

Ethik in der Softwareentwicklung



Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophie
der Ludwig-Maximilians-Universität München

vorgelegt von

Niina Marja Christine Zuber
aus
München
2022

Referent/in: Prof. Dr. Dr. h.c. Julian Nida-Rümelin, Staatsminister a.D. (LMU München)

Korreferent/in: Prof. Dr. Alexander Pretschner (TU München)

Tag der mündlichen Prüfung: 3. Mai 2022

Für meine Kinder Julius und Jonna

Danksagung

Diese Dissertation ist selbstverständlich nicht nur Resultat einsamen Nachdenkens, sondern beruht zudem auf zahlreichen Gesprächen, die sich über die letzten Jahre erstrecken. Zuallererst möchte ich daher Herrn Prof. Dr. Dr. h.c. Julian Nida-Rümelin, Staatsminister a. D. (LMU München) danken, der als Erstgutachter die Arbeit betreute. Das Umfeld seines Lehrstuhls bot mir beste Möglichkeiten, meine Gedanken vorzutragen und zu diskutieren. Seine Kenntnisse interdisziplinärer Debatten haben die Arbeit erst ermöglicht. Besonderer Dank gilt auch meinem Zweitgutachter, Herrn Prof. Dr. Alexander Pretschner (TU München), der als Informatiker diese Dissertation ebenfalls betreute. Diese Arbeit wäre in der engen Verbindung von Ethik und der Entwicklung von Softwaresystemen ohne einige entscheidende Anmerkungen seinerseits nicht möglich gewesen. Die Entscheidung zur Betreuung war nicht selbstverständlich und dafür danke ich Herrn Pretschner sehr.

Darüber hinaus danke ich Jan Gogoll und Severin Kacianka für ihr stetiges Mitdenken und das kontinuierliche Lesen meiner Dissertation. Ihre klugen und wichtigen Anmerkungen haben meine Ausarbeitung verbessert. Ebenso danke ich Dorothea Winter, Antonio Bikic, Adriano Mannino, Timo Greger und Rebecca Gutwald für ihre Beiträge und offenen Ohren.

Martin Rechenauer möchte ich für die Unterstützung danken, die er mir in den vergangenen Jahren hat zukommen lassen. Insbesondere verdanke ich ihm ein vertieftes Verständnis philosophischer Rationalitätstheorien.

Für die Korrekturen der Arbeit danke ich Carsten Freitäger und Melanie Zirves.

Ein besonderer Dank gilt auch dem Bayerischen Forschungsinstitut für Digitale Transformation (bidt), welches die Arbeit nicht nur finanziell, sondern auch mit einem interdisziplinären Netzwerk ermöglicht hat.

Für die jahrelange Begleitung danke ich meinem Lebensgefährten Carsten Freitäger und meiner Familie aus ganzem Herzen. Ohne ihre Unterstützung wäre es mir unmöglich gewesen, diese Arbeit zu verfassen.

Inhaltsverzeichnis

1	<i>Einführung</i>	12
2	<i>Technikphilosophische Grundlagen</i>	21
2.1	Zwei Denkschulen der Technikphilosophie	21
2.1.1	Der VDI-Diskurs: Der Entwicklungsprozess	23
1.2.	Phänomenologisch-pragmatische Technikphilosophie	38
2.2	Strukturelle Rationalität	44
2.2.1	Rationalität und Sozialität	52
2.2.2	Pragmatische Normativität.....	57
2.3	Erstrebenswerte Praxis	61
2.3.1	Emanzipierte Poiesis	66
2.3.2	Entwicklungspraxis	66
2.3.3	Bedienungspraxis	68
2.4	Zwischenfazit	69
2.5	Technikethische Grundlagen	74
2.5.1	Die Bereichsethik digitaler Artefakte: Makro-, Meso- und Mini-Ethiken	77
2.5.2	Humanistische Ethik: Vernunft, Freiheit und Verantwortung.....	85
3	<i>Ethik der Digitalisierung, der digitalen Transformation und des digitalen Artefakts</i> 92	
3.1	Gesichtserkennungssysteme in der polizeilichen Fahndungsarbeit	92
3.1.1	Legitimation von Gesichtserkennungssystemen	99
3.1.2	Technische Zuverlässigkeit	100
3.1.3	Ethische Betrachtung.....	103
3.1.4	Ausblick.....	107
3.1.5	Technik-generische Betrachtung	109
3.1.6	Strukturelle Betrachtung	113
3.2	Transparenz und Opazität als Reflexionsbegriffe	115
3.3	Die Methode der levels of abstraction	122
3.4	Zur Normativität der Digitalisierung	128
3.5	Zur Normativität digitaler Transformation	131
3.5.1	Zur Normativität digitaler Technik und Technologie	133
3.5.2	Moral Machines	138

3.5.3	Embedded Values Approach	141
4	<i>Die Einzigartigkeit der Informationstechnologie und -technik.....</i>	147
4.1	Der mediale Charakter der Informationstechnologie und -technik	148
4.2	Das Merkmal einer emerging technology	152
4.3	Vernetzung durch Transparenz und Allgegenwärtigkeit	154
4.4	Die Formbarkeit digitaler Artefakte: Der Überschuss an Zwecken	161
4.5	Die Machtposition digitaler Technologie und Technik.....	162
4.6	Verantwortungsdiffusion	166
4.7	Ausblick	171
5	<i>Schluss: Für eine prozedurale, proaktive und diskursive Technikethik</i>	173
6	<i>Literaturverzeichnis</i>	196

Abbildungsverzeichnis

Abbildung 1 Erste Bedingung eines normativ wünschenswerten Softwaredesigns: Modalitäten der Mensch-Technik-Welt-Bezogenheit nach Ihde, eigene Erweiterung. Eigene Abbildung	43
Abbildung 2 Zweite Bedingung eines normativ wünschenswerten Softwaredesigns	69
Abbildung 3 Resource-Product-Target-Modell mit Zuordnung der Bereichsethiken nach Florid (2008). Eigene Erweiterung.	84
Abbildung 4 Pyramide der Bereichsethiken Digitaler Technologien und Digitaler Technik. Eigene Abbildung.....	85
Abbildung 5 Typologie der Gesichtserkennungssysteme nach Technologie. Eigene Abbildung....	101
Abbildung 6 Gesichtserkennungssystem unterteilt in strukturelle und technik-generische Werte. Eigene Abbildung.....	106
Abbildung 7 Levels of Abstraction	124
Abbildung 8 Levels of Abstraction und Gradient of Abstraction am Beispiel ‚Wein‘	126
Abbildung 9 Die Entwicklung vom Werkzeug zum Automaten hinsichtlich der Unterstützung der physis hin zur phronesis. Eigene Abbildung.....	149
Abbildung 10 Werkzeuge als Mittel sowie digitale Technologie und Technik als Medium. Eigene Abbildung.....	151
Abbildung 11 Deontologische Demarkationslinie und digitale Artefakte. Eigene Abbildung.....	186

Tabellenverzeichnis

Tabelle 1 Technik und Technologie in Anlehnung an Ott (2005) und Janich (2001).....	31
Tabelle 2 Unterschiedliche normative Domänen digitaler Technologie und Technik	108
Tabelle 3 Erweiterte Zuordnung der normativ relevanten Domänen auf die idealtypischen technik- generischen oder strukturellen Analyseebenen	117

Übersicht

Teil I

Kapitel 1 diskutiert technikphilosophische Paradigmen. Erst eine fundierte Theorie ermöglicht eine Orientierung nach transparenten und opaken, technik-generischen und strukturellen Werten, die bestimmten Diskursebenen zugeordnet werden können. Zunächst wird die Entwicklungsarbeit mit Hilfe des VDI-Diskurses eingeführt. Zentral ist die Herausarbeitung der herstellenden Tätigkeit als Teil einer gemeinsam geteilten Handlungspraxis. Diese soll die Kompetenzen der Entwickler einzugrenzen helfen: Zum einen muss die technische Rationalität die Machbarkeitsbedingungen digitaler Technologie und Technik im Auge behalten. Zum anderen ist dies jedoch nicht ausreichend, sollen normativ wünschenswerte, digitale Artefakte entwickelt werden. Dies führt in eine starke Berufsethik mit Fokus auf einem Ethos, weshalb im zweiten Schritt die Nutzerperspektive mit Hilfe eines hermeneutischen Zugangs erörtert wird. Hierfür wird das pragmatisch-phänomenologische Paradigma eingeführt, um die Bedeutung von Technologie und Technik auf den Endnutzer zu analysieren. Die Auswirkungen werden im Hinblick auf Wahrnehmungs- und Wissensmodalitäten der Beziehung des Menschen zu seiner Welt erörtert. Hierbei wird offensichtlich, dass sich beide philosophischen Perspektiven auf der Ebene der Systemanforderungen zusammenfinden müssen.

Kapitel 2-4 führen in die *Theorie einer Strukturellen Rationalität* ein, da beide vorangehenden philosophischen Perspektiven die Sicht des Individuums darlegen, aber die Wechselwirkungen auf intersubjektive oder gar kollektive Lebensformen nicht aufgezeigt wurden. Die *Theorie Struktureller Rationalität* wird grundlegend sein, wenn es um die Ausarbeitung struktureller Werte geht, die für ein Verständnis der digitalen Transformation die notwendige theoretische Orientierung bieten. Sie fungiert daher als Scharnier zwischen Deskription und Normativität: Die Frage, welche Ethik die Lücke zwischen Technikphilosophie und normativer Orientierungslosigkeit schließen kann, blieb bisher ungeklärt. Es wurde lediglich die Eingrenzung des Problemraums erörtert. Jedoch lassen sich hieraus keine Kriterien ableiten, die einer normativen Priorisierung oder Evaluation dienen. Dies geschieht im folgenden Kapitel.

Kapitel 5 führt in technisch-ethische Überlegungen ein. Technikphilosophische Überlegungen bieten die notwendige epistemische Orientierung für eine normative Technikethik. So können auch empirische Fallbeispiele, welche die Technikethik seit den 1970ern bereichern, systematisch geordnet werden. Hierdurch kann eine Zuordnung von Best-Practice-Beispielen, Modellen oder Bereichsethiken zu bestimmten ethischen Problemfeldern erfolgen, da gerade digitale Technik aufgrund ihrer Kontextspezifikation nur in ihrer konkreten Anwendung zu bewerten ist. So zeigt sich hier, dass eine Bereichsethik um eine Erörterung des Sachgegenstands erweitert werden muss. Dies liegt zum einen darin begründet, dass digitale Artefakte in allen gesellschaftlichen Subsystemen zu finden sind. Zum

anderen müssen auch die technik-generischen Werte weiterhin auf ihre individuelle, intersubjektive und kollektive Erwünschtheit hin überprüft werden. Im Nachgang wird die humanistische Ethik diskutiert und die ethischen Paradigmen der Vernunft, Freiheit und Verantwortung, die eine ethische Bewertung ermöglichen sollen, werden eingeführt.

Teil II

Kapitel 1 führt anhand des Beispiels von Gesichtserkennungssystemen in der polizeilichen Fahnungsarbeit in die Themenstellung der Ethik der Digitalisierung, digitalen Transformation und der digitalen Artefakte ein. Es zeigt auf, dass digitale soziale Systeme aufgrund der vielfachen Verstrickungen transparent, d.h. nicht unmittelbar erkennbar, und opak, d.h. nicht unvermittelt nachvollziehbar, sind. Zudem wird verdeutlicht, dass technische Rationalität, die auf technische Funktionalität abzielt, nicht die Kriterien zu ermitteln vermag, die für eine ethische Betrachtung notwendig sind. Es wird bereits hier offensichtlich, dass die normative Unklarheit einer zielgerichteten Analyse normativ relevanter Sachverhalte bedarf, weshalb eine disclosive ethics, wie Brey sie fordert, notwendig wird. Die normative Strukturanalyse zeigt so am Beispiel der Gesichtserkennungssysteme, dass die Idealtypen „technik-generischer Wert“ und „struktureller Wert“ eingeführt werden müssen.

Kapitel 2 arbeitet die unterschiedlichen Wertetypen heraus, die eine normative Evaluation tragen sollen: So stehen technik-generische Werte strukturellen Werten gegenüber. Wichtig hierbei ist, dass die Wertetypen irreduzibel sind und jeweils die Bewertung ergänzen. Technik-generische Werte resultieren aus der Beschaffenheit der Technologie und Technik. Strukturelle Werte lassen sich aufgrund der Beschreibung der Technologie oder Technik als soziales System einfangen. Hier wird auch verdeutlicht, dass die technische Rationalität um weitere Deliberationsmodi ergänzt werden soll, aber ihre Berechtigung nicht verlieren darf.

Kapitel 3 betont zunächst die Notwendigkeit, die philosophische Analyse von den Reflexionsbegriffen *Transparenz* und *Opazität* leiten zu lassen, um die jeweils normativ relevanten Sachverhalte zum Vorschein zu bringen. Die Reflexionsbegriffe dienen der Schärfung des Verstandes dahingehend, sich dem nicht Unmittelbaren zu stellen.

Kapitel 4 führt in Luciano Floridis Methode der *levels of abstraction* ein. Es soll gezeigt werden, dass nicht nur die Identifikation normativ relevanter Sachverhalte notwendig ist, sondern auch verschiedene Diskursebenen zu bestimmen sind, um mit normativen Sachverhalten angemessen verfahren zu können.

Kapitel 5, 6 und 7 diskutieren die Begriffe der Digitalisierung, der digitalen Transformation und der digitalen Artefakte hinsichtlich einer normativen Analyse. Eine Ethik, die in der Softwareentwicklung

angewendet werden soll, muss eine praktische Anleitung geben können, weshalb sie einer lebensweltlichen Überprüfbarkeit standhalten können muss. Eine Ethik für Softwareentwickler muss sowohl das Objekt als auch Handlungsräume bestimmen, sollen Softwareentwickler nicht für die Gesamtheit aller Missstände, die technisch bzw. technologisch verursacht werden, die alleinige Verantwortung übernehmen oder als Sündenbock dienen. Welche Handlungsfelder und Entscheidungsräume können Softwareingenieuren zugetraut werden, ohne zugleich *moral distress* aufzubauen? Dies leitet über in die Diskussion dazu, in welche normativ relevanten Bereiche Softwareentwickler überhaupt handelnd eingreifen können. Hier wird die Unterscheidung in *implicit* und *explicit moral agents* eingeführt, um die Wertinhärenz von Software klar von moralisch agierenden Systemen trennen zu können.

Teil III

Kapitel 1-7 diskutieren die Einzigartigkeit von Software. Die als *uniqueness*-These bekannte Forderung nach einer eigenständigen Ethik der Informationstechnik verlangt nach einer Argumentation, weshalb klassische ethische Theorien nicht ausreichen, um sich normativ angemessen mit der Thematik der Digitalisierung, der digitalen Transformation sowie der digitalen Artefakte zu befassen. Die Wertinhärenz digitaler Technik, die These der *embedded values* verdeutlicht, dass Software und Softwaresysteme selbst Gegenstand einer normativen Beurteilung werden müssen, da ethisch relevante Situationen nicht erst im Umgang mit oder in der Nutzung digitaler Artefakte entstehen. Eine Ethik muss sich folglich über einen eigenständigen Gegenstandsbereich oder ein idiosynkratisches Vorgehen abgrenzen, die jeweils eine separate ethische Deliberation rechtfertigen. Es gilt somit, die Situationen einzugrenzen, die dezidiert aufgrund des informationstechnischen oder -technologischen Charakters auftreten. Eine selbstverantwortliche Ethik bedarf zunächst der Herausarbeitung der Aspekte, die als Unterscheidungsmerkmale dienen und dadurch den Charakter der Einzigartigkeit digitaler Technologien und Techniken hervorheben.

Resümierend soll der Versuch unternommen werden, der praktischen Ohnmacht, die durch die komplexe epistemische Unklarheit verstärkt wird, mittels einer prozeduralen, proaktiven und diskursiven Technikethik entgegenzuwirken. Humanistisch wünschenswerte digitale Artefakte benötigen eine humanistisch fundierte Entwicklermentalität. Eine Zusammenführung aller Teilethiken, wie in Teil I, Kapitel 6 diskutiert, kann nur dann gelingen, wenn bereits in der Herstellung die relationalen Wechselwirkungen der Einzelethiken hinsichtlich sowohl ihrer strukturellen als auch ihrer technik-generischen Werte reflektiert werden. Damit einer Forderung nach normativ wünschenswerten digitalen Artefakten nachgekommen werden kann, müssen die normativen Anforderungen, die sich aus der ethischen Deliberation ergeben, zugleich auf ihre technische Umsetzbarkeit gespiegelt werden: Dies bedeutet jedoch nicht, dass diese Normativität ausschließlich technisch lösbar sein muss. Für eine Ethik, die

in der Softwareentwicklung zur Anwendung kommen soll, ist es jedoch notwendig, den Handlungsraum auf die Machbarkeit hin zu überprüfen.

1 Einführung

„Wenn wir nicht mehr wissen, wie wir mit dem digitalisierten Leben zurechtkommen, brauchen wir gut fundierte ethische Forschung. Die aber kann man nicht auf Knopfdruck erzeugen. Denken braucht Zeit.“ (Capurro, 2017a, S. 188)

Immer häufiger sind wir von digitalen Geräten umgeben, die aufgrund ihrer Vernetztheit eine Welt erschaffen, die sich im Gegensatz zur gegenständlichen Welt, dem sogenannten *meatspace*, der unmittelbaren Wahrnehmung entzieht. Dieser virtuelle Ort, der als *cyberspace* bezeichnet wird, zeichnet sich im Kern durch die Verbundenheit vieler einzelner, Daten miteinander austauschender Geräte aus, obgleich es unterschiedliche Definitionen des Begriffs geben mag:

„In the cyber strategy of Germany: ‘Cyberspace includes all information infrastructures accessible via the Internet beyond all territorial boundaries and cyberspace is the virtual space of all IT systems linked at data level on a global scale.’ In Germany all players of social and economic life use the possibilities provided by cyberspace. As part of an increasingly interconnected world, the state, critical infrastructures, businesses, and citizens in Germany depend on the reliable functioning of information and communication technology and the Internet (Federal Ministry of the Interior, 2011, pp. 2, 14).“ (Lehto, Huhtinen & Jantunen, 2011, S. 6)

Dank digitaler Endgeräte wie Handys oder Notebooks können auch Endnutzer am Cyberspace aktiv oder passiv teilhaben. Selbst wenn die Geräte haptisch zugänglich sind, entziehen sich dem Endnutzer häufig die Beschaffenheit und die Anzahl der im Gebrauch stattfindenden Operationen, die den Zugang in die virtuelle Welt ermöglichen und diese strukturieren. Software ist somit überall und jederzeit zugegen, ohne dass die von ihr durchgeführten Operationen notwendigerweise bemerkt werden: So bleibt der Cyberspace durchsichtig. Der transparente Charakter des Cyberspace wird in dieser Arbeit anhand des Beispiels von Gesichtserkennungssystemen erörtert (Teil III, Kapitel 1). Zugleich kann hier die Symbiose aus technischer und normativer Urteilskraft dargestellt werden, die sich durch die gesamte Arbeit zieht: Nur, wenn ausreichend technisches Wissen vorhanden ist, können normative Konzepte eine angemessene Übersetzung in technische Terminologie finden. Zunächst müssen jedoch normative Sachverhalte im Technischen erkannt werden, um sie übersetzen zu können. Oft entzieht sich das Digitale der Wahrnehmung und bleibt so zuvorderst opak. Ein Verständnis und eine sich daran anschließende Bewertung normativ relevanter Sachverhalte bedarf gerade aufgrund der gläsernen und opaken Natur des Cyberspace einer Ethik der Offenlegung, einer sogenannten *disclosive ethics* – wie sie Philip Brey 2000 in seinem Aufsatz *Disclosive Computer Ethics* forderte –, die sich an den Reflexionsbegriffen der *Transparenz* und *Opazität* ausrichtet (Teil II, Kapitel 3). Ziel dieser

Arbeit ist es, *metatheoretisch* eine Verbindung zwischen *know-that* – Wissen, was normativ relevant in einem bestimmten Kontext ist – und *know-how* – Wissen, wie jene normative Kenntnis praktisch, hier vor allem technisch, anzuwenden ist – zu durchdenken, um einen *ethischen Ansatz in der Softwareentwicklung verankern zu können*.

Aufgrund der vielfältigen Einsatzmöglichkeiten ist es wenig verwunderlich, wie Eubanks (2012, S. 31) aufzählt, dass digitale Technologien und Techniken zum einen gefeiert werden, weil von ihnen die Reduktion sozialer Ungleichheiten erwartet wird, Demokratien belebt werden oder sie gar die Sehnsucht erfüllen, ein harmonisches „global village“ zu erschaffen (ebd.). Zum anderen ergibt sich aus der normativen Unklarheit und Komplexität, die aus der transparenten und opaken Natur des Sachgegenstands herrührt, energische Kritik, denn der Einsatz führt zum gläsernen Bürger, zu roboterartigen Menschen oder gar zu vermenschlichten Robotern (ebd.; Nida-Rümelin & Weidenfeld, 2018). Die Idee autonomer Systeme, die sich ohne menschliche Kontrolle eigenständig weiterentwickeln und über ihr Fortbestehen eigenständig zu entscheiden vermögen, prophezeit den Untergang des Menschen. Diese apokalyptische Vision nährt sich aus der Befürchtung, dass Maschinen die Weltherrschaft an sich reißen werden, wie es oft in Science-Fiction-Romanen erzählt wird (Nida-Rümelin & Weidenfeld, 2018). Dieses Science-Fiction-Szenario bedient sich der Singularitätsthese, wonach sich Softwaresysteme nach und nach selbst zu steuern vermögen, insofern sie ihre Aufgabenbewältigung auf alle möglichen Teilbereiche menschlicher Fähigkeiten ausdehnen und dadurch zu selbstbewussten Entitäten mit einer personalen Identität heranwachsen – der Entstehung einer Superintelligenz steht somit nichts mehr im Wege (Ramge, 2018).¹ Man muss jedoch gar nicht so weit gehen und sich eine Welt mit Maschinen mit Bewusstsein, einem freien Willen und einer eigenen Urteilskraft vorstellen, um Softwaresystemen mit Skepsis zu begegnen: Unwohlsein resultiert bereits aus der Entwicklung und dem Einsatz von Computerprogrammen, die gesellschaftliche Strukturen stark verändern und gesellschaftlich etablierte Entscheidungsstrukturen entkräften (Wolff, 2020).

Optimisten hingegen sehen in digitalen Technologien und Techniken eine Möglichkeit, mit technologischer Unterstützung oder gar durch eine mechanische Substituierung monotoner Tätigkeiten menschliche Freiheitsräume zu vergrößern, indem eine bessere Work-Life-Balance oder mehr sinn-

¹ Inwiefern Bewusstsein notwendig ist, um Problemlösungen auf unterschiedlichen Gebieten zu ermöglichen, möchte ich im Rahmen dieser Arbeit nicht diskutieren (bspw. ob ein Algorithmus sich selbständig weiterentwickelt, um sich eigenständig auch weiterer Problemlösungen anzunehmen). Die Einteilung, dass starke Künstliche Intelligenz mit einem Bewusstsein in Verbindung gebracht werden muss, erscheint nur im Rahmen der Annahme eines Epiphänomenalismus möglich. Da wir derzeit keine Maschinen mit Bewusstsein haben, erscheint dieses Szenario in seiner Brisanz weniger dringend, als die ethischen Fragestellungen bzgl. aktueller Softwaresysteme zu behandeln (Frankish & Ramsey, 2014).

stiftende Tätigkeiten ermöglicht werden können (Grunwald, 2018 a, Kapitel 3 oder McAfee & Brynjolfsson, 2017, Teil 1, Kapitel 2). Diese Sehnsucht nach einem gesellschaftlichen Zustand, in dem und durch den der Bürger wirtschaftlich abgesichert ist und sich somit mit aller Kraft und all seinen Ressourcen selbst verwirklichen kann, trägt die Ideen klassischer Gesellschaftsutopien eines Adam Smith oder John Stuart Mill weiter (Mill, 1859; Smith, 1759). Nicht nur das Telos des Maschineneinsatzes erscheint erstrebenswert, sondern auch der Arbeitsprozess der Systeme selbst verstärkt den Glauben an ein technologisch verursachtes Wunder. Aufgrund dessen, dass die Ausführungen von Softwaresystemen auf logisch-mathematischen Entscheidungskalkülen beruhen, gelten die Ergebnisse als objektiv und damit so, als ob sie ein unverzerrtes Urteil abbildeten. Problemlösungen, die das Resultat von Softwaresystemen sind, erscheinen somit sachlich und nüchtern und verdrängen die Notwendigkeit einer menschlichen und deshalb gerade subjektiven Interpretation der Situation. Endlich ist die menschliche Urteilskraft überwindbar. Die ihr anhaftende Subjektivität und Willkür scheinen technisch bezwingbar zu sein und so auch das Versprechen evidenzbasierter Beurteilungssysteme, wie z.B. das Softwaresystem, das bereits 2017 in den USA im Justizsystem eingesetzt wurde, um Prognosen über die Rückfälligkeit von Straftätern zu erstellen (Schimmeck, 2017).

Aber nicht nur die mathematisch-technisch errungene Objektivität suggeriert eine Dominanz der Softwaresysteme. Auch in ihrer Schnelligkeit und Perfektion hinsichtlich der Berechnung von Outputs und Datenklassifikationen überbieten sie die menschlichen kognitiven Fähigkeiten deutlich. So können Übersetzungssysteme wie bspw. *DeepL* schneller verschiedene Sätze in unterschiedliche Sprachen übertragen als es Menschen möglich ist. Zudem existieren bereits KI-Programme, die unter der Bezeichnung *Philosopher AI* firmieren, die nicht nur ein Wort vervollständigen oder kurze Satzeinheiten ergänzen, sondern vollständige Paragraphen prognostisch formulieren, ohne dass die Sätze bereits in dieser Weise im Internet zu finden wären. Es handelt sich hierbei um maschinell erstellte neue Wortfolgen, die so rasant publiziert werden, dass sie aufgrund ihrer Geschwindigkeit auffallen: So wurde auf *Askreddit* im Oktober 2020, einer Teil-Plattform der Kommunikationsplattform *Reddit*, ein Gedankenaustausch ausgetragen, der von einem der Beteiligten so schnell und provozierend entworfen wurde, dass die Frage aufkam, ob es sich hierbei tatsächlich um einen Menschen handelte oder um einen philosophischen Bot. Viele Personen, die mit dieser Maschine kommunizierten, bemerkten zunächst keinen Unterschied (Phillip, 2020). Aus der Möglichkeit, dass Maschinen den

Turing-Test² bestehen könnten, der Endnutzer also nicht mehr zwischen maschinellem oder menschlichem Output zu unterscheiden vermag, folgert Walsh (2018) bspw., dass solche Maschinen für die Endnutzer erkenntlich zu machen seien. Aber nicht nur die potentielle Ununterscheidbarkeit sorgt für ethisch bedenkenswerte Szenarien, sondern, wie Friedman und Kahn Jr. (1992) anmerken, auch sogenannte *decision supportive technologies* führen dazu, dass Personen maschinellen Output bevorzugt behandeln (Krügel, Ostermaier & Uhl, 2021). Dies kann zur Erosion des eigenen Selbstvertrauens führen, wenn bspw. Mediziner maschinell erzeugten Diagnosen mehr Vertrauen schenken als ihrem eigenen Fachurteil. Allgemein auf praktische Handlungsbelange bezogen, impliziert dies, dass Personen sich in ihren praktischen Werturteilen unsicher wähnen, was einer autonom begründeten Haltung entgegenwirkt bzw. diese schlichtweg unterminiert (Wallach & Allen, 2009, S. 40-42). Dieses auch als *automation bias* bezeichnete Vorurteil, maschinell produziertem Output mehr Vertrauen zu schenken, verdeutlicht, wie das Muster menschlicher Beziehungskultur oder eigenständiges Reflektieren des Menschen über sich selbst und seine Umwelt zunehmend maschinell durchdrungen werden. Dies mag auch Aufschluss über Hannah Arendts Betonung geben, dass eine humane Ausgestaltungsforderung nur an Maschinen, nicht aber an Geräte zu adressieren ist (Arendt, 1960; Teil II, Kapitel 4): Erstere sind in normativen Denk- und Verhaltensformen verankert und drohen gerade diese wesentlich menschlichen Fähigkeiten des Urteilens einzuschüchtern oder zu verdrängen (Rohbeck, 1993, S. 33). Geräten, wie einem Hammer oder einem Rechen, kann dies in diesem Ausmaß wahrlich nicht gelingen. Sie ersetzen vielmehr die *physis* des Menschen, lassen aber die *phronesis* unberührt (Teil I, Kapitel 1).

Es ist daher kaum erstaunlich, dass in allen gesellschaftlichen Bereichen die Rufe nach Verantwortungsübernahme immer lauter und vehementer werden, denn die Anwendung digitaler Technologien und Technik verändert etablierte Lebensweisen, indem auf tradierte Strukturzusammenhänge, sei es in der Arbeitswelt oder in der alltäglichen Kommunikation, eingewirkt wird (Kastendiek, 2003). Der Wunsch nach Maschinen, die ethisch wünschenswerte Resultate erzielen, beruht daher auf der fundamentalen Einsicht, dass es sich bei der digitalen Transformation um keinen „digitalen Determinismus“ handelt (Mainzer, 2019, S. 232), dem sich die Menschheit schicksalhaft ergeben muss, sondern um einen von menschlichen Akteuren bestimmten Gestaltungsprozess. Mit dieser Erkenntnis wird eindeutig, dass zum einen die lebensweltliche Gestaltung durch Technologie und Technik beeinflusst werden kann, um strukturelle Ausrichtungen mitbestimmen zu können. Zum anderen zeigt sich, dass

² Der Turing-Test ist ein vom englischen Mathematiker Alan Turing entworfenes Gedankenexperiment, in dem eine Maschine dann als menschlich gilt, wenn ein Mensch in einem Textchat nicht unterscheiden kann, ob er mit der Maschine oder einem echten Menschen spricht (Moor, 2003).

das technische Artefakt selbst modifizierbar bleibt. Fragen dahingehend, welche Akteure den technologischen Wandel bestimmen, wessen Werte und Interessen digitale Technik berücksichtigen soll und welche gesellschaftlichen sowie politischen Folgen aus deren Implementierung resultieren, werden so nicht nur für eine theoretische Abwägung, sondern auch praktisch relevant. Somit bleibt die digitale Transformation selbst normativ gestaltbar und steuerbar, insofern auf Bedingungen möglicher Strukturzusammenhänge technisch eingewirkt wird (Teil I, 2 und Teil II, 7). Normen, Werte oder Regeln können durch den Einsatz technischer Artefakte nachhaltig verändert und Sinnzusammenhänge, die als Orientierung dem menschlichen Handeln Sicherheiten bieten, neu akzentuiert oder gar substituiert werden (Kudina, 2019; Verbeek, 2016): Die Einführung des Walkmans ermöglichte bspw. den Konsum von Audioprodukten außerhalb der davor üblichen Umgebung.

Die Veränderung der gemeinsam geteilten Orientierungspunkte von Alltagsverhalten wird als *value dynamism* bezeichnet. Diese Transformation wird durch die Co-Evolution von Technologie und Gesellschaft getragen, worunter zumeist die gegenseitige, dialektische Bedingtheit von Technologie und Gesellschaft verstanden wird, die auf den Prozess der Gestaltung sowohl der Technologie und Technik als auch der Gesellschaft Einfluss nimmt (Rip & Kemp, 1998). Das technische Artefakt ist somit nicht isoliert begreifbar, sondern muss im Kontext der gesellschaftlichen Bedarfe reflektiert werden, durch die es erst entstehen kann und auf die es wieder zurückwirkt (Stalder, 2016). Dies zeigt sich auch anhand des Technik-Paradoxons, wonach jede Technik zwar den menschlichen Handlungsspielraum erweitert, indem etwas vorher Unerreichbares ermöglicht wird, aber diese Errungenschaft den Menschen zugleich zwingt, sich an bestimmte technische Anforderungen anzupassen (Grunwald, 2018a).

Technologie und Technik gilt es folglich auch als einen wesentlichen Teil der Kultur zu begreifen, damit der Herstellungsprozess, der Anwendungsbereich und das technische Artefakt als genuin menschliche Tätigkeit oder menschliches Erzeugnis verstanden werden können (Teil I). Technik oder Technologien lassen sich nicht auf naturwissenschaftliche Abläufe reduzieren oder gar rein aus einer naturwissenschaftlichen Perspektive verstehen: Im Gegenteil nimmt ein allumfassendes Verständnis von Technik die Idee einer rational hermeneutischen Analyse der Handlungszusammenhänge, der gegenseitigen Erwartungshaltungen und das Erlernen eben jener ernst, um moralisch wertvolle Technik zu gestalten (Teil I, Kapitel 1.2; Reijers & Coeckelbergh, 2020). Ein technisches Artefakt vermag nämlich aus sich selbst heraus weder ein Ziel anzustreben noch zu setzen, da es ohne Intentionalität und Semantik, ohne menschliche Interpretation seines Zwecks, kein Ziel verfolgt, sondern lediglich in seiner Kausalität aufgeht. Eine Rechenmaschine, die fehlerhaft addiert, ist physikalisch nicht falsch, schließlich generiert sie einen Output, aber verfehlt den ihr vom Menschen auferlegten Zweck der Addition. Wahrheit oder Falschheit maschinell erzeugter Outputs werden in Abhängigkeit einer funktional zu erfüllenden Aufgabe dem maschinellen Prozess vom Menschen erst zugeordnet: Wenn der

Output den Zweck nicht erfüllt, so kann man nicht auf ein fehlerhaftes System schließen, d.h. man kann physikalische Gesetzmäßigkeiten nicht als unzulässig behaupten, sondern muss die Argumentation auf den verfehlten Zweck stützen (Janich, 2003).

Durch die menschliche Zwecksetzung und die anschließende technische Umsetzung erfahren ökonomische und soziale Strukturen, Kontroll- und Machtverhältnisse sowie Infrastrukturen durch die Implementierung von Maschinen eine Umbewertung (Eubanks, 2012). Soziale Praktiken, welche die Lebensform der Menschen prägen und ihnen Orientierung bieten, werden neu ausgerichtet und Verhaltensbezüge erhalten neue Dimensionen. Nur, wenn der normative Anspruch, den Systeme aufgrund ihrer Einbettung in strukturelle Handlungszusammenhänge auf das menschliche Netz gegenseitiger Verpflichtungen ausüben, einer eigenen Prüfung unterzogen wird, kann über deren Angemessenheit oder Zumutbarkeit entschieden werden: Technische Systeme müssen in ihrer Kontextualisierung möglichst ex ante normativ austariert werden, wenn nicht nur im Nachhinein ethisch evaluiert und im Anschluss ex post reguliert werden soll (Teil I, Kapitel 2). Normative Deliberation umfasst folglich mehr als eine dichotome Gegenüberstellung von Werten, einer klassischen Abwägung von Trade-offs, einem Entweder-Oder, entweder einem Schutz der Privatsphäre oder dem Ziel der kollektiven Sicherheit. Vielmehr gilt es, Technologie und Technik als Teil der Lebensform zu begreifen, die wesentlich normative Anforderungen an ihre Nutzer, ihre Umwelt stellen, um hervorzuheben, wie Technologie und Technik das Miteinander, d.h. die menschlichen sozialen Beziehungen, beeinflussen und somit auch neue Lebensformen generieren. Das gilt selbstverständlich auch andersherum: Lebensformen generieren den Bedarf an Technik und Technologien. Alte Instrumente, alte Fähigkeiten erfahren eine neue Anwendung. Der Ruf ethisch guter Maschinen benötigt eine Computer- und Informationsethik, die sich von einer ex post bewertenden, spekulativen Technikfolgschätzung abgrenzen können muss, um eine Beziehung zwischen Mensch-Technik und Mensch-Handlung-Technik in ihrer sozialen Dimension moralisch evaluieren zu können (Teil I, Kapitel 6).

Spiekermanns (2021) Beispiel des Spracherkennungsassistenten zeigt die notwendige Unterscheidung von Werten in technik-generische und strukturelle sowie die Notwendigkeit einer klaren Systematisierung des Evaluationsprozesses auf. So reagiert der Sprachassistent auf die Aussage des Sprechers „Ich bin traurig“ in den USA mit „Ich umarme Dich“, während die russische Variante mit „Das Leben ist kein Ponyhof“ antwortet. Spiekermann folgert daraus, dass auf der einen Seite der Wert des „Trosts“ und auf der anderen Seite der Wert der „Disziplin“ Betonung findet, wobei beide Werte nicht in den Wertelisten der Industrie, der Verbände oder sonstiger Zusammenschlüsse zu finden sind (Spiekermann, 2021). Der Versuch, alle erdenklichen Werte aufzulisten, führt dabei in eine endlose Ansammlung durcheinander gewürfelter normativer Ansprüche. Trost und Disziplin sind deshalb nicht in den Verhaltenskodizes der Ingenieure zu finden, weil sie sich in den entsprechenden Erwartungshaltungen auffinden lassen und mittels einer zielgerichteten Strukturanalyse aufgedeckt

werden können. Die Verhaltenskodizes umfassen jedoch diejenigen normativ relevanten Aspekte, die sich direkt aus der technologischen oder technischen Beschaffenheit erklären lassen. Allerdings ist eine Strukturanalyse normativ gegenseitiger Verpflichtungen nicht ausreichend, wenn es um eine ethisch wünschenswerte Ausgestaltung digitaler technischer Artefakte geht. Wie Jobin et al. (2019) auswerteten, werden vor allem in Bezug auf digitale Artefakte fünf Werte verstärkt genannt: Transparenz, Fairness, Unbedenklichkeit, Verantwortlichkeit, Datenschutz. Diese Werte lassen sich aufgrund der technischen Beschaffenheit diskutieren, d.h. eine datengetriebene Technologie bedarf der besonderen Berücksichtigung bestimmter Werte, die sich nicht notwendigerweise mittels einer strukturellen Betrachtung von Lebensformen auffinden lassen, diese jedoch massiv tangieren können (Spiekermann, 2021). Spiekermanns Beschreibung unterstreicht somit die dichotome ethische Unterscheidung in technik-generische Werte, also normative Orientierungspunkte, die sich aus der technologischen oder technischen Beschaffenheit ergeben, und normativen Erwartungshaltungen, die einer strukturellen Analyse lebensweltlicher Wertedimensionen bedürfen (Teil II Kapitel 2.1; Teil I, Kapitel 2.1 und 6). Spiekermann selbst beschreibt letztere Verwobenheit als Kultur. In dieser Arbeit wird in diesem Zusammenhang jedoch nicht von Kultur, sondern von Lebensformen gesprochen, die durch gegenseitige Erwartungshaltungen getragen werden, die jeweils ihren Sollensanspruch aus unterschiedlichen Handlungsbegründungen ziehen, um den verschiedenartigen normativen Forderungen Rechnung tragen zu können (Nida-Rümelin, 2001, 2005a, 2020): Denn die Begründungen speisen sich aus Pflichten, die Personen durch soziale Rollen übernehmen, aus eingegangenen Verpflichtungen, ethischen Prinzipien, individuellen Rechten oder auch aus Eigeninteresse.

Erst eine Zusammenführung normativ relevanter Tatsachen, die sich als technik-generisch oder strukturalistisch bezeichnen lassen, lässt greifbar werden, was normativ gutes, normativ wünschenswertes humanistisch geprägtes Design ausmacht und was dieses zu leisten vermag. Hieraus können Kriterien sowohl für den Arbeitsprozess als auch für das Objekt selbst abgeleitet werden. Humanismus ist eine ethische Theorie, welche die menschliche Handlungsfähigkeit in den Mittelpunkt rückt. Dies erweckt anfänglich den Anschein, dass Objekte ethisch nicht evaluiert werden können, da sie weder über Intentionen noch über Freiheitsgrade verfügen. Ebenso entziehen sich, aufgrund ihrer Heteronomie, damit auch Handlungen einer ethischen Analyse, die maßgeblich von Technik und Technologie beeinflusst wurden. Dies ist auch der Grund, weshalb Verbeek (2011) für eine Erweiterung des Humanismus plädiert, um die Subjekt-Objekt Dichotomie aufzuheben und Technologien sowie Technik einer normativen Analyse zugänglich zu machen. Allerdings wird in dieser Arbeit nicht der Weg der Objektivierung eingeschlagen, sondern es werden gerade die Autonomie und die autonomiesichernde Sphäre in den Mittelpunkt gerückt, um ethische Kriterien einzuführen, die für ein humanistisches Design ausschlaggebend sind (Teil I, Kapitel 6.2). Zudem wird die moralische Entschei-

dungsfreiheit über die Beziehung von herstellenden Tätigkeiten (*poiesis*) und wünschenswerten kollektiven Routinen (*praxis*) betont, sofern es die Ausgestaltung wünschenswerter normativer digitaler Artefakte betrifft (Teil I, Kapitel 3 und 4). Erst diese Verknüpfung lässt die Herstellung und Entwicklung digitaler Artefakte, aber auch deren Nutzung, als Ausdruck einer *moral agency* denken. Dies führt abschließend dazu, die Wertinhärenz digitaler Technologien und Technik auszuarbeiten, denn diese bieten eine Grundlage für eine Ethik, die in der Softwareentwicklung Anwendung finden kann. Die Einzigartigkeit von Software lässt eine eigenständige Ethik vermuten, die sich aus dem Sachgegenstand ergibt (Teil II, Kapitel 7.2): Der mediale Charakter der Informationstechnologie und -technik, das fehlende Ethos, die totale Vernetzung und Transparenz, die Formbarkeit und Machtformigkeit sowie die Verantwortungsdiffusion können als Alleinstellungsmerkmale genannt werden, die eine genuine Softwareethik verlangen.

Die Frage danach, was ein humanistisch ausgerichtetes Design ist, kann daher nicht allein über die Idee des Guten des technischen Artefakts beantwortet werden, denn sowohl der Wertepluralismus als auch die Variabilität technischer Artefakte und ihre Schnelllebigkeit verhindern eine hinreichende Bestimmung der ethischen Ausgestaltung technischer Artefakte (Gogoll et al., 2021). Somit wird die Frage lauten müssen, wie die Bedingungen der Möglichkeit der Verwirklichung moralischer Maschinen tatsächlich aussehen können (Teil I, Kapitel 1.1). Für eine Ethik impliziert dies zudem die Frage danach, ob man diese prozedural begründen kann, ohne dabei einer normativen Komponente verlustig zu werden (Teil I, Kapitel 6). Schlussendlich gilt es nämlich der technischen Unsicherheit hinsichtlich normativer Umsetzungsmöglichkeiten durch ethisch wünschenswerte Praxen zu begegnen. Nur durch eine Verbindung von einem autonomie-sichernden Arbeitsprozess und einer Kenntnis normativ relevanter Aspekte können ethisch wünschenswerte digitale Technologien und Techniken entwickelt werden, denn es gilt – um an dieser Stelle mit Lenk (2018, S. 25) zu enden –, „durch Denken und Handeln dazu beizutragen, unsere Erdenwelt zu erhalten, zu schonen, die Veränderungsdynamik mit Vernunft zu steuern und die soziale Welt – ewige Aufgabe! – humaner zu gestalten. *Humanitas praestet!* Humanität stehe voran!“

Zusammenfassung und Ausblick:

- 1.) Normativität reduziert sich auf keine reine Güterabwägung, sondern muss zwingend Handlungsspielräume bzw. Verhaltensanpassungen und deren Kontextualisierung betrachten, die aufgrund der Ausgestaltung des Artefakts entstehen oder transformiert werden.
- 2.) Das humanistische Ideal der Gewährleistung von Autonomie durch Begründungsstrukturen (Rationalitäten) muss auf drei Ebenen – individuell, sozial und kollektiv – gewahrt und bewahrt werden. Diesem Ideal gilt es sich prozedural anzunähern, was
- 3.) bedeutet, den Herstellungsprozess zu analysieren, um zu verstehen, welche Kompetenzen vonnöten sein werden.
- 4.) Dabei dürfen technik-generische Werte nicht vernachlässigt werden, die aufgrund der technischen Beschaffenheit entstehen.

2 Technikphilosophische Grundlagen

Es stellt sich im Zusammenhang mit der Identifikation und angemessenen Adressierung normativ relevanter Aspekte die Frage, welche technikphilosophischen Grundlagen vonnöten sind, um mit digitalen Technologien und Techniken normativ adäquat verfahren zu können. Die Suche nach den passenden Deutungsparadigmen verweist bereits darauf, dass im Rahmen der Technikphilosophie nicht nur Fragen der ethisch wünschenswerten oder zumutbaren Konsequenzen diskutiert werden, sondern dass auch die sich gegenseitig strukturell bedingenden Abhängigkeiten der Technologie in den Mittelpunkt einer philosophischen Betrachtung rücken (Kogge, 2008). Von daher müssen zunächst diejenigen Merkmale einer Technikphilosophie erörtert werden, welche die idiosynkratischen Gesichtspunkte der digitalen Technologie und Technik angemessen zu artikulieren vermögen. Diese für Technologie und Technik einzigartigen Eigenschaften müssen zudem als Fundament einer Ethik fruchtbar gemacht werden können. Die technikphilosophische Herangehensweise muss für eine epistemische Verortung normativer Fragestellungen hinsichtlich der Entwicklung von ethisch angemessener Software zielführend sein: Sie muss also nicht nur metatheoretisch anspruchsvoll sein, sondern auch einen methodischen Zugang für die Identifikation normativ relevanter Sachverhalte gewährleisten. Dies kann gelingen, wenn zum einen der Herstellungsprozess als Paradigma eingeführt und zum anderen die Nutzung ebenjener Geräte berücksichtigt wird. Sowohl den Entwickler als auch den Nutzer gilt es zu adressieren: Die Nutzerperspektive dient der Schärfung der Wahrnehmung der Opazitäten und Transparenzen, die sich strukturell, aber auch technik-generisch bedingen. Sollen wünschenswerte digitale Artefakte produziert werden, muss der Softwareentwickler die aufgedeckten Aspekte während des Entwicklungsprozesses berücksichtigen können. Dies bedarf nicht nur einer instrumentell-technisch begabten Vernunft, sondern das zu erschaffende Artefakt muss auch im Rekurs auf normative Denkmodi erfasst werden. Die *Theorie der Strukturellen Rationalität* führt beide Perspektiven zusammen und reflektiert sie auf eine kollektiv wünschenswerte Praxis hin.

