/2), 303–314.
- Diebel-Fischer, Hermann u.a. (in Vorbereitung): *CareGPT? – Exploring the Requirements for and the Capabilities of LLM-based Counseling Bots*, München.
- Felgner, Ulrich (2020): *Philosophie der Mathematik in der Antike und in der Neuzeit*, Berlin.
- Fischer, Johannes (2012): *Verstehen statt Begründen*, Stuttgart.
- Graf, Friedrich Wilhelm (²2004): *Die Wiederkehr der Götter. Religion in der modernen Kultur*, München.
- Hagendorff, Thilo (2020): The Ethics of AI Ethics. An Evaluation of Guidelines, in: *Minds and Machines* 30 (2020/1), 99–120.
- Himmelreich, Johannes (2022): Against “Democratizing AI”, in: *AI & Society* 38 (2022/4), 1333–1346.
- Hornik, Kurt (1991): Approximation Capabilities of Multilayer Feedforward Networks, in: *Neural Networks* 4 (1991/2), 251–257.
- Luhmann, Niklas (1998): *Die Gesellschaft der Gesellschaft*. Bd. 1 und 2. Frankfurt/M.
- Müller-Mall, Sabine (2020): *Freiheit und Kalkül*, Ditzingen.
- Maurage, Pierre/Heeren, Alexandre/Pesenti, Mauro (2013): Does Chocolate Consumption Really Boost Nobel Award Chances? The Peril of Over-interpreting Correlations in Health Studies, in: *The Journal of Nutrition* 143 (2013/6), 931–933.
- Messerli, Franz (2012): Chocolate Consumption, Cognitive Function, and Nobel Laureates, in: *New England Journal of Medicine* 367 (2012/16), 1562–1564.
- Noble, David Franklin (1999): *The Religion of Technology*, New York.
- Rendtorff, Trutz (³2011): *Ethik*, Tübingen.
- Schwarke, Christian (2017): Technik und Christentum. Anmerkungen zu einem verkanteten Verhältnis, in: Böhmer, Sebastian/Breuer, Constanze/Müller-Bahlke, Thomas/Tanner, Klaus (Hg.): *Technologien des Glaubens. Schubkräfte zwischen technologischen Entwicklungen und religiösen Diskursen (Acta Historica Leopoldina 71)*, Stuttgart, 131–142.
- Seising, Rudolf (2024): Geist, Intelligenz, Information und Daten – Artificial Intelligence im Wandel der Wissenschaftskulturen. Eine ideengeschichtliche Begriffsverortung, in: Reder Michael/Koska, Christopher (Hg.): *Künstliche Intelligenz und ethische Verantwortung*, Bielefeld, 23–47.
- Snow, Charles Percy (1959): *The Two Cultures*, Cambridge.
- Stalder, Felix (2016): *Kultur der Digitalität*, Berlin.
- Wigner, Eugene (1960): The Unreasonable Effectiveness of Mathematics in the Natural Sciences, in: Mickens, Ronald Elbert (Hg., 1990): *Mathematics and Science*, Atlanta, 291–306.
- te Wildt, Bert/Lauer, Gerhard/Schmidt, Robin (Hg., 2024): *Was machen Digitalisierung und Künstliche Intelligenz mit der Psychotherapie? Einwürfe und Provokationen*, Berlin/Boston.