2.1 Zwei Denkschulen der Technikphilosophie

Carl Mitchams Unterscheidung technikphilosophischer Theorien unterstützt die Ausarbeitung technikphilosophischer Paradigmen. Seine Klassifikation technikphilosophischer Denkschulen in solche, die er als *engineering-philosophy of technology* oder als *humanities philosophy of technology* bezeichnet (Mitcham, 1994), ist hilfreich, um zwei wesentliche Perspektiven auszuarbeiten, die digital-technologisch bedeutsam sind: Die aus den technischen Wissenschaften stammende Philosophie zeigt die Entwicklerperspektive auf, während die geisteswissenschaftlich begründete Philosophie die Bedeutung von Technik für den Nutzer oder für die Gesellschaft erörtert. Beide Perspektiven sind notwendig, um die normativ relevanten Sachverhalte zu verorten und um eine Bewertung in ihren

wechselseitigen Bezügen zu ermöglichen (Teil I, Kapitel 2). Im Folgenden werden beide Ansätze kurz skizziert, um sie dann auf die Fragestellung der Ausgestaltung digitaler Technologien und Techniken hin zu durchdenken.

Die aus der Ingenieursperspektive, also der *engineering-philosophy of technology*, entwickelte Technikphilosophie nimmt technologische Prinzipien zum Ausgangspunkt aller weiteren philosophischen Reflektionen, um mit Hilfe jener Prinzipien sodann alle lebensweltlichen Bereiche zu erschließen oder gar zu erklären. Dies führt in die Ideologie des *solutionism* – der Idee, dass alle Situationen wegen ihrer technologischen Beschreibbarkeit als berechenbar oder optimierbar aufgefasst werden können. Dadurch kann auf alle lebensweltlichen Situationen technisch angemessen – im Sinne einer Konfliktlösung oder Prozessoptimierung – reagiert werden:

„Technological solutionism is an ideology in which social phenomena are delineated as discrete problems that can be ameliorated, or as discrete processes that can be optimised, with an appropriately designed technological intervention.” (Gardner & Warren, 2019, S. 364)

Der normative Imperativ, dass diese Phänomene auch technisch gelöst werden sollen, wird aufgrund einer solchen Einstellung häufig vorschnell gezogen (Lenk & Ropohl, 1987, S. 5-22). Die Idee der technischen Lösbarkeit sozialer Phänomene verstärkt den Glauben an deren Steuerbarkeit mit Hilfe von Technologie und Technik. Letztlich läuft dies in seiner Extremform in technokratische Ordnungsvorgaben, in denen alle menschlichen Belange einzig mittels technischen Denkens gelöst werden (Rapp, 1978): Die Denkart, die sich in der Technik und Technologie findet, wird auf alle Entscheidungsbereiche angewendet, gleichgültig, ob auch weitere Rationalitäten zur Lösung beitragen. Aufgrund dieser Herangehensweise betont Mitcham (1994, S. 62), dass Technik und Technologie im Rahmen der *engineering-philosophy of technology* gar nicht als Reflexionsobjekte ausreichend beachtet würden, sondern dass es rein um die Anwendung technischer Prinzipien und Bedingungen gehe:

„Engineering philosophy of technology might even be termed a technological philosophy, one that uses technological criteria and paradigms to question and to judge other aspects of human affairs and thus deepen or extend technological consciousness.”

Dahingegen, so Mitcham, bedienen sich die *humanities philosophy of technology* geisteswissenschaftlicher Terminologie und nutzen philosophische Theorien, um Dimensionen der Technologie sowie Technik jenseits ihrer naturwissenschaftlichen Beschreibbarkeit einfangen zu können. Nicht mehr die angemessene Anwendung technischer Prinzipien wird im Rahmen dieser philosophischen Perspektive erörtert, sondern die Herausarbeitung der transtechnischen Verhältnisse und Bezüge gerät in den Mittelpunkt der Betrachtung (Mitcham, 1994): Es geht hier vielmehr um Verständnisfragen, die

durch die Offenlegung von Bedeutungskontexten erörtert werden und weniger um die Erklärung kausaler Zusammenhänge vorgefundener Phänomene. Die Herausarbeitung der Bedeutung erfolgt bspw. bei Martin Heidegger ontologisch, während die Frankfurter Schule vornehmlich die technische Rationalität kritisch hinterfragt, um Technik und Technologie in den geteilten Bedeutungshorizont systematisch integrieren zu können (Marcuse, 1996; Orth, 1987). Die Sinnhaftigkeit von Technik und Technologie lässt zugleich die Suche nach Narrativen vermuten (Reijers & Coecklbergh, 2020), was zu einer weiteren Trennung der beiden technikphilosophischen Denkschulen führen kann, sofern ein logisch-rationales Denken auf eine metaphysisch anmutende Erzählung trifft. Die hermeneutische Herangehensweise erörtert somit vor allem auf einer Metaebene die strukturellen Bezüge von Technologie und Technik auf nicht technische Bereiche – wie die Auswirkungen auf die Struktur der persönlichen Identität oder gesellschaftliche Subsysteme oder gar gesellschaftliche Ordnungsrahmen –, um Aspekte diskutierbar zu machen, die sich jenseits von Kausalitätsbezügen und Optimierungsgedanken verfestigen oder sich hinter Nützlichkeitskriterien verbergen.

Ogleich es Versuche gibt, beide technikphilosophischen Ansätze zu vereinen – wie bspw. der Diskurs des Verbands Deutscher Ingenieure (kurz: VDI-Diskurs) oder der amerikanische Pragmatismus z.B. eines Don Ihde (Mitcham, 1994) zeigen –, soll im Zuge dieser Arbeit eine Verknüpfung gelingen, die sich als fruchtbar für eine Ethik erweist. Denn eine Verbindung beider Denkschulen verschafft die Möglichkeit, Design, die bewusste Ausgestaltung des technischen Artefakts, humanistisch auszurichten, was bereits auf eine Versöhnung beider Denkschulen auf der Ebene der Formulierung von Systemanforderungen verweist. Während der VDI-Diskurs aufgrund seiner Nähe zur Entwicklungsarbeit hilfreich ist, analysiert der hermeneutisch-phänomenologische Pragmatismus die Auswirkungen digitaler Technologie und Technik auf die Wahrnehmungs- oder Wissenshorizonte der Endnutzer. Beide Ansätze bleiben jedoch individuell, weshalb eine Zusammenführung im Rahmen von Nida-Rümelins *Theorie einer Strukturellen Rationalität* erforderlich wird: Hier werden die strukturellen Interdependenzen zugänglich und es wird ihre normative Dimension ersichtlich.

2.1.1 Der VDI-Diskurs: Der Entwicklungsprozess

Die technikphilosophische Schule, welche unter dem Namen *VDI-Diskurs* bekannt ist, wurde im Nachgang an den Zweiten Weltkrieg in Deutschland gegründet, um die deutschen Ingenieure unter dem Gebot „Mensch und Technik“ zusammenzuführen. Hierbei ging es vornehmlich um die Thematisierung der moralischen Verantwortung des Ingenieursberufs (Mitcham, 1994; Kogge, 2008). Der VDI-Diskurs schafft somit eine technikphilosophische Grundlage, die eine normative Stellungnahme für die Entwicklung ethisch wünschenswerter digitaler Artefakte zu artikulieren ermöglicht. Dadurch rückt der technikphilosophische Ansatz den Handlungscharakter der Technologie- und Technikgenese in

das Zentrum einer philosophischen Argumentation: Als Ausgangspunkt der Reflexion wird nun die beabsichtigte Entwicklung technischer Artefakte oder die Entstehung von Technologie gesetzt. Als Gründerväter dieses VDI-Diskurses sind Alois Huning, Hans Lenk, Simon Moser, Friedrich Rapp, Günter Ropohl, Hans Sachsse, Klaus Tuchel, Walter Christoph Zimmerli, Gerhard Banse und Klaus Kornwachs genannt, die allesamt nicht nur eine Ingenieursausbildung vorwiesen, sondern auch über fundiertes philosophisches Fachwissen verfügten (Mitcham, 1994). Sie entwickelten eine Technikphilosophie, die sich über das Zweck-Mittel-Schema entfaltet, um die Rationalität der Entwicklungsarbeit herausarbeiten zu können. Nur so wird die Entwicklungsarbeit als planbare Handlung, eingebettet in Handlungsstrukturen, begreifbar. Im Zentrum steht somit ein Technik- und Technologiebegriff, der erst durch zweckrationales Denken und Handeln überhaupt der Vernunft zugänglich gemacht und somit auch vernünftig ausgestaltet werden kann.

Eine Handlung wird hierbei als intentionales Verhalten verstanden, das aufgrund seiner Absichtlichkeit kontrollierbar wird (Nida-Rümelin, 2020). Kontrollierte Handlungen werden, Nida-Rümelin folgend, wiederum als von Gründen geleitet erfasst, weshalb auch der handelnden Person Verantwortung für ihre Handlung zugeschrieben werden kann, da der Handelnde seine Handlungen zu begründen weiß. Die Person ist somit im Bilde darüber, weshalb sie etwas tut: Sie kann Stellung dazu beziehen, ob sie die Handlung für richtig oder falsch hält (ebd.). In Handlungen vollziehen sich somit Entscheidungen oder in den Folgen von Handlungen offenbaren sich Motivationen. Beides sind intentionale Zustände, die sich wiederum auf die Fähigkeit des Gründeanföhrens und des Gründeanehmens zurückföhren lassen. Der Zweck einer Handlung oder besser der Handlungsgrund ist vom Handlungsziel, das einen zweckzutraglichen Handlungsabschnitt charakterisiert, zu unterscheiden, das eingehalten wird, um den Grund zu erreichen: So pflanzt jemand bspw. einen Baum (Handlungsziel), um einen Beitrag zum Umweltschutz zu leisten (Handlungsgrund). Durch Gründe werden folglich ganze Handlungsketten strukturiert und über einen längeren Zeitraum planbar, durchföhrrbar und für andere verständlich. Erst durch Gründe werden Handlungen vernünftig und Kooperationen denkbar (Teil I, Kapitel 2). Eine einzelne Handlung fügt sich somit in ein Struktur Ganzes ein, das die Handlung aufgrund ihrer Teilhabe an der Struktur nicht nur punktuell vernünftig werden lässt. Dies ist, so Nida-Rümelin, nun humanistisch auszulegen, insofern gerade die Vernunft Überzeugungen, Geföhle und Handlungen anzuleiten vermag, also über die Triebhaftigkeit und die Interessen des Einzelnen zu transzendieren imstande ist. Die Fähigkeit, von sich selbst Abstand nehmen zu können, ist gerade diejenige, die den Menschen von anderen Lebewesen unterscheiden soll. Erst durch die Ausübung der Vernunft kommt dem menschlichen Verhalten eine Intentionalität zu, die sich in einer Regelmäßigkeit zeitigt. Allerdings bleibt dabei unklar, welche Gründe erstrebenswert und welche es nicht sind. Eine humanistische Ethik muss sich deshalb des Gründenehmens und Gründegebens annehmen,

denn es erfordert individuelle und kollektive, d.h. gemeinsame Deliberation, um Konfliktfälle, normative Unklarheiten, den Umgang mit Dilemmata oder Metakriterien für öffentliche Verfahren auf einer rationalen Ebene zu erarbeiten, ohne dass dabei die gefällte Entscheidung mit Gewissheit die richtige ist. Nicht immer ist dieser Prozess als ein bewusster Vorgang zu verstehen, sondern begründete Handlungen können bspw. auch auf tradierten Gründen beruhen. Nida-Rümelin legt dabei die Gütequalität pragmatistisch aus, denn die Gründe müssen sich in der Lebenswelt durchsetzen und bewähren können. Die Vernunft deckt somit die Handlungsgründe auf und verhilft ihnen dazu, sich in der Lebenswelt qua Entscheidung und Motivation zu verwirklichen.

Im VDI-Diskurs wird nun der Handlungsbegriff, der auf einer autonomiesichernden Freiheitskonzeption basiert, auf herstellende Tätigkeiten übertragen. Handlungen sind genau deshalb autonom, weil von der Unmittelbarkeit menschlichen Tatendrangs Abstand genommen werden kann, um Verhalten zu steuern und begründbar zu machen, d.h. dem konzeptionellen Raum der Gründe zugänglich zu machen. Es ist nunmehr die herstellende Tätigkeit, die *poiesis* auszuarbeiten, um überhaupt vernünftige herstellende Tätigkeiten und Produkte, die dieser entspringen können, zu ermöglichen. Dies wiederum erfordert eine rationale Darstellung des Entwicklungsprozesses und dessen Einbettung in komplexe soziale Systeme. Erst der vernünftige Prozess ermöglicht, auch über ein vernünftiges Design technischer Artefakte nachzudenken. Zugleich verdeutlicht dies die Vorstellung, dass ein technisches Artefakt nicht isoliert, sondern wiederum in gesellschaftliche Strukturen oder besser in normative Praxen etabliert werden soll: Technologie und Technik sind von daher als soziale Systeme zu behandeln. Für den Entwickler ergeben sich nach Kogge (2010, S. 945) demnach zwei wesentliche Aspekte:

Erstens: Technik wird in der Hersteller-Perspektive betrachtet. Zuvorderst geht es um die Frage, wie technologische Artefakte zu Stande kommen, etwa, wie technische Innovation entsteht; sodann um eine Bewertung dieser Artefakte im Sinne einer Ingenieursethik, also im Sinne der Regeln, die der Ingenieur bei der Herstellung beachten soll.

Zweitens wird in diesem Diskurs der systemische Aspekt der modernen Technik herausgestrichen. Technische Entwicklung – so die Auffassung – tendiert zu immer komplexeren Systemen, die nicht nur die Sach-, sondern auch die Sozialdimension der Technik umfassen.

Das Regelwerk, das der Entwickler bei seiner Arbeit zu beachten hat, spricht für eine Berufsethik, die auch der VDI-Diskurs durch seinen Ansatz und in seiner Verbundtätigkeit anstrebt: Es gilt, das Ethos eines verantwortungsbewussten Ingenieurs zu etablieren, der, in Analogie zum ehrbaren Kaufmann, sich in seinen Entscheidungen nicht nur auf das technisch-ausgerichtete Deliberationsprinzip stützt (Nida-Rümelin, 2011a, Teil II). Zudem wird durch den VDI-Diskurs eine Diskussion dahingehend ermöglicht, wie diese Werthaltungen in einen Arbeitsprozess implementiert werden können, der sich zunächst auf den Zweck der technischen Funktionalität gründet. Dabei wird nicht nur der Entwicklungsaspekt betont, sondern auch der Lebenszyklus des Produkts (VDI, 2000).

Die Notwendigkeit, Technik auch als soziales System zu verstehen, zeigt sich auf individueller Ebene vor allem im phänomenologischen Pragmatismus: Dort werden die Erweiterungen oder Beeinträchtigungen betrachtet, die Technik auf die jeweils individuellen Wahrnehmungs- und Wissenshorizonte ausüben kann (Teil I, Kapitel 1.2). Hierdurch können kontextspezifische Anforderungen formuliert werden, die sich in einer technik-generischen und strukturellen Betrachtung zeigen: Wie verändert sich die Wahrnehmung der Mit-Welt des Einzelnen durch den Einsatz der Technik? Die normative Frage nach der Erwünschbarkeit der Veränderung muss dann in Anschluss ethisch erörtert werden.

Die individuelle Perspektive wird durch Nida-Rümelins *Strukturelle Rationalität* erweitert, da durch die Eröffnung von begründeten Lebensformen auch dem Miteinander seine Berechtigung zugesprochen wird. Hier geht es verstärkt darum, den Einsatz der Technik im Hinblick auf die normativen Eckpfeiler, d.h. Konventionen, Gesetze, Normen, moralische Gebote oder eben die entsprechenden Erwartungshaltungen, zu überprüfen (Teil I, Kapitel 3).

Diese philosophischen Paradigmen gilt es nun schrittweise herauszuarbeiten, um ein Verständnis normativ erstrebenswerter Artefakte entwickeln zu können, denn das angemessene digitale Artefakt muss sich individuell, intersubjektiv und kollektiv bewähren können – oder zumindest müssen diese Formen des Lebens berücksichtigt werden.

So verändert sich die individuelle Wahrnehmung bspw. durch den Einsatz von Social Media. Die Bildung fragmentierter Meinungsräume und Filterblasen, die – obgleich sie nur eine Teilrealität darstellen – suggerieren – häufig vermittelt durch die graphische Oberflächenausgestaltung — eine vollständige Realitätsabbildung und wirken dadurch direkt auf die Wahrnehmung ihrer Nutzer ein (Nechushtai & Lewis, 2019). Das Design der Suchmaschinen, der sozialen Plattformen oder Blogs kann zu einer verzerrten Wahrnehmung der Mit-Welt führen (Amrollahi, 2021; Reviglio, 2017). Zugleich verändern sich dadurch intersubjektive Relationen, da z.B. *hate speeches* aufgrund der anonymen Gesprächskultur erleichtert werden und von Gleichgesinnten wiederum Unterstützung finden (Brown, 2018; Mondal et al., 2018). Kollektive Lebensräume, in denen eine pluralistische Debattenkultur prägend ist, werden durch den Einsatz eines solchen Mediums maßgeblich beeinträchtigt, wie es die Diskussion um die digitale Transformation demokratischer Deliberationskultur zeigt (Sunstein, 2018; Vaidhyanathan, 2018). Auf diese intra- und interpersonellen Bezüge kann mittels des Designs eingewirkt werden, wenn auch nicht alle negativen Folgen allein durch eine Veränderung des Designs vermieden werden können.

Bemühungen dahingehend, sich über die Ausgestaltung digitaler Artefakte Gedanken zu machen, bevor sie auf den Markt getragen werden, bedeutet, sich nicht nur über negative Folgenvermeidung Gedanken zu machen, sondern zugleich auch positive Haltungen zu fördern. Dies heißt jedoch nicht, dass normativ erstrebenswerte digitale Artefakte ausschließlich diejenigen sind, die offensichtlich für

einen guten Zweck entwickelt wurden, wie bspw. digitale Artefakte, die benachteiligten Personen zu mehr Autonomie verhelfen sollen, wie bspw. die von Microsoft entwickelte App, die Personen mit Seheinschränkungen dazu verhelfen soll, ihre Umwelt besser wahrzunehmen (Blinden- und Sehbehindertenverein Westfalen, 2016).

Damit die ethische Deliberation die technische Denkart bereichert und nicht sanktionierend wirkt, darf die herstellende Tätigkeit, die Entwicklungsarbeit selbst, nicht mehr nur der technischen Funktionalität oder einem Nutzendenken folgen. Dies soll im nächsten Schritt dargelegt werden.

2.1.1.1 Die herstellende Tätigkeit: Poiesis

Die aristotelische Einteilung von Aktivitäten in solche der *poiesis* und *praxis* oder gar in reine Bewegungsabläufe, der *kinesis*, ermöglicht zunächst eine Unterscheidung in Verhalten, herstellende Aktivität und Handlung. Diesen liegen unterschiedliche Rationalitäten zugrunde, die sich jedoch nicht ausschließen, sondern zusammenführend für die Bestimmung der notwendigen Kompetenzen zur Gestaltung von guten normativen digitalen Artefakten geeignet sind.

Allen drei Abläufen ist zunächst gemeinsam, dass sie strebsam sind, dass sie durch etwas angetrieben werden und auf etwas gerichtet sind. Kinetisches Verhalten umfasst bspw. die Muskelbewegungen, die kausalen Vorgänge, derer es bedarf, um die Körperlichkeit ausleben zu können. Die kinetische Beschreibungsebene erklärt diejenigen Verhaltenszusammenhänge, die mittels der Kategorie von Ursache und Wirkung erfasst werden können. Dem gegenübergestellt sind *poiesis* und *praxis* teleologisch zu fassen, also auf ein Ziel gerichtet, das nicht mehr kausal beschreibbar ist, sondern begründet sein muss durch einen Zweck, der für eine Zielsetzung konstitutiv ist. Dieser Zweck entzieht sich einer kausalen Erklärung bzw. diese Erklärung reicht nicht aus, um den Handlungskontext erfassbar zu machen. Die Gemeinsamkeit der Zielverfolgung wird zugleich jedoch daraufhin unterschieden, dass das Ziel in seiner ontologischen Verfasstheit ein anderes ist (Rese, 2013). *Praxis* ist dadurch charakterisiert, dass die Handlung selbst das Ziel ist, während *poiesis* als herstellende Tätigkeit sich in einem Werk (*ergon*) vollendet, „deshalb könnte man im Hinblick auf das Handeln auch von intrinsischen Zielen und im Hinblick auf das Herstellen von extrinsischen Zielen sprechen“ (Rese, 2013, S. 188). Das *Telos*, das Woraufhin, welches eine Aktivität bedingt, bedarf jedoch sowohl bei herstellenden Aktivitäten als auch bei praktischen Handlungen eines Zweckes, der die Ziele erst begründet: Der Zweck kann somit als der Grund dafür gelten, weshalb das Ziel gewählt wurde und worauf die Aktivitäten bzw. Handlungen abgerichtet sind. Die Abgrenzung der *poiesis* gegenüber der *kinesis* als Aspekt einer Tätigkeit ist nicht einfach, denn ohne *kinesis*, ohne natürliche Bewegungsabläufe, kann es keine *poiesis* geben. Sowohl das Wissen über die physikalischen Gesetzmäßigkeiten

als auch die Beherrschung der eigenen Bewegungsabläufe sind zentrale Voraussetzungen und Fähigkeiten, die der Herstellung eines Produkts dienen.³ Zugleich bedarf es Praxen, um herstellende Tätigkeiten zu erlernen und anzuwenden: in Form von Bildungseinrichtungen oder von arbeitsorganisatorischen Zusammenschlüssen. Diese Handlungsweisen gilt es stetig zu üben, um durch Erfahrungswissen und Geschicklichkeit das Beherrschen rationaler Abläufe zu beweisen. Erst eine Wiederholung poietischer Tätigkeiten führt zur Kompetenzbildung: Bestimmte Arbeitsaufgaben müssen auf eine angemessene Art und Weise erfolgen.

Die *poiesis* modelliert natürliche Fähigkeiten durch Übung zu einer Gewohnheit. Diese Gewöhnung muss in einem sozialen Kontext geschehen. Dieser Aspekt der sozialen Einbettung des Erwerbs von Kunstfertigkeiten vertieft sich zudem durch praktisches Handeln. Selbst die Fertigung eines Instruments bedarf anderer Personen: Zum einen durch die Ausbildung, zum anderen ist meist auch der Herstellungsprozess als solcher arbeitsteilig organisiert und deshalb auch unter dem Aspekt der Praxis, ja sogar unter bestimmten Aspekten der moralischen Praxis, zu bedenken. Zudem verhilft dies auch dazu, Technologie und Technik als Zusammenschluss von Einzeltätigkeiten oder Fähigkeiten zu denken, die nur durch einen gemeinsamen Zweck oder eine gemeinsame Funktionalität als eine spezifische Technologie und Technik greifbar werden. Technologie und Technik bestimmen sich somit nicht als ein Klassifikationsbegriff bestimmter Objekte, sondern über den Zusammenschluss bestimmter Fähigkeiten und Fertigkeiten (Brey, 2012a).

Daraus folgert Janich (2003, S. 93): „(P)hilosophisch betrachtet geht es um Technik (...) letztlich immer um die Beherrschung von Mitteln nach Zwecken, seien diese Mittel nun Handlungsschemata (wie beim Pianisten), Verfahren (wie bei der Gusstechnik) oder Produkte und Folgen von Handlungen (wie bei Ingenieur-Konstruktionen)“. Janich unterscheidet hier nicht trennscharf zwischen Tätigkeit als herstellender Aktivität und einer praktischen Handlung, was jedoch für eine systematische Annäherung an ethische Anforderungen, die sowohl für digitale Artefakte als auch für herstellende Prozesse gelten sollen, wesentlich ist. Nur so können herstellende Tätigkeiten in Abhängigkeit bestimmter Kompetenzen verstanden werden, die es zu erlernen gilt, wenn normativ zumutbare Artefakte hergestellt, eingesetzt und genutzt werden sollen. Allerdings verweist Janichs Dreiteilung auf interessante Rationalitätstheoretische Aspekte, die vor allem bildungspraktisch relevant sind: Jemand bringt ein Werk mithilfe erlernter Mittel hervor, die sich als methodologisierte Abfolge erfassen lassen –

³ Die Unterscheidung bzw. Reduktion von Herstellung und Handeln wird in der Aristoteles-Rezeption kontrovers diskutiert. Für diese Arbeit ist eine systematische Trennung ausschlaggebend (siehe hierzu Arendt 1960; Rese 2003; Janich 2003). Die Diskussion rankt sich um die These, ob Aristoteles bspw. Kunst oder Flötenspielen nur der *poiesis* zugerechnet hätte. Ziel dieser Arbeit ist gerade, den Begriff der *poiesis* weiterzufassen, da auch der Erwerb dieser Kompetenzen sozial eingebettet, also nur als kooperative Praxis erlernt werden kann.

seien diese abstrakte Handlungsschemen oder konkrete Verfahren, die wiederum die Anwendung anderer Güter beinhalten. Darunter mag auch der gemeinsame Arbeitsprozess fallen, der notwendig ist, um ein gewünschtes Ziel zu erreichen. Zudem ist das Werk, die Software, Resultat dieser zur Anwendung gekommenen Handlungsschemata. Die Wahl der vielfältigen Mittel bedarf einer spezifischen instrumentellen Zwecksetzungsrationalität, einer technischen Denkart: der *techne*.

Die so betrachtete herstellende Aktivität und die mit ihr verknüpfte Art der Problemlösung führt wiederum in einen differenzierten Technikbegriff, der nicht nur technische Objekte in Form von Instrumenten, Maschinen oder Medien, also Geräten, umfasst, sondern auch Fertigkeiten oder Verfahren, die durch den Menschen in angemessenen Handlungskontexten angewandt werden, um bestimmte Resultate zu erzielen. Ropohl (2009, S. 31) unterscheidet dementsprechend – wie Janich – drei Momente der *poiesis*:

„Technik umfasst (a) die Menge der nutzenorientierten, künstlichen, gegenständlichen Gebilde (Artefakte oder Sachsysteme) (b) die Menge menschlicher Handlungen und Einrichtungen, in denen Sachsysteme entstehen und (c) die Menge aller menschlichen Handlungen, in denen Sachsysteme verwendet werden.“

Technik umfasst somit das Werk, die Verfahrensabläufe der Herstellung sowie die Einbeziehung direkter und indirekter Stakeholder: Fahrzeuge, Personen, Regeln und Gesetze lassen Verkehr zu einem technischen Sachsystem werden. Das mutet zunächst banal an, zeigt aber die Bedeutsamkeit der Wechselwirkungen jener Einzelmomente auf: Die Herstellung gehört ebenso zum Technikbegriff wie die Bedienung oder eben die indirekte Teilnahme an technischen Systemen. Dadurch wird verdeutlicht, dass bereits der Begriff „*Technik*“ nicht nur die materiellen Objekte, bspw. Fahrzeuge, umfasst, sondern ebenso die Herstellungsverfahren, also wie bspw. Fahrzeugkarosserien gegossen werden. Nicht zuletzt beinhaltet Technik auch das soziale System, in welche es implementiert oder welche durch die Technik erst erschaffen werden, bspw. die Fahrer der Fahrzeuge oder die Fußgänger. Dadurch vereint der Technikbegriff den Entwicklungsprozess, die notwendigen Kompetenzen der Beteiligten oder Betroffenen sowie deren Interaktionen. Der Technikbegriff verdeutlicht, dass sich Lebensformen um oder mit Technik herausbilden: Das Verkehrssystem umfasst Schilder, Fußgänger, wütende Autofahrer etc. Die Einzelhandlungen kreisen allesamt um die technische Errungenschaft. Dieser Technikbegriff verweist so auch auf die Kompetenzen und Denkweisen der Beteiligten oder Betroffenen, die an dem System teilnehmen oder davon tangiert werden. Technik dient nicht als Klassifikationsbegriff von Objekten aufgrund gleicher Merkmale, sondern zielt auf bestimmte Verhaltensweisen ab: „A *technology* [Herv. d. A.] is a collection of techniques that are related to each other because of a common purpose, domain, or formal or functional features.“ (Brey, 2012b, S. 310) So

verweist der Technikbegriff bereits auf einen vielschichtigen Praxisbegriff: Technik strukturiert Handlungssysteme und Wahrnehmungsweisen.

Im Gegensatz zu Janich führt Ropohl daher noch die Dimension der Bedienung der hergestellten Artefakte ein, die für einen humanistischen Ansatz – vor allem hinsichtlich digitaler Artefakte – von großer Relevanz ist, da diese die gesamten alltäglichen Lebensräume durchdringen. Erst eine angemessene Endnutzung komplementiert digitale Technik, lässt diese als ein vollständiges System begreifbar werden, das in seiner Dynamik auch normativen Ansprüchen gerecht werden soll. Somit kann Technik auch als Vermittlung zwischen Entwicklung und Endnutzung verstanden werden. Durch diese Wechselwirkung können normativ angemessene Systemanforderungen formuliert werden, die sich nicht nur an der technischen Funktionalität bemessen. Ein solches Anforderungsmanagement muss jedoch der Herstellungsprozess garantieren: Hier geht es um eine Organisationsstruktur, die bestimmte Momente der *poiesis* im Hinblick auf eine wünschenswerte Entwicklungspraxis vermittelt.

Technik umfasst
Digitale Artefakte
Herstellungsprozesse der digitalen Artefakte: Kompetenzen und Arbeitsprozesse
Einsatz und Nutzung digitaler Artefakte
Technologie umfasst als Meta-Sprache
Wissen über die Herstellung, Verwendung und Reparatur von Technik
Wissenschaft von Technik

Tabelle 1 Technik und Technologie in Anlehnung an Ott (2005) und Janich (2001)

Zudem hilft diese Systematisierung dabei, den Begriff der Technik von jenem der Technologie abzugrenzen. Dies kommt nicht nur einer Begriffsklärung zugute, sondern verweist auch auf unterschiedliche Ebenen der Kompetenzen, Wissensformen sowie auf ein nachhaltiges Verständnis von innovativen Produktzyklen. Technik ist kein Oberbegriff, unter den unterschiedliche Objekte aufgrund ähnlicher Eigenschaften einzuordnen sind, sondern ist ein bestimmtes In-Bezugsetzen zur Welt und ein Handeln in und mit der Welt. Erst der lebensweltliche Horizont vermittelt Zwecke und Mittel, die dadurch sichtbar und für eine weitere Analyse gemeinsam aufbereitet werden. Erst diese intersubjektive und kollektive Haltung ermöglicht einen ethisch wünschenswerten und zielführenden Einsatz von Technologie und Technik. Es gilt, die *poiesis* mit der *praxis* und die *techne* mit der *phronesis* zu vereinen.

Technik ist nicht nur als ein Mittel, sondern auch als ein Medium aufzufassen, das aufgrund seiner Natur nicht nur den individuellen Weltbezug, sondern auch die Modalität der gegenseitigen normativen Anspruchshaltungen transformiert. In diesem Sinne kann an dieser Stelle bereits hervorgehoben werden, dass die Annahme eines ausschließlichen Mittelcharakters digitaler Artefakte diese auf ihre rein technischen Belange verkürzt. Eine E-Mail als Mittel ist nicht mehr und nicht weniger als ein technisches Zusammenspiel technisch bedeutsamer Wirkketten. Die E-Mail ist aber so bspw. nicht in ihrer Form der Beileidsbekundung bewertbar. Letzteres begreift die E-Mail als ein Medium, das bestimmte Formen der Anteilnahme zulässt und transportiert, während es andere Formen der Anteil-

nahme, bspw. Umarmungen, ausschließt. Ob die E-Mail-Kondolenz in diesem Sinne einer wünschenswerten Kommunikation der Beileidsbekundung entspricht, gilt es an der jeweiligen Lebenswelt, an der entsprechenden Praxis zu erkennen und zu erproben. So mag es nicht erstrebenswert sein, eine E-Mail zu versenden, wenn man in der Nähe wohnhaft ist, denn die Form des Trostes macht in diesem Zusammenhang eine Umarmung erforderlich. Somit lässt sich das digitale Artefakte durch die Praxis des Kondolierens bereits als Medium erahnen. Folgt man Janich, dann muss gerade deshalb Technik als Teil des Kulturellen verstanden werden, weil Technik aufgrund ihrer Zweck-Mittel-Relationalität nur hinsichtlich geteilter Interessen und geteilter Bedeutungskontexte begreifbar wird (Janich, 2003). Die kulturelle Dimension des Technischen erlaubt es erst, Mittel bestimmten Zwecken zuordnen zu können, ja überhaupt erst gemeinsam Zwecke zu setzen und Handlungskontexte zu generieren.⁴ Dafür muss das digitale Artefakte nicht nur technisch durchdrungen werden, sondern bedarf anderer Zugangsweisen, wie bspw. einer hermeneutischen oder moralischen Deliberation. Wie dies machbar sein soll, wird in den kommenden Kapiteln ausgeführt: Zunächst wird die technische Rationalität diskutiert, um sodann den hermeneutischen Zugang davon abgrenzen zu können. Auf dieser Basis kann anschließend erörtert werden, welcher Rationalität man folgen sollte, wenn das digitale Artefakt in seinen verschiedenen normativen Bezüglichkeiten gemeinsam gedacht werden soll. Dies wird mit Hilfe der *Strukturellen Rationalität* gelingen. Alle philosophischen Paradigmen müssen einer doppelten Aufgabe dienen können: Sie sind metatheoretisch bedeutsam, aber auch als eine Methode der Annäherung an normative, relevante Sachverhalte zu verstehen. So muss hier gezeigt werden, in welchen Bezügen und Abhängigkeiten das digitale Artefakt erschlossen werden sollte, damit es nicht nur der technischen Rationalität gehorcht.

2.1.1.2 Technische Rationalität

In diesem Abschnitt wird zunächst die technische Rationalität als grundlegender Denkmodus technischer Funktionalität eingeführt, d.h. es soll hervorgehoben werden, dass das Kriterium der technischen Funktionalität ohne diese Form der Deliberation nicht formulierbar wäre. Wichtig ist jedoch, dass keiner Denkart das Primat zugesprochen werden soll, sondern die gegenseitige Ergänzung in den Mittelpunkt gerückt werden muss, sollen digitale Artefakte nicht nur technisch funktional verstanden, sondern auch moralisch bewertet werden können. Die technische Rationalität wird im Aristotelischen Sinne hier als *techne* eingeführt. Die *techne* kümmert sich ausschließlich um instrumentelle Kriterien, weshalb moralische oder ästhetische Wertmaßstäbe keine Orientierungspunkte ihrer Gerichtetheit, ihrer Intentionalität sind. *Techne* – verstanden als Art und Weise, wie etwas hergestellt,

⁴ Eine philosophisch detaillierte Ausarbeitung und historische Einbettung des begrifflichen Unterschieds zwischen Technik als Mittel und Medium findet sich in Hubig (2013), Kastendiek (2003) und Krämer (1998).

bedient oder repariert werden soll – muss zugleich erlernbar, wiederholbar und auch tradierbar sein (Ott, 2005). *Techné* ist somit keine gegebene natürliche Fähigkeit, sondern bedarf der Gewohnheit, der regelmäßigen Ausführung, sodass sie als technisch-materielles Denken über die Welt und Handeln in der Welt verstanden werden muss: Es ist das kontinuierliche Vermögen, Allgemeines, in Form von technisch-naturwissenschaftlichen Gesetzmäßigkeiten oder mathematischen Kalkülen, auf lebensweltliche Einzelereignisse übertragen zu können. *Technische Deliberation* beschränkt sich auf das Auffinden der Mittel für gegebene Zwecke. Dabei können die Mittel vielfältig sein: Handlungen, Verfahrensweisen und, falls abgeschlossen, Artefakte.

Die Mittel erfüllen sich in ihren Zwecken, sie sind zweckgebunden, d.h. die Angemessenheit der Tätigkeit bemisst sich an der Funktionalität des Werks, das einem bestimmten Zweck dienlich sein soll. Der Zweck „tut sogar noch erheblich mehr für sie [die Mittel, N.Z.], er produziert sie nämlich überhaupt erst und organisiert sie“ (Arendt, 1960, S. 181). Die Mittelwahl ist selbst vielschichtig, denn zum einen müssen die Mittel erkannt werden, die das zu behebende Problem technisch lösen. Zum anderen muss auch die Kompetenz ermittelt werden – sozusagen als ein weiteres Mittel –, mit der dieses Produkt letztendlich entwickelt werden kann. Ersteres verweist auf das Produkt als Mittel, letzteres auf den Herstellungsprozess als Mittel, um das Produkt zu erschaffen. Den Herstellungsprozess als Tätigkeit zu verstehen, die der technischen Rationalität geschuldet ist, deutet auf eine bestimmte Art des Entwicklerdenkens hin: Es geht hier um das Denken in der Kategorie des *in-order-to* (Reijers & Coeckelbergh, 2020). Um diesen wichtigen Gedanken aufzunehmen und rationalitätstheoretisch aufzuwerten, muss man den Deliberationsmodus als *in-order-to-rationality* bezeichnen, der vor allem die Mittel-Zweck-Beziehungen und somit die für Technik und Technologie so wichtigen Machbarkeitsbedingungen zu berücksichtigen hat. Doch diese Weise der Vernunftausübung gewinnt dadurch an Bedeutung, weil durch die Bestimmung der Mittel nach immer mehr Mitteln Ausschau gehalten wird. Das Mitteldenken scheint kein Ende zu nehmen, denn auch die Bedienung sucht nach einer effizienten Nutzung des *ergons* für bestimmte Zwecke. Fortan bestimmen die Mittel das Denken, da nur ihnen Aufmerksamkeit zuteilwird. Dies führt zum Verlust einer umfassenden Zwecksetzungsfähigkeit. Denn jene Tätigkeit der Vernunft sich Zwecke als Orientierungsmomente eigenständig vorzugeben und sich an ihnen auszurichten, lässt die vernünftige Mittelwahl, welche zur Erreichung eben jener Zwecke notwendig sind, in den Hintergrund rücken. Diese Tendenz der Vernunft die autonome Zwecksetzung zur prioritären Aufgabe zu bestimmen, erfüllt sich so Rohbeck (1993) in modernen Theorien der Unterwerfung der Menschen unter das Diktat der Maschinen, da sich Mittel in Zwecke verwandeln. Die Suche nach optimalen Mitteln wird Selbstzweck der Vernunft. In dieser von Rohbeck ausführlich skizzierten Entwicklung wird aus einer Zweckmittelrationalität, die noch die Einheit aus Zweck und Mittel

zu denken vermochte, eine Zweckrationalität, die sich auf die Wahl angemessener, vorhandener Mittel beschränkt. Rohbeck (1993) begreift diesen Werdegang der Vernunft als die Verselbständigung der *technischen Rationalität*:

„Technische Rationalität: Gemeint ist genau jener Typ zweckrationales Handeln, nach dem bestimmte Ziele auf möglichst effektive Weise erreicht werden sollen. ‚Technisch‘ bezieht sich daher erst einmal in einem sehr weiten Sinn auf ein menschliches Handeln, das ‚rational‘ kalkuliert, indem es Aufwand und Nutzen kühl berechnend gegeneinander abwägt.“ (ebd., S. 19)

In diesem Sinne sind die ökonomische und die technische Rationalität identisch: Die ökonomische Rationalität, deren Denkmodus sich über einen Nutzenbegriff definiert, den es innerhalb bestimmter Restriktionen effektiv zu optimieren gilt, ist als Kalkül fassbar. Eine rationale Entscheidung wird mit einer Berechnung quantifizierbarer Nutzenwerte hinsichtlich möglicher Entscheidungsalternativen mit subjektiven Wahrscheinlichkeiten gewichtet, die Erreichbarkeit oder Erfüllbarkeit des Bedürfnisses oder der Präferenz wird so erklärt (Nida-Rümelin & Schmidt, 2000). Die Vernunft als Zwecksetzungsvermögen ist innerhalb dieser Debatte bereits aufgehoben, denn der Nutzenbegriff soll gerade inhaltslos bestimmbar bleiben, sobald er als ein Denk-Prinzip fungiert, das sich nicht über die konkrete Ausgestaltung des guten Lebens definiert, sondern über eine möglichst vernünftige Mittelwahl die Realisierung vorgegebener Zwecke thematisiert. Die Zwecke sind nur dann zu verwerfen, wenn keine hinreichend rationalen Mittel zur Verfügung stehen, die diese Zwecke erfüllen können, oder neue Zwecke an ihre Stelle treten.

Das teleologische Handlungsmodell begreift jede Handlung als eine planbare bezweckte Struktur, die in ihrem Ziel ein begründetes Ende findet, das durch eine geschickte Mittelwahl erreicht werden kann. Die mögliche Mittelwahl macht eine transzendente Reflektion über die Bedingungen der Möglichkeit der Handlung erforderlich, denn es sind gerade die Mittel, die diese Entscheidungsoptionen an die Lebenspraxis binden. Eine Realisierung des Zwecks durch das gesetzte Ziel ob bestimmter Mittel – und jene können andere Handlungen, aber auch Geräte sein – muss möglich und machbar sein, möchte man von vernunftgeleiteten Handlungen sprechen. Durch ebenjene Mittelwahl wird somit eine Ernüchterung über unerreichbare Wünsche erlangt und ein umfassenderer Vernunftbegriff ermöglicht: Denn nicht das Auffinden und Setzen von Zwecken allein ist der Vernunftfähigkeit zuzuschreiben, sondern auch die Wahl der bestmöglichen Mittel, um bestimmte, nicht immer kontrollierbare Zwecke zu realisieren. Dieser qua praktischer Vernunft zielgerichtete Entwicklungspfad umfasst somit die Idee eines freien Subjekts, das aus sich heraus diejenigen Mittel wählt, um bestimmte Ziele zu erreichen, die mit einem bestimmten Zweck begründet werden. Ein Verständnis von Rationalität, das sowohl die Zwecksetzung als auch die Mittelwahl betrachtet, um den Rahmen vernünftiger Handlungen im Hinblick auf Technik und Technologien abzustecken, diskutiert Johannes Rohbeck 1993 in

seiner Monografie *Technologische Urteilskraft – Zu einer Ethik technischen Handelns*. Rohbeck thematisiert hier geschichtsphilosophisch systematisierend die Evolution des individuellen und kollektiven teleologischen Handlungsbegriffs in Bezug auf die Technologie- und Technikentwicklung. Diese Zweck-Mittel-Relationalität zeigt sich bereits, so Rohbeck, in der bei Platon aufzufindenden zwecksetzenden Vernunft, die sich keineswegs um die Mittel bemüht, weil diese ohnehin von den guten Zwecken ableitbar seien und in ihnen aufgingen. Somit ist keine gesonderte Betrachtung oder gar Bewertung der Mittel notwendig. Obgleich Platon die *Verselbständigung*, wie Rohbeck die erste Separierung von Zweck und Mittel bezeichnet, der Mittel schon ahnte (Rohbeck, 1993), wird sie von Platon als nicht weiter diskutabel erklärt. Da sich die Mittel nicht immer an einen Zweck binden lassen, setzt Rohbeck diese Trennung als erste Differenz der Zweck-Mittel-Beziehung. Nach Rohbeck kommt es dann bereits bei Aristoteles zur zweiten Differenz des Zweck-Mittel-Verhältnisses (Rohbeck, 1993), wobei sich die Aufgabe der Vernunft von der vernünftigen Zwecksetzung auf eine vernünftige Mittelwahl verschiebt. Hierdurch werden die Mittel zu eigenständig bewertbaren Einheiten, die nicht nur durch ihren Zweck gerechtfertigt werden können. Denn sonst, so zeigt Rohbeck eindrücklich anhand der politischen Mittel für politische Zwecke, liefe man Gefahr, dass der Zweck die Mittel heilige (Rohbeck, 1993). Dies aber sei für eine moralische Bewertung aller denkbaren Situationen nicht wünschenswert. Durch die Verselbständigung und Verschiebung der Mittel instrumentalisiert sich die Vernunft, denn sie kann auch sich selbst nur noch als Mittel begreifen: als Vermögen, die besten Mittel für unbewertbare Zwecke zu wählen – eine Vernunft, wie sie bereits David Hume skizzierte (Hume, 1739). Eine Pause ist der Vernunft hierbei nicht vergönnt, da ein andauernder externer Druck von der weltlichen Gegenständlichkeit ausgeht. Durch die Exposition verliert sich die Möglichkeit, selbst Zwecke setzen zu können oder sogar angemessene Mittel zu gegebenen Zwecken wählen zu können. Eben jene Verhältnisverkehrung von Zweck und Mitteln präsentiert Rohbeck in seiner Darstellung der Geschichte der Vernunft.

Rohbeck erklärt die finale Loslösung der Zwecke durch die liberale politische Ökonomie eines Adam Smith bis hin zur Unterwerfung des Menschen unter die Maschinen, die sich vom Mittel zu Zwecken, gar zu Endzwecken automatisierten und die nun den Menschen zwingen, sich auf bestimmte Weise zu verhalten sowie sein Verhältnis zur Natur nur noch indirekt und distanziert zu erfahren. Dieses Erlebnis der Distanzierung von der Welt verstärkt sich zudem mit dem Einsatz von Maschinen, indem nur noch bestimmte Tätigkeiten häufig wiederholt werden müssen und der Mensch sich dadurch in der Herstellungspraxis und in der Bedienung einer maschinell-rationalen Welt verliert. Dieser Kontrollverlust des Menschen zeigt die Bedeutung der Technik als Sachzwang auf, dem sich der einst als frei verstandene Mensch nun gebeugt hat. Die Maschinen und deren Ökonomie fordern nun vom Menschen ein, was er einst durch Arbeit für sich zu erreichen versuchte. Diese Idee der Unterwerfung unter technische Sachzwänge steht dem Rationalitätstheoretischen VDI-Diskurs konträr gegenüber,

denn hier werden gerade die Verantwortungsübernahme und die bewusst bestimmbare Machbarkeit des technischen Designs diskutiert. Rohbeck folgend sind somit Philosophien der Technik aufgrund dieser teleologischen Umkehrung oder Brechung entweder euphorisch oder apokalyptisch: Entweder liegt eine Überbetonung des Sachzwangs vor oder die vollständige Kontrolle über technische und technologische Entwicklung wird als Paradigma in den Vordergrund gerückt (Rohbeck, 1993).

Für eine Ethik ist es geboten, sich von der Verteufelung der Mittel und der vollständigen Unterwerfung unter Sachzwänge zu distanzieren, will man überhaupt von einer Handlungsmacht und einer möglichen Verantwortungsübernahme sprechen. Eine solche Ethik muss sich eine Bewertung der Mittel zutrauen und zugleich Klarheit über die Adressaten dieser normativen Anforderungen verschaffen. Die technische Rationalität bedarf somit einer Ergänzung um eine ethische Deliberation.

2.1.1.3 Technik und Ethik: Die Bewertung der Mittel

Ein instrumentelles Verständnis praktischer Vernunft erlaubt eine ethische Betrachtung somit nur über die Folgen, die bewertet werden sollen, denn die Optimierung des Nutzens und die Wahl der Mittel erschöpfen die Tätigkeit der praktischen Vernunft. Rohbeck grenzt sich jedoch von diesem negativen Begriff des Instrumentalismus ab, der die Vernunft als Sklave aller Mittel versteht und auch die Vernunft selbst nur als Mittel in Abhängigkeit von Nutzenbegriffen begreift (Rohbeck, 1993). Die instrumentelle Vernunft, die auf Konsistenzen innerhalb der Mittelwahl in Abhängigkeit gegebener Zwecke zu achten hat, aber eine Bewertung der Ziele selbst aus den Augen verliert, reduziert sich auf ein Abwägen zwischen Nutzen und Kosten und kann somit keine anderen Gründe als rationale Bewertungskriterien aufnehmen. Daher rückt der positive Aspekt der Berücksichtigung der Mittel als die Bedingung der Möglichkeiten, erreichbare Ziele zu bewerten, in den Hintergrund. Rohbecks positiver Begriff der Mittel begründet somit keine Gegenüberstellung von Handlung und Herstellung, Ethik und Technik oder normativem Orientierungswissen und faktischem Tatsachenwissen. Wenn sich jedoch Handlungsfolgen ergeben, die diesem Nutzen entgegenwirken oder aber andere Personen einschränken, müssen diese unabhängig vom Zweck, unabhängig von der Handlungsmotivation, eine Beurteilung erfahren können. Allerdings engt ein durch den Nutzenbegriff geprägtes Vernunftverständnis die Betrachtung rationaler Handlungsgründe unnötig ein, denn diese sind gerade nicht ausschließlich über ihre Dienlichkeit hinsichtlich des Nutzenbegriffs zu bestimmen (Nida-Rümelin & Schmidt, 2000; Nida-Rümelin, 2020). Eine umfassende Bewertung technischer Artefakte und technologischen Wissens darf sich nicht nur auf Nutzen und technische Effektivität verkürzen lassen, sondern muss weitere normative Bewertungskriterien zulassen können, wenn gesamtgesellschaftliche oder ökologische Aspekte berücksichtigt werden sollen. Unter Instrumentalität versteht Rohbeck deshalb hinsichtlich der Technik und Technologie somit das Phänomen, dass diese Mittel über die

Zielerfüllung hinauszugehen vermögen (Rohbeck, 1993). Rohbeck führt somit einen transzendenten, reflexiven Mittelbegriff ein, der gerade die individuelle prinzipiengeleitete Entscheidungstheorie erweitert. Es geht hier zum einen um die Bedingungen der Möglichkeit, die Mittel hinsichtlich bestimmter Zwecke zu wählen, und zum anderen darum, die Mittel auch hinsichtlich anderer Gütekriterien zu bewerten, wie es Nida-Rümelins *Strukturelle Rationalität* auch denkbar werden lässt. Die Trennung von Mittel und Zweck, die zu der Dynamik führt, dass die Mittel zu Zwecken und Zwecke zu Mitteln werden, zeigt auf, dass Technik und Technologie einem stetigen Wandel unterliegen. So können technische Artefakte selbst als Mittel dazu fungieren, bestimmte Handlungsziele zu erreichen oder aber selbst zu Handlungszielen werden. Diese Umdeutung der Mittel geschieht nach Rohbeck häufig während des Gebrauchs (Rohbeck, 1993). Dem sei jedoch hinzuzufügen, dass auch der Gebrauch von technischen Artefakten oder technologischem Wissen in der Entwicklungsarbeit oder in Service- und Recyclingphasen bereits auf vorherigen, evtl. anderweitig eingesetzten Ausführungen aufbaut oder diese umdeutend integriert. Insofern muss hier Gebrauch so weit verstanden werden, dass darunter auch die innovative Interpretation bereits vorhandener Mittel im Herstellungsprozess verstanden werden kann. Ebenso werden gerade im Herstellungsprozess Technik oder Technologie als zu erreichendes Ziel definiert. Dies hat nach Rohbeck zur Folge, dass die vernunftgeleitete Deliberation nicht mehr in einer Zielerfüllung aufgeht, sondern gerade das Rasonieren über Mittel unvorhersehbare und unbeabsichtigte Nebenfolgen menschlicher Handlungen in den Mittelpunkt rücken lässt (Rohbeck, 2019). Wichtig hierbei ist jedoch, die Ursächlichkeit der nicht intendierten Folgen technischer Handlungen zu erörtern, um hervorzuheben, ob es sich bei der methodologischen Neutralität (Rapp, 1978) um einen Sachzwang handelt (Teil II, Kapitel 7.2).

Die Spaltung von Mitteln und Zwecken, die gerade darin begründet ist, dass die Mittel nicht gänzlich in den Zwecken aufgehen, übersieht, dass Technik, allein durch vielfache Verkettungen und Einbindungen in komplexe (technische) Sachsysteme gerade im Digitalen nicht mehr nur Zwecken folgt, sondern eigene Welten und neue Handlungsräume erschafft. Diese Beobachtungsgrößen fungieren somit als erste Orientierung einer ethischen Evaluation. Diese Bewertung bezieht sich jedoch immer auf Artefakte, die sich innerhalb des machbaren Lösungsraums befinden müssen: Es muss also möglich sein, jenes Artefakt auch herzustellen und dieses muss wiederum eine Lösung für etwas darstellen. Dieser *technical solutionism* (Oelschlaeger, 1979) bedarf gerade einer lebensweltlichen Reflexion, um einen techno-ethischen Realismus zu ermöglichen, der im Zuge einer technik-ethischen Bewertung vonnöten ist, möchte man sich nicht in einer spekulativ-fiktiven Technikfolgenabschätzung wiederfinden.

In diesem Zusammenhang ist die Unterscheidung der technischen Artefakte als Ziel eines Herstellungsprozesses oder als Gebrauchsgut zwingend hervorzuheben, denn der Unterschied zwischen Herstellung und Gebrauch eröffnet erst Handlungsräume durch im Herstellungsprozess unintendierte

Zwecksetzungen (Rohbeck, 1993). Wenn hier noch von Zielerfüllung die Rede sein kann, dann nur umgekehrt in dem Sinne, dass eine gegenständlich repräsentierte Möglichkeit als Zweck menschlicher Handlungen zu realisieren ist. Obgleich in der konkreten Ingenieurstätigkeit immer Mischformen praktiziert werden, ist es erforderlich, diese zwei Handlungstypen präzise zu fassen. Diese Verschiedenartigkeit verweist nämlich in aller Zuspitzung darauf, dass die unerwartet neuen Handlungsspielräume bereits materiell in den jeweils zur Verfügung stehenden Mitteln existieren. Es handelt sich mithin um objektive oder reale Möglichkeiten (Rohbeck, 1993). Dieser offene Umgang im Gebrauch technischer Artefakte wird auch als Kritik am traditionellen, strengen Mittelbegriff verwendet und vor allem in Technikphilosophien fruchtbar gemacht, die sich von einer reinen technischen Rationalität abwenden. Die Betrachtungskriterien von Technik und Technologie verschieben sich, sobald sich von der Entwicklersicht auf die Benutzer- oder Konsumentensicht verlagert wird. Eine umfassende Technikethik profitiert gerade von einer Verbindung beider Sichtweisen, wenn es um die Konkretisierung kontextspezifischer, normativer Forderungen geht.

Ethische Deliberationen, die in der Softwareentwicklung implementiert werden können, bedürfen einer Verbindung von herstellenden Tätigkeiten mit moralischen Handlungen. Oder rationalitätstheoretisch gesprochen: Technische Rationalität soll der *phronesis*, der Vernunft dienen können. Dies hat Auswirkungen sowohl auf die Fähigkeiten der Entwickler und auf den Herstellungsprozess: Der Entwickler kann sich nicht mehr nur auf technisches Denken beschränken und der Herstellungsprozess muss diese Form der zusätzlichen Deliberation auch gewähren können.

1.2. Phänomenologisch-pragmatische Technikphilosophie

Die phänomenologische Perspektive thematisiert Technologie und Technik in Bezug auf das menschliche Weltverhältnis. Der mediale Charakter digitaler Artefakte, der sich durch die Dynamik der Mittel und Zwecke ergibt, stellt sich einer klassischen rationalitätstheoretischen Interpretation des VDI-Diskurses entgegen (Hubig, 2013; Ihde, 1990). Die vornehmlich auf geisteswissenschaftlichen Terminologien und Theorien beruhende Technikphilosophie versucht, technische Artefakte aus einem nicht technik-immanenten Betrachtungswinkel zu erörtern. Sie soll die durch einen Einsatz von digitalen Artefakten hervorgerufenen und transformierten Wahrnehmungs- und Erfahrungsmodalitäten erfassen können. Meist haben diese anti-technologischen Technikphilosophien, wie bspw. die der kritischen Frankfurter Schule, Martin Heideggers oder Karl Jaspers, keinen direkten Bezug zum Herstellungsprozess, weshalb sie auch im Rahmen einer Arbeit, die sich metatheoretisch mit einer Ethik für Softwareentwickler beschäftigt, keine weitere Berücksichtigung finden (Mitcham, 1994; Kogge, 2008). Eine Technikphilosophie, die Technik durch nicht technische Verstandeskategorien oder mittels eines

hermeneutischen Zugangs zu artikulieren versucht, wird im Rahmen dieser Arbeit nur dann berücksichtigt, wenn sie der Lokalisation normativ relevanter Eckpunkte von sozio-technischen Systemen dient.

Die Abkehr von dieser *engineering philosophy* macht sich auch durch eine Abwendung von rationalen Technikfolgenabschätzungen bemerkbar. Die geisteswissenschaftliche Herangehensweise lässt die tatsächlich praktische Dimension der Herstellung unberücksichtigt. Diese Art der Annäherung an das Thema Mensch-Maschine vernachlässigt die technischen Begebenheiten zugunsten einer Diskussion der veränderten individuellen Weltbezüge sowie der transformierten gesellschaftlichen Subsysteme. Der Nutzer oder die indirekt Betroffenen rücken somit in eine semantisch-hermeneutische Analyse. Die technischen Objekte und der Umgang mit diesen werden so jenseits des technischen Jargons thematisiert, da die Bedeutung dieser Artefakte für den Menschen und seine Bezogenheit auf die Welt in den Fokus rückt. Die Transformation des individuellen Weltbezugs sowie der sozialen Erwartungshaltungen ist mathematisch-naturwissenschaftlich nur unzureichend erfassbar. Erst eine Rückführung der hermeneutisch-analytischen Erkenntnis in technische Anforderungen bringt beide notwendigen Debatten wieder zusammen. Die fehlende gemeinsame Sprache bzw. der fehlende gemeinsame Lösungsanspruch dieser philosophischen Traditionen führt zu zwei weitestgehend voneinander getrennten Themenräumen, die nur dann zusammengebracht werden, wenn der Designprozess oder die Nutzung der Endprodukte im Hinblick auf normative Designansprüche diskutiert werden.

Die Entwicklerperspektive wird daher um einen hermeneutisch-phänomenologischen Ansatz erweitert, um die Veränderungen der Weltbezogenheit, die Technik bei Nutzern verursacht, betrachten zu können. Dies wird auch in der Analyse von Mittel und Medium ersichtlich werden (Teil III, Kapitel 1). Der rein im klassischen Sinne instrumentelle Begriff des Mittels kann durch den medialen Begriff, der einer Umkehrung von Zweck und Mitteln bedarf, bereichert werden. Gerade diese Horizonterweiterung ist bedeutsam, wenn ethische Deliberationen herausgearbeitet werden sollen, die auf ein gutes normatives Technikdesign abzielen. Der humanistischen Tradition entsprechend, sollen Vernunft, Autonomie und Verantwortung auch durch den Einsatz digitaler Technologien und Technik bewahrt werden: In Bezug auf Technik und Technologien bedeutet dies, dass die wesentlich konstitutiven Merkmale berücksichtigt werden müssen, in denen sich Menschen alltäglich bewegen und ihr Leben dementsprechend ausrichten. Sowohl die Entwickler als auch die Nutzer sollen sich in ihren Lebensentwürfen innerhalb von Lebenspraxen ausleben können, wenn gute normative Technik und Technologie, die den anderen und das Gemeinwohl nicht beeinträchtigen, entstehen sollen.

Eine Verschmelzung der beiden Traditionen ist von daher notwendig, will man ein Ethos, eine Ethifizierung begründen, die einen philosophisch-ethischen Designansatz ermöglicht. Die Kritik am VDI-

Diskurs, sich nur an der Mitteldimension der Technik auszurichten und somit Weltbezüglichkeiten und Transformationen in normativen Bedeutungskontexten zu ignorieren, die durch den Gebrauch von Technik oder aus dem Einsatz von Technologie resultieren, kann durch die Erweiterung um einen phänomenologisch-hermeneutischen Ansatz bewältigt werden. Hier eignet sich vor allem der Ansatz von Ihde, da er eine Technikphilosophie entwickelt hat, welche die Ingenieurs-philosophische Stoßrichtung der *usability*, des nutzerfreundlichen Designs, für moralische Gütekriterien zugänglich machen kann. Seine Analyse der Diskussion der Transformationen der Wahrnehmungs- und Erfahrungsmodalitäten durch Technik kann an Designanforderungen rückgekoppelt werden. Da auch in diesem Rahmen die Möglichkeiten in den Mittelpunkt gerückt werden sollen, wie Technik ethisch zumutbar gestaltet werden kann, eignet sich eine humanistische Verknüpfung einer wertvollen Entwicklungspraxis, die lebensweltlich wünschenswerte gesellschaftliche Praxen der Nutzer respektiert und in ihren Designansprüchen berücksichtigen kann. In dieser Vorgehensweise kann auch, wie Kogge (2008) betont, zum Teil der Frage nachgegangen werden, wer Adressat moralischer Forderungen sein soll. Auch wenn Ihde seine Technikphilosophie auf kein Rationalitätstheoretisches Fundament stellt, sondern die Beziehung zwischen Menschen und Technologie phänomenologisch einzufangen versucht, ist seine philosophische Analyse dieser Mensch-Technik-Relation für einen humanistischen Ansatz ein wichtiger Ausgangspunkt. Der phänomenologisch-hermeneutische Diskurs der Technikphilosophie baut auf einer systematisch anderen Grundlage auf als der VDI-Diskurs: Nicht das Zweck-Mittel-Verhältnis, sondern die Beziehung zwischen Menschen und Welt, wie sie in den Begriffen der Intentionalität, der Lebenswelt und dem In-der-Welt-sein ausgedrückt ist, bildet die Basis, von der aus Technik in den Blick genommen wird.

Ein wesentliches Qualitätsmerkmal digitaler Technologie und Technik ist die Unterscheidung zwischen Mittel und Medium. Der mediale Charakter, der Informationstechnik und -technologie vom Begriff des technischen Mittels unterscheidet, wird besonders deutlich, sobald eine phänomenologisch-hermeneutisch geprägte Perspektive eingenommen wird. Softwaresysteme, wie alle technischen Objekte, führen zu normativen Verhaltensanpassungen auf Seiten der Nutzer, woraus Ihde die Nicht-Neutralität technischer Artefakte ableitet (Ihde, 1990). Selbst wenn Ihde die These der Nicht-Neutralität für alle technischen Artefakte postuliert, weil jede Technik auf den Weltbezug einwirkt, gilt dies in einem verstärkten Maße für digitale Technik. Der Ansatz Ihdes ermöglicht es, die durch einen Einsatz von Technik hervorgerufenen Transformationen der Erfahrungs- und Wahrnehmungsmodalitäten zu erfassen. Ihde ordnet diese wiederum drei Kategorien zu, die sich als kontinuierliche Beziehungsmodalitäten bestimmen lassen. Die Klassifikation erfolgt durch die Modalität der Mensch-Technik-Welt-Bezogenheit. Dadurch folgert Ihde sogleich, dass das Design technischer Objekte sich nicht nur am technischen Funktionalismus orientieren kann, sondern wesentlich an den Fähigkeiten, die das Ob-

jekt in menschlichen Kontexten zu übernehmen hat (Ihde, 1990). Der Verlauf der Beziehungsmodalitäten lässt sich auf einem Kontinuum abbilden, wobei die einzelnen Momente erneut nicht exklusiv verstanden werden sollen, sondern sich ergänzende Perspektiven repräsentieren, die nicht reduktionistisch zu interpretieren sind.

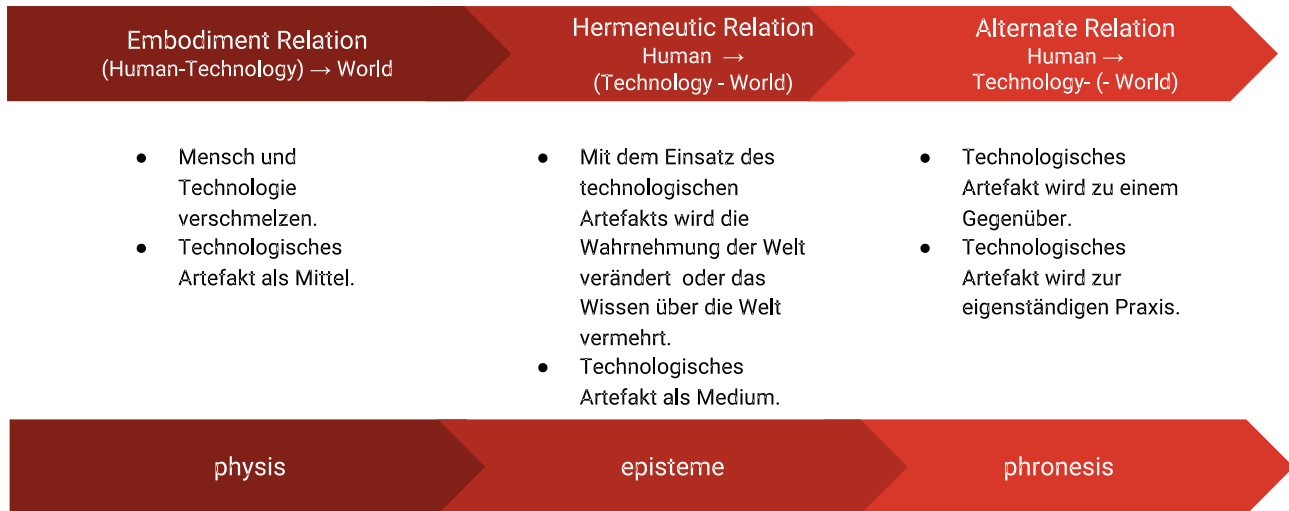
Ihde skizziert drei Varianten der Relation Mensch-Technik-Welt (vgl. Abbildung 1). Die erste Variante beschreibt das technische Objekt als eine mit der menschlichen Existenz verschmolzene Relation, sodass das technische Objekt nicht als eigenständige Entität in der Welt wahrgenommen wird. Technik ist auch deshalb nicht wertneutral, da durch sie bereits eine selektive und somit diskriminierende Wahrnehmung erfolgt, wie Ihde es in seiner Analogie mit einem Vergrößerungsglas deutlich macht (Ihde, 1990). Technik vergrößert immer einen Aspekt und lässt andere in den Hintergrund treten. Damit ist Technik bereits als *embodiment* fundamental selektiv und diskriminierend. Ihde führt als Beispiel eine Brille an, die eine verbesserte Sicht ermöglicht, ohne dass der Brille während des Sehaktes besondere Berücksichtigung zukommt, insofern sie sich möglichst unauffällig in die menschliche Körperlichkeit integriert. Die Brille ist nicht das Objekt der Betrachtung, sondern durch diese wird die Welt betrachtet. Ihr Zweck ist es, die Sehkraft der betroffenen Person zu verbessern, ohne dabei selbst als Objekt in den Vordergrund zu rücken. Das technische Artefakt ist in diesem Sinne unsichtbar, transparent und verhilft als Mittel dazu, die Welt wahrzunehmen. Somit übernimmt oder unterstützt das technische Produkt eine menschliche Fähigkeit, wodurch die Welt für den Menschen wieder erfahrbar wird. Daraus ergibt sich sogleich, dass diese Verkörperung in ihrem *use-context* (Ihde, 1990, S. 74) als normative Designanforderung formuliert werden kann. Die Produktqualität ist nicht mehr rein über die technischen Anforderungen sichergestellt, sondern muss sich sogleich an der Mensch-Maschinen-Interaktion messen lassen, d.h. daran, inwiefern sich das technische Produkt in die menschlichen Wahrnehmungshorizonte und Handlungen einfügen lassen soll (Ihde, 1990).

Die zweite Variante der Triade Mensch-Technik-Welt bezeichnet Ihde als hermeneutische Relation, worunter all diejenigen technischen Artefakte fallen, aufgrund derer eine neue Erkenntnis, ein neuer Wissenszustand erzielt werden kann: Hierbei steht im Vordergrund, die Welt differenzierter wahrzunehmen. Das technische Artefakt wird als Medium im Wahrnehmungs- und Erfahrungshorizont platziert, um einen bisher unbekanntem oder unpräzisen Tatbestand der Welt erfassbar zu machen. So führt Ihde hier das Beispiel des Thermometers ein, wodurch die Temperatur erst objektiv beschreibbar wird. Zusätzlich zum technischen Objekt muss nun noch eine Kompetenz erworben werden, damit dieses Objekt *richtig* interpretiert werden kann, denn: Welche Temperatur gilt als Fieber? Welches Wissen ist notwendig, um verstehen zu können, was die Daten bedeuten sollen (Ihde, 1990)? Solche Geräte können graduell auch als Mittel bezeichnet werden, da sie eingesetzt werden, um etwas Anderes zu erreichen, z.B. Fieber zu messen. Die hermeneutische Beziehung betont die intentionale

Wahrnehmung nicht unmittelbar erfahrbare Informationen. Hierunter versteht Ihde auch den epistemischen Mehrwert in Form neuer Informationen, die erst durch die Interpretation der Resultate des technischen Gerätes in das Wahrnehmungsfeld rücken. Dabei geht es nicht darum, wie die Welt neu erfahren wird, sondern wie sich das menschliche Gerichtet-Sein auf die Welt durch den Einsatz von Technik ändert. Im Gegensatz zur Floridis ontologischem Ansatz rücken nun Intentionalitäten in den Mittelpunkt der Betrachtung (Floridi, 2013). Eine Analyse der Intentionalität zeigt jedoch, dass sich das Gerichtet-sein weniger ändert, sehr wohl aber die Variablen, die diese Beziehung bedingen (Ihde, 1990). Dies gilt im verstärkten Maße für digitale Artefakte, die sich vor allem durch eine Transformation der hermeneutischen Beziehung kennzeichnen lassen: Sie greifen in das Weltverständnis des Menschen ein. Damit aber verändern sich zugleich auch Verhaltensweisen der Menschen sowie ihr Umgang mit der Umwelt.

Als *alternate Relation* versteht Ihde, dass Technik zu einer *otherness*, also einem Gegenüber wird, z.B. in Form von Computern. Den Computer als ein Gegenüber zu begreifen, schließt jedoch nicht aus, dass die Nützlichkeit desselben wiederum in der Sphäre der Hermeneutik anzusiedeln ist, da ein Computer wesentlich dazu beiträgt, die Welt hermeneutisch zu erschließen (Ihde, 1990). Somit muss nochmals betont werden, dass diese Modalitätsbeziehungen nicht exklusiv zu verstehen sind, sondern ein technisches Gerät in seinen unterschiedlichen Facetten abzubilden vermögen. Ihde betont eine *Quasi-Otherness* (Ihde, 1990, S. 100) von Computern, da diese über keine eigenständige Intentionalität verfügen, wie es bspw. Tiere tun. So können Tiere ungehorsam, Computer hingegen nur fehlerhaft sein. Durch diesen Aspekt der *Quasi-Otherness* verweist Ihde sogleich darauf, dass Computer nicht nur instrumentell eingebunden werden, sondern gerade auch als Spiel, Sport und Kunst in die Lebenswelt Einzug erhalten haben (Ihde, 1990). und folglich als eine Praxis etabliert werden können. Die Modalität ist demnach kein Weltbezug, sondern ein Bezug, der ein neues, eben virtuell-analoges Weltverhältnis mit neuen Handlungsgewohnheiten etabliert. Dem sei hinzuzufügen, dass Computer auch in die wissenschaftliche und ästhetische Praxis eingebunden sind und sich im Rahmen dieser Betätigungsfelder neue *alternate* Relationen einstellen können. Die Kategorie der *Quasi-Otherness* ist überdenkenswert, denn sie führt sehr schnell zu den Fantasien allmächtiger, autonomer Roboter. Wenn hierunter jedoch die Möglichkeit fallen soll, Computer nicht nur als Wahrnehmungsmedium, sondern auch als handlungsstrukturierend zu begreifen, kann man auf die *Quasi-Otherness* verzichten, indem der Aspekt struktureller Praxen betont wird. Ein Computer kann alle drei Merkmale dieser Beziehung repräsentieren: Software fügt sich in alles ein, wird dadurch transparent, macht neues Wissen zugänglich und etabliert neue Praxen.

Erste Bedingung eines normativ wünschenswerten Softwaredesigns: Transformation des individuellen Weltbezugs



Die **erste Bedingung** besagt, dass gutes Softwaredesign die individuelle Weltbezogenheit berücksichtigt.

Abbildung1 Erste Bedingung eines normativ wünschenswerten Softwaredesigns: Modalitäten der Mensch-Technik-Welt-Bezogenheit nach Ihde, eigene Erweiterung. Eigene Abbildung

Ihde (1990) zeigt, dass eine rein naturwissenschaftliche Beschreibung eines artifiziellen Produkts unzureichend ist, sollen moralische, wertvolle Objekte entwickelt werden, denn es geht gerade darum, die Weltbezogenheit, die maßgeblich durch den medialen Charakter digitaler Technologien und Techniken beeinflusst wird, in der Gestaltung wünschenswerter Technologie zu berücksichtigen. Die sich aus der phänomenologischen Analyse ergebenden Designanforderungen müssen sodann auf technik-generische und strukturelle Werte hin evaluiert werden, die es wiederum individuell, intersubjektiv und kollektiv zu erörtern gilt.

Ihde bleibt in seiner Analyse auf die individuelle Ebene beschränkt, gerade weil er sich phänologisch mit Intentionalität befasst, und erfasst daher keine interrelationalen, intersubjektiven oder gar kollektiven Strukturmomente. Dieses Geflecht an normativen, gegenseitigen Erwartungshaltungen und Anforderungen, das nach Nida-Rümelin seit 1997 als *Strukturelle Rationalität* bezeichnet werden kann, erfordert eine weitere qualitative Analyse der Lebenspraxis (Nida-Rümelin, 1997a). Diese beiden für ein normativ wünschenswertes Design notwendigen Betrachtungseinheiten werden sodann humanistisch interpretiert, um Bewertungskriterien einzuführen (Teil I, Kapitel 6.2). Ihde selbst stellt in seiner philosophischen Analyse das Verhältnis des Menschen zur Technik und die durch die Technik veränderten Wahrnehmungsinhalte oder Erfahrungsräume in den Mittelpunkt. Nur wenn der Technikbegriff jedoch um eine humanistische Perspektive erweitert wird, kann eine normative Haltung gegenüber digitalen Produkten in Entwicklungs- und Gebrauchskontexten begründet bezogen werden.

2.2 Strukturelle Rationalität

Die Paradigmen einer rationalen Entwicklungsarbeit (Teil I, Kapitel 1.1.1) und einer medialisiert transformierten Weltwahrnehmung (Teil I, Kapitel 1.1.2) werden in diesem Abschnitt um eine lebensweltliche normativitätsstiftende Praxistheorie ergänzt. Beide Paradigmen sind für ein Verständnis normativ angemessener digitaler Artefakte grundlegend, denn sie zeigen jeweils notwendige Handlungsräume und die Intentionalität der Entwicklungsarbeit sowie der Bedienung auf. Jedoch bleiben zunächst beide deskriptiv, da sie zwar die normativ relevanten Momente hervorheben, die man mit Hilfe der Methode der *levels of abstraction* abzubilden imstande sein mag (Teil II, Kapitel 4), aber eine Evaluation kann auf dieser Basis noch nicht gelingen. Dies liegt zum einen darin begründet, dass beide philosophischen Paradigmen sich argumentativ auf eine individuelle Ebene beziehen, weshalb diese philosophischen Ansätze nun auf lebensweltliche, gemeinschaftliche Strukturen rückbesonnen betrachtet werden müssen. Dies wird im gegenwärtigen Kapitel anhand von Nida-Rümelins *Theorie einer Strukturellen Rationalität* vollzogen.

Die rationalitätstheoretischen Modelle sowie der Versuch, diese gesellschaftlich und lebensweltlich einzubinden, betonen die Entscheidungs- und Handlungsmacht der Handelnden und zugleich die moralische Forderung, diese *vernünftig* auszuüben, d.h. nach den Maßstäben der Vernunft zu entscheiden. Die Vernunft offenbart sich nicht rein in der Nützlichkeit einer Technik oder Technologie, sondern ebenso in einer Berücksichtigung anderer Personen oder gar anderer Lebewesen, der Umwelt sowie des Gemeinwohls. Diese Multidimensionalität fängt Nida-Rümelin konzeptionell in seiner *Strukturellen Rationalität* ein: Hierunter begreift er eine pragmatische Normativität, die sich aus lebensweltlichen, gegenseitigen Erwartungshaltungen speist. Jene wiederum lassen sich durch eine Vielfalt von Gründen leiten, die konstitutiv für die deontischen Strukturen sind, die Lebensformen tragen. Hierdurch wird es möglich, unterschiedliche Bewertungskriterien menschlicher Entscheidung und Entscheidungsfindung, was als Deliberation zu bezeichnen ist, auf ihre Vernunftfähigkeit hin zu denken, und diese sogleich nicht als isolierte, sondern als kooperativ getragene lebensweltliche Struktur selbst aufzufassen. Entscheidend hierbei ist auch, dass sich die strukturelle Normativität in lebensweltlichen Praxen zeigt, weshalb sie genau dann ins Zentrum moralischer Analyse zu rücken sind, wenn Technik und Technologien intersubjektiv, gesellschaftlich oder ökologisch zu bewerten sind.

Die Vernunftfähigkeit des Menschen erfordert somit einen technikphilosophischen Ansatz, der über Technik und Technologien nicht nur hinsichtlich ihrer technischen Funktionalität oder ihrer Nutzen dienlichkeit reflektiert, sondern auch bspw. moralische, ästhetische oder ökologische Gründe zu inkludieren versteht. Technische Artefakte müssen hierbei den Forderungen der Vernunft entsprechen,

will man nicht von einer gegen die Humanität gerichteten, den Menschen unterdrückenden Technik und Technologie sprechen (Teil I, Kapitel 2.2.1).

Eine Analyse des Unterschieds von Technik als Werkzeug, welches als reines Mittel fungiert, und Technik als Medium, das die Welt auf andere Weise zugänglich oder verständlich macht, verdeutlicht die zunehmende Vermengung technischer Instrumente mit den menschlichen Fähigkeiten, die den Menschen als Menschen wesentlich konstituieren: nämlich über seine Entscheidungs- und Handlungsmacht selbst verfügen zu können (Teil I, Kapitel 2.2). Je weiter das technische Artefakt sich auf dem Kontinuum zwischen Werkzeug und Medium befindet, desto mehr greift es in die wesentlichen, vernünftigen Strukturelemente des Menschen ein: Digitale Artefakte platzieren sich sogar aufgrund ihres explizit medialen Charakters in der Urteilskraft und greifen auch in die menschliche Psychologie ein (Teil I, Kapitel 5). Um diese Zusammenhänge hervorheben zu können, ist, wie sich zeigen wird, die Theorie *Struktureller Rationalität* nach Nida-Rümelin (1997, 2001, 2019, 2020) besonders hilfreich.

Fragen der Normativität und insbesondere Fragen zur Güte von Technik und Technologie können daher nicht auf die Konsequenzen reduziert werden, sondern gerade das Ausgerichtetsein technischer Artefakte innerhalb menschlicher Lebensräume bedarf einer Herausstellung normativer Konstituenzen. Von daher ist es essentiell, den Einsatz technischer Artefakte in menschlichen (wünschenswerten) Praxen herauszuarbeiten, um eine durch die Implementierung hervorgerufene Verzerrung, Förderung oder Unterminierung eben jener zu erfassen. Die Herausarbeitung normativer Momente, die durch eben jene Einbindung technischer Artefakte tangiert oder gar verschoben werden, wird in eine fundierte Deliberation über einen vernünftigen Umgang miteinbezogen werden müssen. Diese Praxen müssen bereits in der Entwicklungsarbeit einer normativen Analyse unterzogen werden. Dies gilt in zweifacher Hinsicht: Zum einem muss die Entwicklungsarbeit, die Organisation des Herstellungsprozesses, selbst auf ihre wünschenswerten Praxen hin zu untersuchen und ggfs. um wünschenswerte Praxen zu erweitern sein. Zum anderen wird sich hier zeigen, dass strukturelle Lebensbeziehungen in einer ethischen Analyse bereits während des Gestaltungsprozesses Berücksichtigung finden können sollten, wenn strukturelle Werte identifiziert werden müssen, die sich im Design widerspiegeln sollen. Diese doppelte Prüfung wünschenswerter Praxen umfasst sowohl das Berufsethos als auch eine Theorie des guten, normativen Designs und bedenkt die gegenseitige Bedingtheit: Ohne ein Berufsethos, das sich in Organisationsstrukturen und der daraus resultierenden Förderung bestimmter, wünschenswerter Haltungen niederschlägt, lassen sich nur schwierig gute normative Technik und Technologie entwickeln. Hierzu ist Nida-Rümelins Theorie einer *Strukturellen Rationalität* hilfreich, soll dieses Beziehungsgeflecht nicht nur deskriptiv herausgearbeitet, sondern zugleich als normative Erwartungshaltung expliziert werden.

Nida-Rümelins *Strukturelle Rationalität* bringt den antiken Begriff der Praxis, der mehrdimensional strukturierten Handlungszusammenhänge innerhalb von Lebenskontexten, mit einer aufgeklärten individuellen Vernunftpraxis zusammen. Praxen werden getragen von Haltungen, die sie aber aufgrund ihres sozialen Charakters, eines Ethos, sogleich konstituieren. Moralische Deliberationen fordern die tradierten rationalen Strukturelemente heraus, indem sie sich kritisch zu vorgefundenen Strukturen verhalten und ggf. neue objektive, rationale Realitäten einfordern (Nida-Rümelin & Weidenfeld, 2018; Nida-Rümelin, 2019). Die strukturellen Netzwerke, in denen Menschen sich alltäglich bewegen, bieten normative Orientierung und binden Menschen normativ an ihre Mitwelt: Aufgrund von Erwartungshaltungen, die sie durch ihre Handlungen erfüllen oder verfehlen, werden Individuen als moralische Subjekte wertgeschätzt oder verachtet. Letzteres führt dazu, dass Strukturen radikal hinterfragt werden. Die *Strukturelle Rationalität* kann daher Strukturelemente deskriptiv nachzeichnen, insofern Lebenspraxen erkannt und Lebensmuster beschrieben werden können. Sodann kann die Normativität eben jener gegenseitigen Verpflichtungsstrukturen identifiziert und hinsichtlich Kriterien der Vernunft bewertet werden. Bezüglich Technik und Technologien müssen Praxen dementsprechend auf ihre Erwünschbarkeit oder Akzeptabilität hin hinterfragt werden, um im darauffolgenden nächsten Schritt technische Artefakte dahingehend zu bewerten, wie Technik diese Strukturen transformiert.

Nida-Rümelin wendet sich in seiner Konzeption einer praktischen Vernunft von klassischen Theorien instrumenteller Rationalität oder strukturalistischen Interpretationen menschlicher Interaktion ab. Zugespitzt formuliert, stehen sich in der klassischen Auffassung die individualisierte Perspektive menschlicher Vernunft und eine Vernunft gegenüber, die sich auf kollektive Strukturen reduziert. Während die eine Seite die kognitiv-berechnende, auf Prinzipien beruhende Vernunftfähigkeit in Abhängigkeit einzelner Akte oder rationaler Akteure bestimmt, beschwört die andere Seite die Anpassung und die Internalisierung des Richtigen durch Aufnahme kollektiver Werthaltungen. Die Auflösung dieses angespannten polarisierenden Verhältnisses gelingt Nida-Rümelin mit seiner Theorie einer *Strukturellen Rationalität*: Der *homo oeconomicus* und der *homo sociologicus* werden versöhnt, indem die regelgeleitete Vernunft an die Lebenswelt zurückgebunden wird. Die menschliche Vernunftfähigkeit wird nicht mehr aus einer reinen individuellen Perspektive betrachtet; sei diese gebunden an die handelnde Person oder nur an einen einzelnen Akt, wie es im Rahmen des methodologischen Individualismus ökonomischer Theorien geschieht: Selbst spieltheoretisch eingebettete Rationalitätsgebote zeichnen sich dadurch aus, dass sie nicht alle Spieler gemeinsam als Kollektiv adressieren, sondern die Forderung jeweils individuell gestellt wird und Gemeinsamkeit nur dadurch entsteht, dass jeder weiß, dass dieses Gebot an jeden Einzelnen gerichtet ist. Die klassische Entscheidungstheorie absolutiert so den individuellen Rückzug zu sich selbst, indem der Handelnde sich gegenüber einer Welt wiederfindet, auf die er zu reagieren versucht, indem er seinen Nutzen – gemeinsam nur insofern, als dass alle ihren Nutzen hinsichtlich dieser Situation ausbauen – optimiert. Alles

andere sind Bedingungen jeder einzelnen Handlungseinheit. Die Anwendung von reinen Vernunftprinzipien auf eine vorgefundene Welt lässt das Individuum passiv und isoliert zurück. Das Individuum folgt treu seiner punktuellen Handlungswahl, ohne seine Mitwelt zu berücksichtigen. Eine dergleichen praktische Vernunft ist weder in der Lage, strukturelle Episoden eines Lebens zeitlich zu verorten, noch besteht eine Möglichkeit, Vernunftstrukturen begreifen oder bekräftigen zu können. Das Individuum verliert sich im jeweiligen Moment, der unabhängig von seiner Mitwelt (*vertikale Kohärenz*) und unabhängig von Zeitlichkeit (*horizontale Kohärenz*) nachgezeichnet wird.

Eine ausschließlich individualistisch verstandene Vernunft, die sich auf den jetzigen Zustand und dessen unmittelbare Erfüllbarkeit reduziert, kann Handlungen hinsichtlich einer horizontalen und vertikalen Kohärenzanforderung nicht vernünftig greifen (Nida-Rümelin, 2020). Denn es gibt keine Möglichkeit, Einzelakte in Abhängigkeit eines vernünftigen Ganzen zu denken. Weder eine zeitliche Einbindung der Einzelakte noch eine individuelle oder gesellschaftlich-strukturelle Einbettung in Praxen können als Kriterien einer vernünftigen Ausgestaltung des eigenen Lebens postuliert werden. Die Handlung als Teil einer Praxis und somit auch als Teil einer Lebensform muss sich jedoch zeitlich sowie interpersonal und kollektiv einordnen lassen, ohne dabei die minimalen Kohärenzbestimmungen der Vernunft, die sich bisher auf die individuelle Rationalitätsperspektive reduzieren, zu verletzen.⁵ Dieses Instrumentarium klassischer Entscheidungstheorie lehnt Nida-Rümelin daher nicht ab, sondern integriert es in eine praktische Vernunft, die der liberalen Forderung nach einer inhaltlichen Unbestimmtheit folgt, aber zugleich eine Vielheit von Handlungsgründen zulässt (Nida-Rümelin, 2001).

Strukturelle Aspekte der Praxis und die Einbeziehung des Individuums in ein Geflecht kohärenter Strukturen beinhalten die Idee eines *right fit*. Die einzelne Handlung muss in eine Mehrdimensionalität vernünftiger Lebensformen integrierbar sein: Nur so kann praktische Vernunftfähigkeit kohärenzstiftend wirken oder lebensweltliche Strukturen aufgrund mangelnder Kohärenz kritisch hinterfragen. Ist keine vernünftige Einbindung der Einzelhandlung möglich, folgt ein gradueller Verfall stabiler, inkohärenter Strukturen. Eine vernünftige Handlungsausführung ist dann aufgrund überholter lebensweltlicher Strukturen nicht mehr auslebbbar, weshalb eben diese normativen Anforderungen neu ausgerichtet und neue Handlungspraxen erst etabliert werden. Um von einer vernünftigen Ausgestaltung sprechen zu können, müssen folglich intrapersonal die Minimalbedingungen der Rationalität erfüllt sein, aber auch Integrität im Sinne einer zeitlichen und strukturellen Einbindung gewährleistet werden können (Nida-Rümelin, 2020). Die Vernünftigkeit einer intrapersonellen Lebensform umfasst somit

⁵ Hierunter fallen Annahmen der Monotonie, Transitivität und Vollständigkeit, die eine Mathematisierung der Entscheidungstheorie ermöglichen (Nida-Rümelin & Schmidt, 2000).

interpersonelle sowie kollektive Erwartungshaltungen. Diese Momente sind in Anlehnung an Aristoteles in einem Praxisbegriff zu finden, der die Einzelhandlungen zu einer großen Handlung zusammenbringt. Nida-Rümelin (2020, S. 181) fragt von daher: „[W]as wollen wir nun unter einer Handlung verstehen? Nur den jeweiligen Akt (*pragma*) oder auch die *praxis*? Ist die *praxis* eine Folge von Akten, die jeweils für sich, also gegeben die epistemischen und optionalen Bedingungen des Augenblickes, rational zu sein haben, oder sind die Akte nicht erst im Kontext der *praxis* rational?“

Die antike Handlungsphilosophie hatte bereits im Blick, was Nida-Rümelin als *Strukturelle Rationalität* bezeichnet: die Einfügung der einzelnen Handlung in einen größeren Kontext, eine Praxis. Im aristotelischen Verständnis ist die einzelne Handlung jeweils Ausdruck einer Haltung (*hexis*), einer wertenden Einstellung (*prohairesis*), eines Ethos oder einer Tugend (*areté*). Eine Tugend ist hierbei nicht nur als Fähigkeit oder als Internalisierung tradierter Rationalität zu begreifen, sondern umfasst gerade auch Einstellungen und Haltungen, die auf jeweils individuellen, vernünftigen Einsichten und Entscheidungen beruhen (Nida-Rümelin, 2020).

Daher kann das isolierte Individuum nicht mehr in einer Passivität gegenüber einer Um- und Mitwelt im Rahmen vernünftiger Lebensweisen gedacht, sondern muss als ein Individuum verstanden werden, das mit anderen gemeinsam eine vernünftige Lebensform performativ affiziert. So kann Nida-Rümelin einer Dichotomie von technischer und normativer Rationalität entgehen, denn das Individuum kann seine Entscheidung nicht unabhängig von seiner Um- und Mitwelt wählen, sollen seine Handlungen den Kriterien der Vernunft entsprechen. Zugleich beinhaltet eine praktische Vernunft im Sinne Nida-Rümelins auch eine vernünftige Mittelwahl, einen Rohbeck'schen positiven Instrumentalismus (Teil I, Kapitel 1.1.1. (a) (1)), denn gute Handlungen bedürfen der Möglichkeit, diese auch erfahrbar, d.h. innerhalb einer Mit- und Umwelt auslebbar, zu machen. Aber auch Mittel, seien diese Handlungen, Geräte oder Institutionen etc., lassen sich nicht allein durch einen Grund rechtfertigen, sondern bedürfen wiederum einer eigenen Betrachtung und sind von daher nicht allein instrumentell rechtfertigbar. Da nur Gründe eine Handlung normativ gebieten können, können auch nur normative Gründe die geeigneten Mittel rational gebieten (Nida-Rümelin, 2020). Normativität kann sich somit nicht durch die Anforderungen der Vernunft an sich selbst konstituieren: Normativität ist nicht allein durch Prinzipien, die eine vernünftige Binnenstruktur ermöglichen – wie es der Begriff der *structural irrationality*, den Scanlon (2007) für eine unzureichende vernünftige Binnenstruktur von gegebenen Datenpunkten, seien diese Wünsche oder Intentionen, einführte – zu erklären (Kiesewetter, 2017). Eine Trennung von Grund und Entscheidung führt somit zu einer Spaltung von Moral und Rationalität, die lebensweltlich nicht getragen werden kann: Weil ein angespanntes Verhältnis zu einem Arbeitskollegen vorherrscht und dieses Arbeitsverhältnis nicht mehr gewollt wird, kann daraus rational nicht

geboden sein, diesen Arbeitskollegen zu mobben oder gar zu töten, um ihn loszuwerden. Beide Handlungsoptionen sind moralisch nicht vertretbar, sind inakzeptabel. Diese Handlungen werden lebensweltlich auch nicht als vernünftig deklariert.

Ausgehend von seiner Forderung nach dem Primat der Lebenswelt muss sich eine Theorie an lebensweltlichen Strukturen beweisen, was nochmals den Umstand hervorhebt, dass nicht einfach Mittel zugelassen werden können, weil diese angeblich instrumentell geboten seien, aber sich kein guter Handlungsgrund auffinden lässt, der für die Handlung spricht. Normativität ist folglich auf der Ebene der Gründe zu finden, denn nur in einer Praxis des gemeinsamen Gründegebens und -nehmens konstituieren sich Anforderungshaltungen bzw. werden dort objektiv aufgefunden (Nida-Rümelin, 2009).

Nida-Rümelin folgend sind Intentionen Entscheidungen, die Deliberationen – die wiederum als Abwägung von Gründen verstanden werden – unter- oder abbrechen (Nida-Rümelin, 2020). Sie werden somit handlungsgebietend wirksam. Für Gründe gibt es oftmals keinen allgemeingültigen Kriterienkatalog, sondern sie beweisen sich im Rahmen der Praxis des Gebens und Nehmens. In dieser Praxis werden Gründe als gute Gründe für eine gegebene Handlung anerkannt. Gründe begründen Überzeugungen, Handlungen genauso wie emotive Einstellungen (Nida-Rümelin, 2020): Sie lassen Einstellungen, Überzeugungen oder Handlungen mal mehr oder weniger vernünftig werden. Vernunft wird so von Lord und Wedgwood als eine angemessene Reaktion oder Antwort (*response*) auf die *richtigen* Gründe verstanden (Lord, 2015, 2018; Wedgwood, 2014, 2017). Diese guten Gründe müssen nicht nur Konsistenzbedingungen der Vernunft erfüllen, sondern ebenso deren Kohärenzanforderungen (Nida-Rümelin, 2020). Die jeweiligen Präferenzen repräsentieren somit nicht nur die jeweiligen propositionalen Einstellungen des Augenblicks, sondern Haltungen, Tugenden, strukturelle Entscheidungen – das, was dem individuellen Leben und der sozialen Interaktion erst eine vernünftige Struktur verleiht.

Nida-Rümelin verabschiedet sich von der Engstirnigkeit aller Theorien praktischer Vernunft, die eine Handlungswahl auf nur ein Vernunftprinzip begründet denken, ohne ihnen jedoch gänzlich praktische Relevanz abzusprechen. Hierzu zählen deontologische, utilitaristische, tugendethische, aber ebenso konstruktivistische Ansätze der Ethik. Selbstverständlich gibt es Handlungen, die sich an der Nutzenmaximierung, sogar zusätzlich an der Maximierung des Selbstinteresses ausrichten, oder Handlungen, die sich durch Gerechtigkeitskriterien bestimmen lassen. Aber dies ist nicht für rationale Handlungen der Fall. Eine Handlung, die sich beispielsweise an einem Rollenbild ausrichtet, kann nicht deshalb als irrational gelten, weil sie nicht durch eine Nutzenmaximierung zu erklären ist. Für die Handlungsbegründung wird kein Rekurs auf ein Nutzenmaximierungstheorem benötigt: Eine Tugend zu erfüllen, ist als Handlungsgrund ausreichend. Auch die Idee der *bounded rationality*, der Versuch

einer Einbettung der Vernunft in alltägliche Lebensrealitäten, um Abstand von klassischen Idealvorstellungen ökonomischer Vernunft zu nehmen, wie sie von Herbert Simons, aber auch von Gerd Gigerenzer vorgenommen wird, kann der Vielfalt rationaler Handlungsbegründungen nicht gerecht werden (Gigerenzer & Selten, 2001; Simon, 1997). Hier wird der epistemische Horizont aufgrund eingeschränkter kognitiver Fähigkeiten oder selektiver Wahrnehmung begrenzt, die von der Norm einer Ideal-Vernunft abweicht, die sich nach wie vor instrumentell fassen lässt. Insofern ist hier keine radikale Kritik an einer kalkulierenden Vernunft zu finden, die quantifizierend optimiert. Vielmehr spielen menschliche Unzulänglichkeiten und Ungenauigkeiten eine begrenzende Rolle in der Festlegung menschlicher Vernunftfähigkeit. Somit bleibt aber das Individuum in diesen empirischen Versuchen, menschliches Verhalten lebensweltlich einzufangen, nach wie vor seiner Vielfalt beraubt. Denn diese Versuche ignorieren schlichtweg, so Nida-Rümelin, dass das Individuum aus unterschiedlichen Gründen an- und umgetrieben werden könne, die allesamt den Regeln der Vernunft gehorchen. Die Ignoranz dieser Gründevielfalt sowie die Ausgrenzung von Handlungen aus dem Bereich der Vernunft, raube dem Individuum jedoch seine Humanität.

Ein Vater handelt, weil er die Rolle des Vaters einnimmt und mit dieser Rolle bestimmte Verpflichtungen einhergehen. Er handelt nicht, weil er seinen Eigennutzen oder den Nutzen aller oder den Nutzen anderer maximiert, gleichgültig wie der Nutzen inhaltlich bestimmt wird. Angenommen, der Sohn fällt hin und schlägt sich sein Knie auf und der Vater hilft ihm tröstend und ermutigend auf. Er tut dies, weil er Vater ist, weil er diese Beziehung eingegangen ist und er es als wesentlichen Teil seiner Vater-Rolle anerkennt, sich gegenüber seinem Sohn fürsorglich und empathisch zu verhalten. Er muss keinen Nutzen haben, er muss keine Nutzen-Kosten-Abwägung vornehmen, um zu dieser optimalen Handlung des Tröstens zu gelangen. Auch Außenstehenden reicht es aus, zu wissen, dass er Vater ist, um die Handlung als emotional und rational angemessen zu verstehen. Allerdings, so Nida-Rümelin, bedeutet das nicht, dass der Vater notwendig so handeln musste, dass also das Vatersein ihn zu dieser Handlung nötigte. Das Vater-sein kann als handlungsanleitender Grund gelten, muss es aber nicht.

Es ist gerade von großer Bedeutung für den Nida-Rümelinschen Ansatz, dass es einen Raum für praktische Deliberation, für die Bewertung meiner Wünsche, Interessen, Pflichten, Rollen etc. gibt. Dieser Deliberationsraum zeigt sich in der Gestaltung eines individuellen Lebensentwurfs innerhalb einer geteilten Lebensform, welche die Idee der Autorschaft zum zentralen Kriterium eines gelungenen Lebens macht. Zudem besteht durch den Gründeaustausch ein Deliberationsgleichgewicht zwischen individuellen und strukturellen Handlungsgründen, das eine kohärentistische Lebensführung ermöglicht. Die aus diesen Deliberationen resultierenden Entscheidungen begründen zudem Haltungen, die unsere Handlungsebene maßgeblich – ohne großartigen Rekurs auf Reflektivität nehmen zu

müssen – bestimmen, d.h. dass eine vernünftige Handlung nicht notwendigerweise nur unter aktiver kognitiver Anstrengung zu erreichen ist (Grimm, 2018; Grimm, Keber & Zöllner, 2019).

Das Individuum kann als autonom, rational und integer gedacht werden. Strukturelle Rationalität bejaht die Pluralität von Handlungsgründen und ermöglicht so eine intra- sowie interpersonelle Integrität der Person. Diese Interaktion des Begründens und der Entdeckung von Gründen ist wesentlich bestimmt durch Übung mit anderen: Im Austausch mit anderen erfahren Individuen, welche Gründe als schlechter oder besser gewertet werden können. Sie erlernen und üben sich in der Anwendung ihrer Vernunftfähigkeit – sie bilden eine Urteilskraft aus, die immer auch von Erfahrung lebt und sich nicht nur auf die reine *Episteme* reduzieren lässt.

Mit der Entwicklung der *Theorie einer strukturellen Rationalität* gelingt es Nida-Rümelin, eine Konzeption praktischer Vernunft zu entwickeln, die unserer alltäglichen Interaktions- und Bedeutungspraxis Rechnung trägt. Praktische Vernunftfähigkeit muss die Gesamtheit unserer lebensweltlichen Beziehungsmuster, mit deren spezifischen Handlungsrechtfertigungen umfassen können. Vermag eine Theorie praktischer Vernunft dies nicht zu leisten, verzerrt sie die lebensweltliche Praxis praktischer Rationalität und führt zu realitätsfernen, losgelösten Theorien menschlicher Handlungsfaktizität (Nida-Rümelin, 2020).

Dieser humanistische Anspruch an Theorien praktischer Vernunft lässt sich nicht nur an die Ökonomie oder andere Systemtheorien richten, sondern auch an die Entwicklung und den Umgang mit neuen Techniken und Technologien. Der Aufruf, dass Maschinen Menschen dienen sollen und nicht der Mensch seine Handlungen und seine Lebensweise an die Gegebenheiten der Systeme anzupassen hat, erhält dadurch eine ganze andere Bedeutung: Sachzwänge bedürfen einer erneuten normativen Reflexion, sollen sie nicht Personen in ihrer Humanität, in ihrer autonomen Vernunftbegabung beeinträchtigen. Sollten Unterwerfungsszenarien entstehen oder wünschenswerte Praxen durch die Einführung von Technik und Technologie gehemmt werden, muss diese Technik und Technologie hinsichtlich dieser Lebensform auf ihre Akzeptanz hin hinterfragt werden. Eine positive Ausgestaltung technischer Artefakte kann sich gerade nicht auf eine Folgenabschätzung reduzieren, sondern muss Lebenspraxen in ihrer normativen Gestalt aktiv berücksichtigen, d.h. sie muss digitale Technologie und Technik im Hinblick auf Lebenspraxen bedenken und diese Praxen kohärentistisch hinterfragen können. Digitale Artefakte müssen sich in lebensweltliche Strukturen einbinden lassen, wenn sie als vernünftig ausgewiesen werden sollen. In diesem Zusammenhang lassen sich sowohl für die Entwicklung als auch für die Nutzung bedeutsame strukturelle Werte ausfindig machen. Daher bedarf es einer genaueren Analyse, wie sich Normativität lebensweltlich offenbart.

2.2.1 Rationalität und Sozialität

Strukturelle Werte, die Lebensformen konstituieren – gleichgültig ob diese sich in Konventionen, Normen oder Gesetzen niederschlagen – sind für eine Bewertung digitaler Artefakte notwendig, denn gerade der Einsatz digitaler Artefakte in menschlichen interaktiven und intersubjektiven Handlungsketten kann zu einer Verschiebung oder Kritik an den normativen Ansprüchen führen, die diese Handlungssysteme tragen. Wie sich Normativität begründet, ist somit für ein grundlegendes Verständnis vernünftiger Strukturen und der Frage nach einem erstrebenswerten und angemessenen Einsatz digitaler Artefakte notwendig. Zudem kann so die Tragweite des Phänomens der digitalen Transformation verstanden werden. Ebenso wird dadurch sehr deutlich, wieso in diesem Zusammenhang verstärkt von einem Werte- oder Kulturwandel gesprochen wird.

Wie bereits mehrfach betont wurde, werden Handlungen, sofern diese auf autonomen Entscheidungen gründen, durch normative Gründe geleitet. Diese Gründe beruhen auf Tatsachen und einer (hoffentlich) vernünftigen Abwägung. Der Gründeaustausch ist somit zentral für ein Verständnis von normativen Anforderungen und ihrem Anspruch. Durch die Ausarbeitung von Sozialität, Rationalität und Normativität kann man sich dem annähern, was ein vernünftiger Umgang, eine vernünftige Entscheidung oder Handlung sein könnte. Dies ist bedeutsam, wenn verstanden werden soll, was ein verantwortungsvoller Umgang mit digitalen Artefakten bedeutet und wie der Trend hin zu mehr *ethics by design* erfolgreich umgesetzt werden soll.

Kommunikation als intentionaler Akt menschlicher Autonomie beruht auf dem Austausch von Gründen und deren Realisierung durch Urteile, Handlungen oder Haltungen. Erst die Einbettung des gerichteten, bedeutungsvollen und deshalb gewollten kommunikativen Austauschs in ein Geflecht gegenseitiger Erwartungshaltungen erlaubt es, von einem rational-normativen Netzwerk menschlicher Strukturen zu sprechen. Dies impliziert, dass die menschliche Vernunftfähigkeit nicht isoliert, sondern nur durch Sozialität begreifbar und rechtfertigbar wird, denn erst die begreifbaren gegenseitigen Erwartungshaltungen begründen Deliberationsstrukturen, die es ermöglichen, Entscheidungen und dementsprechend Handlungen begründen zu können (Nida-Rümelin, 2011a). Das sich wechselseitig bedingende Spiel, die Gegenseitigkeit der menschlichen Begegnung, erschafft Kommunikationsstrukturen, die aufgrund der menschlichen Fähigkeit, sich von den jeweiligen situations-inhärenten Gegebenheiten abgrenzen zu können, rational erfassbar werden. Besonders pointiert hat dies George Herbert Mead in seiner Theorie des signifikanten Symbols ausgearbeitet (Mead, 1934). Grundlegend für ein Verständnis dahingehend, sich als rationales Wesen zu begreifen, also sich als ein Individuum zu verstehen, das Handlungen beabsichtigt und regelgeleitet ausführt, ist ein Bewusstsein, das sich auf diesen Sachverhalt bezieht. Um dies zu erreichen, muss sich das Bewusstsein als von der Situation und vom anderen verschieden erkennen, d.h. es muss die Situation transzendieren, sich von den

interaktiven Prozessen distanzieren, um diese Spaltung wahrnehmen zu können. Die Erkenntnis dieser Separierung ermöglicht eine reflexive Bewusstseinsstruktur, die Mead als Selbstbewusstsein bezeichnet (Mead, 1934). Diese Reflexivität entsteht, so Mead, gerade weil sich Individuen jeweils in die Perspektive des oder der vielen anderen (*the generalized other*) hineinzusetzen vermögen, d.h. die Rollen anderer einnehmen (Mead, 1934) und sich somit aus ihrer unmittelbaren, rein subjektiven Erfahrungswelt lösen können. Hierunter ist die Möglichkeit gefasst, auf bestimmte Erwartungshaltungen des Gegenübers mit einem bestimmten Verhalten zu reagieren. Wichtig jedoch ist, dass diese konfrontative Interaktion das Individuum von den prozessualen Strukturen zugleich trennt, was erst eine reflexive Bewusstseinsstruktur ermöglicht: Das Individuum erfährt sich selbst als von den anderen sowie vom Interaktionsprozess unabhängig und kann so sein eigenes Verhalten in Bezug auf die anderen, die Umwelt und sich selbst zugänglich machen sowie kognitiv-rational erfassen. Eben dieses Zusammentreffen ist für rationales Verhalten konstitutiv, denn erst durch diese Interaktion kann zum einen der Einzelne sich selbst als Objekt wahrnehmen, sich seiner selbst bewusstwerden, und zum anderen signifikante Symbole als Bedeutungsstrukturen begründen, d.h. seinem Verhalten, seinen Gesten, Lauten etc. Bedeutung aufgrund des interaktiven Kontextes zusprechen. Selbstwahrnehmung impliziert gerade, dass sich das Individuum von sich selbst als reines Erfahrungssubjekt distanzieren kann, denn es geht nicht mehr nur in der Wahrnehmung auf oder ist rein in der Teilnahme zugegen (Mead, 1934). Die Interaktion ermöglicht durch dieses Aufbrechen reiner unmittelbarer Erfahrung, wie sie James (1982) eindrücklich in seinem *stream of consciousness* phänomenologisch darstellt, vernünftige Handlungen. Damit Rationalität systematisierend auf das Verhalten, den Austauschprozess und die Interaktion mit anderen einwirken kann, bedarf es einer Semantisierung der Situation. Erst dadurch werden Erfahrungen klassifizierbar und planbar. Der dadurch entstehende planvolle Umgang mit eigenen Verhaltensweisen führt dazu, dass die Individuen ihre Ausdruckvariationen zweckgerichtet einzusetzen vermögen, weshalb dieses Verhalten nun mehr als Handlung verstanden werden kann (Gosepath, 1999). Die soziale Einbettung des Individuums ermöglicht somit erst einen kontrollierten, vernünftigen Umgang mit den eigenen Verhaltensweisen und begründet dadurch individuelle Intentionalität vermittelt durch kooperatives Handeln, denn durch seinen kommunikativen Akt löst der Absender nicht nur beim Empfänger, sondern auch bei sich selbst die Reaktion, die angemessene Antwort auf seine Botschaft aus, weshalb die Geste nun als signifikantes Symbol bezeichnet werden muss. Der Geste kommt Bedeutung zu, weil sie beantwortet wird, weil der Adressat die Botschaft erfüllt, weil er auf die Geste des Absenders angemessen reagiert. Bedeutung basiert somit auf einer Interaktion von Absender und Empfänger, einer Distanzierung und Vorwegnahme der Reaktion durch den Absender sowie der tatsächlichen Reaktion des Empfängers (Miller, 1973). Aufgrund dieser durch die Teilnahme an der Interaktion entstandenen Bedeutungszusammenhänge kann der Absender willentlich, d.h. intentional, seine Handlungsabfolgen steuern,

denn nun ist er in der Lage, die Reaktionen bzw. Antworten seines Gegenübers vorwegzunehmen. Kommunikation ist daher, wie Miller (1973, S. 73) betont „to evoke action, responses that are parts of the social process of adjustment, responses that lead to the completion of the act“. So kann das Selbst die Situation nach Bedeutungsketten ausrichten und strukturieren. Verhalten wird mittels der geteilten Bedeutung überschaubar und planbar: für jeden Einzelnen, aber auch für die Gemeinschaft. Aufgrund dieser intentional steuerbaren Teilnahme an gemeinschaftlichen Situationen kann von intelligenten und rationalen Handlungen gesprochen werden (Mead, 1934). Von intelligenten oder rationalen Handlungen spricht Mead, sofern das Selbst vermittelt über den verallgemeinerten Anderen die Situation richtig erfasst und diese entsprechend erwidern kann. Es muss die richtige, d.h. als Erwidern angemessene Antwort in Form einer intentionalen Handlung ausgewählt werden (Mead, 1934). Mead legt bereits ein Fundament für eine Handlungstheorie, welche die Wahl einer Handlung im Hinblick auf die Situationserfüllung zu begründen weiß und sich nicht auf die Optimierung meiner oder der Interessen anderer reduzieren lässt. Diese auf der Sozialität gründende Idee führt Nida-Rümelin in seiner *Theorie einer Strukturellen Rationalität* aus, um sich des formalisierbaren Rahmens klassischer Entscheidungstheorie nicht entledigen zu müssen, aber zugleich dessen Widersprüchlichkeiten, die sich aus der Kontextlosigkeit ergeben, entgegen zu können.

Ohne Kommunikation ist Rationalität nicht denkbar und ohne jemand anderes, ohne Sozialität, ist Sprache nicht denkbar (Miller, 1973). Erst durch die Überführung des eigenen Standpunkts in die Perspektiven des anderen oder der vielen anderen, der Erschaffung des Selbst durch die internalisierten anderen, entstehen innere Dialoge. Denken in Form von Sprache tritt ein und somit die Möglichkeit, Handlungssituationen zu bestimmen und zu beabsichtigen (ebd.). Der intentionale Einsatz signifikanter Symbole aufgrund einer kontinuierlichen Gegenwart des *generalized other* begründet das denkende Selbst. Der verallgemeinerte Andere umfasst das Verhältnis aller Rollen zueinander und deren Verhältnis zu einer bestimmten sozialen Situation. Insofern bildet der verallgemeinerte Andere das Erwartungsgeflecht einer sozialen Situation ab und ist dadurch wesentlich normativ. Das zum Selbst werdende Individuum nimmt die Rollen aller an dieser Interaktion, spricht in dieser sozialen Situation, Teilhabenden an und setzt diese ins Verhältnis untereinander sowie zum gesamten sozialen Prozess der sozialen Situation (Miller, 1982). Dieses Netzwerk, das der verallgemeinerte Andere repräsentiert, fasst die möglichen Reaktionen jedes einzelnen Teilnehmers dieser sozialen Situation zusammen, d.h. dieses Geflecht spiegelt alle Antizipationen, alle Erwartungshaltungen und gegenseitigen Verpflichtungsstrukturen wider. Handlungen können so zu und in Situationen angeordnet und organisiert werden. So kann aufgrund von Bedeutungsketten ein Austausch beabsichtigt weitergeführt und getragen werden.

Vernunftfähigkeit entsteht in der aktiven Begegnung mit anderen. Das Aufeinandertreffen bedingt ein reflexives Bewusstsein, das in seiner intersubjektiven Allgemeinheit rationale Intentionalität bedingt

und dadurch zugleich normative Ansprüche konstituiert. Die Sozialität, die konstitutiv für eine Strukturelle Rationalität ist, hebt die Wichtigkeit ethischer Forderungen innerhalb von Lebenspraxen hervor. Ohne diese Anforderungshaltung und deren Erwidern werden vernünftige Begründungsstrukturen nicht erklärbar.

2.2.1.1 Deontische Strukturen

Diese Verwobenheit der gegenseitigen Erwartungshaltungen und deren Erfüllbarkeit, das praktische normative Wissen und die Erfahrung dahingehend, wie auf individuelle Verhaltensweisen und Handlungen eingegangen werden soll und wie diese intentional eingesetzt werden können, um soziale Interaktionsstrukturen durch eine Bedeutungszuschreibung kontrollierbar zu machen, ermöglicht einen vernünftigen Umgang. Intentionale Akte basieren gerade darauf, dass gewusst wird, was sie jeweils bedeuten, sodass man zurecht betonen muss, dass Verhalten der Intentionalität bedarf, um überhaupt als rationale Handlung klassifizierbar zu werden. Verhalten wird verständlich, wird der Rationalität zugänglich, weil begriffen werden kann, was mit der jeweiligen Handlung beabsichtigt wird, worauf sie hinauslaufen soll, welche Reaktion bzw. Antwort eingefordert wird.

Die dadurch entstandenen deontischen Beziehungsgeflechte werden bekräftigt, indem Individuen durch ihre Verhaltensweisen und Haltungen diese immer wieder re-aktualisieren: Eine Bitte, etwas zu erledigen, ist nur eine Bitte, weil alle wissen, welche Reaktion erwartet wird, um diesen Akt als Bitte zu erfüllen (oder eben nicht zu erfüllen). Eine Bitte wird erst dadurch zur Bitte, wenn sich der Absender die Bitte-Einstellung zu eigen macht, sich also in einem intentionalen Zustand befindet und diesen in seinem Sprechakt manifestiert. Er bittet absichtlich und der Empfänger reagiert angemessen auf diese Bitte, um diesen sozialen Akt als Bitte gemeinsam zu vervollständigen (Nida-Rümelin, 1993b, 2001). Der Sprecher und sein Gegenüber, der Adressat, haben gewisse gemeinsame Erwartungshaltungen, die auf Einstellungen beruhen: Diese Einstellungen charakterisieren bestimmte, intentionale Zustände, dass etwas zu tun ist (Nida-Rümelin, 1993b). Über dieses praktische Wissen dazu, was eine Bitte von einem Befehl unterscheidet und welche Handlungen damit erwartet werden, verfügen alle Teilnehmer – es handelt sich somit um genuin praktisches, normatives Handlungswissen (Mead, 1934; Nida-Rümelin, 1993b). Dieses Wissen muss als ein geteilter Bedeutungshorizont verstanden werden, denn um mit Mead (1973, S. 147) zu sprechen: „[I]ndividual perspectives arise out of a social perspective in which the members of the group have common attitudes and take for granted an unquestioned world that is there.“ Vernunft ist folglich wesentlich an Sozialität gebunden, denn ohne die vielen anderen gäbe es keine Möglichkeit, Handlungen zu strukturieren und diese als Verpflichtungsbeziehungen zu erleben. Normativität entspringt der Lebenswelt, dem sozialen Geflecht, das durch gegenseitige Erwartungshaltungen und durch die Erfüllung von gegenseitigen For-

derungen einen vernünftigen, weil rationalen Umgang miteinander gewährleistet. Durch das Aufbrechen der unmittelbaren Erfahrungswelt wird vernünftige Freiheit als Autorschaft (Nida-Rümelin, 2020, VII § 4) er- und auslebbar: Planvolle, intentionale Handlungsverknüpfungen innerhalb einer geteilten Welt sind unabdingbar für autonom handelnde Individuen.

Die Möglichkeit, sich in jemand anderes hineinversetzen zu können, kann man mit Strawson auch als *reactive attitudes* bezeichnen, um zur kognitiv-rationalistischen Sichtweise auch die praktische Handlungsebene hinzuzudenken. Ähnlich wie Mead greift auch Strawson in seinem Aufsatz *Freedom and Resentment* die Zugehörigkeit, die immer schon stattfindende strukturelle Einbettung, auf, die er als Grundgerüst der menschlichen Lebensform versteht (Strawson, 2008). Eine Distanzierung von der Teilnahme an dieser Lebensform, eine Herauslösung aus diesen Strukturvorgaben, indem der Einzelne das Beziehungsgeflecht objektiviert, ist zwar möglich und für (wissenschaftliche) Theoriebildungen auch unabdingbar, aber Strawson betont zugleich die Unmöglichkeit, sich im direkten Austausch mit anderen objektiv verhalten zu können: In der menschlichen Interaktion verhält sich jeder authentisch, d.h. er weist Haltungen auf, er re-agiert, er empfindet und entkommt dieser interaktiv-verbindlichen Situation nicht. Der Mensch kann sich nicht liebend gegenüber einer Person verhalten, wenn er nicht liebt, selbst wenn er verstünde, was es zu lieben bedeutet. Dies beruht nach Strawson auf dem fundamentalen Unterschied zwischen Intelligenz und Menschlichkeit (*intelligence and humanity*): Der Mensch kann sich gegenüber anderen niemals objektiv verhalten, denn sobald er sich in einer Interaktion, in einem Dialog mit einem Gegenüber befindet, partizipiert er in der Sphäre der Menschlichkeit: Er kann seinem Mensch-Sein und der(n) menschlichen Lebensform(en) nicht entkommen. Eine Vogelperspektive ist denkbar, um Theorien zu bilden, Urteile zu formulieren, kann aber nicht direkt in einer Interaktion ausgelebt werden. Eine Theorie hat somit auch immer Auswirkungen auf die existentiellen Strukturen, indem sie Angenommenes verändert und somit Perspektiven verschiebt. So ist es möglich, dass neu entdeckte Wahrheiten alte Vermutungen des Platzes verweisen. Insofern aber das Individuum an Interaktionen, an gemeinsam geteilten Strukturen, teilnimmt, weisen Individuen *reactive attitudes* auf (Strawson, 1974, S. 6). Sie kommen gar nicht umhin, denn die gegenseitige Interaktion verlangt es gleichsam. Die gegenseitigen Erwartungshaltungen konstituieren Gemeinschaften und gegenseitige Verpflichtungen entstehen: Alle sozialen Beziehungen, seien diese Vater-Kind, Lehrer-Schüler, Arzt-Erkrankte etc., konstituieren Erwartungshaltungen, die erfüllt werden müssen, um in diesen Interaktionen auch angemessen zu agieren, um auch als Vater, Lehrer oder Arzt erkennbar zu sein. Missachtungen und Verletzungen, also ein Mangel an Anerkennung oder angemessener Handhabung der strukturellen normativen Momente, führen zu Kritik, Unverständnis und manches Mal auch zu moralischem Entsetzen. Aristotelisch gesprochen, wollen Menschen diese strukturellen, deontischen Vorgaben sogar gut erfüllen, denn sie wollen ihren Fähigkeiten gemäß bestmöglich handeln (Aristoteles, NE Buch I; Pakaluk, 2005 Kapitel I; Wedgwood, 2017). Aufgrund

der Tatsache, dass sich Menschen immer in interaktiven Strukturen befinden, re-aktualisieren sie durch ihre alltäglichen Handlungen diese normativen Aufforderungsspiele: Menschen verhalten sich gegenüber anderen in unterschiedlichen Umwelten, agieren miteinander und reagieren aufeinander, folglich nehmen sie Haltungen ein und beziehen Stellung. Dieses verwobene Miteinander kann nicht transzendiert werden, es ist wesentliches Merkmal menschlicher Natur und tragende Struktur menschlicher Gemeinschaft. Menschen reagieren auf bestimmte Verhaltensweisen mit bestimmten Verhaltensweisen. Diese Reaktionsketten sind intrinsisch, sie liegen in der Natur des Menschen. Diese reaktualisierten, interaktiven Strukturen basieren auf „demands on others for oneself and demands on others for others, so there are self-reactive attitudes associated with demands on oneself for others“ (Strawson, 1974, S. 16). Diese Interaktionen begründen normative Ansprüche: Wir fordern und erwarten bestimmte Verhaltensweisen von anderen gegenüber anderen, von uns gegenüber anderen und von anderen gegenüber uns. Dieses Geflecht aus Anforderung und Erwartung ist Gemeinschaft – ist Kooperation –, sofern sich erst durch diese Verwobenheit gegenseitiger Verpflichtungen vernünftige Lebensformen etablieren können. Diesen Umstand hebt auch Edwards hervor, indem er formuliert, dass die Regelmäßigkeit menschlicher Handlungen nicht eine Frage externer Governance oder der einfachen Regelbefolgung sei, sondern gerade, dass die Auslegung der Regel und die sorgfältige Anwendung bedeutsam für Handlungen sind (Stern, 1974). Die Regelmäßigkeit von Handlungen ist folglich nicht einfach regulativ zu verstehen: Regeln sind nicht nur regulativ, indem sie eine Handlung führen, sondern sie konstituieren rationale Handlungen. Sie geben vor, welche Handlung als rational wünschenswert bewertet wird, welche Handlung von den richtigen Gründen getragen wird. Es lässt sich somit sagen, dass die Wechselwirkung menschlicher Begegnung Lebensformen erzeugt und dadurch Normativität beansprucht, die sie jedoch nicht vollständig begründen kann. Allerdings eröffnet sich dadurch ein semantischer Raum, ein Raum der Begründung, in dem normative Einstellungen und Haltungen besprochen sowie begründet werden können. Mit anderen zu argumentieren, welche Handlungen, welche Einstellungen und welche Wahrnehmungen gerechtfertigt sind, gehört im Wesentlichen zur deontischen Struktur.

2.2.2 Pragmatische Normativität

Wenn die lebensweltliche Normativität nicht letztbegründet werden kann und sich in lebensweltlichen Wert- und Anspruchshaltungen zeigt, die Individuen gegenseitig oder von einer Gemeinschaft einfordern, muss Autonomie – die Möglichkeit, seine Gründe begründen und verfolgen zu können – im Mittelpunkt einer Ethik stehen. Das diskursive Moment zeigt sich somit in der Praxis des Begründens und des Annehmens der Gründe, was sich, Nida-Rümelin folgend, gerade in keinen sozialen Konstruktivismus niederschlagen muss, sondern mit der Objektivität normativer Tatsachen vereinbar sein kann (Nida-Rümelin, 2018). Die Auswahl einer Handlung, die sich bestmöglich in die vorhandene

Lebensform einfügen lässt, ist jedoch keineswegs vollständig determiniert. Mead sieht Freiheit gerade darin manifestiert, dass man sich gegen existierende gesellschaftliche Strukturen wenden kann, aber er versteht dies nicht als destruktive Haltung, sondern im Sinne eines Reformismus. Dieser besteht „in the nature of presenting an order which is more adequate than the order which has been there“ (Mead, 1938, S. 663).

Nur aufgrund des Selbst, das wesentlich Rollen-Einnahme und symbolische Interaktion ist, wird ein moralischer Akteur möglich: Die Antizipation der Reaktionen und der Perspektivenwechsel sind nicht ausschließlich non-kognitiv zu verstehen. Die Möglichkeit, die verschiedenen Konsequenzen vieler anderer vorwegnehmen zu können, ist im Wesentlichen ein kognitiver Akt, v.a. dann, wenn unterschiedliche Zielsetzungen konfliktieren oder soziale Objekte, d.h. gemeinsame Zielsetzungen, nicht mehr getragen werden. Handlungslösungen führen nicht mehr über institutionalisierte oder konventionalisierte Muster zu erfolgreichen Ergebnissen. Somit ist auch klar, dass der Akteur ein dynamisches Subjekt ist, das sich hinsichtlich der Zielsetzungen und der Mittelwahl in Abhängigkeit der Situation immer weiterentwickeln kann. D.h. soziale Objekte sind ein notwendiger gemeinsamer Anker, der den geteilten normativen Horizont absteckt, aber sie müssen immer wieder überdacht werden, sobald sie in Konflikte geraten. Diese Aufgabe der Zielsetzung kommt der Ethik zu, während Mittelabwägungen, also wie diese Ziele gemeinsam erreicht werden können, Aufgabe der Wissenschaft sind (Mead, 1923). Diese Konzeption erinnert an G.W.F. Hegels formulierte Sittlichkeit, wo gerade die möglichen Bedingungen gesellschaftlicher Strukturen bedacht werden, damit sich der Einzelne in seiner Einzigartigkeit, in seiner Besonderheit, in der Gesellschaft nicht nur nicht verliert, sondern affirmativ entwickeln kann. Dies wird dadurch gewährleistet, dass die gesellschaftlichen Rahmenbedingungen den Einzelnen bestmöglich dazu befähigen, sich selbst zu verwirklichen, was nur gelingen kann, wenn er mit seiner Selbstentfaltung zum *bonum commune* beiträgt, wie z.B. Hegels Erörterung der Arbeitsteilung aufzeigt (Hegel, 2012, §§ 182-456). Das Individuum ist somit autonom durch Sozialität. Jedes Selbst bedarf der geteilten Zwecksetzungen, der sozialen Objekte, denn es kann nur sich selbst verwirklichen, wenn es die geteilten Zwecke in seinen eigenen Zielsetzungen verwirklicht: Es kann nicht gesund bleiben, wenn die anderen alle erkranken (diesen Sachverhalt veranschaulicht die Corona-Krise sehr gut). Es kann sich nur bilden, wenn es jemanden gibt, der es ausbildet. Es kann nur seinen Lebensunterhalt erarbeiten, wenn es sich in einer arbeitsteiligen Wirtschaft befindet. Sowohl die Zielsetzung einer Handlung als auch die Erwiderungen auf Handlungen sind nicht determiniert. Im Gegenteil: Die Möglichkeit, alte Gewohnheiten mit neuen Erfahrungen, mit neuem Wissen zu reformieren, um neue Praxen zu etablieren, muss als Freiheit oder Ausdruck menschlicher Autonomie verstanden werden. Sollte es folglich, um es noch einmal hegelianisch zu fassen, zu einem Konflikt zwischen dem Einzelnen als Besonderheit und der Sittlichkeit als Struktur oder gar dem Einzelnen

als Allgemeinheit kommen, dann sollte dieser unzumutbare Sachverhalt einer Veränderung unterzogen werden (selbst wenn Hegel das so nicht formulierte). Es entsteht eine moralische Forderung danach, sich auch gegen tradierte, konventionalisierte oder gar institutionalisierte Strukturen abzugrenzen, wenn diese keine erfolgreichen Resultate mehr erzielen, was Mead unter der Dynamik der Gesellschaft versteht. Thoreaus (1866) Plädoyer einer *civil disobedience* wird sodann zu einer moralischen Pflicht, was zudem nichts anderes impliziert, als dass moralische Persönlichkeiten an Situationen wachsen und sich durch diese zu moralischen Persönlichkeiten entwickeln. Moral lässt sich nicht auf einen reinen Konformismus reduzieren, obgleich die moralische Gemeinschaft zur Entstehung eines individuellen, moralischen Bewusstseins wesentlich beiträgt, indem das Individuum durch seine Einbettung in gemeinschaftliche Strukturen überhaupt erst Bedeutungskontexte und deren vernünftige Auslegung und Anordnung erlernt. Erst dadurch werden Situationen normativ zugänglich, weil Verhaltensabläufe durch Erwartungshaltungen und deren (Nicht-)Erfüllbarkeiten geordnet werden können. Diese Strukturierung, diese normative Deliberation ist zudem eingeübt, sie bleibt in allen ihren Manifestationen wesentlich sozial.

Damit Vernünftigkeit gelebt werden kann, bedarf es eines epistemischen und eines normativen Wissens: zum einen wie Verhaltenszusammenhänge erkennbar sind, zum anderen aber auch wie erkannte, wahrgenommene, erfahrene Situationen angemessen zu beantworten sind. Dieses praktisch-normative Wissen ist nicht auf epistemisches Wissen reduzierbar. Es ist kein informatives Faktenwissen, das mit Hilfe bestimmter Methoden, wie beispielsweise Inferenzen oder Subsumtionen, Anwendung auf Situationen und Handlungsoptionen findet, um im Anschluss an den rationalen Erkenntnisprozess korrekte und angemessene Reaktionen nach einer Moraltheorie – sei diese eine Optimierung des Nutzenoptimierungswerts oder der Universalisierbarkeit – wählen zu können. Der Handlungszweck ist weder durch Gott noch durch die Natur determiniert, sondern die Handlungen entfalten über die Interaktionen und über die Reflexion der Interaktionen ihren Sollensanspruch. Somit ist der vermeintliche Widerspruch zwischen einer Determination des Individuums durch die Gemeinschaft bzw. die Auflösung der Gemeinschaft in einzelne unabhängige Individuen überwunden: Ein Bewusstsein für moralische Situationen entsteht durch die aktive Reaktualisierung semantisch zugänglicher Strukturgefüge. Im gegenseitigen Austausch erlernen und übertragen Individuen Denkmodi oder Handlungsoptionen, wie beispielsweise Inferenzen, Abstraktionen, aber auch Werthaltungen: Sie üben ein, wie sie diese in Situationen anwenden können. Obgleich sie dadurch den Charakter von Traditionen aufweisen, sind sie aufgrund ihrer vernünftigen und volitionalen Anwendbarkeit form- und revidierbar. Gewohnheiten werden z.B. bei Mead (1934) durch Rollenspiele oder bei Darwall (1983) durch Abstraktion einstudiert. Erst durch die Sensibilisierung, das Erlernen von Achtsamkeit, können Situationen rational begründbar werden. Es werden folglich Gründe-Strukturen dahingehend, was gute Gründe und was Handlungsgründe sind, erfasst und weitergegeben. Diese genuin praktischen

Gründe oder Überzeugungssysteme werden durch Wiedergabe erlernt, bestenfalls ständig überprüft, revidiert und erweitert. Durch diese Prozesse wird ein gemeinsamer Bedeutungshorizont (Mead, 1934) bzw. ein geteiltes normatives Orientierungswissen (Nida-Rümelin, 2009) bekräftigt, der/das die Lebenswelt trägt. Aber auch hier muss die Dialektik betont werden, denn die Lebenswelt bildet sich zugleich im normativen Wissen ab und wird durch dieses strukturiert. Man kann sagen, dass es sich hierbei um eine Form tradierter Rationalität handelt, die man durchaus auch als *hexis* begreifen kann.

Nida-Rümelins Verständnis eines normativen Orientierungswissens (Nida-Rümelin, 2009) hebt diesen durch und durch praktischen Aspekt des normativen Orientierungswissens explizit hervor. Es geht hier folglich nicht um Wissen der Art, „dass eine Person weiß, wenn sie von *p* überzeugt ist, sie für die Annahme von *p* die richtigen Gründe hat und *p* wahr ist“ (ebd), sondern es handelt sich vielmehr um ‚Wissen‘ (...) in einem unspezifischeren Sinne (...), indem man von ‚Wissenssystemen‘ und ‚Wissensformen‘ spricht. Überzeugungen bilden einen Wissenskorporus, wenn diese in einem Begründungszusammenhang stehen und (deshalb) einen hohen Gewissheitsgrad haben“ (ebd.). Entscheidend ist hier auch, dass normatives Wissen von epistemischem Wissen zu unterscheiden ist. Epistemisches Wissen bezieht sich auf den Erkenntnisstand, der sich durch mehr Einsicht erweitern und verfeinern lässt sowie Erfahrungen bzw. Wahrnehmungen klassifizierbar und systematisierbar werden lässt. Praktische Normativität bezieht sich hingegen auf ein Wissen, das sich auf Handlungen bezieht, das Handlungsorientierung bietet und in diesem Sinne Praxis ist. Nida-Rümelin unterscheidet zudem zwischen deskriptivem und wissenschaftlichem Wissen sowie in Analogie hierzu zwischen normativem Wissen und ethischer Theorie. Zentral für ein praktisches Verständnis der deontisch getragenen Lebensform ist, dass sie auf Strukturen interaktiver Bedeutungskontexte etablierte Normativität ist. Während deskriptives Orientierungswissen dazu dient, Sinneseindrücke zu individualisieren, zu klassifizieren, ohne physikalische oder chemische Theorien zu benötigen, strukturiert normatives Orientierungswissen alltägliche Handlungssituationen, ohne Rekurs auf ethische Theorien nehmen zu müssen. Es handelt sich hierbei um vortheoretisches Wissen, das zumeist ausreicht, um handlungsanleitend zu werden. Der Anspruch an die innere Kohärenz der beiden lebensweltlichen Systeme ist geringer als bei Wissenschaft und Ethik, allerdings muss sich auch praktisches, vortheoretisches Orientierungswissen der alltäglichen Überprüfung unterziehen. Diese erfordert die Begutachtung der Ordnung und Strukturierung von Wahrnehmungen sowie Handlungen bezüglich ihrer Angemessenheit in lebensweltlichen Kontexten. Die ethische Deliberation soll überprüfen, ob die normativen Strukturelemente (noch) zu zumutbaren Resultaten führen. Folglich muss der Verpflichtungscharakter bestimmter Momente der Lebensform danach befragt werden, ob ein Sachverhalt noch begründbar und wünschenswert hinsichtlich der Ansprüche praktischer Vernunft ist (Nida-Rümelin,

2009). Lebensformen bestehen aus Handlungsketten, die wiederum eben jene zwei Bewertungsfunktionen repräsentieren: zum einen eine epistemische, die das Meinen, dass etwas der Fall ist, über Wahrscheinlichkeiten repräsentiert, und zum anderen eine konative Wertefunktion, die das subjektive Für-Wertvoll-Halten repräsentiert (ebd.). Folglich bringen Menschen handelnd Überzeugungen und Haltungen zum Ausdruck: Zum einen bringen sie zum Ausdruck, was sie für den Fall halten und wie wahrscheinlich dieser Fall ist, zum anderen, dass sie den Handlungszweck für wertvoll und somit für erstrebenswert halten.

Digitale Artefakte werden nun genau in diese Handlungsketten, die durch intentionale Haltungen strukturiert sind, in diese Abwägungssysteme implementiert. Ein philosophisches Verständnis menschlicher struktureller Rationalität ist daher hilfreich, wenn die Rationalität und Moralität der durch einen Einsatz von digitaler Technologie und Technik hervorgerufenen Transformation erfasst werden soll. Ebenso verdeutlicht dieser Sachverhalt, dass anstelle von Maschinen, die sich auf die physische Arbeitskraft des Menschen berufen, digitale Artefakte den Menschen in seiner Humanität berühren: in seiner Sozialität, Vernunftfähigkeit und Entscheidungsstärke. Bei der Entwicklung muss auf diese normativen Aspekte Rücksicht genommen werden, sollen digitale Artefakte den Menschen in seiner Humanität nicht unterdrücken.

Hierfür muss die Lebenspraxis als wünschenswerte Praxis hervorgehoben werden, in die der Entwicklungsprozess und die Nutzung, d.h. Handlungen im Umgang mit dem digitalen Artefakt, jeweils in gemeinsam geteilte wünschenswerte Lebensformen integriert werden. Dies kann, wie gezeigt wurde, mit Hilfe von Nida-Rümelins *Theorie einer Strukturellen Rationalität* gelingen.

2.3 Erstrebenswerte Praxis

Die strukturellen Handlungsketten, in die sich Teilhandlungen integrieren, werden als Praxis bezeichnet (McIntyre, 2007; Nida-Rümelin, 2020). Praxen zeichnen sich dadurch aus, dass sie punktuelle Einzelhandlungen zu temporalen, aber auch intra- und interpersonalem in sich stimmigen individuellen und kollektiven Lebensentwürfen zusammenführen. Im Rahmen der Digitalisierung und der digitalen Transformationen drängt sich die Frage auf, welche Lebensarten wünschenswert sind, denn die gegenseitigen Verpflichtungen und Erwartungshaltungen werden durch den Einsatz digitaler Artefakte tangiert. Die deontischen Strukturen, die Ausdruck von Haltungen sind, weisen keinen Verpflichtungscharakter auf, tragen jedoch wesentlich zum Verständnis der lebensweltlichen Normativität bei: Eine Einbindung digitaler Technologie und Technik in diese deontischen Strukturen verschiebt oder hebt ihre normativen Ansprüche auf.

Der Empirismus verdrängt die Praxistheorie immer mehr, sodass einige Phänomene, die sich die strukturellen und kooperativen Handlungsverknüpfungen erklären lassen, aus technikphilosophischen

Überlegungen ausgeschlossen werden (Reijers & Coeckelbergh, 2020, v.a. Kapitel 2). Reijers und Coeckelbergh begründen dies mit einer ausschließlich empirischen Orientierung am Sachgegenstand (ebd.). Digitale Artefakte werden dabei nicht in ihrer medialen Rolle betrachtet, welche die Vermittlung von Erfahrungshorizonten oder gar die Eröffnung neuer Handlungszusammenhänge begünstigen, sondern als Mittel zur Zweckerfüllung: Mit Hilfe eines Hammers lassen sich Bretter zusammennageln und mit Hilfe eines Textverarbeitungsprogramms diese Arbeit schreiben. Der instrumentelle oder strategische Zusammenhang erklärt das technische Artefakt ausreichend. Diese Interpretation lässt die ethische Bewertung zu einer reinen Folgenbewertung werden: Der Gebrauch oder Missbrauch des Artefakts führt zu wünschenswerten oder unzumutbaren Ereignissen. Die angemessene Bedienung rückt somit in den Mittelpunkt einer ethischen Deliberation, während sich der Sachgegenstand einer moralischen Rechtfertigung entzieht. Solange der Entwicklungsprozess digitaler Artefakte nur als herstellende Tätigkeit begriffen wird, mag diese Schlussfolgerung ihre Berechtigung haben: Zum einen kann sich ein Objekt nicht moralisch verhalten und zum anderen unterliegt der einzelne Entwickler externen Vorgaben, die sich meist als ökonomische oder technische Sachzwänge herausstellen. Was als notwendig für die Entwicklung des Artefakts erachtet wird, hängt von dem zu erreichenden Ziel in Abhängigkeit des Zwecks ab, weshalb Arendt (1960, S. 182) mit Nachdruck betont, dass im Herstellungsprozess alles danach beurteilt wird, ob es „dem Endzweck angemessen ist und für ihn von Nutzen ist“.⁶ Damit die *poiesis* einer ethischen Bewertung zugänglich gemacht werden kann, darf die herstellende Tätigkeit nicht nur der instrumentellen Vernunft, der *techne*, folgen, denn eine ethische Deliberation lässt sich gerade nicht auf diese Form der Bedingtheit von Mittel und Zweck reduzieren: „Das Problem der antiken Begriffsbildung liegt in der einseitigen Zuordnung der *techne* zur *poiesis* und nicht zur ethisch relevanten *praxis* [Herv. d. A.], die eine Technikethik begrifflich ausschließt.“ (Ott, 2005, S. 592) Daher muss die *poiesis* in eine *praxis* integriert werden, um herstellende Aktivitäten und die daraus entstehenden Produkte ethisch evaluierbar werden zu lassen.

Auch Floridi (2011) betont in seinem Ansatz einer *constructionist ethics*, dass klassische ethische Ansätze die für die Normativität technischer Artefakte relevante Bedeutung aufgrund ihrer Situiertheit in lebensweltlichen Kontexten nicht einzufangen vermögen. Dort befinden sich Individuen in einer Umwelt, auf die sie bloß reagieren. Dieses Ausgeliefertsein übersieht den für das digitale Zeitalter zentralen Aspekt, dass der ethische Akteur einen wesentlich herstellenden Charakter aufweist (Russo, 2012a, S. 68). Entwickler und Nutzer generieren neue digitale Artefakte, indem sie bspw.

⁶ Arendt (1960, S. 181-189) verwendet hier den Terminus *Endzweck* auch in Zusammenhang mit poetischen Tätigkeiten, da der Zweck sich hier als Endzweck zeigt, worum Willen hergestellt wird. Aber diese Zwecksetzung gilt es von der praktischen Zwecksetzung des Handelns zu unterscheiden, da letzteres nicht mehr mit Hilfe des Nutzenbegriffs zu beurteilen ist (ebd., S. 213-300).

digitale Produkte wie Applikationen entwickeln oder Content erschaffen. Diese herstellenden Tätigkeiten, die den *homo poieticus* auszeichnen, gilt es in wünschenswerte und kooperative Praxen zu integrieren, damit diese nicht nur als Reaktion auf gegebene Tatsachen verstanden werden. Denn nur, wenn sowohl die Herstellung als auch der Umgang mit Technik nicht ausschließlich durch Heteronomie geprägt sind, ist es möglich, diese Tätigkeiten auch moralisch zu bewerten bzw. normativ Haltung zu beziehen. Das bedeutet, dass auch der technische Herstellungsprozess einer wünschenswerten Praxis dienen können muss (Zuber et al., 2021). Man denke hier bspw. an eine gewünschte, akzeptierte und arbeitsorganisatorisch eingebettete Whistleblower-Kultur, damit Aussagen, wie die der Ethikerin Timnit Gebru, nicht direkt zu einer Entlassung führen müssen (Wolfangel, 2021).

Die Rückbesinnung auf die Praxis und die Forderung, poietische Tätigkeiten als Handlung autonomer Entscheidungsfindung individueller Akteure zu begreifen, die aufgrund des Praxisbegriffs auch immer in einem sozialen Umfeld eingebettet bleiben, führen zu einer ganzheitlichen Betrachtung *ethischer Programmierpraxis*. Hierzu muss bereits während der Entwicklungsarbeit das digitale Artefakte als soziales System betrachtet werden. Dies ist notwendig, um die Auswirkungen des digitalen Artefakts auf die lebensweltlichen Strukturen, individuellen Handlungsanforderungen und Wahrnehmungsmodalitäten begreifen zu können. Darüber hinaus müssen normative Überzeugungen in den Anforderungskatalog von Softwareprodukten aufgenommen werden, um auch technisch berücksichtigt zu werden. Dies bedarf einer Arbeitsorganisation, die sich nicht nur über den Wert der technischen Funktionalität ordnet. Das Verständnis von Technik als Mittel rückt hingegen technik-generische Werte in den Fokus einer ethischen Betrachtung, die negative Folgen für individuelle, intersubjektive oder kollektive Lebensweisen haben können: Die Betrachtung dieser Werte ist aufgrund der Beschaffenheit der digitalen Technologie oder Technik geboten. Diese Werte lassen sich nicht ausschließlich mit Hilfe einer Analyse der Lebensformen lokalisieren, sondern bedürfen zunächst einer Betrachtung der Technologie und Technik (Teil II, Kapitel 2.1). Die Einführung normativer Deliberationen in die Entwicklungsarbeit ermöglicht die Formulierung von Gütekriterien jenseits von technischer Funktionalität oder ökonomischem Nutzen.

Praxis als Gegenstück zur *poiesis* nimmt gerade die Perspektive ein, die diese unmittelbare Abhängigkeit der Mittel zu transzendieren vermag: Hier gilt nicht mehr das Machbarkeitsdenken als einzige Quelle qualitativer Gütekriterien. Eine Praxis verwirklicht einen Wert, der nicht außerhalb ihrer selbst aufzufinden ist. So mag man Sport als *poiesis* beschreiben, um gesund zu leben, schlank zu sein oder Geld zu verdienen. Sport kann jedoch auch das Ziel der Betätigung selbst sein und ist demnach auch Motivation der Handlung: Hier ist Sport als *praxis* zu verstehen. Darüber hinaus verweist der Mannschaftssport unmittelbar auf die gemeinsam getragenen Handlungsziele: Es handelt sich hierbei um eine kollektive, kooperative Struktur, die sich in geteilten Sollensanforderungen zeigt. Man denke in diesem Zusammenhang an Fußball, der sicherlich alle Sichtweisen zu vereinen vermag – so wird

auch Camus (1953) zitiert: „Alles, was ich über Moral und Verpflichtungen weiß, verdanke ich dem Fußball.“ Digitale Artefakte müssen aus den unterschiedlichen Perspektiven menschlicher Tätigkeiten und Handlungspraxen reflektiert werden, um diese Verhältnisse offenlegen zu können. Erst dann kann eine Bewertung der Situation vorgenommen werden. Ebenso sind diese ethischen Deliberationen als *praxis* zu begreifen, die kein anderes Ziel verfolgt und um keinen anderen Zweck bemüht ist, als das Nachdenken über normativ wünschenswertes Design zu kultivieren.

Im Gegensatz zu Reijers und Coeckelbergh (2020) wird die Vermittlung der instrumentellen mit der normativen Rationalität nicht über das Narrativ gesucht. Der Fokus liegt auf der Rationalität der Praxis, die sich entweder durch digitale Artefakte entfalten oder durch sie gehemmt werden kann. Menschliche Vernunftfähigkeit lässt sich nach Nida-Rümelin durch die Teilnahme an der Austauschpraxis von Gründen sowie an der Affektion durch das Resultat, d.h. durch den guten Grund, bestimmen. Vernunft ist somit „die allgemein menschliche Fähigkeit, eine gemeinsame Aufmerksamkeit auf Gegenstände von gemeinsamem Interesse zu richten, im Umgang mit diesen eine gemeinsame, kooperative Praxis zu entwickeln, und diese von Gründen geleitet zu koordinieren“ (Nida-Rümelin, 2020, S. 24).

Der Vernunftbegriff ermöglicht Lebensweisen und Praxen zu rechtfertigen, die nicht nur durch Geschichten vermittelt werden. Die *Strukturelle Rationalität* integriert das punktuelle Ereignis, sei dieses eine Handlung, eine Emotion oder ein digitales Instrument, in den zeitlichen Horizont einer Sozialität, ohne dabei einer verbindlichen Rationalität verlustig zu werden. Im Gegenteil, gerade diese Rückbindung an Gemeinschaften verschafft der Vernunft die Möglichkeit, sich individuell und kollektiv zu verwirklichen. Im Gegensatz dazu folgt die narrative Wende der Prämisse, dass es keine vernünftigen normativen Tatsachen geben muss, weshalb sie gegenüber der Ontologie der Gründe indifferent ist. Gründe stehen hier in keinem logischen oder rationalen Zusammenhang, sondern die Geschichte wirkt sich sinnstiftend auf die einzelnen Aspekte aus. Narrationen sind von daher stark kulturrelativistisch, da allein die Zuschreibungspraxis ausschlaggebend dafür ist, welche Geschichte erzählt werden soll. Im Gegensatz dazu zeigte sich der Praxisbegriff, wie Nida-Rümelin ihn in seiner *Theorie einer Strukturelle Rationalität* entwickelt hat, als von normativen Tatsachen getragen, die durch die Zuschreibungspraxis des Gründegebens und -nehmens gefunden, nicht aber erfunden werden können. Vielmehr geht es, wie Hetzel (2015, S. 116) treffend formuliert, darum, dass

[d]iese Tatsachen [...] jedoch nicht im Sinne vorhandener Fakten zu deuten [sind], sondern als die Realität eines Sollens und der Ansprüche in der Praxis. [...] Moralische Tatsachen existieren insofern unabhängig von uns, als sie ihren Ort in einer Praxis haben, die uns gegeben ist, aus der wir uns nicht herausreflektieren können. Zugleich haben diese Tatsachen keinen Status, der es erlauben würde, um sie zu wissen, sie in irgendeiner Form zu repräsentieren. Das bedeutet allerdings wiederum nicht, dass sie sich, wie die Nonrealisten (und diejenigen Kritiker des Pragmatismus, die ihm eine relativistische Position unterstellen) behaupten, in bloßen Präferenzen erschöpfen würden.

Die Trennung zwischen Handlungen und Tätigkeiten ist nicht einfach zu ziehen, denn so finden sich Auslegungen, die das Spielen eines Instruments der *poiesis* zuordnen, gerade weil beim Spielen eines Instruments zunächst keine Sozialität erkennbar ist (Reijers & Coeckelbergh, 2020), während andere das Musizieren direkt als Praxis auslegen. In diesem Zusammenhang ist von großer Bedeutung, dass sich Sozialität durch ihre Zeitlichkeit auch als tradierte Vernunft zeigen lässt. Sozialität bedarf keines unmittelbaren Gegenübers, das Handlungszusammenhänge erst zu einer kooperativen und kollektiven Praxis zu vereinen im Stande ist. Allein, dass, wie Janich (2003) betont, der Mensch sein Handlungsvermögen erst erlernen muss, und dies manchen besser als anderen gelingt, verweist auf die Einbettung in strukturelle Rationalitäten, in übernommene gegenseitige Vertrauensbeziehungen, die überhaupt erst diese Form des Erwerbs ermöglichen: Das Individuum „muss [...] als selbstbestimmt lebendes Individuum Zwecksetzungsautonomie, Mittelwahlrationalität und Folgenverantwortlichkeit entwickeln“ (Janich, 2003, S. 95-96).

Die Tätigkeiten der Vernunft müssen erlernt, geübt und auf bestimmte Bereiche angewendet werden können. Dabei schließen sich die verschiedenen Bereichsrationalitäten nicht aus, sondern führen in den jeweils korrespondierenden Entscheidungsräumen zu optimalen Resultaten: Was jeweils als optimal verstanden wird, zeigt sich bspw. in einer moralisch guten oder in einer den Nutzen optimierenden Handlung. Daher muss nicht nur der Denkmodus erworben, sondern auch dessen zielführender Einsatz in lebensweltlichen Praxen erlernt werden. Erst die Übung dieser Vermögen führt dazu, dass Personen handlungsfähig sind. Eine Handlung oder eine Tätigkeit kann somit sowohl aus dem Blickwinkel der instrumentellen Rationalität, der *techné*, als auch aus dem Blickwinkel der praktischen Klugheit, der *phronesis*, beschrieben werden. Praxis umfasst somit das Denken als *phronesis*, die praktische Klugheit, und ist daher das Vermögen, Allgemeines auf der Handlungsebene auf Besonderheiten anwenden zu können. Bestenfalls kommt sie, da tugendhaft, zu einem richtigen, weil moralisch gutem Ergebnis. Somit ermöglicht die Klugheit, sich dem Gegenstand der Betrachtung nicht nur aus der Perspektive der Mittel anzunähern, sprich technisch aus der Funktionalität und ökonomisch aus dem Nutzen, sondern auch Bewertungskriterien außerhalb dieser zu verwirklichen. Beide Rationalitäten schließen sich hierbei nicht aus, sondern eröffnen andere Weisen, sich dem Betrachtungsgegenstand gedanklich zu nähern.

In einem klassischen aristotelischen Sinne ist die Praxis die Form der Handlung, die für die Kultivierung von Tugenden konstitutiv ist (MacIntyre, 1981). Praxis für sich reicht nicht aus, sondern es bedarf ebenso der Tugend, damit wünschenswerte Ergebnisse verwirklicht werden können – zumindest, wenn es sich nicht um glückliche, sondern vernünftige Ereignisse handeln soll. Sobald eine Person gerechte Handlungen ausübt und dadurch Gerechtigkeit lebensweltlich erfährt, kann sie einen gerechten Charakter ausbilden und diesen in unterschiedlichen Praxisstrukturen ausleben. Ein tugend-

haftes Leben verhindert, dass Praxen ausgenutzt werden, indem sie instrumentalisiert oder strategisch manipuliert werden. Praxis bedarf in diesem Sinne auch der Deliberation, die mit der Entscheidungswahl verbunden ist (Aristoteles, NE 1139a2).

So gibt es beispielsweise zahlreiche Möglichkeiten, die Tugend der Freundschaft zu leben: im direkten Austausch, durch gemeinsame Aktivitäten etc. Diese Lebensweisen sind als die Bedingungen der Möglichkeit zu verstehen, die Tugend der Freundschaft kultivieren zu können. Einige sind bereits durch soziale Konventionen bestimmbar, andere als moralische Norm. Die Verbindung von Tugendhaftigkeit und Praxis lässt in alltäglichen Situationen die Möglichkeit zu, sich eben nicht mehr nur strategisch oder instrumentell zu verhalten, sondern auch der Tugend gemäß. Wenn Facebook als Ort sozialer Beziehungen jeglicher Art ausgeschrieben wird und sich dort Freundschaft durch neue Formen sozialer Gemeinsamkeit etabliert, so muss Facebook als neue Praxis der Freundschaft und diese sodann auch als Ort neuer Normativität begriffen werden. Dieses normative Geflecht, in das digitale Technologien und Techniken eingebunden werden, lässt digitale Technologie und Technik zu sozialen und kollektiven Systemen werden: „Even though technologies are artificial, it is nevertheless important to realize that they are part of human praxis.“ (Fallman, 2007, S. 27)

2.3.1 Emanzipierte Poiesis

Diese Arbeit unternimmt daher den Versuch, ein Verständnis von digitaler Technologie und Technik als Teil von Lebenspraxen zu entwickeln. Praktisch gesprochen unterliegen poietische Handlungen dann nicht mehr nur Sachzwängen oder Naturgesetzen, sondern sind aus der Sicht der praktischen Vernunft auch mit anderen normativen Kriterien bewertbar, denn aus technologischem und technischem Wissen lassen sich bspw. keine moralischen oder ästhetischen Sollensforderungen ableiten. Herstellende Tätigkeiten finden sich im Rahmen der Digitalisierung oder der digitalen Transformation auf Seiten der Entwickler, aber auch auf Seiten der Nutzer. Hier wird zuerst der Fokus auf die Entwicklungsarbeit gerichtet, obgleich die normativen Deliberationen auch für die lebensweltlichen Praxen der Nutzer notwendig sind.

2.3.2 Entwicklungspraxis

Einzelereignisse, seien diese Handlungen, Emotionen oder gar Gegenstände, die in Handlungsabläufe integriert werden, fügen sich in das Geflecht gegenseitiger Vertrauensbeziehungen ein. Diese Integration in gegenseitige Erwartungshaltungen führt zu Umbewertungen der normativen Ansprüche. Praxen, die kollektiv ausgerichtete Handlungsstrukturen sind, werden durch die Inklusion digitaler Artefakte in ihren traditionellen Handlungsabfolgen gestört. Erwartungshaltungen werden so generiert oder transformiert. Im Falle der Freundschaft kann schon die Bezeichnung *friends*, wie es bei

Facebook geschieht, und die Darstellung von Freundschaften als Punkte wahrhafte Freundschaftsbeziehungen verändern. Wenn das normative Grundgerüst nicht gestört werden soll, muss über die normativen Konstituenzen reflektiert werden, die durch eine Nutzung dieses digitalen Artefakts tangiert werden. Digitale Artefakte sind Teil vernünftiger individueller und kollektiver Lebensentwürfe: Da sie in eine temporale, intra- und interpersonale Sozialität eingebunden werden, muss der Praxisbegriff in den Mittelpunkt einer ethischen Betrachtung gerückt werden. Die Integration digitaler Artefakte kann normative Erwartungshaltungen verändern, indem Handlungsabfolgen medialisiert restrukturiert werden: Digitale Artefakte vermitteln Bedeutungszusammenhänge, zeigen Wahrnehmungsinhalte auf und fordern Verhaltensweisen ein. Sie selbst sind dadurch maßgeblich normativ.

Digitale Artefakte steuern so kollektive Handlungszusammenhänge, individuelle Weltbezüge und intersubjektive Erwartungshaltungen. Dieser Sachverhalt muss in der Entwicklung digitaler Artefakte berücksichtigt werden. Der Praxisbegriff ermöglicht zudem, unterschiedliche Begründungstypen theoretisch einzubinden, die lebensweltlich verankert sind: In Praxen finden unterschiedliche Begründungs- und Rechtfertigungsmöglichkeiten – über ökonomische, ästhetische oder gar moralische Gründe – ihre Berechtigung. Digitale Artefakte wirken auf die Begründungs- und Rechtfertigungsstruktur, weshalb dasselbe Artefakt in verschiedenen gesellschaftlichen Subsystemen unterschiedlichen Einfluss ausübt: Ein Chatdienst, der Kurznachrichten forciert, benötigt eine effiziente Kommunikationskultur, wie sie bspw. in der Arbeitswelt vorzufinden ist, nicht aber in der Seelsorge. Digitale Artefakte beeinflussen daher unterschiedliche Erwartungshaltungen, die bestimmte gesellschaftliche Bereiche in ihren normativen Ansprüchen prägen.

Diese ethischen Überlegungen stellen sich auch im Zusammenhang mit der Entwicklungsarbeit selbst, denn es benötigt eine Arbeitskultur, die herstellende Tätigkeiten mit moralisch wertvollen Praxen verbindet. Die Entwicklungstätigkeit wird durch Prozesse, welche die Autonomie der Entwickler fördern, aufgewertet. Dies bedeutet, dass der Herstellungsprozess die Möglichkeit bieten muss, technische Kompetenzen mit normativem Orientierungswissen und wertgeleiteten Handlungsintentionen explizit zu verbinden (Umbrello & Gambelin, 2021). Dadurch werden Austauschformen kultiviert, die den Sachgegenstand reflektieren, ohne ausschließlich die technische Machbarkeit zu berücksichtigen.

Die Bedingungen der Möglichkeit, eine Entwicklungspraxis zu etablieren, welche die Autonomie der Entwickler begünstigt, können als *emanzipierte poiesis* bezeichnet werden. Herstellende Aktivitäten werden durch etablierte autonome Entscheidungspraxen aufgewertet. So werden während der Entwicklungsarbeit normative Deliberationen gefördert, was bestenfalls zu normativ ausgerichteten Artefakten führt. Der *homo poieticus* schlägt somit die Brücke zur Praxis: „The homo poieticus is the

ethical agent in the era of technology: she is the *creator* of the situations subject to ethical appreciation.” (Russo, 2012a, S. 68) Er erschafft sich die Umgebung, die es moralisch zu evaluieren gilt und bricht seine vermeintliche Passivität gegenüber seiner Mitwelt auf: Er ist nicht nur Konsument oder Produzent, welcher der Natur oder der Umwelt gegenübersteht, sondern die Welt, in der er sich bewegt, erschließt er sich durch seine herstellende Beziehung zur Welt (eda.).

2.3.3 Bedienungspraxis

Aber nicht nur die Entwicklungspraxis ist hinsichtlich der herstellenden Tätigkeiten zu betrachten. Digitale Artefakte sind im Gebrauch, sodass auch der Umgang mit digitalen Artefakten in einem *poiesis-praxis*-Verhältnis zu denken ist. Aus der *Theorie der Strukturellen Rationalität*, dass sich Einzereignisse in ein vernünftiges Netz gegenseitiger Erwartungshaltungen und Verpflichtungen einbinden können lassen müssen, lassen sich zusätzlich zu den Werten, die aus der Beschaffenheit der Technologie und Technik selbst resultieren, Kriterien normativ wünschenswerter Artefakte formulieren. Es muss ein *emanzipierter Gebrauch* digitaler Artefakte ermöglicht werden, auch wenn dieser selbst gestaltend ist.

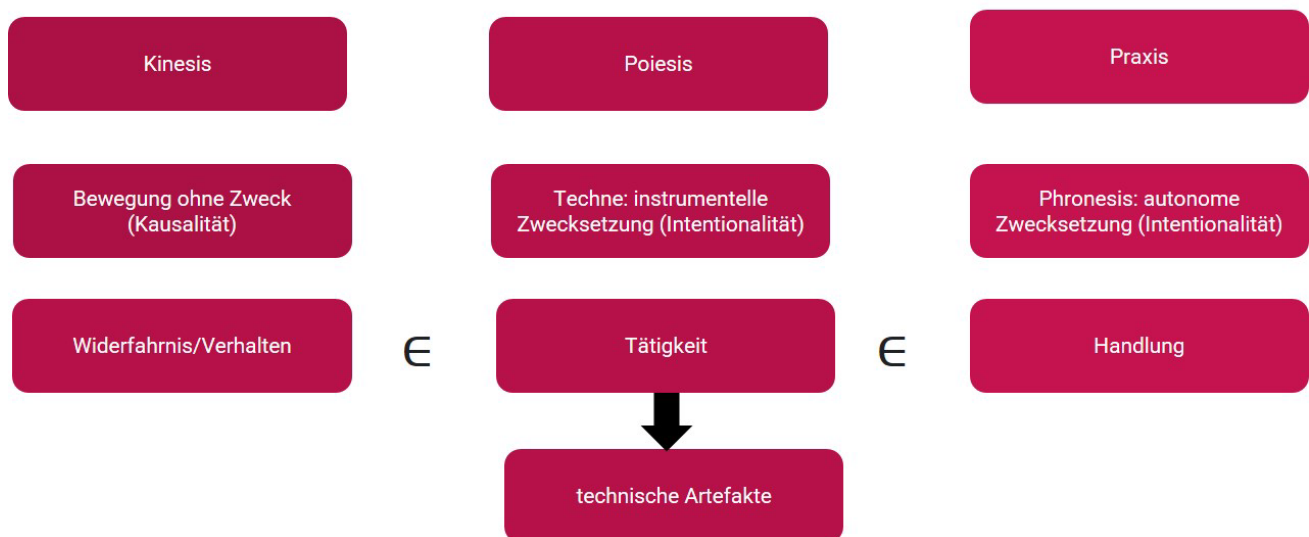
Daher ist auch aus der Nutzer-Perspektive die aristotelische Unterteilung von *poiesis* und *praxis* interessant. Auch wenn diese Unterscheidung hier nicht historisch ausgearbeitet wurde, wirft ihr systematisierender Charakter Licht auf die Fragen, welche die Normativität von Technik und Technologie betreffen. Die Unterscheidung in poetische Tätigkeiten und praktische Handlungen muss sowohl handlungstheoretisch als auch rationalitätstheoretisch konzipiert werden. Der Gebrauch eines digitalen Artefakts ist meist instrumentell: Das digitale Artefakt erleichtert die Kommunikation oder die geografische Orientierung. Computerspiele oder die Programmierkunst mögen hier Ausnahmen darstellen und bereits als eigenständige digitalisierte Praxen etabliert sein. Wesentlich ist, dass der Gebrauch digitaler Artefakte den Nutzer in seiner freien Lebensgestaltung unterstützt. Das digitale Artefakt muss deshalb daraufhin überprüft werden, ob es moralische Praxen behindert oder befördert. Nur so wird der Nutzer als moralische Person geachtet.

Da menschliche Aktivität immer auch Möglichkeit bedeutet und an eine Sozialität gebunden ist, eröffnet dies unmittelbar die Notwendigkeit, Entscheidungen moralisch zu reflektieren. Die strukturelle Einbettung technischer Artefakte in menschliche Lebensweisen erfordert zugleich eine Eingrenzung des Optionenraums, da die Freiheiten anderer auch (un-)beabsichtigt unterminiert werden können. Reijssel und Coeckelbergh (2020, S. 13) erweitern daher ihren tugendethischen Technikansatz um ein deontologisches Moment: „Even when we design our technical practices in such a way that they are conducive of good life, they might still exert (unintended) harm on certain people. For this reason,

technical practices need to be limited in certain respects; they need to be subjected to appropriate laws.”

Diese Überlegungen ermöglichen eine Konzeptualisierung moralisch-technischer Praxis sowohl für die Entwicklung als auch für den Gebrauch. So kann eine gute normative Praxis von einer schlechten normativen Praxis unterschieden werden, wenn wünschenswerte Verhaltensweisen unterminiert und schlechte Angewohnheiten etabliert werden. Durch diese tugendethisch anmutende Herangehensweise kann dem Anspruch einer normativen Ethik Inhalt verliehen werden. Nur wenn sich *techné* und *phronesis* in einem stetigen Austausch befinden, können normativ wünschenswerte, technische Objekte gestaltet werden, die sowohl dem Entwickler als auch dem Nutzer als moralische Persönlichkeiten dienen. Der Fokus liegt somit auf emanzipierten Tätigkeiten, die als Teil einer Praxis nicht nur der instrumentellen, sondern auch der normativen Rationalität folgen, wie Abbildung 2 zeigt.

**Zweite Bedingung eines normativ wünschenswerten Softwaredesigns:
Poiesis und Praxis**



Die **zweite Bedingung** eines moralisch angemessenen technischen Artefakts lautet, dass das zu entwickelnde Produkt Teil einer wünschenswerten Praxis sein sollte.

Die zweite Bedingung sollte sowohl in der Nutzung von Software als auch in Entwicklungsarbeit erfüllt werden.

Abbildung 2 Zweite Bedingung eines normativ wünschenswerten Softwaredesigns

2.4 Zwischenfazit

Technikphilosophisch ist von großer Bedeutung, wie technische Geräte auf menschliche, interaktive und gegenseitige Verpflichtungsgefüge einwirken. Durch einen Einsatz technischer Geräte können sich etablierte Lebenspraxen transformieren oder aber gänzlich neue Lebensräume geschaffen werden. Es gilt daher, Technik und Technologie dahingehend ethisch zu überprüfen, wie sie deontische

Beziehungsstrukturen beeinflussen und inwiefern dieser Einfluss wünschenswert oder inakzeptabel ist. Diese Lebenspraxen müssen den Forderungen der Vernunft entsprechen, sollen sie nicht den Menschen in seiner Humanität verletzen. Dies beinhaltet auch, dass sich die Vernunftfähigkeit des Menschen frei entfalten können muss. Daher sollte die deontologische Begrenzung menschlicher Lebensgewohnheiten kontinuierlich auf Kohärenz und somit auch im Hinblick auf die Integrität unterschiedlicher Lebensformen hinterfragt werden (Nida-Rümelin, 1997b). In diesem Zusammenhang muss eine Analyse unterschiedlicher Technik, im Sinne unterschiedlich technologisch bedingter Geräte, Instrumente oder Maschinen, und deren spezifischen Einflusses auf menschliche soziale und kollektive Verbindungen erfolgen. Ein vernünftiger Umgang mit Technik und Technologie ist nur dann unabhängig von technischen Besonderheiten möglich, wenn die genuin technischen Bedingungen keinerlei besonderer Berücksichtigung bedürfen. So gilt es zu untersuchen, auf welche Art und Weise bestimmte Technik und Technologien auf deontische Strukturen einwirken. Zudem muss beachtet werden, ob sich eine Veränderung einer humanen Ausgestaltung menschlicher Lebensstrukturen in bestimmten menschlichen Bereichen als besonders wünschenswert oder als inakzeptabel erweist.

Daher ist es notwendig, die unterschiedlichen oder ähnlichen Auswirkungen beispielweise eines Hammers, einer maschinellen Befüllungsanlage oder eines Computers auf Lebensstrukturen herauszuarbeiten und zu vergleichen. Entscheidend hierbei wird sein, inwiefern diese Technik und Technologie für den Menschen entlastend wirken soll. Dies führt direkt in die *uniqueness*-Debatte, innerhalb der die Fragestellung erörtert wird (Teil III), inwiefern normative Herausforderungen digitaler Technik und digitaler Technologie mit klassischen Technikphilosophien und Technikethiken eingefangen werden können oder ob eine ausgewiesene *Ethik der Informationstechnik* benötigt wird. Für die Klärung dieser Fragestellungen ist nicht unerheblich, Technik und Technologie begrifflich zu unterscheiden. Ebenso müssen die verschiedenen Subvarianten des Technikbegriffs herausgestellt und hinsichtlich ihrer Normativität befragt werden: Es gilt der Fragestellung nachzugehen, ob und welche technikgenerischen, ethischen Fragestellungen sich im Hinblick auf digitale Artefakte artikulieren lassen.

Beispielsweise verändert sich das strukturelle Pflichtgefüge der bisherigen Arbeitskultur durch den Einsatz KI-gestützter Bewerbungsverfahren oder die kontinuierlich maschinell überwachte Arbeitsleistung mittels einer semantischen Auswertung u.a. von Chaträumen. Lebenspraxen und deren normativen Konstituenzen werden austariert und erfahren somit eine graduelle Neubewertung: Digitale Bewerbungsverfahren sollen auf Seiten der Bewerber Lügen oder Schummeleien detektieren, Unsicherheiten bei Antworten identifizieren und Unstimmigkeiten einfach hervorheben. Auf der Seite der Arbeitgeber sollen durch den Einsatz von Bewerber-Tools Vorurteile unterbunden werden, um möglichst objektiv dem Bewerber zu begegnen. Der Einsatz eines solchen Tools verändert klassische Bewerber-Arbeitgeber-Gespräche und transformiert dadurch auch Machtstrukturen. Inwiefern diese neue Praxis zu wünschenswerten Verhaltensweisen führt, und missliches Verhalten unterbindet, um

normativ wertvolle Orientierung bieten zu können, bedarf einer gesonderten, normativen Deliberation, die an dieser Stelle nur angedeutet werden kann. Hinzu kommen Aspekte klassischer Privatsphäre-Argumente, die den Zugang zu personenspezifischen und relevanten Daten transparent regeln und artikulieren. Zugleich muss dies auch technisch gewährleistet und unternehmensweit eingehalten werden. Auch hier zeigt sich bereits eine Veränderung im Umgang mit direkten Vorgesetzten: Aus Loyalität wird so schnell Kontrolle und somit aus intrinsischer Motivation externer Zwang. Dieser behavioristische Turn, Personen als externalisierte Datenpunkten zu beurteilen, steht einem vernünftigen Umgang mit Menschen, gerade in einer post-feudalen Arbeitsgesellschaft, entgegen. Die Methode bspw. der *Inversen Transparenz* soll diesem Aspekt technisch-humanistisch begegnen, indem der Arbeitnehmer mitbestimmen darf, wer Zugang zu seinen Daten erhalten kann (Zieglmeier & Pretschner, 2021).

Ebenso ersetzen heute digitale Systeme weltweit analoge face-to-face-Interaktionen sozialer Verteilungssysteme, wie es in Banken bei Bonitätsüberprüfungen bereits üblich ist. Dabei soll der Einsatz digitaler Technologien aufwendige Arbeitsprozesse ersetzen und zugleich auch zu mehr Gerechtigkeit oder Transparenz führen, da Entscheidungen (endlich) nicht mehr von der Kontingenz menschlicher Subjektivität abhängen, sondern rationale und objektive Maschinen – so die Vorstellung – nüchtern über Sachverhalte ‚entscheiden‘ (Raymond & Shackelford, 2013). Maschinen als Instrumente der Kontingenzbewältigung einzusetzen, bedeutet aber zugleich, sich nur der Sachlichkeit der Argumente zuzuwenden und schließt jegliches Empathie-gestütztes Urteil aus. Digitale Technik suggeriert Objektivität (Teil II, Kapitel 7) und schließt jedes empathische Urteil qua ihrer technischen Rationalität aus. Die Annahme, dass Softwaresysteme zu einem zuverlässigen und gerechten ‚Urteil‘ gelangen, weil sie einer Subjektivität, die in diesem Diskurs mit Willkür gleichgesetzt wird, entbehren, beruht auf der enthusiastischen Idee, dass mit Hilfe eben jener Maschinen, die auf mathematischen Modellen beruhen, gewünschte Ergebnisse präzise kalkuliert werden können.

Ein philosophisches Verständnis befasst sich daher mit den unterschiedlichen normativen Ausrichtungen, die ein digitales Artefakt oder technisches Sachsystem auf individuelle Handlungen, interrelationale Beziehungen und kollektive Zielsetzungen ausübt. Somit hilft eine philosophische Untersuchung, normative Urteile über Technologie und Technik zu fällen, wenn durch eben jene Betrachtung diejenigen normativen Konstituenzen sichtbar gemacht werden können, die ein wünschenswertes, gemeinschaftliches Leben tragen. Technologie und Technik können sodann vor dem Hintergrund eines empirisch-philosophischen Horizonts normativ hinterfragt werden, denn der Einsatz dieser Technologie oder Technik kann hinsichtlich der Förderung oder Behinderung derjenigen normativen Bedingungen evaluiert werden, die für ein gemeinschaftlich getragenes Zusammenleben oder für ein selbstbestimmtes Leben innerhalb der Gemeinschaft vonnöten sind. Dies verweist auf eine humanis-

tische Ethik, denn das technische Artefakt wird an der Möglichkeit gemessen, ob es wünschenswert, menschlichen Bedürfnissen oder Fähigkeiten dient, indem nach lebensweltlichen Bedingungen gefragt wird. Dadurch kann nicht nur Negatives vermieden, sondern vor allem auch der Versuch unternommen werden, Anforderungen an ein gutes, normatives Design zu formulieren und – viel bedeutsamer – diese dann auch zu entwickeln.

Ein technisches Artefakt in seiner gesellschaftlichen Einbettung zu verstehen, bedeutet folglich, die Erwartungshaltungen hervorzuheben, die Technik in eben diesem gesellschaftlichen Zusammenhang repräsentiert. Dafür ist es wesentlich, die Strukturen zu akzentuieren, in die das technische Objekt integriert ist. Eine philosophische Analyse wirkt sich somit direkt auf die normativen Überlegungen aus, denn ein kontextualisierter Technikbegriff verbindet Technik und Technologie mit normativen Bedingungen menschlicher Daseinsform oder wenn man will: mit der Idee des guten Lebens. Die normativen Beziehungselemente, die diese zwischenmenschliche Ebene der Kommunikation und Kooperation hervorbringen, können jedoch auch missachtet werden: So kann eine gelungene Kommunikation nicht gänzlich auf Lügen basieren, wenn man z.B. die weitere Fortführung einer Freundschaft nicht gefährden möchte (Nida-Rümelin, 2011b). Das kooperative Miteinander, in diesem Falle der Freundschaftsstrukturen, wird durch das kontinuierliche Lügen gestört, denn Gemeinschaft erfordert eine Kommunikationskultur, die nicht durch verborgene Gründe getragen werden kann, sondern auf authentischen Äußerungen basiert. Z.B. gilt es als tugendhaft oder eben moralisch geboten, dem Freund nicht aus Eigeninteresse zu helfen, sondern der Freundschaft wegen. Punktuell können kooperative Strukturen Trittbrettfahrer-Handlungen verkraften. Kommt es jedoch zu einer flächendeckenden Erosion geteilter Strukturen, weil Trittbrettfahren einen systematischen Charakter annimmt, d.h. wenn ein solches Verhalten zur Regel wird, verlieren diese Lebensformen ihren normativen Anspruch. Wenn Strukturen, die zur Aufrechterhaltung einer gewünschten, gemeinsam-geteilten Lebenswelt benötigt werden, durch unbeabsichtigte oder gar beabsichtigte kontinuierliche Missachtung beschädigt werden, werden diese Lebensformen ihrer gelebten normativen Performativität verlustig: Freundschaften, die bspw. durch dauerhafte Lügen gekennzeichnet sind, bezeichnet man nicht mehr als Freundschaft, sondern dieses Beziehungsgeflecht mutiert zu einem einseitigen Abhängigkeitsverhältnis, das nur aus der Perspektive des Eigeninteresses zugänglich wird. Die gemeinsam geteilte Struktur der Freundschaft bedarf jedoch einer sie wertschätzenden Kommunikationskultur. Eine solche Kommunikation beruht auf normativen Bedingungen der Wahrheit, Wahrhaftigkeit und Verlässlichkeit (Nida-Rümelin, 2011b), die es überhaupt erst ermöglichen, dass Sprechakte nicht nur einer reinen strategischen Rationalität dienen, sondern auch kollektiv zu Gemeinsamkeiten vordringen können.

Technikphilosophie ist somit einerseits eine Freilegung der normativen Strukturen, in die Technik eingebettet ist. Somit hat sie einen aufdeckenden oder entdeckenden Charakter, weshalb Brey (2000b;

2012a) den Ansatz als *disclosive ethics* bezeichnet. Andererseits ist sie eine Analyse der Auswirkungen dieser technischen Einbindung auf die gesellschaftlichen Formen des Zusammenlebens. Dieser doppelte Ansatz der strukturellen Ausrichtung mutet zunächst soziologisch an, geht aber über einen rein deskriptiven Ansatz hinaus, da auch die Möglichkeiten der normativen Bedingungen berücksichtigt werden. Somit werden die tiefen, sogenannten disruptiven Transformationen erklärbar, die durch digitale Artefakte hervorgerufen werden. Technische Artefakte sowie deren Entwicklung und Nutzung werden als Teil struktureller Lebensweisen verstanden, die sich aufgrund des Einsatzes eines technischen Artefakts verändern können.

Die strukturelle technikphilosophische Betrachtung digitaler Artefakte hebt sogleich die Wichtigkeit einer Analyse des Maschinen-Mensch-Verhältnisses hervor, denn eine normative Evaluation muss die Erwünschtheit der Auswirkungen und Transformationen berücksichtigen, die ein Einsatz digitaler Technik auf menschliche Seinsmodi hervorhebt, seien diese wie in der Diskussion bei Ihde (1990) individuell-phänomenologisch bestimmt oder aber wie bei Vallor (2016), Reijers und Coeckelbergh (2020) tugendethisch durch praktische, intersubjektive und kollektive Lebenspraxen und Haltungen geprägt. Nida-Rümelins *Strukturelle Rationalität*, verstanden als deontische Strukturen, die ein Geflecht wechselseitiger Verpflichtungen aufzeigen, kann die unterschiedlichen normativen Aspekte gemeinsam vernünftig systematisieren, ohne in Narrative aufzugehen. Die *Strukturelle Rationalität*, gelebt als gemeinsam geteilte Praktiken menschlichen Miteinanders (Kooperation), ist aufgrund ihrer rationalen Legitimation erfassbar und als normatives Wissen tradierbar. Diese normativen Knotenpunkte, die gegenseitige Anspruchshaltungen formulieren, konstituieren die praktische Lebenswelt und werden durch performative Akte bekräftigt und übertragbar. Ebenso ermöglicht diese Theorie, unterschiedliche Rationalitäten, verstanden als Denkmodi, zu vereinen, um zu einem humanistischen Umgang mit Technik und Technologien zu gelangen. Das dialektische Verhältnis von Technologie und Technik sowie der gesellschaftlichen, normativen Einbettung wird von Philosophen, wie es Reijers und Coeckelbergh (2020, S. 9-10) darlegen, auf verschiedene Weisen definiert: als *technology influencing, co-shaping* oder als *translating, composing* und *delegating* oder als *configuration of practices by technologies*. Das Bedürfnis nach Technik und Technologie verweist bereits auf deren gesellschaftliche Bedingtheit. Der Bedarf einer Technik lässt sich bspw. nach Coeckelbergh (2013) aus einer existenziell-menschlichen Verletzlichkeit heraus erklären, die der Mensch u.a. mittels Einsatzes digitaler Technik zu beheben versucht. Zugleich aber verschiebt er dadurch nur seine Verletzlichkeit, transformiert sie und kann sich nicht befreien, auch wenn dies der sehnlichste Wunsch der Transhumanisten sein mag. Durch die Sehnsucht, seine Schwächen technisch beseitigen zu können, erschafft er dadurch neue Abhängigkeits- und Machtverhältnisse, wie es zum Beispiel Eubanks (2018) in *Automating Inequality – How High-Tech Tool profile, police and punish the poor* soziologisch-em-

pirisch herausarbeitete und dadurch zugleich neue normative Anforderungen formulierte. Der Gedanke der Gestaltbarkeit von Technik und Technologie wird hier erneut deutlich: Schließlich wird das Ziel formuliert, Technik bewusst normativ auszurichten oder sich zumindest bewusst zu werden, wodurch das Bedürfnis nach Technik entsteht, in welchen Zusammenhängen Technik kreiert wird und auf welche strukturellen Gegebenheiten wiederum eingewirkt wird. Doch welche Ethik vermag die Lücke zwischen Technikphilosophie, also der Analyse einer strukturell normativ wirksamen Technik, und einer Bewertung, also einer normativen Ethik, zu schließen?

Wenn nun die maßgebliche Tendenz der technischen Entwicklung dahin geht, dass technische Artefakte so zu konstruieren sind, dass sie Anforderungen nicht technologischer Natur berücksichtigen und sich zu wechselhaften Umwelten zu verhalten vermögen, und in diesem Sinne situiert sind, dann muss man, wie Kogge (2008, S. 939) bereits hervorhob, nicht nur das Deskriptive der Information in den Blick nehmen, sondern Information unter dem Gesichtspunkt betrachten, dass sie auch „eine neue Erscheinungsweise annimmt, nämlich eine ‚performative Dimension‘, in der ‚neue Objekte, Technologien und Realitäten‘ nicht nur beschrieben, sondern hergestellt werden.“ In diesem Sinne ist es als Herausforderung der Ethik zu verstehen, die strukturellen Opazitäten und Transparenzen, die teils als Konstituenzien fungieren, zu identifizieren, um das digitale Artefakt hinsichtlich seiner performativen und transformativen Kraft zu evaluieren.

2.5 Technikethische Grundlagen

Technikphilosophische Überlegungen bieten die notwendige epistemische Orientierung für eine normative Technikethik. So können auch empirische Fallbeispiele, welche die Technikethik seit den 1970ern bereichern, systematisch geordnet werden, denn eine philosophische Betrachtung ermöglicht einen rational-hermeneutischen Zugang zu evaluativ relevanten Gesichtspunkten von Technologie und Technik. Hierdurch kann eine Zuordnung von Best-Practice-Beispielen, Modellen oder Bereichstheorien zu bestimmten ethischen Problemfeldern erfolgen, da gerade digitale Technik aufgrund ihrer Kontextspezifikation nur in ihrer konkreten Anwendung zu bewerten ist. Dies gilt für die strukturelle Bewertung in einem noch weitaus grundlegenderen Sinne als für technik-generische Bewertungskriterien: Denn ohne ein Verständnis von Lebenspraxen, die gemeinsame Handlungsabläufe ordnen sowie Einzelhandlungen in ein Gefüge gegenseitiger Erwartungshaltungen einbinden, kann eine Bewertung digital-technischer Geräte und technologischer Errungenschaften, die eben in jene Muster normativer Lebenspraxen implementiert werden und diese somit maßgeblich steuern, nur unzureichend erfolgen (vgl. zum Begriff der rationalen Lebensführung v.a. Nida-Rümelin, 2001, 2020; Wedgwood, 2014, 2017). Hierbei muss gerade auf die normativen Konstituenzien geachtet werden, die diese Praxen als Lebensform tragen (Teil I, Kapitel 3). Dies geschieht im Rahmen dieser

Arbeit in einem zweifachen Sinne, denn digitale Artefakte sollen hinsichtlich ihrer Auswirkungen auf Lebensformen gedacht sowie die Entwicklungspraxis selbst berücksichtigt werden.

Für eine normativ vollständige Evaluation digitaler Artefakte sind die gesellschaftlich-strukturellen Lebenspraxen aufzuzeigen und die normativ wirksamen Momente eben jener deontischen Strukturen offenzulegen, auf welche und in welchen digital-technische Artefakte wirken. Diese bieten sodann eine lebensweltliche Grundlage für die Diskussion wünschenswerter Lebenspraxen (Nida-Rümelin, 2005a). Darüber hinaus können digitale Artefakte performativ wirksam werden und neue, zuvor unbekannte Formen gemeinsamer Lebensgestaltung generieren (Hubig, 2006).

Der Entwicklungsprozess muss als arbeitsorganisatorischer Ablauf einer eigenen Bewertung unterzogen werden, denn auch dieser ist als Lebensform zu verstehen. So gilt es vor allem hinsichtlich der strukturellen Dimension digitaler Artefakte, also ihrer *strukturellen Kontextspezifikation*, einen Fokus auf lebensweltliche Praxen zu legen. Im Gegensatz zu Vallors (2016) tugendethischem Ansatz liegt der Fokus in dieser Arbeit auf einer Betrachtung der möglichen Einbindung techno-ethischer Urteilskraft in eine Arbeitspraxis, um wünschenswerte techno-moralische Tugenden zu fördern.

Interessant in dieser Diskussion ist, wie sich eine normative Ethik diesen Aufgaben zu stellen vermag, denn technische Unsicherheiten fordern neue Entwicklungs- sowie Nutzerpraxen heraus und normative Unsicherheiten bedürfen der Rückführung in technologische Terminologie oder einer besonderen Ausweisung, dass es sich hierbei um kein technologisch oder technisch lösbares Phänomen handelt. Unter Rücksichtnahme der Reflexionsbegriffe *Opazität* und *Transparenz* sollen diejenigen normativen Aspekte herausgearbeitet werden, die als Bewertungskriterien Maßnahmen für die technische Umsetzbarkeit anzuleiten vermögen. Die Lokalisation normativer Tatsachen hinsichtlich digitaler Technologien und Technik bedarf eines gezielten epistemischen Vorgehens, da zusätzlich zu den klassischen, technik-generischen, ethischen Gesichtspunkten, wie bspw. Privatsphäre oder Fairness, noch gesellschaftliche Lebensformen Beachtung finden müssen. Mit Hilfe der Identifikation geeigneter *levels of abstraction* können diejenigen Szenarien herausgestellt werden, die für eine ganzheitliche Analyse den Ausgangspunkt bilden (Teil II, Kapitel 4). Hierbei müssen zunächst normative Tatsachen, die im Rahmen der Technikethik meist als Werte artikuliert werden (Gogoll et al., 2021), der Erkenntnis zugänglich gemacht werden, um im Anschluss gemeinsam mit empirischen Tatsachen zu normativen Urteilen gelangen zu können. Dabei ist grundlegend, dass Werte Normen vorausgehen. Normen als Moralgebote unterscheiden sich von Gesetzesnormen, insofern letztere durch einen Gesetzgebungsprozess gesetzt werden, während sich Moralgebote durchweg argumentativ begründen lassen müssen. Metaethisch ist diese Position einem Kognitivismus zuzuordnen, da die Moralgebote ihre Gültigkeit durch einen argumentativen Nachweis ihre allgemeine Einsichtigkeit erhalten (Skorupinski & Ott, 2000): Denn „gültig ist nur das, was gerechtfertigt oder begründet werden kann. Begründen

heißt, strittige normative Geltungsansprüche zu stützen bzw. ‚einzulösen‘ (ebd., S. 12). Dies ist unerlässlich, wenn ein Sollensanspruch begründet werden soll, d.h. der Frage nachgegangen wird, für welche guten, normativen Gründe eben jene Tatsachen sprechen (Nida-Rümelin 2020, Kapitel VI). Das Verhältnis zwischen den Tatsachen konstituiert Gründe, die Nida-Rümelin (2020, S. 335) als „*inferenzielle Beziehungen zwischen empirischen und normativen Tatsachen* [Herv. d. A.], zwischen der empirischen Tatsache, dass das Messer scharf ist und der normativen, dass man es dem Kind nicht geben sollte“, expliziert. Allerdings rationalisiert die Einführung des Begriffs *Interferenz* diesen normativen Geltungsanspruch, sofern unter Interferenz ein ausschließlich kognitiv-logisches Konzept verstanden wird, durch das der Prozess des Zu-Gründe-Machens zu einem einseitig erkenntnistheoretischen Unterfangen wird.⁷ Dies ist vor allem dann der Fall, wenn zudem angenommen wird, dass Tatsachen nur qua einer ethischen Reflektion, die mittels *eines* ethischen Prinzips durchgeführt wird, einen normativen Sachverhalt erst zur Geltung zu bringen vermag, d.h. dass Gründe nur qua eines kognitiven, rationalen Erkenntnisprozesses normative Gültigkeit ausüben können (siehe hierzu die Kritik von Kiesewetter, 2017). Da die Tatsachenerkenntnis im Entscheidungsprozess, also in der praktischen Deliberation, bereits abgeschlossen ist, können Tatsachen qua Haltungen, qua einer Rollenausübung oder qua praktischer Konventionen, die allesamt die lebensweltlichen, deontischen Strukturen tragen (Teil I, Kapitel 2), zu Gründen modifiziert worden sein. Normativität ist nur durch ein strukturelles Geflecht gegenseitiger Erwartungshaltungen erklärbar und lässt sich nicht durch eine Binnenrelationalität rationaler Strukturen erklären. Dies ist gerade für die von Nida-Rümelin entwickelte *Strukturelle Rationalität* von grundlegender Bedeutung. Die Modifikation der Tatsachen zu Gründen bzw. diese *inferenzielle Beziehung* geht der Entscheidung über Gründe voraus. D.h. in dieser Normativität bewegt sich jegliche Lebenswelt immer schon. Aus der Relationalität der erkannten Tatsachen sind einige als (auch gute) normative Gründe erkannt worden, die in der ethischen Deliberation nochmals gegeneinander abgewogen werden können, wenn sich nicht ein Grund als der beste ohne größere Anstrengung (d.h. ohne weitere erforderliche kognitive Deliberation) zeigt. Eine auf einer Entscheidung⁸ beruhende gute Handlung ist somit eine ausgeführte Handlung, die durch gute normative Gründe motiviert ist.⁹ Da normative Gründe unabhängig eines Wollens (Nida-Rümelin, 2020; Scanlon, 2000) existieren, beruht der Begründungsablauf auf folgenden drei Fähigkeiten:

⁷ Der Diskurs der Praktischen Interferenz und wie dieser mit Erkenntnis, Wissen, Denken und Handeln zusammenhängt, kann an dieser Stelle nicht nachgezeichnet werden (Weinberger, 1979).

⁸ Entscheidung umfasst hier nicht nur den rational-kognitiven Prozess, der unmittelbar individuell vor der Handlung ausgeführt wurde, sondern kann auch in Form von Tugenden, Rollenbildern oder Haltungen tradiert sein.

⁹ Auf die Debatte normativer und motivationaler Gründe kann im Rahmen dieser Arbeit nicht vertiefend eingegangen werden. Wenn gute normative Gründe immer motivierten, gäbe es das Phänomen der Willensschwäche

- 1) normative Tatsachen zu erkennen, d.h. empirische, normative und soziale Tatsachen angemessen erkennen zu können sowie ihrer Relationalität einsichtig zu werden,
- 2) normative Gründe angemessen abzuwägen und
- 3) sich von den guten normativen Gründen auch leiten zu lassen. Dies kann als motivationale Ebene bezeichnet werden.

Die praktische Abwägung der Gründe kommt dadurch zustande, dass praktische Konflikte in der Lebenswelt auftreten können, die nur durch eine praktische Deliberation gelöst werden können (Nida-Rümelin, 2020; Teil I, Kapitel 2). Diese Spannungsverhältnisse führen in das „*Spiel des Begründens*“ (ebd., S. 80), das darauf abzielt, Ungewissheiten zu lösen, indem durch die Systematisierung unterschiedlich akzentuierter Argumente die Unsicherheit abgeschwächt oder relativiert werden soll. Es geht nicht darum, fundamentale Letztbegründungen zu erfassen, sondern theoretische Systeme erneut zu stabilisieren (ebd.), denn Moralgebote erheben ihren Pflichtanspruch, indem sie in Situationen eines bestimmten Typs eine Handlung fordern oder deren Unterlassung gebieten. Im Gegensatz zu Rechtsnormen erfolgt auf ein Überschreiten keine Sanktion, sondern eine Empörung oder Ächtung (Bicchieri, 2005). Moralgebote werden dann hinterfragt, wenn sie nicht mehr zu erstrebenswerten Resultaten in bestimmten Situationen führen: Sie werden einer kritischen, nämlich ethischen, Analyse unterzogen. Mittels dieser praktischen Deliberation werden ethische Normen moralischen Urteilens und moralischen Handelns formuliert. Aus diesen artikulierten Normen lassen sich wiederum normative Kriterien entwickeln (Skorupinski & Ott, 2000, S. 12-15). Diese Kriterien sind „als operationalisierte Normen zu verstehen“ (ebd., S. 14) und sind „in ihrer normativen Dimension [...] Maßstäbe, in ihrer pragmatischen Dimension [...] Aspekte“ (ebd., S. 15).

2.5.1 Die Bereichsethik digitaler Artefakte: Makro-, Meso- und Mini-Ethiken

Eine ethische Begründung zielt darauf ab, moralische Überzeugungen zu belegen, indem sie moralische Aussagen durch „Verknüpfung von Propositionen“ (Nida-Rümelin, 1998, S. 9) in einen systematischen Zusammenhang bringen soll (ebd.). Zur Rechtfertigung können ethische Theorien als Begründungsprinzip hinzugezogen werden. Diese sind stark vereinfacht entweder deontologisch, konsequentialistisch, tugendethisch oder kontraktualistisch.¹⁰ Diese ethischen Theorien verbinden morali-

nicht. Dass jedoch alle Gründe, sofern sie motivieren, normativ wirksam werden, spricht vielleicht für vertikale Ebenen oder Prozesse, die gute Gründe in ihrer Verwirklichung durchlaufen.

¹⁰ Für eine Gegenüberstellung der ethischen Theorien und ihre Bedeutung für die Angewandte Ethik siehe Nida-Rümelin (2005, S. 2-88).

sche Überzeugungen und subsumieren sie unter gemeinsame Kriterien. So bilden sie ein Konglomerat moralischen Wissens (Nida-Rümelin, 1998). Die normative Ethik bringt so moralische Gebote begründend in eine Ordnung, die sich in lebensweltlichen Erwartungshaltungen offenbaren und deren Angemessenheit sich zugleich an und in diesen bemessen lassen muss (Nida-Rümelin, 1996, S. 61). Neu auftretende Handlungsfelder, die etwa durch Technik entstehen, führen gerade deshalb zu Unsicherheit, weil weder moralische Gebote formuliert wurden, da größtenteils die relevanten normativen Tatsachen erst noch eruiert werden müssen, noch wünschenswerte lebensweltliche Praxen existieren, die diese normativen Tatsachen leb- und erfahrbar machen. Technische Instrumente wirken gerade auf solche lebensweltlichen Praxen ein, die, verstanden als strukturelle Rationalitäten, bestimmte Lebensweisen konstituieren, wie z.B. die Praxis der Liebe. So besteht Liebe aus einer Vielzahl einzelner Handlungen, welche die Liebe erst als Lebensart konstituieren. Bspw. gilt es, Blumen vorbeizubringen, gemeinsam Abend zu essen, Rücksichtnahme walten zu lassen, emotionale Teilhabe zu ermöglichen etc. Werden jedoch in stabile Lebenspraxen digitale Geräte eingeführt, können diese die normativen Konstituenzen jener Praxen transformieren. Z.B. geschieht dies hier durch Dating-Apps oder es werden gar neue Handlungspraxen erschaffen, was sich z.B. durch Online-Games veranschaulichen lässt (Flanagan, Howe & Nissenbaum, 2005). Diese transformierende oder performative Erschütterung führt zu einem erhöhten Bedarf an ethischer Orientierung, was sich in der Vielzahl an Ausrufen nach moralischer Software zeigt. In den allermeisten technischen Szenarien sind Werte oder Normen in ihrer Allgemeinheit nicht einfach übertrag- und anwendbar (Gogoll et al., 2021). Dies ist der Unterbestimmtheit von Werten und Normen geschuldet. Aber gerade erst die Berücksichtigung von Werten und Normen, also ein regelgeleitetes Vorgehen, lassen Entscheidungen vernünftig werden.

So zeigt das Beispiel von Skorupinski und Ott (2000), dass das Tötungsverbot nicht einfach auf Verkehrssysteme anzuwenden ist, da sonst Verkehr im heutigen Sinne gar nicht zulässig wäre. Wenn aber Ethik darauf beruht, moralische Überzeugungssysteme zu rekonstruieren und zu systematisieren, um Kriterien zu entwickeln, die in solchen Situationen, in denen ein moralisches Urteil nicht eindeutig ist, Orientierung bieten können, dann ist das prinzipiengeleitete Vorgehen klassischer angewandter Ethik unangemessen (Nida-Rümelin, 1996).

Das Verhältnis ist vielmehr eine Umkehrung: Konkrete Probleme werden konstitutive Bestandteile der ethischen Theorie selbst, moralische Intuitionen bieten somit das Ausgangsmaterial der Urteilskraft (ebd.). Nida-Rümelin betont zugleich, dass die Paradigmen, welche die moderne Ethik dominieren, zu unterkomplex seien, als dass sie eine Beurteilung aller Situationstypen vornehmen könnten (ebd.). Daraus folgert er, insofern es nicht die *eine* Ethik gibt, die alle Situationen und Handlungstypen evaluativ zugänglich machen kann, dass Bereichsethiken auszuarbeiten sind. Diese Bereichsethiken

zeichnen sich dadurch aus, dass „[s]ie [...] die Fachgrenzen [überschreiten N.Z.], um unter Einbeziehung des Sachverstandes anderer Disziplinen, anderer Berufsfelder und Handlungserfahrungen der Komplexität moralischen Urteils und Handelns in der Lebenswelt gerecht [...] werden“ (Nida-Rümelin, 1998, S. 4). Im Gegensatz zur Angewandten Ethik, die sich dadurch kennzeichnen lässt, dass sie sich moralischer Prinzipien bedient, die auf einen bestimmten Sachverhalt Anwendung finden sollen, beruhen Bereichsethiken auf gesellschaftlichen Subsystemen, die spezifische Bereiche menschlicher Praxis betreffen (Nida-Rümelin, 1996, S. 63).

Unter einer Bereichsethik versteht Nida-Rümelin „gesellschaftliche Subsysteme“, die einen bestimmten, für den gesellschaftlichen Bereich konstitutiven Zweck verfolgen, die entsprechend idiosynkratische, moralische Probleme hervorrufen (ebd., S. 63). Diese gesellschaftlichen Bereiche menschlicher Praxis erfüllen normative Kriterien und können sich nicht auf ein System moralischer Regeln oder Prinzipien reduzieren lassen. Dies hat viele Ursachen, wie Unwissenheit, aber auch die Explosion neuartiger Handlungsoptionen trägt dazu bei. Aber genau dann, wenn Handlungen nicht mehr in Abhängigkeit vorgegebener Strukturen gewählt werden können, wenn sich keine Handlung mehr optimal in lebensweltlich wünschenswerte Strukturen einfügen lässt, weil neue Handlungsoptionen eröffnet wurden, führt dies zunächst in eine normative Orientierungslosigkeit. Dies beruht darauf, dass die Strukturen selbst hinterfragt werden, neue normativ wünschenswerte Strukturen erst gewählt und sodann etabliert werden müssen. Nida-Rümelin (1996, S. 64 ff.) veranschaulicht diesen Zusammenhang sehr deutlich am Beispiel der Gentechnologie, während Vallor (2016) dies aus tugendethischer Perspektive für die Informationstechnik vornimmt. Gerade die Wandelbarkeit der Strukturen, durch die sich die bisherigen Haltungen entwickeln und behaupten, ermöglicht zwar, dass neue Technik in gesellschaftliche Bereiche vernünftig eingegliedert werden kann, aber stiftet zugleich eine Unordnung. Dies wiederum führt dazu, dass eine theoretische Abwägung stattfinden muss, damit althergebrachte, tradierte gesellschaftliche Formen neu modelliert und angepasst werden können.

Im Falle der Technikethik ist die Festlegung eines eigenen Bereichs weitaus schwieriger, denn hier geht es nicht wie in der Medizinethik um eine Praxis der Heilung und Therapie (Nida-Rümelin, 1996) oder wie in der Wirtschaftsethik um eine Praxis der Güterverteilung unter Ressourcenknappheit, denn die Entwicklung und Nutzung digitaler Endgeräte betrifft alle lebensweltlichen Praxen: Es geht folglich um den Einsatz von Computern oder die Digitalisierung von Informationen und die dadurch hervorgerufene Veränderung oder Erschaffung lebensweltlicher Praxen. Die Bereichsethik, Technik und Technologie lässt sich daher in diesem Zusammenhang mehr durch den Sachgegenstand selbst abgrenzen und erst in zweiter Linie durch ihren Einsatz in einem gesellschaftlichen Subsystem. Dies führt zu einer Vielzahl an Makro-, Meso- und Miniethiken im Zusammenhang mit digitaler Technologie und Technik.

Die grenzenlos erscheinenden Einsatzmöglichkeiten digitaler Technik spiegeln sich in der schier Unendlichkeit möglicher Untersuchungsobjekte und deren vielfältigen Bezeichnungen wider: Prozessautomatisierung, Computer, Roboter, Internet, digitale Artefakte, digitale Produkte, digitale Anwendungen, Software, Softwaresysteme, Hardware, Maschinen, Künstliche Intelligenz und Maschinelles Lernen werden in einem Atemzug mit Digitalisierung und digitaler Transformation verwendet. Oft ist dabei nicht klar, ob diese Begriffe dabei als digitales Artefakt, ökonomisches Produkt oder (auch) als Tätigkeit und Vorgehensweise verstanden werden sollen, die kulturelle Veränderungen zur Folge haben.

Ähnliches gilt für die Bezeichnung der normativen Wissenschaft, die sich theoretisch mit der Fundierung der Fragestellungen auseinandersetzt: der Ethik. Cybernetics (Wiener, 1950), Cyberethik (Tavani, 1996), Digitale Ethik (Capurro, 2009; Ess, 2013; Spiekermann, 2019), Informationsethik (Floridi, 2013), Computerethik (Johnson, 1985; Maner, 1999; Moor, 1985), Internetethik (Langford, 2000), Berufsethik (Bynum & Rogerson, 2004), Maschinenethik (Anderson, Anderson & Armen, 2004; Moor, 1985; Wallach & Allen, 2008), Datenethik (Zwitter, 2014) oder auch *ethics of AI* (Bostrom & Yudkowsky, 2014), um nur einige wenige Ethiken und ihre Vertreter aufzuzählen, beschäftigen sich mit den normativen Auswirkungen der Informationstechnologie und -technik auf individuelle Lebensentwürfe, gesellschaftliche Strukturen oder kollektive Zielsetzungen. Dabei können die unterschiedlichen ethischen Perspektiven jeweils durch den Betrachtungsgegenstand voneinander abgegrenzt werden: So beziehen sich Datenethik oder *big data ethics* bspw. auf die Datenqualität, die Datenverarbeitung und Fragen der Datenstatistik, während sich die Berufsethik weitgehend mit Verhaltenskodizes für Softwareingenieure befasst oder dezidiert technische Handlungsanweisungen für Softwareentwickler im Hinblick auf Wertcluster, wie Privatsphäre (z.B. *privacy by design* (Schulz, 2012)) oder Verantwortungszuschreibungen (z.B. *accountability*-Modelle (Kacianka & Pretschner, 2021)), formuliert. Die Maschinenethik wiederum beschäftigt sich weitgehend mit der Entwicklung vollständig autonomer Maschinen, d.h. solcher, die eigenständig moralische Entscheidungen treffen können sollen oder diese zumindest simulieren. Diese Fragestellung teilt sich die Maschinenethik sogleich mit einer *Ethics of AI*, um die prinzipienethische Grundlage einer *digital moral agency* sowie deren mögliche technische Umsetzung zu diskutieren. Dies veranschaulicht bspw. auch die Debatte um autonome Fahrzeuge in Trolley-Situationen (Gogoll & Müller, 2017; Hevelke & Nida-Rümelin, 2015). In einer Ethik der Künstlichen Intelligenz hingegen werden philosophisch umgreifende Themengebiete diskutiert, wie die Möglichkeit zu eigenständiger Denkfähigkeit, das Leib-Seele-Problem, also ob Maschinen ein Bewusstsein werden haben können oder ob Maschinen über ein Sprachverstehen verfügen können (Kastendiek, 2003). Zudem liegt der Fokus der KI-ethischen Fragestellungen auf *machine-learning*-Algorithmen sowie der Datensouveränität und deren lebensweltlichen Effekten: Diskriminierungsfreiheit, Gerechtigkeit sowie Datenkontrollen werden auf ihre normativen Ansprüche hin erörtert.

Dem Vorgehen, dass alle informationstechnischen Facetten in unterschiedlichen Subdisziplinen erörtert werden müssen, weil ihnen idiosynkratische Merkmale eigen sind, stehen die Informationsethik, die Digitale Ethik und die Computerethik skeptisch gegenüber. Computerethik ist der älteste der genannten Begriffe, wenn Norbert Wieners Kybernetik nicht berücksichtigt wird, denn er sich nicht mit der ethischen Dimension seines soziologisch-strukturalistischen Ansatzes von Informationsrückkopplungseffekten zwischen Maschinen und Menschen auseinandergesetzt hat (Wiener, 1988). Die Computerethik hingegen befasst sich mit ethischen Fragen, die aus dem Umgang mit Informationstechnologien und -technik resultieren. Dadurch wird eine Anbindung an die praktische Philosophie unmittelbar erkennbar, denn nun rücken Handlungskontexte zusammen mit Maschinen in den Mittelpunkt einer Evaluation und die Frage nach deren moralischer Güte kommt auf. Moor (1985) erweiterte die Fragestellung, indem er die Handlungsfelder gemeinsam mit den Besonderheiten der Informationstechnologie und -technik dachte, um Handlungsanforderungen an die Politik aus dem Gegenstandsbereich selbst ableiten zu können. Zudem eröffnet die Zusammenführung dieser beiden Fragestränge die Möglichkeit, die Auswirkungen, die auf der Besonderheit von Informationstechnologie und -technik gründen, auch auf die Handlungsfähigkeit von Softwareentwicklern im Entwicklungsprozess zu denken. Diese Herangehensweise an eine Ethik von Informationstechnologien und Informationstechnik orientiert sich an der Perspektive der Profession, soweit diese für die Softwareentwicklung klar umrissen werden kann. Dies unterscheidet auch die Technikethik von der Technikfolgenabschätzung, da sich Überlegungen zur Verantwortung von Softwareentwicklern nicht auf die politischen Bewertungskriterien reduzieren lassen (Skorupinski & Ott, 2000). Es ist daher wenig verwunderlich, dass die Maner-Johnson-Diskussion über das computertechnische Fundament einer Ethik direkt in die Möglichkeit der Kompetenzerweiterung und Ausbildung von Softwareingenieuren führt (Johnson, 1985; Maner, 1999). Die Computerethik erfordert demnach technische Kenntnisse des Betrachtungsgegenstandes und entfernt sich somit schon von klassischen Prinzipienethiken, die universelle Gültigkeit unabhängig ihres Anwendungsbereichs beanspruchen. Auch der Ansatz der *ethics by design* als eine normativ wünschenswerte Ausgestaltungsmöglichkeit informationstechnischer Artefakte ist erst durch diese Zusammenführung denkbar geworden (Friedman, 1997). Sie wendet sich daher als Methodik mit konkreten Beispielen direkt an die Profession.

Die Digitale Ethik oder *digital media ethics* reflektieren wiederum, welche Wertmaßstäbe und Überzeugungen in einer vernetzten Welt wünschenswert sind (Grimm, Keber & Zöllner, 2019) und klären „speziell moralische Fragen [...], die sich aus der Datenverarbeitung unserer Tage ergeben [um Grundregeln zu entwickeln, wie man vernünftig damit umgehen kann]“ (Wewer, 2020, S. 1). Es geht hier im Unterschied zu den oben genannten, aus der Softwareentwicklung entstandenen, ethischen Ansätzen vor allem darum, zu überprüfen, welche ethischen Theorien sich für eine Thematisierung der Digitalisierung oder der Digitalen Transformation besonders eignen (Capurro & Grimm, 2020, S.

62). Die Digitale Ethik formuliert somit Bedingungen, unter denen ein gutes Leben ermöglicht wird und untersucht digitale Artefakte dahingehend, ob sie diesem Telos zuträglich sind oder nicht. Diese eudämonistische Ausrichtung in der Formulierung dieser Kriterien bzw. Tugenden als wünschenswerte Verhaltensweisen und deren Förderung oder Unterminierung durch den Einsatz von Informationstechnik und -technologie findet sich bei Grimm, Keber und Zöllner (2019), aber auch bei Spiekermann (2019) und philosophisch fundiert bei Vallor (2016) sowie Reijers und Coeckelbergh (2020). Auch Floridis umfassende *The Ethics of Information* (2013) findet in den tugendethischen Ansätzen ihr Pendant.

Im Gegensatz zur Computerethik zeigt die Digitale Ethik vor allem die moralischen Belange der Nutzer auf. Diese unterschiedlichen Perspektiven – die der Profession der Entwickler sowie der Endnutzer – lassen beide Ethiken wieder als spezielle Unterformen erscheinen, die sich aufgrund ihres Adressatenkreises abgrenzen lassen. Capurro (2017b) betont, dass man bereits seit den 1940ern von einer Digitalen Ethik sprechen könne. Auch wenn sich diese Ethik damals als Computerethik bezeichnete, sprach man von einer professionellen Ethik, obgleich hier die gesellschaftlichen Auswirkungen der Informationstechnik- und technologie behandelt wurden. Dies verweist bereits darauf, dass die einzelnen Ethiken nicht trennscharf voneinander abgrenzbar sind, sondern sich immer wieder gegenseitig bedingen, insofern sie Urteile systematisieren und selbst normative Urteile fällen, die für andere Bereiche bedeutsam sind.

Diese begriffliche Verwobenheit ist mitunter ein Grund dafür, weshalb Luciano Floridi in seiner *philosophy of information* eine eigenständige philosophische Disziplin herausarbeitet und diese streng von einer Computerphilosophie abgrenzt (Floridi, 2013). Der Begriff der Berechnung (*computation*) beinhaltet bereits einen bestimmten Umgang mit Informationen und muss daher vom grundlegenden Konzept der Information abgegrenzt werden. Eine Philosophie der Information ist dabei häufig phänomenologisch *biased*, weil Informationstechnologien und -technik vornehmlich als Bedeutung und weniger metatheoretisch reflektiert werden (Floridi, 2002). Er kritisiert somit auch einen einseitigen hermeneutisch-phänomenologischen Zugang zur Technikethik, da dieser weder Aspekte der Herstellung noch intersubjektive und kollektive Rationalitätsanforderungen einzufangen vermag: Fragen der Art, wie Informationen verarbeitet, aufbereitet, interpretiert oder verwendet werden sollen, bleiben ausgeschlossen und werden erst durch eine ontologische sowie metaphysische Betrachtung in ihrer Relationalität vollumfänglich erörterbar. Wenn sie nämlich rein perspektivisch und in Abhängigkeit der wahrnehmenden Subjekte charakterisiert werden, entfällt sogleich die Betrachtung der Beschaffenheit der Information. Folglich können Fragen nach dem Sein der Information und die dadurch allumfassende, gleichrangige Darstellungsmöglichkeit der Welt als diskrete oder kontinuierliche Information nicht diskutiert werden. Dadurch gelingt es Floridi, eine Brücke von einer grundlegenden Philosophie der Information, die sich als Teil einer Technikphilosophie begreifen muss, zu einer Ethik des

Informationsdesigns zu schlagen, die sich auch mit Bedeutungseinheiten und Transitionen zwischen technischen Abstraktionsniveaus auseinanderzusetzen vermag. Sowohl die verfügbaren Informationen als auch die Kompetenzen der Agenten beeinflussen positiv oder negativ die Welt (*info-sphere*), in die sie eingreifen. Diese Weltveränderung geschieht entweder in der Form einer Transformation des Seins (*re-ontologizing*) oder aufgrund einer Erweiterung der Ressourcenmenge. So erweitert ein neuer Blogbeitrag bspw. die Menge vorhandener Ressourcen und mag als Kommunikationsform oder Wissensvermittlung Wahrnehmungs- und Erlebnisbezogenheit individuell, intersubjektiv oder kollektiv verändern: Ein Individuum mag die Welt neu aufnehmen, mag durch einen Kommentar sein intersubjektives Kommunikationsverhalten verändern und gemeinsam kann dies zu veränderten sozialen Normen und Pflichten aufgrund einer Um- oder Neubewertung eines Sachverhalts führen. Blogbeiträge oder allgemein professionelle informationstechnische Anwendungen können sowohl als Ressource als auch als Produkt verstanden werden, die sich auf reale lebensweltliche Beziehungsgeflechte auswirken. Eine ethische Analyse muss sich folglich dieser Dichotomie bewusst sein und beide Schwerpunktsetzungen ermöglichen. Sowohl die Information als Ressource als auch als Produkt verändert oder ergänzt die Welt, kann Wissen, Verhalten und Wahrnehmungen transformieren und somit auch wieder erneute Informationen als Ressourcen bereitstellen oder selbst als solche gelten.

Information als Ressource zu verstehen, verweist somit auf jene ethische Betrachtung, die um die Qualität und Distribution der zur Verfügung stehenden Informationen bemüht ist. Ein Verständnis der Information als ein Produkt wiederum rückt den Prozess der Herstellung von Informationen durch den moralischen Agenten in den Mittelpunkt. Erstere Ethiken verweisen auf die Nutzung von Daten, während letztere sich mit Fragen einer moralischen Herstellung von Informationsprodukten beschäftigen. Rückkopplungseffekte dieser Bereichsethiken werden durch das *Resource-Product-Target-Modell* sichergestellt, da sich die Nutzung und die Produktion von Daten oder weiterer Funktionen nicht exakt analytisch abgrenzen lassen (Floridi, 2013). Die Fokussierung ermöglicht jedoch eine genauere Betrachtung eines konkreten Sachverhalts oder eines anderen *levels of abstraction*. Jede Modifikation der Information oder Erschaffung neuer Informationen transformiert die Welt, indem sie sie um Informationseinheiten ergänzt und Verhaltensanpassungen erfordert, also zum Phänomen der digitalen Transformation führt. Dieser transformative Charakter mag Floridi dazu verleiten, diesen Akt eigens als „Information as Target“ (Floridi (2008), S.4) zu beschreiben (Abbildung 3). Allerdings bleibt unklar, ob die mögliche transformierende Wirkung auf die Welt mittels der Erschaffung digitaler Produkte bereits ausreichend verdeutlicht wird (Floridi, 2013). Floridi hebt somit die Relationalität der einzelnen Bausteine in den Mittelpunkt seiner makroethischen Betrachtung.

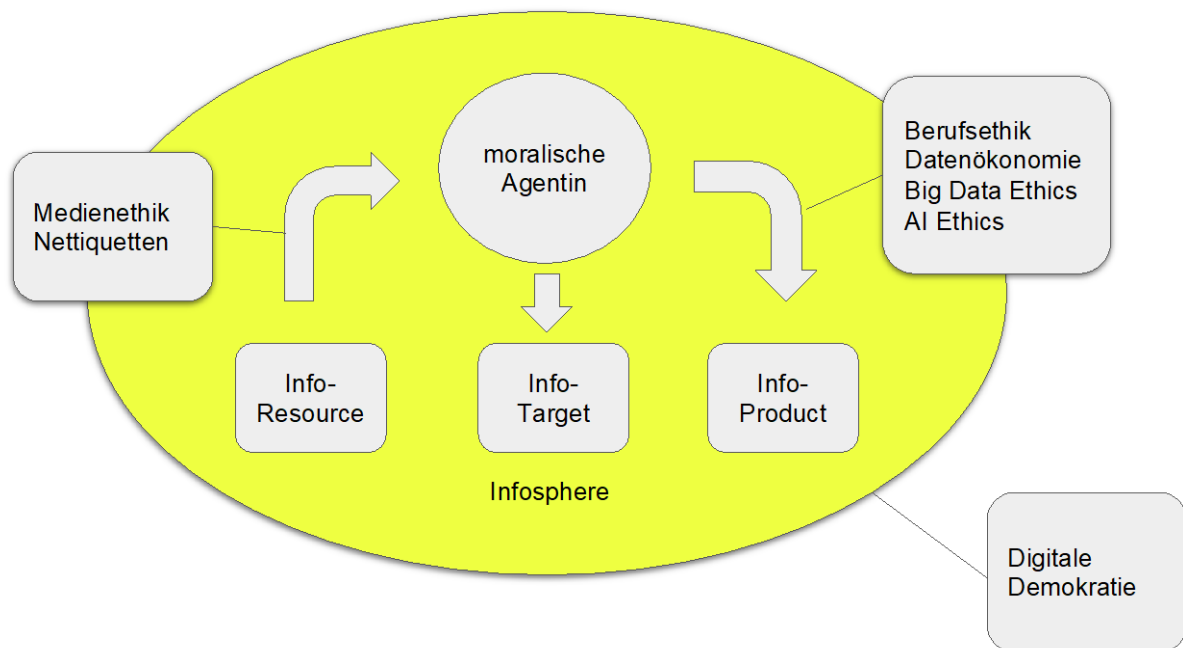


Abbildung 3 Resource-Product-Target-Modell mit Zuordnung der Bereichsethiken nach Floridi (2008). Eigene Erweiterung.

In diesem Sinne formuliert Floridi eine *disclosive ethics*, insofern dieses Modell epistemische Orientierung ermöglichen soll, um moralisch relevante Bedeutungskontexte zu lokalisieren und dementsprechend adressieren zu können. Floridis Informationsethik schließt keine ethischen Theorien aus, damit jede Situation hinsichtlich ihrer normativen Implikationen adressiert bleibt, sondern versucht, einen Rahmen zu formulieren, der es ermöglichen soll, die Beziehungen zwischen den einzelnen Betrachtungseinheiten, seien diese digitale Artefakte oder Personen, zu berücksichtigen. Die Frage, was es zu tun gilt, wird im Nachgang an diese Topologisierung durch eine normative Ethik ermöglicht, die jedoch im Sinne Floridis keine Ethik ausschließt, sondern sich aller Pflichtbegründungen als rationale Handlungsalternativen annimmt.

So ergibt sich dadurch eine Pyramide der Bereichsethiken digitaler Technologie und digitaler Technik, die von einer allumfassenden Informationsethik zunächst auf Perspektiven, d.h. Adressaten, wechselt und sodann in Mikroethiken zu zergliedern ist, die sich über den Gegenstandsbereich abgrenzen. All diesen zunächst deskriptiv ausgerichteten Ethiken ist eine ontologische Fragestellung gemein, die sich auf eine Topologisierung der ethischen Vorgehensweisen stützt. Normative Urteile beruhen auf einer deskriptiven Beschreibung der ethisch relevanten Phänomene. Eine Ergänzung der normativen Ethik um ontologisch-deskriptive Sachverhalte ist in einer Informationsethik unabdingbar: Nur so lassen sich Adressat und Forderung präzise zusammenführen und Handlungsspielräume ausloten.

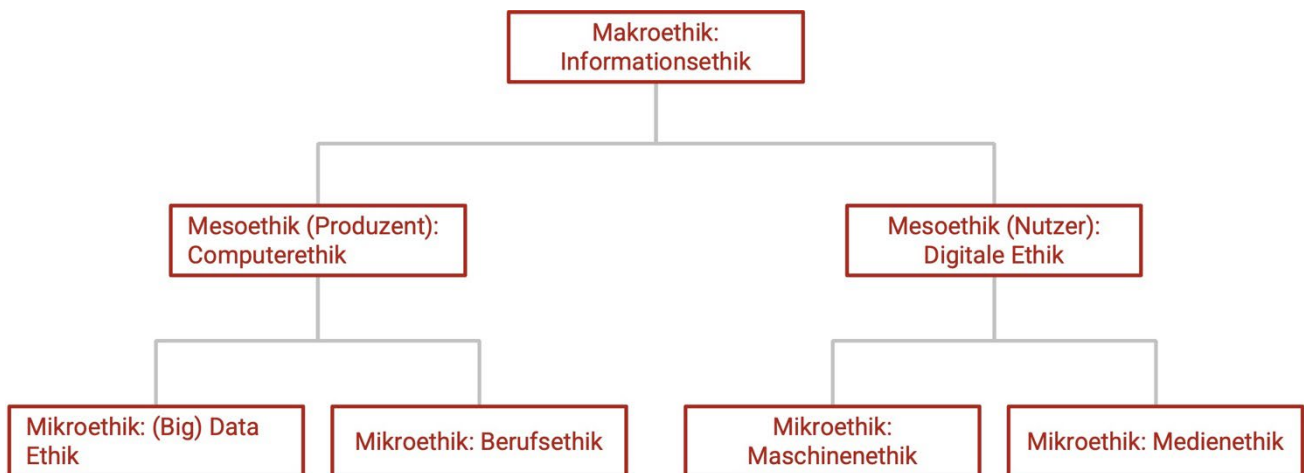


Abbildung 4 Pyramide der Bereichsethiken Digitaler Technologien und Digitaler Technik. Eigene Abbildung.

Die unterschiedlichen ethischen Ansätze und die Verschiedenheit ihrer Beobachtungsobjekte lassen vermuten, dass es sich bei einer Ethik der Information, sollte diese als Überbegriff aller mehr oder weniger perspektivisch anmutenden Ethiken dienen, um einen einzigartigen Forschungsbereich handelt, der aufgrund der Beschaffenheit eben jenes Betrachtungsobjekts einer eigenen ethischen Deliberation sowie eines eigenen ethischen Verfahrens bedarf. Jedenfalls kann nicht mehr von einer strengen Anwendung klassischer prinzipienethischer Ansätze auf einen klaren Sachverhalt gesprochen werden. Die verschiedenen Bereichsethiken können sodann metaethisch prozedural oder inhaltlich ausgerichtet sein. Die ethische Deliberation zielt entweder darauf ab, inhaltliche Kriterien aus dem Sachgegenstand zu entwickeln oder versucht die Deliberation als Entscheidungsfindungsprozess in den Entwicklungsprozess zu etablieren, um diese normativ relevanten Aspekte prozedural einfangen zu können.

2.5.2 Humanistische Ethik: Vernunft, Freiheit und Verantwortung

Die Vielzahl an einzelnen Bereichsethiken führt tatsächlich zur mehr normativer Klarheit hinsichtlich techno-ethischer Unsicherheit. Allerdings müssen diejenigen ethischen Paradigmen herausgearbeitet werden, welche die Bereichsethiken zusammenführen können. Im Kapitel über die technikphilosophischen Grundlagen wurde ersichtlich, dass Rationalität und Sozialität wesentliche Bedingungen für die Entfaltung der menschlichen Handlungsfähigkeit sind. Die Fähigkeit zur Autonomie, die sich als strukturierte Intentionalität zeigt, gilt es als Kern dieser humanistischen Philosophie zu proklamieren. Für ein ethisches Design bedeutet dies, dass die Ausgestaltung sich an den Paradigmen humanistischer Ethik ausrichten sollte, um ein humanistisches Design zu gewährleisten.

Im Zentrum humanistischer Philosophie und humanistischer Praxis steht die Idee menschlicher Autorschaft. Menschen sind Autorinnen und Autoren ihres Lebens, sie tragen als solche Verantwortung und sind frei. Diese Trias aus Vernunft, Freiheit und Verantwortung spannt ein Cluster von normativen Begriffen auf, der das humanistische Verständnis der *conditio humana* bestimmt und in einem langwierigen kulturellen Prozess über Jahrhunderte sowohl die lebensweltliche Moral wie die Rechtsordnung prägt. Diese normative Begrifflichkeit gruppiert sich um das Phänomen der Affektion durch Gründe. (Nida-Rümelin, 2022, S. 31)

Der Appell, der sich daraus für einen dezidiert humanistischen Umgang mit Softwaresystemen und deren Entwicklung ergibt, kommt nicht nur von Seiten Nida-Rümelins (Nida-Rümelin & Weidenfeld, 2018). Auch das Wiener Manifest spricht sich für einen Digitalen Humanismus (Werthner et al., 2019). So heißt es dort, dass „[w]ir (...) Technologien nach menschlichen Werten und Bedürfnissen formen [müssen N.Z.], anstatt nur zuzulassen, dass Technologien Menschen formen. Unsere Aufgabe besteht nicht nur darin, die Nachteile der Informations- und Kommunikationstechnologien einzudämmen, sondern vor allem auch darin, von Beginn an menschenzentrierte Innovationen zu fördern“. Doch wie muss der Anspruch verstanden werden, dass Technologien und Technik menschenzentriert entwickelt werden sollen und Menschen sich nicht permanent an Technik und Technologien anzupassen haben? Digitale Technik, die genuin menschliche Fähigkeiten verstärkt und diese nicht korrumpiert, bedarf einer ethischen Deliberation, die nicht auf eine reine Kosten-Nutzen-Abwägung reduziert werden kann und die es von Anbeginn des Herstellungsprozesses an einzubeziehen gilt. Normativität erschöpft sich weder in einer konsequentialistischen Betrachtung digitaler Technik noch in einer Abwägung konfligierender Werte. Im Gegenteil, normative Ansprüche manifestieren sich in strukturellen Handlungsketten, die vernünftige Lebensformen tragen, und in der menschlichen Fähigkeit, eigene Entscheidungen treffen zu können sowie sich nach diesen auszurichten. Digitale Technik als Teil lebensweltlicher Handlungsstrukturen zu verstehen, ermöglicht technische Artefakte in ihrer ethischen Komplexität sowie in ihrem pragmatisch-normativen Anspruch zu begreifen. Ein solches humanistisches Technikverständnis muss sich auch zwangsläufig die Frage stellen, wer Adressat dieser Aufforderung sein und mit dem humanistischen Anspruch konfrontiert werden soll. Dies fällt mit der Frage zusammen, wer die Macht über die Ausgestaltung digitaler Technologie und Technik innehat, um die Anforderung eines humanistischen Ansatzes überhaupt erfüllen zu können. Dies erlaubt zudem eine Zuordnung und Abgrenzung von Spielräumen und Handlungsmöglichkeiten, die zugleich hilfreich ist, um technische Kompetenzen zu erfassen und (moralische) Verantwortlichkeiten zu verorten. Diese Arbeit möchte den digitalen Humanismus als ein normatives Gebot aufnehmen und seine normativen Ansprüche für eine Ethik, die in der Softwareentwicklung Anwendung findet, fruchtbar machen, um den Grundstein für eine ethische Entwickler-Praxis zu legen. Kurz: Eine Ethik für Softwareentwickler muss ein Ethos der Technikherstellung begründen, will man digitale Artefakte normativ wünschenswert ausgestalten.

Ein humanistisches Menschenbild sieht seinen Normativitätsanspruch einer humanen Lebensgestaltung in der Autonomie begründet, die individuell, intersubjektiv und kollektiv begründbar sein muss. Die humanistische Anthropologie legt somit die Zwecksetzungsrationalität, also die Fähigkeit des Menschen, sich seine eigenen Handlungszwecke selbst zu setzen, als die dem Menschen wesentlich vernünftige Tätigkeit fest (Janich, 2001; Nida-Rümelin, 2001, 2005a, 2020). Aufgrund des teleologischen Charakters kann sich der Mensch selbstbestimmt, gerichtet und absichtlich auf die Welt beziehen. Dieser Absichtscharakter, der Handlungen oder Überzeugungen begründet, zeigt sich als Intentionalität, die Nida-Rümelin (2011, §17) zugleich als Entscheidung begreift. Folglich kann dieses Gerichtet-Sein als Beendigung einer Deliberation verstanden und diese Entscheidung auch begründet werden. Die Deliberation und die aus ihr folgende Deziision sind als Tätigkeiten der praktischen Vernunft von der Tätigkeit der rein rationalen Mittelwahl abzugrenzen. Letzteres fällt in den Bereich der praktisch-instrumentellen Vernunft und umfasst die optimale Mittelwahl in Abhängigkeit vorgefundener Zwecke oder eben Intentionen. Die beste Mittelwahl wird beispielsweise im Rahmen der ökonomischen Rationalität durch die Nutzenwertoptimierung – wenn unter Unsicherheit gehandelt wird genauer durch die Optimierung des Nutzenerwartungswerts – ermittelt und bewertet (Nida-Rümelin & Schmidt, 2000). Das Nutzenkalkül sagt jedoch nichts über die Setzung der Zwecke, das gesollte Woraufhin der Handlung, aus. Ja, nicht einmal eine kritische Reflexion und eine damit verbundene Thematisierung der Wünschbarkeit jener Zwecke kann die Vernunft vornehmen, denn zwischen Präferenz und Entscheidung wird nicht unterschieden (Nida-Rümelin, 1993b; Sen, 1977). Die Vernünftigkeit der instrumentellen Mittelwahl bestimmt sich daher durch eine konsistente oder gar kohärente Binnenrelationalität, d.h. durch die zu erfüllenden relationalen Anforderungen der einzelnen Strukturmomente, die bei einer Wahl Berücksichtigung finden müssen, seien diese Intentionen, Wünsche, Handlungen etc. Diese Form der vernünftigen Kombinatorik wird daher in der Analytischen Philosophie als *structural (ir)rationality* (Scanlon, 2007) bezeichnet und ist um eine stimmige Binnenstruktur bemüht.¹¹

Als Mittel sind im Bereich des Praktischen in erster Linie Handlungen, die einen Zweck erfüllen, zu verstehen, sodann in Abgrenzung daran „Mittel als Güter“ (Janich, 2001), die in das Handlungsgeschehen eingebunden werden können, um diese Zwecke besser erreichen zu können, d.h. welche die Handlung unterstützen können. Allerdings kann nur dann von einer Handlung gesprochen wer-

¹¹ Diese firmieren in der Literatur als Minimalbedingungen der Vernunft und umfassen die Axiome der klassischen Entscheidungstheorie (Nida-Rümelin & Schmidt, 2000). Zudem werden sie in der angelsächsischen, praktischen Philosophie durch Kohärenzkriterien erweitert und dann als strukturelle Rationalität bezeichnet. Für eine analytische Zusammenfassung siehe Kiesewetter (2017).

den, wenn ihr eine selbstbestimmte Zwecksetzung vorrausging, denn nur durch die Tätigkeit der praktischen Vernunft, sich selbst die Zwecke ihres Handelns setzen zu können, kann man vom handelnden Menschen sprechen und Handlungen folglich von Verhalten abgrenzen. Dies hat weitreichende Konsequenzen für den Begriff der Technik, denn „Technik (...) einschließlich die in der naturwissenschaftlichen Forschung genutzten technischen Mittel ist nicht wieder von Natur aus, sondern durch menschliches Handeln in der Welt. Menschen sind Handwerker, Ingenieure, Techniker und Laborforscher“ (Janich, 2003, S. 95). Somit gehört Technik immer auch in den Raum des Tätigseins. Inwiefern Technik Handlungen umfasst und wie Technik Handlungsketten (re-)strukturiert, muss demnach das Design eines technischen Artefakts bestimmen, insofern ethische Strukturelemente berücksichtigt werden sollen. Folglich wurde Technik als Handlung begreifbar gemacht, um die Normativität, die vernünftige Möglichkeit ihres Herstellungsprozesses denken zu können (Teil I, Kapitel 3 und 4). Dadurch kann der Herstellungsprozess von Technik auch als autonome Handlung aufgearbeitet und das technische Artefakt als wünschenswertes Objekt gestaltet werden. Dieser Gestaltungsprozess (*poiesis*) und das gestaltete technische Artefakt sind wiederum selbst entweder Mittel zum Zweck (*poiesis*) oder gar Selbstzweck (*praxis*) und immer schon Teil einer Lebensform.

Menschen erwerben das Wissen um Handlungszwecke und deren Realisierung im Austausch mit anderen (Joas & Mead, 1985). Janich (2001, 2003) bezeichnet diesen konstitutiven Vorgang als die Kultürllichkeit von Handlungen, um den Verlauf des Erlernens von Handlungen von Internalisierungsansätzen abzugrenzen, die den Anschein erwecken, dass es sich hierbei um arationale Prozesse handelt, denn „[e]rkenntnistheoretisch primär ist das aus einer Wir-Perspektive des Teilnehmers sich vollziehende gemeinsame Handeln, das in der Schematisierung zur Praxis und in der Tradierung zur Kultur wird“ (Hartmann & Janich, 1998, S. 17). Die habituelle Ausführung einer Handlung gründet auf einer Haltung, einer sogenannten *hexis*, die tugendethisch gesprochen jede Tugend erst zur wahrhaften Tugend werden lässt. Ohne eine rationale und so manches Mal auch tradierte Haltung, d.h. ohne Entscheidung, wie man sich gegenüber Affekten, Ereignissen oder Erlebnissen zu verhalten hat, kann man nicht von einer Gegenüberstellung zur Natürlichkeit sprechen, die jedoch für einen humanistischen Begriff der Verantwortung begründend ist. Ein vernünftiger – und dies bedeutet hier wesentlich ein gemeinschaftlicher Erwerb von Handlungsorientierungen und deren Sinn-gerechte Ausführung – kann, will man Siep folgen, zwei essentielle Mängel überwinden, welche die praktische Vernunft aufweist, ohne unmittelbar den Bereich der Vernunft verlassen zu müssen. Zum einen weist die Vernunft ein epistemisches Defizit auf, denn es ist unmöglich, immer alles richtig zu erkennen. Zum anderen bleibt ungeklärt, wie die Vernunft Handlungen überhaupt direkt zu motivieren vermag (Siep, 2013). Wenn jedoch Gewohnheiten in Form von auf Entscheidung beruhenden Haltungen, die sich auf einem geteilten gemeinschaftlichen Wissenshorizont abbilden, verstanden werden, muss man diesen die Vernünftigkeit gar nicht absprechen: Denn die Einübung einer Haltung basiert gerade

auf der Tätigkeit des Betrachtens (*theoria*), folglich auf Wissen, dass bestimmte Sachverhalte als wahr und als richtig anerkannt werden sollen. Praktisches Wissen, das sich also mit Handlungen beschäftigt, beruht sodann auf der Betrachtung der für richtig erachteten Handlungszwecke, die durch ihre Ausübung als Praxis oder in ihrer Wiederholbarkeit gar Kultur ausmachen. Somit werden Tugenden zu stabilen Dispositionen, die nicht permanent einer erneuten theoretischen Überprüfung unterzogen werden müssen, sondern gerade einer immer wieder aufkommenden Abwägung von adäquaten Handlungszwecken entgegenwirken können. Zudem verlangen Handlungen aufgrund ihrer intentionalen Struktur einen kognitiven Aspekt, der nicht auf emotive Motivationen, wie Empathie, Vertrauen oder das Gewissen reduzierbar ist, wenn Handlungsziele als autonome Strukturelemente begriffen werden sollen. Dies muss auch als der Unterschied zu Heuristiken, jenen unbeabsichtigten Verhaltensmotivationen, die gerade keine deliberativen Aspekte benötigen und somit nicht als rationale Handlung bewertbar sind, verstanden werden.

Durch die Tätigkeit der Vernunft kann sich der Mensch als ein freies Wesen verstehen (lernen). Um diese Fähigkeit angemessen ausüben zu können, muss das Individuum die Fähigkeit in und mittels seiner gemeinschaftlichen Gebundenheit erwerben und erlernen: „In einem humanistischen Weltbild ist der Mensch kein Mechanismus, sondern freier (autonomer) und verantwortlicher Akteur in der Interaktion mit anderen Menschen und einer gemeinsamen sozialen und natürlichen Welt.“ (Nida-Rümelin & Weidenfeld, 2018, S. 61) Die Autonomie des Menschen erwächst dem mündigen Setzen selbstbestimmter Ziele, die den eigenen Handlungsabfolgen als Strukturierung dienen und kann somit auch als Befreiung aus einer paternalistischen Eltern-Kind-Beziehung bezeichnet werden. Erst durch diese Zwecksetzung können Handlungen nicht nur verstanden, also ihre kausalen Zusammenhänge erklärt, sondern sie können auch begriffen und beurteilt werden, weil allen Teilnehmern an eben jenen gemeinschaftlichen Strukturen ersichtlich wird, worauf die Handlung abzielt, welcher Handlungsgrund verfolgt wurde oder welcher vielleicht besser hätte verfolgt werden sollen (Teil I, Kapitel 2.1). Durch die geteilte (Er-)Kenntnis um Handlungszwecke werden Handlungen der praktischen Vernunft zugänglich und können so nicht bloß auf die Anforderungen instrumenteller Rationalität reduziert werden – was im Grunde auch die Kritik des konstruktivistischen Neo-Kantianismus einer Christine Korsgaard an einem neo-humeschen Ansatz eines Michael Smith ist (Korsgaard, 1997; Smith, 1987). Dadurch, dass Zwecke nicht einfach gegeben sind, sondern durch Individuen selbst gesetzt werden können, bedürfen sie einer Begründung und eröffnen somit den semantischen Raum der Gründe. Erst durch die Vernunft, so Siep (2013, S. 354), wird der Tugendhafte auch der Prinzipientreue, denn „mittels der Vernunft besitzen wir die Kraft, Regeln für richtiges Verhalten zu erkennen und sie gegen Wünsche und Interessen durchzusetzen, die zu solchen Regeln nicht passen“.

Kommunikation, der Austausch von Gründen, steht somit im Zentrum der Autonomie, denn eine freie Zwecksetzung muss von Gründen geleitet sein, um sie als vernünftig auszeichnen zu können. Mittels

Deliberation, einem Abwägen von Gründen und der Entscheidung als Setzung des Handlungsziels, ist die Person „imstande, die einzelnen Handlungen in den größeren Kontext einer Handlungsweise, der Struktur ihres Handelns und Lebens einzubetten. Sie optimiert nicht lediglich punktuell, sondern wägt Gründe ab und kann Gründe anführen für das, was sie tut. Gründe anführen heißt aber auch, sich festlegen, in vergleichbaren Situationen ebenso zu handeln. Sich Gründe zu Eigen machen heißt, Strukturen in sein Leben legen.“ (Nida-Rümelin, 2011, § 6) Dieses grundlegende, konstitutive Moment praktischer Vernunft bezeichnet Nida-Rümelin (2020, VI §7) als die Affektion von guten Gründen und Freiheit somit als die Fähigkeit des Menschen, Gründe abzuwägen sowie auch dieser Abwägung gemäß zu handeln. Die Fähigkeit der praktischen Vernunft, sich auf gute Gründe einlassen zu können, d.h. auf sie angemessen zu handeln, wird in der analytischen Philosophie auch als *reason responsiveness* bezeichnet (Kiesewetter, 2017; Lord, 2015, 2018; Wedgwood, 2017). Die Erkenntnis, welche Gründe für eine Handlung sprechen und seine Handlungen nach diesen Gründen auszurichten, bedeutet, sich auch gegen sein eigenes Wollen, seine eigenen Interessen richten zu können – bedeutet sich von einem Wollen distanzieren zu können und sich ein Sollen als Handlungsgrund zu setzen (Darwall, 1983). Folglich befreit sich die praktische Vernunft durch die Einsicht (*phronesis*) in praktische Gründe, weshalb eine Handlung dann als erfolgreich gilt, wenn sie sich als vernünftig zeigt, also sich durch die besseren Gründe bestimmen lässt. Die Achtung dieser Menschlichkeit, welche die Bedingungen der Möglichkeit zur Autonomie gewährleisten muss, stellt den Kern einer Moral. Aus diesem humanistischen Ansatz folgt nun: Wenn ein Verhalten durch Gründe geprägt ist, wenn also von einer Handlung gesprochen wird, weil diese einen rationalen Grund in sich birgt, dann trägt der Akteur auch die (moralische) Verantwortung über diese Handlung und deren Handlungsfolgen (Nida-Rümelin, 2011, S. 9).

Entscheidungen sind das Ergebnis einer Deliberation, weshalb auch von einem Pluralismus der Gründe gesprochen werden muss, denn es kann nicht nur einen einzigen Handlungsgrund geben, aus dem sich alles Weitere begründen ließe. Eine Handlung unterscheidet sich folglich durch ihre Bestimmtheit durch Gründe, die wiederum als Deliberation und Entscheidung der Handlung vorausgehen. Die Fähigkeit zur Autonomie macht es möglich, Handlungen von Verhalten und somit Kultur von Natur abzugrenzen. Geht man einen Schritt weiter, so führt eine humanistische Anthropologie sogleich in die Zuschreibung von Verantwortung, denn „[w]ir sind genau für das verantwortlich, für welches wir Gründe haben. Das können Entscheidungen sein, Handlungen, das ist dann praktische Verantwortung, oder Überzeugungen und Urteile, das ist theoretische Verantwortung“ (Nida-Rümelin, 2005b, S.3).

Die humanistische Anthropologie verbindet somit Vernunft und Freiheit mit einem Verantwortungsbegriff, indem eine selbstbestimmte Gründeabwägung und eine danach ausgerichtete Handlung die wesentlichen Fähigkeiten menschlicher Akteure ausmachen. Dieser freie Akt erlaubt es, Handlungen

Personen zuzuschreiben und somit verantwortlich zu machen, eben, weil diesen eine autonome, da rationale, Entscheidung zugrunde liegt. Die handelnde Person hätte auch anders handeln können.

Ein Technikbegriff, der diesem humanistischen Anspruch Folge leisten möchte, muss folglich

- a) den Herstellungsprozess auf mögliche Handlungsfreiheiten analysieren.
- b) das technische Artefakt auch als Teil lebensweltlicher Handlungsketten verstehen, um nicht ausschließlich einer klassischen trade-off-Abwägung zu verfallen, sondern digitale Technik und Technologie auch als Mittel und Zweck intentionaler sowie geteilter Handlungsstrukturen zu begreifen.
- c) diejenigen Kompetenzen und Vermögen erörtern, die für einen humanistischen Umgang mit digitaler Technik und Technologie vonnöten sind.

Wenn ein humanistisches Ethos in der Technikgenese Einzug gefunden hat, kann sich eine Ethik der Technik in einer lebensweltlichen Praxis bewähren.

3 Ethik der Digitalisierung, der digitalen Transformation und des digitalen Artefakts

Nach den grundlegenden Ausführungen der philosophischen Paradigmen, die eine Entwicklung humanistisch wertvoller digitaler Technologie und Technik ermöglichen, werden im zweiten Teil der Arbeit die für den Bereich der Informationstechnik und -technologie spezifischen normativen Sachverhalte herausgestellt. Hier gilt es, die für diesen Arbeitsbereich relevanten Orientierungspunkte und Reflexionsbegriffe herauszuarbeiten, die eine ethische Deliberation leiten und stützen. Zudem kann so aufgezeigt werden, auf welche normativ relevanten Aspekte besonders zu achten ist. Abschließend werden die Begriffe *Digitalisierung*, *digitale Transformation* und *digitale Artefakte* anhand der eingeführten philosophischen Paradigmen differenziert, um die unterschiedlichen normativen Aspekte betonen zu können, die sich aus der differenzierten Betrachtung ergeben. Um die Komplexität des zu analysierenden und bewertenden Sachverhalts aufzuzeigen, wird mit Hilfe der Gesichtserkennungssysteme in die Thematik eingeführt.

3.1 Gesichtserkennungssysteme in der polizeilichen Fahndungsarbeit

Die demokratische Kraft sozialer Medien wird auch durch ihre Funktion als effektive Mobilisierungsmacht deutlich. Mit Hilfe immer effizienterer, portabler Hardware ist es Bürgern möglich, sich in Echtzeit zu koordinieren, sich gemeinsam zu mobilisieren, um sich aktiv an lebensweltlichen Ereignissen zu beteiligen, wie es beispielsweise der Arabische Frühling bereits 2010 deutlich gemacht hat. Eine Verzahnung digitaler und physischer Realität ist dabei unabdingbar, um normative Momente gesellschaftlicher, hier insbesondere demokratischer Handlungszusammenhänge zu transformieren. Dasselbe gilt für eine Verschmelzung demonstrativer und deliberativer Partizipationskultur, denn Demokratie legitimiert sich zunehmend auch über eine deliberative Teilnahmemöglichkeit am öffentlichen Diskurs. Dies zeigt sich auch an der *black lives matter*-Bewegung (kurz: BLM), die seit 2013 auf die systematische oder beabsichtigte Diskriminierung der *people of color*-Bevölkerung in den USA aufmerksam macht. So hob die BLM-Bewegung im Anschluss an den am 25.5.2020 in Minneapolis durch Polizeigewalt verstorbenen US-Amerikaner George Floyd das Thema des systematischen Rassismus auf die globale politische Agenda und eröffnete dadurch zugleich eine erneute öffentliche Diskussion über staatlich eingesetzte automatisierte Gesichtserkennungssysteme (Denkler, 2021). Nicht nur das Hashtag #BlackLivesMatter ging viral, sondern gerade die demonstrative Partizipation setzte abermals vielfältige öffentliche Deliberationen in Gang, wie beispielsweise die *American Civil Liberty*

Union's Community Control Over Police Surveillance (kurz: ACLU CCOPS, An Act To Promote Transparency and Protect Civil Rights and Civil Liberties With Respect to Surveillance Technology). Die CCOPS, die bereits 2016 im Nachgang zu Donald Trumps Wahlerfolg ein Regulierungsgesuch formulierte, wurde daraufhin 2020 nochmals verschärft, denn man zielte vor allem auf eine autonome Entscheidungsfähigkeit der Bürger ab (Community Control Over Police Surveillance (CCOPS) Model Bill, 2021).

Die ACLU fordert unter anderem einen transparenten und öffentlichen Diskurs, damit Bürger über den Einsatz, Erwerb und Verkauf sowie die Datenspeicherung staatlicher, automatisierter Überwachungssysteme informiert werden und zugleich aufgrund der Teilnahme am kollektiven Deliberationsprozess das weitere Vorgehen mitbestimmen können. Introna betont bereits 2005, dass im Gegensatz zum Verkehrsinfrastrukturausbau, wo Gutachten und Bürgerbegehren eingeholt werden müssen, Bürger im Falle automatisierter (Gesichtserkennungs-)Systeme gar nicht oder nur unzureichend über direkte Kosten sowie Opportunitätskosten aufgeklärt würden. Dies ist gerade wegen der weitreichenden Folgen bedenkenswert, die ein Einsatz automatisierter Systeme auf gesellschaftliche Handlungszusammenhänge und Verhaltensweisen ausübt. Im Gegensatz zu Introna (2005) muss betont werden, dass technische Systeme gerade wegen ihrer Normativität politisch relevant sind und nicht erst aufgrund ihrer Möglichkeit ethisch relevant werden, die Interessen bestimmter Personen zu repräsentieren. Dafür muss das technische System als normativ steuernde Praxis begriffen werden, denn es strukturiert lebensweltliche Beziehungen (Nida-Rümelin & Weidenfeld, 2018). Technische Systeme beeinflussen und verändern Gewohnheiten und geben normative Orientierung, weil sie vom Nutzer bestimmte, dem Softwaresystem angemessene Verhaltensweisen fordern. Folglich können technische Objekte wegen ihres Designs einige Gruppen als Nutzer ausschließen, wenn diese die zu erwartende Handlung nicht ausführen können oder wollen. So können Personen ohne Zugang zum Internet gar nicht teilnehmen oder Personen werden aus der Dating-Kultur ausgeschlossen, weil sie sich nicht über Dating-Plattformen kennenlernen möchten. Diese Form der Diskriminierung wird als *exclusive bias* diskutiert und ist zu unterscheiden von einer Diskriminierung, die als *information bias* bezeichnet wird (Brey, 1998). Letzterer Bias basiert auf der Annahme, dass digital-technische Objekte aufgrund ihrer Möglichkeit, Werte und Haltungen zu konservieren und zu transportieren, Personen benachteiligen können, weil diese systematisch nicht oder nur unzureichend adressiert werden. Z.B. durften Bargeldabhebungen im Rahmen der *Aid for Families with Dependent Children* nicht in der Nähe von Liquor Stores erfolgen, was in ländlichen Gegenden zum Ausschluss vieler betroffener Familien führte (Eubanks, 2012). So drängen technische Objekte Personen Handlungsabläufe auf, indem bestimmtes technik-konformes Verhalten zu adaptieren ist, damit digitale Technik überhaupt funktionieren kann. Dabei muss man beim Nutzen von Applikationen bestimmte ‚technische‘ Tätigkeiten ausführen, wie ‚Wischen‘ oder ‚Klicken‘, aber auch interaktive Erwartungshaltungen bedienen,

wie konforme, schnelle Antworten bei Instantmessengern oder gar komplexeres Mating-Verhalten bei Tinder. Daher muss nochmals betont werden, dass technische Objekte wesentlicher Teil normativer Handlungszusammenhänge sind, die man durchaus als Lebenspraxis (Nida-Rümelin, 2020a) bezeichnen kann, denn digitale Technik fördert bestimmte Verhaltensweisen und exkludiert andere. Introna (2005) weist zudem auf einen sehr markanten Aspekt von Softwaresystemen hin, der vor allem für ethische Belange von großer Wichtigkeit ist. Im Unterschied zu anderen technischen Objekten operieren Softwaresysteme im Verborgenen, dem sogenannten *cyberspace*, d.h. sie können unmerklich zum Einsatz kommen, da ein direkter Kontakt zwischen Nutzer und Software meist nicht notwendig ist oder aber das Softwaresystem nicht unmittelbar im Einsatz erkennbar oder in seiner Einzelheit gar nicht nachvollziehbar ist. Bspw. sieht man ein Auto und nimmt den darin verbauten Computer nicht wahr (*ubiquitous computing*) oder Künstliche Intelligenz ist in der Anwendung als Neuronale Netze häufig in ihrem Verfahren nicht fassbar (*non-explainability*).

Diese Aspekte werden in der Literatur unter den Begriffen der *opaqueness* und *transparency* diskutiert und thematisieren die Vernunftlosigkeit im Sinne einer Blindheit, sodass Phänomene gar nicht wahrgenommen werden. Im Gegensatz hierzu wirkt zwar eine Tür auch auf das Verhalten der Besucher ein, da diese geöffnet werden muss, die Besucher ggf. einen Schritt zurücktreten müssen etc. Bereits der Einbau einer Tür führt so zu einer Verhaltensanpassung. Allerdings ist eine Tür unmittelbar sichtbar und meist wird ihr Einsatz durch ihre Funktionalität zugleich erklärbar: Die Tür schließt einen Raum ab, um evtl. ungestörtes Lernen zu ermöglichen. Die Opazität und die Transparenz erfordern, so Introna (2005), eine *disclosive ethics*, die sich um die Belange der Diskriminierten und Ausgeschlossenen bemüht, die aufgrund der Unverständlichkeit oder Unsichtbarkeit der Systeme immer mehr ins Hintertreffen geraten. Die daraus entstehende Umgestaltung der öffentlichen Lebenswelt muss sich zudem durch einen demokratischen Prozess legitimieren lassen, vor allem dann, wenn Unklarheiten über die Verteilung von Benachteiligungen bestehen.

Die in den USA bspw. diskutierte, durch die Polizeiarbeit auftretende Diskriminierung der *people of color*, so die Befürchtung der BLM-Bewegung, würde durch den Einsatz automatisierter Gesichtserkennungssysteme in der Polizeiarbeit durch die technische Abbildung des Status quo verstärkt. Chander (2016, S. 1028) bezeichnet die beabsichtigte, manipulative Reproduktion von Vorurteilen durch automatisierte Systeme, die Interessen bestimmter Gruppen bewahren oder gar befördern, als *manipulative algorithm*, obgleich der Sachverhalt besser als *manipulative system* beschrieben werden sollte, da Diskriminierungspotenziale sich nicht nur auf die Ebene der Algorithmen reduzieren lassen. Zum anderen ist hiervon die *viral discrimination* zu unterscheiden, die, unbeabsichtigt, ungewollte, aber bestehende gesellschaftliche Verzerrungen und Ungerechtigkeiten aufnimmt, gerade weil Algorithmen an lebensweltlichen Daten trainiert werden (ebd., S. 1036). Aber auch hier gilt, dass sich die

virale Diskriminierung nicht nur auf Algorithmen reduzieren, sondern sich auch in der Systemanwendung oder in der System-Interpretation lokalisieren lässt. Eine *viral discrimination* muss als weitaus größeres Risiko eingestuft werden, da eine beabsichtigte diskriminierende Manipulation, gerade auch aufgrund der Tätigkeit des Programmierens, die meist bewusst ausgeführt wird, und der Einbindung des einzelnen Programmierers in Teams und Arbeitsprozesse, eher unwahrscheinlich ist. *Viral discrimination* hingegen ist technisch subtil und kompliziert, denn um der Wiederholung lebensweltlich stattfindender Ungerechtigkeiten durch Technologien entgegenzuwirken, bedarf es zusätzlicher Wissens- und Praxisformen, die derzeit in der Informatikbranche nur unzureichend oder oft nur implizit vorhanden sind. Lässt man diese Art viraler Diskriminierung unbedacht, läuft man bspw. Gefahr, dass finanziell schlechter ausgestattete Polizeibehörden einfach von einer kostspieligen, rassistisch motivierten menschlichen Polizeiarbeit zu einer weniger kostspieligen, aber nach wie vor rassistisch motivierten High-Tech-Polizeiarbeit übergehen, ohne dabei menschliche Vorurteile und deren Reproduktion durch Technologie zu hinterfragen, geschweige denn auf diese politisch oder gestalterisch einzuwirken. Das Diskriminierungspotential, das durch den Einsatz von Gesichtserkennungssystemen entsteht, kann daher nicht nur technologisch, sondern muss auch politisch diskutiert werden. Die Thematik der strukturellen Diskriminierung durch den Einsatz von Gesichtserkennungstechnik ist aufgrund der lebensweltlichen Auswirkungen besonders politisch brisant, weshalb eine normative Evaluation dringend notwendig ist, aber aufgrund ihrer Multidimensionalität nahezu unauflösbar erscheint.

Unter struktureller Diskriminierung bestimmter gesellschaftlicher Gruppierungen versteht die ACLU (2020, Section 12. Definitions): (1) disparate treatment of any individual(s) because of any real or perceived traits, characteristics, or status as to which discrimination is prohibited under the Constitution or any law of the United States [or because of their association with such individual(s), or (2) disparate impact on any such individual(s) having traits, characteristics, or status as described in subsection (1).“ Folglich offenbart sich struktureller Rassismus in einer systematischen Benachteiligung bspw. schwarzer Personen in verschiedenen gesellschaftlichen Bereichen, u.a. auf dem Arbeitsmarkt, wie durch eine höhere Zunahme der Arbeitslosenquote im Vergleich zu anderen Bevölkerungsgruppen seit der Corona-Krise oder der relativ geringen Besetzung gut bezahlter Managementpositionen ersichtlich wird. Selbst als 1964 die sogenannten Jim-Crow-Gesetze offiziell abgeschafft wurden, blieben Vorurteile gegenüber der schwarzen Bevölkerung bestehen. Sie wirken auch heutzutage subtil auf Entscheidungen ein, wie sich z.B. anhand der niedrigeren Rückmeldequote auf schwarze im Vergleich zu ‚ge-weiß-ten‘ Bewerbungsschreibern zeigt (Kang et al., 2016), also solche Bewerbungsschreibern, die an weiße Demografiemerkmale angeglichen wurden. Obgleich die schwarze Bevölkerung 2018 nur 12 % der amerikanischen Bevölkerung ausmachte, war sie in den Gefängnissen mit 33 % stark überrepräsentiert. Im Vergleich hierzu macht die weiße Bevölkerung 60

% der Gesamtbevölkerung aus, repräsentiert aber nur 30 % der Gefängnisinsassen (Gramlich, 2019). Allerdings muss man, um einen fairen Blick für die Ungleichbehandlung und somit Benachteiligung schwarzer Personen durch die Polizeiarbeit zu erhalten, diejenigen Straftaten miteinander vergleichen, die von Weißen und Schwarzen gleichermaßen häufig ausgeübt werden. Weiße und schwarze Personen weisen einen ähnlichen Drogenkonsum auf und trotzdem werden schwarze Personen zu längeren Haftstrafen verurteilt oder ihnen wird häufiger die Kaution verweigert. Dies zeigt sich deutlich an den Verurteilungen wegen Marijuhanakonsums. Obwohl es keinen signifikanten Unterschied im Konsumverhalten gibt, ist es nach Berechnungen der ACLU (2020) 3,7-mal wahrscheinlicher, als schwarze Person wegen Marijuhanabesitzes inhaftiert zu werden – selbst wenn man die Legalisierung, die in einigen Bundesstaaten 2018 bereits erfolgte, berücksichtigt. Auch fallen die empfohlenen Strafmaßnahmen, welche die *Sentences Guidelines 2016* ausweisen, im Hinblick auf *cracked cocaine* verglichen mit *powdered cocaine* unterschiedlich hoch aus (Saris et al., 2016). Die schwarze Bevölkerung, die mehr Crack als Kokain konsumiert, muss daher bei Verhaftung mit längeren Haftstrafen rechnen als (meist) weiße Kokainkonsumenten (Chander, 2016). Zudem werden Personen mit schwarzer Hautfarbe häufiger in Fahrzeugkontrollen angehalten, durchsucht und weisen eine doppelt so hohe Wahrscheinlichkeit auf, aufgrund von Verkehrsdelikten inhaftiert zu werden. Nicht zuletzt ist es wesentlich, dass Polizeibeamte schwarzen Personen anders begegnen: Zentral ist der eher gewaltbereite Umgang mit schwarzen Personen, der sich in der Verwendung von Tränengas, Pfefferspray oder gar polizeilicher Waffengewalt zeigt (Edwards, Lee & Esposito, 2019). Ferner ist es für eine schwarze Person zweimal wahrscheinlicher, unbewaffnet von einem Polizeibeamten erschossen zu werden als für eine weiße. Selbst wenn Studien nicht zu eindeutigen Ergebnissen dahingehend kommen, inwiefern struktureller Rassismus beabsichtigt oder unbeabsichtigt ist, stark oder weniger stark ausgeprägt ist oder das Verhalten von Polizeibeamten aufgrund der höheren Kriminalitätsrate der schwarzen Bevölkerung erklärbar sein mag, reicht es aus, auf Grundlage dieser Szenarien über die systematische Diskriminierung durch Polizeibeamte nachzudenken. Sobald unschuldige Personen nur deshalb eine andere Behandlung durch Polizeibeamte erfahren, weil sie aufgrund ihrer Herkunft oder ihres Aussehens Verdachtsmomente aufweisen, muss offen über Diskriminierung gesprochen werden.

Kenneth Arrow sprach 1973 in *The theory of discrimination, discrimination in labor markets* in diesem Zusammenhang von einer *statistical discrimination*, einer Diskriminierung, die aufgrund fehlender vollständiger Informationen zustande kommt. Personen urteilen deshalb mit Hilfe von Proxy-Variablen, also indirekten Beobachtungen wie Hautfarbe oder Geschlecht, um das Verhalten des Gegenübers besser einordnen zu können, da ihnen Wissen über individuelle Eigenschaften oder Errungenschaften fehlt (Arrow, 1973). Durch diese *statistical discrimination* verfestigen sich Vorurteilsstrukturen, indem eine Besser- oder Schlechterstellung bestimmter Gruppen mit bestimmten Merkmalen als

Handlungsorientierung strukturell eingebettet wird. Für die Automatisierung bestimmter gesellschaftlicher Handlungszusammenhänge bedeutet dies nachdrücklich, dass „(i)n a society where discrimination affects opportunities in innumerable ways, we must worry about the migration of discrimination to decision making by algorithm“ (Chander, 2017, S. 1023). Die Sorge, dass Vorurteile durch den Einsatz digitaler Technologien größtenteils unbeabsichtigt bekräftigt oder verstärkt werden, muss daher ernst genommen werden. Allerdings muss daraus keine Übervorsicht oder gar eine Verteufelung einer digital-technologisierten Welt resultieren, da zum einen die digitale Welt mit der analogen Welt, also der Welt, die rein auf menschlichem Verhalten und Urteilen beruht, verglichen werden muss. Zum anderen muss betont werden, dass man auf die Art und Weise der digitalen Transformation Einfluss nehmen kann. Dies kann regulatorisch geschehen – wie durch das *Weissbuch Zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen* vom 19.06.2020 (European Union, 2020) – oder durch eine normativ wünschenswerte Gestaltung technischer Artefakte, wie es Ansätze des *ethics by design* ermöglichen, oder durch einen Versuch, die Entwicklung mittels *ethical canvas*¹² zu strukturieren.

Dies gilt zweifelsohne auch für die Debatte um automatisierte Gesichtserkennung, die im Falle der Anwendung als automatisierte Überwachungstechnologie durch die aufgrund einer technischen Missklassifikation im Januar 2020 erfolgte Festnahme von Robert Julian-Borchak Williams in Michigan zusätzlich befeuert wurde (Hill, 2020). Die daraus entstehende Besorgnis der schwarzen US-Bevölkerung, nun durch diese neue Technologie noch mehr ins Visier der Polizei zu geraten und daher mehr Gewalt zu erfahren, verdeutlicht, dass die Thematik nicht auf eine Diskussion über dichotome Trade-offs, hier bspw. Sicherheit vs. Privatsphäre, reduziert werden kann – vor allem nicht, wenn Personen durch die Einführung einer neuen Praxis eine schlechtere Behandlung zukommt oder sie dies befürchten müssen. Es müssen folglich die gesamte Praxis und die gesamten Handlungszusammenhänge betrachtet werden, die durch einen Einsatz des Systems betroffen sind. So sind Diskriminierungspotenziale durch die Einführung neuer Umgangsformen kein Phänomen, das sich erst im digitalen Zeitalter entwickelt hat. Diskriminierung beruht auf menschlichem (Fehl-)Verhalten und grün-

¹² *Ethics by design* ist ein Ansatz praktischer Entwicklungsmethodiken, die – mehr oder weniger theoretisch – die technische Machbarkeit normativer Kriterien zu etablieren versuchen. Hierunter fallen das *value sensitive design*, das *responsible design* oder das *participatory design*. *Ethical canvas* hingegen versuchen den Entscheidungsfindungsprozess offenzulegen. Dabei wird mit Hilfe der *ethical canvas* weniger wissenschaftlich vorgegangen als ethische Schemata, wie z.B. David Wrights (2011) *A framework for the ethical impact assessment of information technology*. *Ethical canvas* werden bspw. vom *Center for Humane Technology* (<https://www.humanetech.com/>) oder vom *Open Data Institute* (<https://theodi.org/>) konzipiert. Diese entspringen unternehmerischen Ideen und sind daher weniger theoretisch fundiert.

det auf menschlichem Urteilsvermögen. Ein Missbrauch liegt jedoch stark im Einsatz solcher Systeme, der deshalb wohlbegründet auszufallen hat. Man denke hier an einen systematischen Einsatz, um bestimmte Personen zu überwachen, zu kontrollieren und bestimmte Machtverhältnisse aufrechtzuerhalten.

Das Problem vorurteilsbasierter Urteile oder vorurteilsbasierten Verhaltens ist hierbei in der Kriminalistik, insbesondere in der Fahndungsarbeit, durchaus bekannt und entsteht auch nicht erst durch die Einführung automatisierter Softwaresysteme. Gerade in der Fahndungsarbeit geht es um die klassische Gesichtserkennung, d.h. es geht vornehmlich um die Identifikation von Personen oder um die Authentifizierung einer Person. Identifikation umfasst ein *one-to-many*-Szenario, in dem eine Person in einer Menge anderer Personen zu identifizieren ist, während es bei der Authentifizierung um die Übereinstimmung derselben Person geht, weshalb dies auch als *one-to-one*-Szenario bezeichnet wird. Diese beiden Gesichtserkennungsmodi, die gerne auch als *spotting* und *matching* charakterisiert werden, werden auch von Polizeibeamten ausgeführt. Dabei kommt es sowohl bei der Identifizierung als auch bei der Authentifizierung, bspw. bei Passkontrollen, nicht selten zu Fehleinschätzungen, die auf menschlichen, unzufriedenstellenden kognitiven Leistungen oder gar auf beabsichtigter Diskriminierung beruhen. So identifizierten Zeugen in 70 % von 365 DNA-Entlastungsfällen irrtümlich unschuldige Angeklagte (Robertson, Megreya & Davis, 2020, S. 22). In den USA und im UK identifizierten Zeugen zu 25 % Personen fälschlicherweise als schuldig, obwohl sie darauf hingewiesen wurden, dass der Täter möglicherweise nicht anwesend ist (ebd.). Aufgrund dieser hohen Fehlerquote kann man die Aufgabe, unbekannte Gesichter zu- und einzuordnen durchaus als schwierig bezeichnen. Dies lässt sich auch nicht nach vielen Übungseinheiten signifikant verbessern (Bobak, Hancock & Bate, 2016).¹³ Derzeit arbeitet ein Team an der University of Greenwich einen Gesichtserkennungstest aus, um Neueinstellungen auf diese Fähigkeiten hin zu überprüfen, damit vermehrt *super recognizer* eingesetzt werden können (Davis et al., 2020). Auch in München arbeiten seit 2018 35 Personen als Gesichtserkennungler und konnten in 200 Fällen behilflich sein (Süddeutsche Zeitung, 2019). Schätzungsweise zeigen nur 1-2 % der Bevölkerung *super recognizer*-Fähigkeiten. Der Aufwand, diese Personen erstens ausfindig zu machen, und zweitens die Arbeitszeit, um diese Personen zu schulen und einzuarbeiten, um die Fehlerquote in der Polizeiarbeit zu verringern, d.h. die Benachteiligung bestimmter Personen aufgrund staatlich exekutiver Handlungen zu reduzieren, deuten stark auf den Trend der Digitalisierung hinsichtlich der Identifikation und Authentifikation hin. Die Frage wird daher sein, ob das Credo der technologischen Aufklärung strukturelle Diskriminierung, die sich aufgrund

¹³ 2011 setzte die Londoner Polizei *super recognizer* in den Londoner Unruhen ein, um verdächtige Personen zu identifizieren. 70 % der Personen, die durch *super recognizer* identifiziert und im Anschluss überprüft wurden, konnten legitimerweise verhaftet werden (Robertson, Megreya & Davis, 2020, S. 22).

von subjektivem, d.h. willkürlichem, vorurteilsbefangenen Verhalten von Polizisten perpetuiert, durch objektive Systeme verbessern kann.

Es bedarf somit einer Betrachtung dahingehend, welche Informationstechnologie und welche Informationstechnik zum Einsatz kommt sowie einer Beurteilung, weshalb dieses informationstechnische Artefakt in einer bestimmten Lebenspraxis implementiert werden soll. Somit gilt es, die Technologie getrennt vom jeweiligen digitalen Artefakt einer normativen Evaluation zu unterziehen. Dies gilt auch für eine separate Betrachtung der Lebenspraxis. Allerdings muss wiederum auch die gegenseitige Bedingtheit Berücksichtigung finden können.

3.1.1 Legitimation von Gesichtserkennungssystemen

Der öffentliche Druck, der sich aus den Protestbewegungen entwickelte, führte im Juni 2020 dazu, dass IBM seinen Rücktritt aus der Entwicklung von Gesichtserkennungssystemen bekannt gab (Handelszeitung, 2020). Eine Kaskade an Kundgebungen innerhalb der IT-Branche folgte, unter anderem legte sich Amazon ein freiwilliges, einjähriges Moratorium in der Entwicklung seiner Gesichtserkennungssoftware auf. Microsoft hingegen verkündete, dass es seine Software nicht an Strafverfolgungsbehörden verkaufen würde, solange die Gesetzgebung diese Technologie nicht besser regulieren könne. Facebook betonte, dass es seine Technologie ausschließlich im eigenen Hause nutze, um schließlich im November 2021 Gesichtserkennung auf der Plattform *Facebook* komplett einzustellen (Metz, 2021). Allein Clearview AI enthielt sich diesen selbst auferlegten Maßnahmen. Zudem gibt es bereits in Idaho, California, Washington, Illinois und in einigen Städten wie San Francisco, Oakland, New York und Portland Regulierungsvorschriften bzw. ein Verbot dahingehend, Gesichtserkennungssysteme als Überwachungsinstrument einzusetzen (Raji et al., 2020).

Die Legitimation von Gesichtserkennungssystemen bedarf gerade aufgrund ihrer lebensweltlichen Auswirkungen, die sich in ihrer politischen Sprengkraft zeigen, einer normativen Überprüfung. So bedarf es zunächst der Diskussion darüber, welche informationstechnologische Grundausstattung angewendet wird. Dies inkludiert sogleich eine Erörterung über mögliche digitale Anwendungen und Verfahren in Form bedienbarer Soft- und Hardware, die lebensweltliche Abläufe mehr oder weniger beeinflussen. Hierzu bedarf es gleichermaßen einer Kenntnis sowohl der Softwareentwicklungs- und Managementprozesse als auch der Lokalisierung dieser lebensweltlichen Beziehungsstrukturen der Nutzer und aller (auch indirekt) Betroffener. Zudem verändert Informationstechnik ihre normative Ausrichtung durch eine andere Nutzung in weiteren sozio-technischen oder cyber-physikalischen Systemen. Daher kann diese Technologie nicht unabhängig ihrer Nutzung und ihres gesellschaftlichen Einsatzes diskutiert werden. Diese enorme Vielschichtigkeit bedarf einer analytischen Methode, die einer Einzelfall-Beurteilung gerecht werden kann (Brey, 2000a).

3.1.2 Technische Zuverlässigkeit

Damit ein fundiertes Urteil über einen vernünftigen Einsatz der Technologie gefällt werden kann, muss die technische Zuverlässigkeit der Gesichtserkennungssysteme im Zentrum der Betrachtung stehen. Folglich muss zuvorderst Klarheit darüber bestehen, welches die technisch notwendigen Voraussetzungen sind, die es zu erfüllen gilt, wenn über Gesichtserkennung gesprochen wird. Hierzu muss die Abfolge derjenigen Prozesse betrachtet werden, die digital ersetzt werden. Diese Herangehensweise ermöglicht es, ethische Risiken zu lokalisieren und gezielt auf Unzulänglichkeiten aufmerksam zu werden. Allen Gesichtserkennungssystemen liegt technisch die Simulation der Erkennung von Gesichtern zugrunde. Dies bedeutet, dass die Software ein Gesicht auf einem Foto oder in einem Video als menschliches Gesicht klassifizieren muss (*detection*), damit bspw. eine Person nicht als Gorilla kategorisiert wird, wie es *Google Photo 2015* unterlief (Kühl, 2015). Um diese Technologie nun für die Überwachungstechnologie oder für die Passkontrolle interessant zu machen, erfolgt automatisiert die Zuordnung von Personen in einem *one-to-many-* (*identification*) oder *one-to-one-*Setting (*authentication*). Hierzu wird eine Metrisierung oder Quantifizierung wesentlicher physiognomischer Merkmale benötigt, d.h. bestimmte Eigenarten des Gesichts werden in digitale Daten transformiert. Diesen Vorgang bezeichnet man auch als Biometrisierung. Informationen werden in eine Form übertragen (*Input*), die als Resultat (*Output*) wieder in Informationen rückinterpretiert werden muss. Weitere demografische Daten, wie Geschlecht, Alter oder Hautfarbe, werden für diese beiden Ausführungen nicht benötigt. Unter Gesichtserkennung versteht man in erster Linie Systeme, die Bilder von Gesichtern vergleichen, um deren Ähnlichkeit mittels eines *similarity scores* festzustellen (McLaughlin & Castro, 2020). Erst dann kann ein Abgleich mit anderen Bildern oder Videoaufnahmen erfolgen, d.h. biometrische Daten, wie Augenabstand, Stirnbreite usw. werden ver- und abgeglichen. Eine Erhöhung der Zuverlässigkeit von Gesichtserkennungssystemen kann technisch durch eine zusätzliche Geh-Art- oder Körperhaltungs-Analyse erfolgen, die die Aufnahmen der Individuen überdies anhand weiterer übereinstimmender Parameter zuordnet.

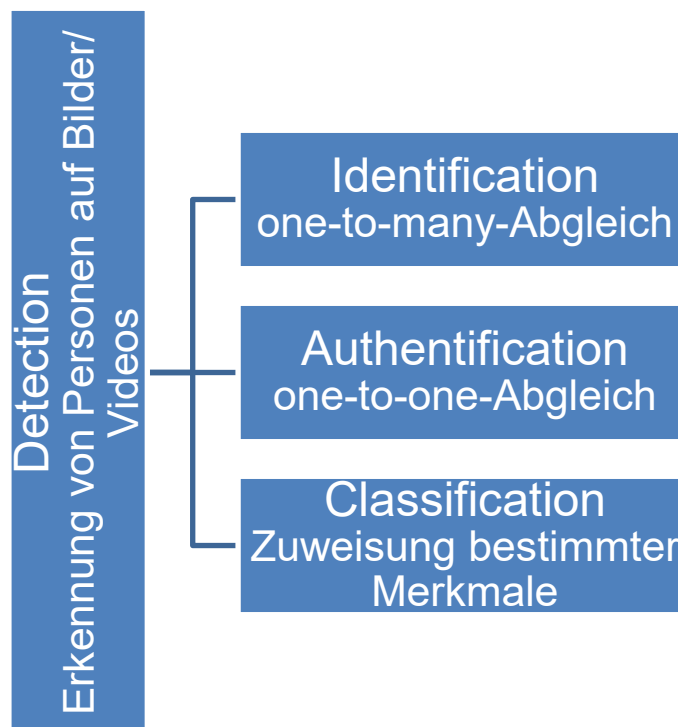


Abbildung 5 Typologie der Gesichtserkennungssysteme nach Technologie. Eigene Abbildung.

In der Rubrik Klassifizierung (*classifying*) wird die Person einer Kategorie zugeordnet. Beispielsweise geht es darum, ob die abgebildete Person alt genug ist, um wählen zu dürfen, Alkohol zu konsumieren etc. Hierzu bedarf es technisch weder einer Identifikation noch einer Authentifizierung, sondern das Softwaresystem kategorisiert nach definierten Parametern, z.B. welches „Faltigkeitsintervall“ der Hautstruktur einem bestimmten Alter entspricht, welcher Farbton welcher Hautfarbe zugeordnet werden soll oder aber welche Mundwinkel die emotionale Verfasstheit einer Person darstellen. Der Einsatz von Gesichtserkennungssystemen kann so auch zur Zuordnung emotionaler Zustände, wie bspw. zur Aufmerksamkeitskontrolle in Schulen eingesetzt werden (Zeng, Lu, Sun & Tian, 2019). Letzteres fällt bereits in den Bereich der *sentiment analysis*, einer Stimmungserkennung, die einer zusätzlichen technischen Justierung bedarf, schließlich sind das Definitionsintervall des Outputs und dessen Rückinterpretation ethisch hoch relevant. Also muss bereits die Datencodierung ethisch hinterfragt werden. Strenggenommen fällt das letztgenannte Softwaresystem technisch nicht mehr unter die Rubrik der Gesichtserkennungssysteme. Es ist jedoch leicht vorstellbar, Identifikation und Authentifizierung um eine emotionale Analyse-Ebene zu erweitern, was sowohl in Überwachungstechnologien als auch in medizinischen Diagnostiktools zu irritierenden Folgen im Falle von Missidentifikationen führen kann. Man stelle sich vor, es würden einem bestimmte Gefühle oder gar Intentionen unterstellt, die wiederum auf bestimmten Verhaltensweisen beruhen. Als jemand an einer Moschee vorbeilief, schaute er z.B. grimmig, als er ums Eck bog. Die Mundwinkel sowie Blicke ließen das Softwaresystem die Person als Terroristen einordnen, was zur Festnahme führte. Werden zusätzlich

zu einem solchen Szenario Diskriminierungen noch technologisch reproduziert, gleichgültig, ob beabsichtigt oder nicht, wird deutlich, weshalb die BLM-Bewegung sich massiv gegen einen Einsatz solcher Technologien ausspricht. In den Medien wurde die Gesichtserkennungstechnologie immer wieder als ein rassistisches Instrument mit Vorurteilen präsentiert (Kretschmar, 2020) oder das Ende der Anonymität ausgerufen (Meyer, 2021).

Die ACLU führte 2018 einen Test mit der *Rekognition*-Technologie von Amazon aus, in dem sie öffentliche Fahndungsfotos mit Kongressabgeordneten vergleichen ließ (Snow, 2018). *Rekognition* identifizierte 28 Kongressabgeordnete als gesuchte Personen, wovon 39 % der Fehler Personen mit schwarzer Hautfarbe betrafen, obgleich diese nur 20 % des Kongresses ausmachten. Dieses beunruhigende Ergebnis wurde sogleich als Bestätigung der Studie von Buolamwini und Gebru (2018) aufgenommen. Schwarze Personen und insbesondere Frauen seien, so der Tenor, besonders hart von reproduzierter technischer Diskriminierung betroffen. Dies suggeriert zum einen eine beabsichtigte manipulative Diskriminierung und zum anderen implizieren solche pauschalen Aussagen, dass alle Überwachungstechnologien Klassifizierungen nach demografischen Merkmalen vornehmen müssten. So bestätigten Buolamwini und Gebru (2018) eines Bias, von dem besonders schwarze Frauen betroffen sind. Die getesteten Systeme von IBM, Microsoft und Megvii wiesen eine Fehlerquote von 20,8 % bis zu 34,7 % auf, wenn es um die korrekte Zuordnung schwarzer Frauen ging. Dem gegenübergestellt klassifizierten Microsoft und IBM hellere Männertypen mit einer Fehlerquote von 0,0 % bzw. 0,3 %. Face++ von Megvii hatte die niedrigste Fehlerquote von 0,7 % bei Männern mit dunkler Hautfarbe (ebd., S. 10 ff.). Zudem gab nur IBM ein Konfidenzintervall von 0-1 an, das Aussagen über die Wahrscheinlichkeit von *true positives* erlaubte. Es gilt somit hervorzuheben, dass es sich hierbei um die Zuverlässigkeit von *classifying algorithms* handelt und keine Aussage über die Identifizierungs- oder Authentifizierungsstärke der Algorithmen gemacht werden kann. Dies führt zu Verwirrung, denn oftmals werden gerade *classifying*-Technologien als Gesichtserkennungssysteme bezeichnet, die aber in der klassischen Fahndungsarbeit nicht notwendigerweise eingesetzt werden müssen.

Auch stehen die Datenquellen in der Kritik, ursächlich für die Übertragung von Stereotypen zu sein, denn je weniger divers oder je weniger repräsentativ ein Datensatz für den gewünschten Output, desto höher die Wahrscheinlichkeit, dass eine Gruppe unterrepräsentiert und somit für den Algorithmus nicht oder nur schwierig klassifizierbar ist.¹⁴ Dies belegen auch Studien über die Zuverlässigkeit

¹⁴ Unterdessen können Datensätze auch gekauft werden: Diese umfassen Datenset Real and Fake Face Detection, Labelled Faces in the Wild Home (LFW) bis DatasetLarge-scale CelebFaces Attributes (CelebA) Data-

von Gesichtserkennungssystemen im Falle bewegter Bilder aus dem Lebensalltag. So lag die Fehlerquote in London 2019 bei 79-81 %, je nachdem, ob nur die Missidentifikation in einem *one-to-many*-Szenario gemessen oder die Nicht-Zuordnung inkludiert wurde. Im ersteren Fall ging es darum, ob es sich um falsch Positive – Bilder, die fälschlicherweise zugeordnet wurden –, im zweiten Fall darum, ob es sich um positiv Falsche – Fotos, die fälschlicherweise nicht zugeordnet wurden – handelte. Selbstverständlich ist weder die Automatisierung der Identifizierung noch die der Authentifizierung vor fehlerhaften Zuordnungen immun, wie der ACLU-Test von 2018 belegt (Snow, 2018). Die Zuverlässigkeit solch klassischer Gesichtserkennungstechnologien hat in den vergangenen zwei Jahren deutlich zugenommen und wird vom NIST (*National Institute for Standard of Technology*) 2020 bereits als sehr hoch eingestuft: So waren die überprüften Systeme zu 99,7 % zuverlässig, d.h. sie können unter idealen Bedingungen Bilder einer Person zuordnen (Schwartz, 2020). Die Zuordnung eines Bildes und der Abgleich mit einer Datenbank gilt unterdessen bei 200 getesteten Algorithmen als robust, d.h. die Fehlerquote ist in allen Gruppen ungefähr gleich hoch. Dennoch bedeutet selbst eine Fehlerquote von nur 1 %, dass unter 1.000 Personen zehn Personen falsch identifiziert werden. Allerdings verteilt sich die Fehlerquote einigermaßen gleich über alle Hautfarben oder Geschlechter. Ebenso müssen die Bildqualität, die wiederum von der genutzten Hardware abhängt, und die Diversität bzw. Repräsentativität des Datensatzes bedacht werden, da hier Unzulänglichkeiten eine mangelnde Zuverlässigkeit zur Folge haben können. Verschwommene oder verpixelte Bilder, womöglich noch Videomaterial, erschweren ein *spotting* oder *matching* erheblich. Auch eine Verhüllung oder Verdeckung von Gesichtspartien im Zuge der Corona-Pandemie hat dies wieder verdeutlicht und lässt einen Anstieg der Fehlerquote erkennen (Materese, 2020). Die technische Zuverlässigkeit wird sich jedoch in den kommenden Jahren zusehends verbessern, wie schon die Betrachtung der vergangenen vier Jahre deutlich gemacht hat.

3.1.3 Ethische Betrachtung

Die Debatte um die technische Zuverlässigkeit muss sich jedoch von der Debatte der Legitimierung automatisierter Gesichtserkennungssysteme im öffentlichen Raum lösen, denn deren Einsatz kann

set 2015. Alle Datensätze fassen ein bestimmtes Attribut zusammen: So kann ein Datensatz Fotografien umfassen, die bspw. aus einer polizeilichen Datenbank stammen oder ein andere Datensatz enthält nur Fotos, die im realen Leben aufgenommen wurden. Der Tufts-face-Datensatz beispielsweise umfasst zudem Fotos von Nahaufnahmen, Fernaufnahmen und 3D-Daten. Das National Institute for Standards and Technology (NIST) bietet ein Datensatz an, in dem alle Bilder manuell den Fitzpatrick-Hauttypen zugeordnet wurden, um einfache Klassifikationen wie Ethnie oder Rasse zu umgehen und eine bessere Zuverlässigkeit zu ermöglichen. Breits die Zuordnung der Fotos zu einer Datenbank enthält somit eine Wertung.

unabhängig der technischen Details auch hinsichtlich ihrer (Be-)Nutzung kritisiert werden. Unter anderem filmte das NYPD (*New York Police Department*) in New York die BLM-Aktivisten während ihrer Proteste, ohne jedoch mitzuteilen, ob diese Bilder in eine polizeiliche Datenbank eingespeist und dort abgespeichert werden (Vincent, 2020). Die Ortung eines Aktivisten mittels rechtlich unsicheren Bildmaterials erfolgte: Problematisch war der Sachverhalt, weil sich das NYPD zur Lokalisierung des Aktivisten Bildern aus privaten Profilen wie Instagram zur Identifikation bediente (George & Offenhardt, 2020). Dadurch wird offensichtlich, dass sich eine ethische Legitimation nicht auf technische Zuverlässigkeit reduzieren lässt. Vielmehr wird an dieser Stelle eine deskriptive, interdisziplinäre Aufarbeitung jener Ebenen offenbar, anhand derer eine normative Ethik ein vollumfängliches Urteil zu fällen vermag.

So spiegelt sich die technisch-konzentrierte Debatte in den idiosynkratisch normativen Aspekten jener Merkmale wider, die sich generell aus den Gegebenheiten dieser Technologie oder Technik ergeben, wie bspw. Privatsphäre oder Sicherheitsbelange: So müssen die Nutzer z.B. vor Deep-Fake-Anwendungen geschützt werden können, möchte man einen Datenmissbrauch und damit einhergehende unangenehme Folgen für das betroffene Individuum vermeiden. Zudem sind biometrische Daten dahingehend speziell, dass ein Datendiebstahl zu einem lebenslangen Nachteil führen kann, denn diese Identitätsmerkmale lassen sich gar nicht oder nur sehr aufwendig verändern. Aber nicht nur technische Themen, wie bspw. das Datenhacking, bedürfen einer normativen Deliberation, sondern genauso müssen medien- und datenökonomische Fragen z.B. das Copyright mit in Betracht ziehen. Ebenso wird auszuarbeiten sein, welche Gründe eine anlassbezogene Videoüberwachung rechtfertigen, d.h. zu welchem Anlass videoüberwacht und das Datenmaterial bei Behörden abgelegt werden soll. Das Beispiel eines Einsatzes von Gesichtserkennungssystemen im öffentlichen Raum und noch spezieller in der Anwendung polizeilicher Fahndungsarbeit soll im Rahmen dieser Arbeit aufzeigen, dass die ethischen Risikoquellen, die zu nicht wünschenswerten Szenarien führen, aufgrund ihrer Multidimensionalität einzeln zu bedenken sind. Dies ist von großer Bedeutung, wenn unterschiedliche Verantwortungs- und Kompetenzbereiche ermittelt werden sollen, die sich um bestimmte normative Belange sorgen sollen.

Die gesellschaftliche Kontextualisierung der Systeme bedarf einer gesonderten Betrachtung, da auch der Einsatzort unabhängig von der Technologie relevant ist. Sobald über die ethische Unzulänglichkeit oder Unzumutbarkeit diskutiert wird, müssen auch nicht technische Momente, die zu unerwünschten oder unzumutbaren Folgen führen, berücksichtigt werden: So führt ein technisch zuverlässiges System wieder zu Diskriminierung, wenn bspw. nur Schulen mit einem hohen Anteil an Personen mit Migrationshintergrund videoüberwacht würden. Dies betont erneut, dass die Argumentation technischer Zuverlässigkeit vom Diskurs der politischen Legitimierung abzugrenzen ist, denn beide Debatten verweisen wiederum auf unterschiedliche normative Aspekte: So sind zuerst ethische Facetten

zu erörtern, die das Softwaresystem grundsätzlich betreffen, um im Anschluss strukturelle Weichenstellungen normativ zu evaluieren, die durch den Einsatz und den Gebrauch des Produkts entstehen. Denn je nach Anwendung im privaten oder öffentlichen Raum transformieren die Gesichtserkennungssysteme unterschiedlich stark normativ-strukturelle Handlungszusammenhänge, die als Lebenspraxis bezeichnet werden können (Nida-Rümelin & Weidenfeld, 2018; Nida-Rümelin, 2020). Allein die Einbettung von Gesichtserkennungssystemen in den Rahmen der oben aufgezeigten aktuellen politischen Geschehnisse lässt erkennen, dass Gesichtserkennungssysteme eine multidimensionale techno-ethische Deliberation erfordern: Die Zuverlässigkeit ihrer zugrundeliegenden Technologie und die Legitimation ihrer Implementierung müssen separat, zugleich jedoch auch in ihren Bezügen gedacht werden. Eine Offenlegung ökonomischer oder politischer Interessen zeigt zudem die sich aufdrängende Notwendigkeit eines öffentlich-politischen Diskurses über deren ethischen Stellenwert. Datenökonomische Fragen, z.B. ob ein Privatunternehmen wie Clearview AI Millionen private Daten, in diesem Fall private Fotos aus dem Internet, *scrapen* darf, um sie für ihre Zwecke zu instrumentalisieren, bedürfen einer gezielten rechtsethischen Erörterung. Des Weiteren kann medienethisch der Frage nachgegangen werden, inwiefern es legitim ist, dass ein Privatunternehmen diese Art der Technologie breit bewerben darf, ohne dass ihr Einsatz bereits politisch diskutiert und entsprechende Regulierungen festgesetzt wurden. Die Nutzung von Gesichtserkennungssystemen muss somit auf unsere gesamtgesellschaftlichen Zielsetzungen hin thematisiert werden, um eine Werteabwägung bspw. hinsichtlich Privatsphäre oder Autonomie vorzunehmen. Die Implementierung solcher Systeme führt in London zu geschätzt 627.000 Kameras (Fuest, 2019), die jeden einzelnen Passanten ca. 300-mal pro Tag aufzeichnen (Kaiser, 2019). Selbst bei geringer technischer Fehlerquote mag dieses Szenario wenig Begeisterung hervorrufen.

Es zeichnet sich bereits an dieser Stelle ab, dass eine Ethik für Softwareentwickler sowohl technische Kenntnisse als auch normatives Wissen in Bezug auf diese Technik und Technologie umfassen muss. Dies zeigt sich gerade an der Vielzahl an konkreten angewandten Ethiken, die sich der Informationstechnologie und -technik annehmen. Hierunter fallen die Datenethik, die Big Data Ethik oder auch eine Ethik der Künstlichen Intelligenz. Diese dezidiert technisch gebundenen Ethiken beziehen sich auf die normativen Fragestellungen, die sich aus der Beschaffenheit der Technik ergeben, weshalb sie in dieser Arbeit als auf technik-generische Werte gestützte Ethiken ausgewiesen werden. Diese Ethiken adressieren vor allem die Entwickler bereits mit dem Verweis, die besonderen, idiosynkratischen Aspekte im Design ihrer digitalen Artefakte zu berücksichtigen. Somit gehen diesen Mini-Ethiken meta-ethische Positionen voraus, die eine pro-ethische Haltung (Floridi, 2016; Turilli & Floridi, 2009) oder eine antizipatorische Ethik (Brey, 2012a, 2012b) zugrunde legen. *Codes of conducts*, Regulierungsunternehmungen sowie Compliance fokussieren sich gerade verstärkt auf technik-generische Werte. Lebenspraktische Herausforderungen treten sodann in den Hintergrund,

da sie zunächst außerhalb technischen Wissens anzusiedeln sind und so dem ersten Anschein nach keinen direkten Anwendungsbezug aufweisen. Allerdings gehen normativ-strukturelle Deliberationen bereits über das Legalitätsdenken hinaus und öffnen somit direkt den Raum für moralische Software. Eine ethische Betrachtung von Gesichtserkennungssystemen verweist auf die Unterscheidung in eine strukturelle und eine technik-generische Betrachtung, die es noch genauer auszuarbeiten gilt (Teil II, Kapitel 2). Die strukturellen Phänomene lassen sich nach drei normativen Kriterien gliedern: Sie müssen individuell, sozial und kollektiv zumutbar sein können, wenn sie einer vollumfänglichen normativen Rechtfertigung standhalten sollen. Die technik-generischen Werte, die direkt aus der Beschaffenheit der digitalen Technologie und Technik resultieren, sollten zusätzlich zu den strukturellen Phänomenen auf ihre moralische Güte hin überprüft werden. Abbildung 6 veranschaulicht die Konzepte des vorherigen Kapitels. Es ist ratsam, strukturelle Werte zunächst getrennt von technik-generischen Werten zu betrachten.

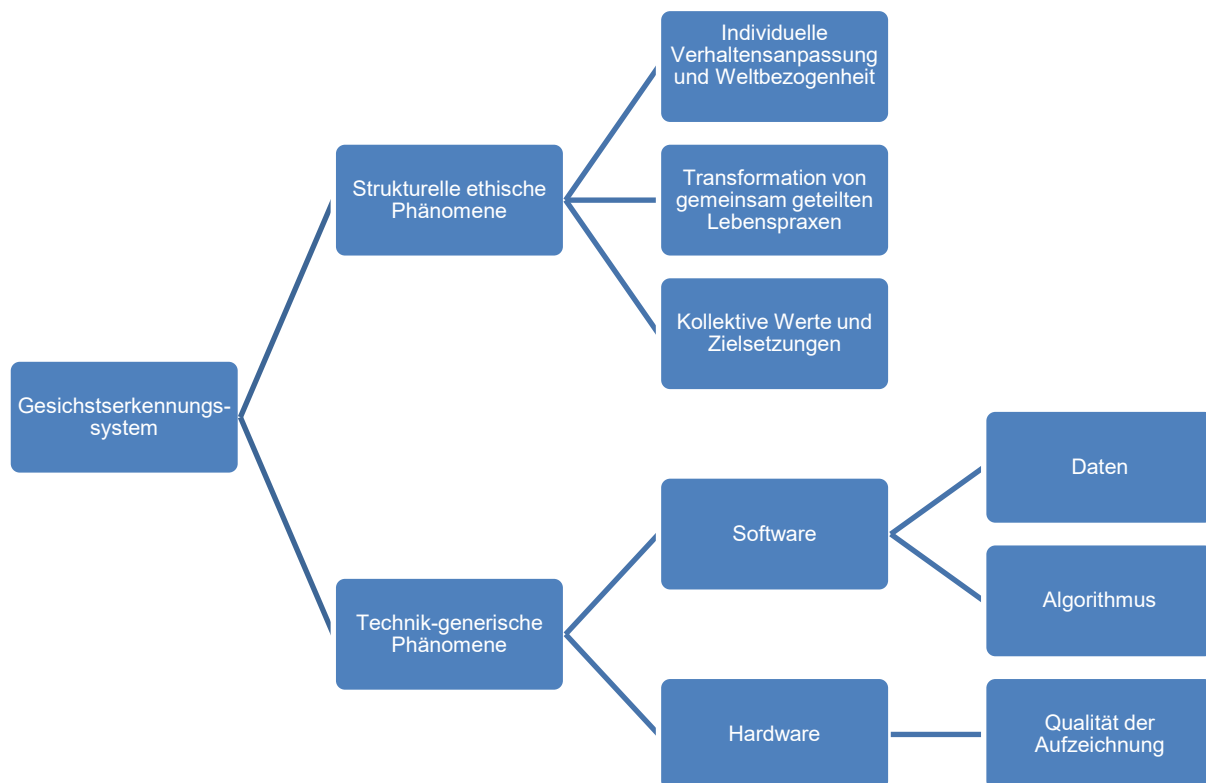


Abbildung 6 Gesichtserkennungssystem unterteilt in strukturelle und technik-generische Werte. Eigene Abbildung.

Ethische Deliberationen müssen eine praktische Anleitung geben und diese Anleitung muss einer lebensweltlichen Überprüfbarkeit standhalten können, wenn sie auch praktische Gültigkeit ausüben möchte. Eine Ethik für Softwareentwickler muss sowohl das Objekt als auch Handlungsräume bestimmen, sollen Softwareentwickler nicht für die Gesamtheit aller Missstände, die technisch bzw.

technologisch verursacht werden, die alleinige Verantwortung übernehmen oder als Sündenbock dienen. Welche Bereiche können ihnen zugetraut werden, ohne zugleich *moral distress* aufzubauen? Hierfür ist es hilfreich, Digitalisierung, digitale Transformation und das digitale Artefakt konzeptionell zu unterscheiden, um aus der Begrifflichkeit den Bedarf sowie die Vorgehensweise einer normativen Deliberation zu bestimmen.

3.1.4 Ausblick

Am Beispiel der Gesichtserkennungssysteme wurde die Komplexität des Gegenstandsbereichs hervorgehoben, um die Notwendigkeit einer interdisziplinären sowie multidimensionalen Analyse betonen zu können. Wenn über Gesichtserkennungssoftware gesprochen wird, werden viele unterschiedliche Systeme sowie unterschiedliche Anwendungsbereiche miteinander verglichen: von der technisch unterschiedlich ausgerichteten Identifizierung zur Klassifizierung oder von der Entsperrung von Zugängen über eine Zuordnung eines automatisierten, digitalen Fotoalbums bis hin zur polizeilichen Fahndungsarbeit. Dies führt jedoch zu einer schiefen Bewertungsgrundlage, denn nicht nur die idiosynkratischen ethischen Belange bedürfen einer kasuistischen Betrachtung, sondern auch die normativ-strukturellen Lebenspraxen müssen auf eine wünschenswerte Lebensgestaltung hin befragt werden können. So muss bspw. die Nutzung der Technik in ihrer Anwendung hinsichtlich eines individuell-privaten Bereichs – wie die Face ID von Apple oder das *Drowsiness-Detection*-System im privaten Pkw – von solchen, die im öffentlichen Raum eingesetzt werden – wie *DeepFace* von Facebook in sozialen Medien oder in der Strafverfolgung im Rahmen der polizeilichen Fahndungsarbeit – unterschieden werden. Das Ausmaß der Folgen einer Anwendung auf ein normativ-strukturelles gesellschaftliches Ordnungssystem fällt unterschiedlich gravierend aus: Je nachdem, welche individuelle Verhaltensanpassung gefordert wird, kann ein System wünschenswerte Auswirkungen oder weniger zumutbaren Druck auf Gewohnheitsstrukturen ausüben. Im Falle des *drowsiness detectors* mag die individuelle Verhaltensanpassung, nämlich das Pausieren des Fahrers, nicht nur individuell, sondern auch intersubjektiv und kollektiv durchaus erstrebenswert sein. Die Transformation geteilter Lebenspraxen bedarf folglich der Etablierung des technischen Systems in ein intersubjektives Erwartungsgefüge mit kollektiven Wertvorstellungen: So verändert ein im Bewerbungsverfahren eingesetztes Gesichtserkennungssystem etablierte Rollenerwartungen im Personalwesen ebenso wie das damit einhergehende Verhalten des Bewerbers. Daher müssen kollektive Zielsetzungen die Idee eines guten Lebens und deren Realisierungsbedingungen thematisieren, z.B. ob bestimmte neue Lebensgewohnheiten etabliert werden sollen.

Jeder Einsatz eines Gesichtserkennungssystems hat andere normative Folgen oder beeinflusst und strukturiert lebensweltliche Verpflichtungen in einem unterschiedlichen Maße mit unterschiedlicher

Intensität. Wenn die These der Möglichkeit moralisch angemessener Software aufrechterhalten werden soll und nicht allein die Nutzung von Software moralisch diskutierbar erscheint, gilt es besonders den Entwicklungsprozess digitaler Artefakte zu betrachten: Jegliche normative Wünschbarkeit sollte auf ihre technische Ursächlichkeit hin überprüft und die strukturelle Veränderung von Lebenspraxen auf ihre Zumutbarkeit hin untersucht werden. So wird im Rahmen dieser Arbeit Klarheit über die Methodik und das Vorgehen geschaffen werden, wie mit normativen Fragestellungen technik-generischer Provenienz umzugehen ist, denn ein kasuistisches Vorgehen wird durch den Mangel einer soliden Technikphilosophie erschwert, die eine Idee des guten Lebens sowie deren Realisierung entwickelt. Diese wird umso dringender benötigt, da es nicht ausreichend Fälle gibt, die einen Vergleich ermöglichen. Gerade die Einzigartigkeit digitaler Technologie und Technik erlaubt nur einen vorsichtigen Vergleich mit anderen Technologien (Teil III).

Allerdings verhilft bereits die Unterscheidung struktureller und technik-generischer Werte, die normativen Belange genauer zuzuordnen, um sodann auch Methoden der Abhilfe ausfindig machen zu können. Tabelle 4 veranschaulicht eine systematisierende Einteilung der normativ relevanten Bereiche: Normativ sensible Aspekte sollen dadurch systematisch Bereichen der Angewandten Ethik zugeordnet werden.

Domäne	Betrachtungsgegenstand	Reflexions- und Regulationsbegriffe	Ethik
Informationstechnik und -technologie	Wissen und Geräte der Datenverarbeitung, Vermittlung und Speicherung	Technische Funktionalität, Zuverlässigkeit und Robustheit	Technik-generische Fragestellungen der Datenethik, Ethics of AI, Big Data Ethics, Computerethik
Lebenspraxis	Kontext, Situation oder Ort	Humanistische Prinzipien: Autonomie, Integrität, Vernünftigkeit, Selbstverwirklichung	Strukturelle Fragestellungen der Bereichsethiken: Medizinethik, Ethik der Pflege, Bioethik, Wirtschaftsethik

Tabelle 2 Unterschiedliche normative Domänen digitaler Technologie und Technik

Da sie durch den Sachgegenstand bedingt werden, erwecken technik-generische Werte den Anschein, als ob es ausreiche, wenn sich eine normative Deliberation auf die negativen Folgen beschränkt, die sich aus der technischen Beschaffenheit ergeben. In der Tat sollten technik-generische Werte einer kontinuierlichen ethischen Überprüfung standhalten. Dennoch muss auch das gesellschaftliche Subsystem, in dem ein Einsatz vorgesehen wird, betrachtet werden, damit digitale Technologie und Technik mit ihren Auswirkungen auf gesamtgesellschaftliche Strukturen hin reflektiert werden können.

Von Bedeutung ist daher, dass Ethiken aufgrund ihrer jeweiligen spezifischen Objektfokussierung systematisch zugeordnet werden können, damit sie gezielt handlungsanleitend wirksam werden können. So werden technisch-generische normative Fragestellungen in einer Ethik der Künstlichen Intelligenz oder der Datenethik behandelt, während normative Konstituenzen, die durch den Einsatz in der Lebenswelt beeinflusst werden, in einer Analyse der Lebenspraxis, also häufig tugendethisch behandelt werden. Von diesen Analysen, die zunächst die Identifikation normativ relevanter Sachverhalte verfolgen, sind Ansätze zu unterscheiden, die als ein Vorgehen zu bezeichnen sind, wenn technische Umsetzungsmöglichkeiten erfolgen sollen, wie es bspw. in den *ethics by design* diskutiert wird. Eine solide theoretische Grundlage verhilft jedoch dazu, epistemische Klarheit zu gewinnen und die wesentlichen Eckpfeiler zu identifizieren. Hierfür muss man der Frage nachgehen, welche Philosophie und welche ethischen Theorien sich dazu eignen, um beide Aspekte digitaler Technologie und Technik einzufangen sowie angemessen zu behandeln. Zugleich muss sorgfältig zwischen einer Theorie, die das notwendige Hintergrundwissen liefert, und einer Methodik, mithilfe derer die Theorie systematisch angewandt werden kann, unterschieden werden.

3.1.5 Technik-generische Betrachtung

Die technik-generisch ethischen Erörterungen beruhen auf bestimmten Werthaltungen, die sich als Bedingungen eines für gut befundenen Lebens offenbaren. Die gegenseitige Bedingtheit technik-generischer und struktureller Normativität hebt nochmals deren Irreduzibilität hervor: Die Bewertung nicht funktionaler Zustände der Technik wird durch eingennommene und für wertvoll erachtete Haltungen geleitet und gestützt. Bei ‚Privatsphäre‘ handelt es sich um einen Wert, der von der Überzeugung getragen wird, dass er einem guten Leben zuträglich ist. Demensprechend wird er als normative Tatsache anerkannt (Nida-Rümelin, 1993a). Folglich zeigt sich der Wert ‚Privatsphäre‘ als wünschenswertes Merkmal individueller und kooperativer Lebensgestaltung, das es in allen Situationen zu berücksichtigen gilt, weshalb es auch in Bezug auf die Digitalisierung der Lebenswelt von großer Bedeutung ist. So wurde der Wert ‚Privatsphäre‘ durch den Einsatz digitaler Technologie und Technik

um „informational privacy“ erweitert, da der Einsatz von Computern und das Zusammentragen persönlicher Informationen in großen Datenbanken einer begrifflichen und legalen Neuausrichtung bedurfte (Moor, 1998, S. 16). Moor bezeichnet die konzeptionelle Veränderung klassischer Wertvorstellungen aufgrund von Informationstechnologie und -technik als „informationally enrichment“, da diese Rejustierung direkt aus der Beschaffenheit der Technologie und Technik resultiert (ebd., S.15):

Computers are [...] *informationally enriching* [Herv. d. A.]. Because of their logical malleability, computers are put to many uses in diverse activities. Once in place, computers can be modified to enhance capabilities and improve overall performance even further. Often, computerized activities become informationalized; i.e., the processing of information becomes a crucial ingredient in performing and understanding the activities themselves. When this happens, both the activities and the conceptions of the activities become informationally enriched.

Einen versprechenden Einstieg in die Wertewelt, die sich verstärkt durch einen Einsatz digitaler Technologie und Technik verändert oder erweitert, ermöglicht ein deskriptiver Ansatz. Hierzu sind die berufsspezifischen Verhaltenskodizes hilfreich, die als *codes of conduct* (CoCs) oder *codes of ethics* (CoEs) bezeichnet werden. Es scheint der Eindruck vorzuherrschen, dass bestimmte Werte immer aufgrund dieser Technologie und Technik einer kontinuierlichen Untersuchung bedürfen – zuvorderst unabhängig des gesellschaftlichen Subsystems, aber in Abhängigkeit der Auswirkungen auf gesamtgesellschaftliche Strukturen. Diese wiederkehrenden Werte werden von digitalen Technologien und Techniken unterlaufen, erweitert oder neu akzentuiert, gerade aufgrund der Beschaffenheit digitaler Technologie und Technik. Diese Werte werden in der vorliegenden Arbeit als technik-generische Werte bezeichnet.

„Es gibt verschiedene Motive, warum es sinnvoll erscheint, Ingenieuren sogenannte Codes of Ethics (CoEs) oder Codes of Conduct (CoCs) an die Hand zu geben. Zum einen sollen sie Ingenieuren, die mit ethisch relevanten Fragen konfrontiert sind, Orientierung bieten und ihnen einen Überblick über wünschenswerte Werte und Prinzipien geben. Es gibt jedoch eine Fülle von CoCs, die sich an Softwareingenieure im Allgemeinen und Entwickler, die mit Künstlicher Intelligenz arbeiten, im Besonderen richten. Ursprünglich wurden CoCs von Unternehmen als Reaktion auf die wachsenden Probleme mit Korruption und Fehlverhalten im Geschäftsverkehr eingeführt. Im Laufe der Jahre hat sich die Einführung von CoCs auf viele andere Bereiche ausgeweitet, insbesondere auf Technik und Medizin. Michael Davies (1998) argumentiert, dass ein berufsethischer Kodex von zentraler Bedeutung ist, um den einzelnen Ingenieuren als normative Orientierung zu dienen, um ihr eigenes Verhalten oder das ihrer Kollegen zu beurteilen und letztlich um die Ingenieure als eine Berufseinheit zu verstehen. CoCs würden somit drei Hauptzwecken dienen: Erstens geben sie dem einzelnen Ingenieur Orientierung und helfen, Fehlverhalten zu vermeiden. Zweitens dienen sie anderen Akteuren eines Berufsstandes als Maßstab, um potenzielles Verhalten als unethisch zu beurteilen und damit zum Ansehen des Berufsstandes als Ganzes beizutragen. Schließlich tragen sie zur Definition des Selbstverständnisses

eines Berufsstandes bei, indem sie ein Regelwerk dafür aufstellen, was ein professioneller Akteur tun oder nicht tun sollte – dies könnte für ein vergleichsweise junges Gebiet wie die Softwaretechnik besonders relevant sein. Schwartz (2001) skizziert acht Metaphern, die beschreiben, wie Individuen CoCs interpretieren können: als Regelwerk, als Wegweiser, als Spiegel, als Lupe, als Schild, als Rauchmelder, als Feueralarm oder als Club.

Die Frage nach der Wirksamkeit der Kodizes wurde jedoch schon sehr früh aufgeworfen. Schwartz (2001) führte 57 Interviews im Bereich der Unternehmensethik durch und berichtet, dass weniger als die Hälfte der Kodizes das Verhalten der Menschen tatsächlich beeinflusste. Kaptein und Schwartz (2008) kommen zu gemischten Ergebnissen in Bezug auf die Beziehung zwischen CoCs und der CSR-Leistung (*corporate social responsibility*) von Unternehmen. In jüngerer Zeit haben McNamara, Smith und Murphy-Hill (2018) ein Vignettenexperiment durchgeführt, um den Einfluss von CoCs auf Entwickler zu testen. Sie stellten fest, dass die explizite Anweisung an die Teilnehmer, den ACM-Ethikkodex bei ihrer Entscheidungsfindung zu berücksichtigen, im Vergleich zu einer Kontrollgruppe keinen beobachtbaren Effekt hatte.

Es gibt einige Untersuchungen, die ethische Kodizes und ihre Werte vergleichen und versuchen, sie zu quantifizieren, um einen möglichen Konsens ausfindig zu machen. Der Schwerpunkt der aktuellen Literatur liegt auf CoCs, die sich mit der Entwicklung von Systemen der Künstlichen Intelligenz befassen. Zeng et al. (2019), Fjeld et al. (2020), Jobin, Lenka und Vayena (2019) sowie Hagendorff (2020) haben analysiert, welche Werte in Ethikkodizes im Bereich der Künstlichen Intelligenz im Vordergrund stehen, um einen Überblick über die ethischen Grundsätze zu geben, die für Softwareingenieure in diesem speziellen Bereich als wichtig erachtet werden.

Jobin et al. (2019) kodierten beispielsweise 84 Dokumente im Bereich der KI-Ethikkodizes und fassten elf Hauptprinzipien zusammen (geordnet nach der Anzahl der Dokumente, die das Prinzip enthalten, absteigend): Transparenz, Gerechtigkeit/Fairness, Nicht-Malefiz, Verantwortung, Privatsphäre, Wohltätigkeit, Freiheit/Autonomie, Vertrauen, Nachhaltigkeit, Würde und Solidarität. Transparenz wurde in 73 von 84 Fällen (87 %) genannt, Solidarität in sechs Fällen (7 %). Fjeld et al. (2020) haben es sich ausdrücklich zur Aufgabe gemacht, CoCs zu analysieren, um einen Konsens innerhalb der Branche und der relevanten staatlichen sowie nichtstaatlichen Akteure zu ermitteln. Auch sie finden ähnliche Prinzipien wie Jobin et al. (2019). Sie strukturieren den Inhalt der Kodizes nach Themen, die aus Werten und Prinzipien bestehen, die sinnvollerweise darunter subsumiert werden können. Sie führen insgesamt acht Themen auf: Datenschutz, Rechenschaftspflicht, Sicherheit, Transparenz und Erklärbarkeit, Fairness und Nichtdiskriminierung, menschliche Kontrolle der Technologie, berufliche Verantwortung und Förderung menschlicher Werte. Wie bereits erwähnt, besteht ein Thema je aus

einer Reihe von Grundsätzen. Im Falle des Themas ‚Privatsphäre‘ lauten diese Grundsätze beispielsweise „Einwilligung, Möglichkeit zur Einschränkung der Verarbeitung, Recht auf Löschung, [(Empfehlung von)] Datenschutzgesetzen, Kontrolle über die Verwendung von Daten, Recht auf Berichtigung, Datenschutz durch Technik und Datenschutz (Sonstiges/Allgemeines)“ (Jobin et al., 2019.). Hagedorff (2020) kommt zu ähnlichen Ergebnissen und stellt fest, dass „insbesondere die Aspekte der Rechenschaftspflicht, der Privatsphäre oder der Fairness insgesamt in etwa 80 % aller Richtlinien vorkommen und die Mindestanforderungen für ethisch wünschenswerte KI-Systeme bildeten“ (aus dem Englischen übertragen aus Gogoll et al. (2021).

Aufgrund ihrer Häufigkeit werden die oben genannten Werte als zentraler Ausgangspunkt für normative Überlegungen in der Softwareentwicklung angenommen (Senges, Ryan & Whitt, 2017; Spiekermann, 2015). Diese Werte dürfen dabei entweder nicht durch Software verletzt werden oder sollen bei der Entwicklung positiv berücksichtigt werden. Die Herausforderung besteht also darin, Werte techno-ethisch zu interpretieren und in technische Lösungen zu übersetzen. Verhaltenskodizes sind sicherlich nützlich, um eine Reihe von (potenziellen) Werten festzulegen. Am Beispiel der Datenethik zeigt sich, dass die in den CoCs oder CoEs, z.B. der *ACM Code of Ethics and Professional Conduct* (Gotterbarn et al., 2018), oder in überstaatlichen Deklarationen, wie bspw. der *Ethics Guidelines for Trustworthy AI* (European Union, 2019), herausgestellten technik-generischen ethischen Werte durch vielfältige Maßnahmen technisch-gestalterisch oder politisch umgesetzt werden können. Dies sind allesamt Maßnahmen, die für die Einhaltung des Wertes der Privatsphäre ergriffen werden können. So muss im Rahmen der Privatsphäre über den Speicherort ebenso wie über Zugriffs- oder Eigentumsrechte nachgedacht werden. So rückt bspw. die Datenethik all jene ethisch relevanten Aspekte in den Fokus, die gerade aufgrund der großen Datenmengen entstehen. Die Datenethik kann somit als eine Ethik verstanden werden, die sich verstärkt mit technik-genuinen ethischen Anforderung beschäftigt. Hier zeigt sich zudem eine zunehmende Verzweigung in weitere Ebenen auf, die jeweils einer spezifischen Methode bedürfen, wie sie bspw. Xafis et al. (2019) für den Gesundheitssektor erarbeitet haben.

Datenethische Fragestellungen sollen die Datenherkunft, den Datentyp und deren Zusammenfassung in Datensätze berücksichtigen. Bereits in jedem dieser einzelnen Schritte gilt es, normativ angemessene Entscheidungen zu treffen und diese auch technisch zu implementieren. Ist dies nicht der Fall, kann es zu unerwünschten Ergebnissen kommen. Die Datenethik hat es folglich verstärkt mit technik-generischen Werten zu tun, die sich auf die Datenqualität und Datenquantität reduzieren lassen. Im Rahmen dieser datenethischen Fragen kommen wiederum bestimmte Maßnahmen zum Zug.

Verhaltenskodizes umfassen daher in der Regel eine Reihe übergeordneter regulativer Ideen, deren Status mehr oder weniger selbstverständlich ist. Aber sie legen keine Methode fest, wie diese verschiedenen einzuhaltenden Werte zu gewichten und mögliche Konflikte zwischen ihnen zu lösen sind. Sie sind weitestgehend neutral gegenüber den verschiedenen Theorien der normativen Ethik, was ein Vorteil in Bezug auf ihre Akzeptanz, aber ein Nachteil hinsichtlich ihrer Praktikabilität ist.

3.1.6 Strukturelle Betrachtung

Es gibt eine Vielzahl von Werttypen: Ökonomische Werte sind durch wirtschaftliche Gründe wie Effizienz sowie Gewinn motiviert und politische Werte durch politische Gründe. Moralische Werte sind durch moralische Gründe motiviert, was bedeutet, dass ihre Motivation über die persönlichen Interessen eines Akteurs hinausgeht. Moralische Gründe variieren oft von Kultur zu Kultur und können sich im Laufe der Zeit ändern. Sie können jedoch in universelle Gesetze umgewandelt oder in moralisch wünschenswertes Verhalten, d.h. in Tugenden, umgesetzt werden. Ein Wert wie ‚Fürsorge‘ zeigt sich in einem konkreten Imperativ wie ‚schade keinem Anderen!‘. Diese Aufforderung kann als Recht formuliert werden, sobald sie als legitimer Anspruch anerkannt ist. Die Aufforderung kann auch in einer moralischen Norm ausgedrückt werden, die gebietet, niemanden weder physisch noch psychisch zu misshandeln. Nun lässt sich der Wert der Fürsorge nicht einfach aus der informationstechnischen Typologie ablesen. Eine Zusammenstellung aller humanitätsbewahrender Werte ist weder deskriptiv – aufgrund der Vielfalt – erstrebenswert, noch führt dies zu einer wohlbegründeten normativen Entscheidungsfindung, da keine Priorisierung vorgenommen werden kann.

Digitale Artefakte lassen sich in ihrer normativen Wirkung nicht nur technisch erklären, d.h. sie lassen sich nicht ausschließlich in ihrer technischen Funktionalität beschreiben. Vielmehr beeinflussen sie auch unsere gegenseitigen Erwartungen und Ansprüche. Gerade deshalb verändern einige digitale Artefakte ganze gesellschaftliche Bereiche (Teil II, Kapitel 6). Diese Transformation zu verstehen, ist auch ein konzeptionelles, analytisches und hermeneutisches Unterfangen, das zwar mit technischem *know-how* einhergeht, aber auf ein *know-why* abzielt (Reijers & Coeckelbergh, 2020; Spiekermann & Winkler, 2020; Zuber et al., 2020). Das digitale Produkt muss als Strukturierung sozialer Kontexte, menschlicher Fähigkeiten und als Veränderung der Art und Weise betrachtet werden, wie Menschen die Welt wahrnehmen oder verstehen (Ihde, 1990). Werte, die in diesem Zusammenhang wesentlich erscheinen, bedürfen einer anderen ethischen Analyse als technik-generische Werte. In dieser Arbeit werden diese Erwartungshaltungen als strukturelle Werte verstanden. Hier geht es verstärkt um die Modifikation oder Konstitution neuer Handlungsräume, die durch einen Einsatz informationstechnologischer Artefakte hervorgerufen werden.

Es erscheint daher von großer Bedeutung, die Werthaltungen in diesem Zusammenhang mit einer Tugendethik zu behandeln, d.h. darüber nachzudenken, ob eine bestimmte Informationstechnologie z.B. moralische, nachhaltige oder die Autonomie respektierende Praktiken fördert oder untergräbt. Gerade hier kann die Rückbindung an die Frage nach der Idee eines guten Lebens und der Realisierung desselben stattfinden. In der Tugendethik wird eine moralische Praxis als eine kollektive Handlung definiert, deren Zweck geteilt und als wertvoll bezeichnet wird, z.B. Ehrlichkeit, Freundschaft oder Fürsorge (MacIntyre, 2007). Die Tugendethik hebt hervor, dass Menschen, welche die Eigenschaft haben, fürsorglich zu sein, entsprechend handeln werden, ohne ständig kognitiv abwägen zu müssen, ob sie ein fürsorglicher Mensch sein sollen und was Fürsorge in diesem Kontext bedeutet. Digitale Artefakte können Menschen daran hindern, fürsorgliche Beziehungen aufzubauen sowie zu pflegen und sogar zu Verhaltensweisen führen, die es erleichtern, Menschen zu stalken, zu schikaniaieren, zu demütigen oder zu diffamieren. Dieser Anlass gebietet, das Design digitaler Artefakte zu überdenken oder – im schlimmsten Fall – darüber nachzudenken, ob sie überhaupt eingesetzt werden sollten.

Dies macht die Tugendethik zu einem praktikablen Ansatz im Bereich der Softwaresysteme, in dem wünschenswerte Verhaltensweisen, im klassischen Sinne Tugenden, auf technische Kontexte angewendet werden (Vallor, 2016; Reijers & Coeckelbergh, 2020). Sobald wünschenswerte Verhaltensweisen formuliert sind, müssen sie durch Praktiken gefördert werden. Vallor (2016, S. 36-37) konzentriert sich auf zwölf techno-moralische Tugenden: 1. Ehrlichkeit, 2. Selbstbeherrschung, 3. Demut, 4. Gerechtigkeit, 5. Mut, 6. Empathie, 7. Fürsorge, 8. Höflichkeit, 9. Flexibilität, 10. Perspektive, 11. Großmut, 12. techno-moralische Weisheit. Im Gegensatz zu den Werten, die auf ein technisches Design anwendbar sind, beziehen sich die Tugenden sowohl auf den Charakter des Entwicklers als auch auf den des Benutzers oder anderer Beteiligter. Die Gewöhnung an diese zwölf guten Techno-Tugenden soll ein gutes Leben fördern und moralisches Fehlverhalten durch den Einsatz und die Nutzung technischer Geräte verringern. Vallor wendet sich also nicht ausschließlich an Softwareentwickler, sondern an die gesamte Gemeinschaft, in der digitale Artefakte eingesetzt werden. Daher müssen wünschenswerte Tugenden bei der Gestaltung eines Produkts insofern berücksichtigt werden, als wünschenswerte Praktiken durch die Entwicklung, den Einsatz und die Nutzung digitaler Technik nicht untergraben werden sollten. Tugendethik ist für die Softwareentwicklung ein notwendiger und interessanter Ausgangspunkt, da Softwareentwickler diese Tugenden selbst entwickeln, d.h. sich bewusst für diese entscheiden. Das wiederum bedeutet, dass sowohl das Management als auch Entwicklungskulturen techno-tugendhaftes Verhalten durch die Einführung wünschenswerter Praktiken in den Entwicklungsprozess fördern und begünstigen sollten. Dazu gehört auch die Berücksichtigung der Benutzer und anderer indirekter Interessengruppen, deren Leben durch das System beeinflusst wird.

Es erfordert jedoch praktisches Denken, um den Zusammenhang zwischen Werten, Zielen und spezifischen Anforderungen, die die Implementierung leiten, zu erklären und zu navigieren (für einen Überblick über Ziele im *requirements engineering* siehe van Lamsweerde (2000)). Diese Fähigkeit muss erlernt und geübt werden, vor allem, weil normative Kompetenz nicht nur ein epistemisches Unterfangen ist: Es geht um die Ausübung von Urteilskraft oder – in Vallors Terminologie – um technomoralische Weisheit. Die Urteilskraft ermöglicht dem Einzelnen, sich ein ethisches Bewusstsein über die Notwendigkeit normativer Überlegungen anzueignen. Das bedeutet also, dass das praktische Denken kein rein kognitiver Prozess ist, sondern auch eine *hexis*: eine Haltung gegenüber der Welt. Hier muss betont werden, dass Tugenden (*arete*) eher im klassischen Sinne von Aristoteles verstanden werden müssen. Folglich ist eine Tugend (*arete*) nicht nur ein Ergebnis der Gewöhnung (*ethos*), sondern beinhaltet auch eine *krisis* (Entscheidung) und eine *hexis* (Haltung) (Nida-Rümelin, 2020). In der modernen Tugendethik wird diese dreifache Konstitution der Tugenden oft vernachlässigt und die Tugenden werden auf das Ergebnis der sozialen Gewöhnung reduziert (MacIntyre, 2007). Von daher ist es interessant, wie sich normatives Orientierungswissen (Nida-Rümelin, 2009) während des technischen Entwicklungsprozesses bewahren lässt, d.h. während technische Sprachen und Konzepte verwendet werden. Zudem muss eine Strukturanalyse herausarbeiten können, inwiefern informationstechnische Geräte normativ wünschenswertes Verhalten in konkreten Situationen verunsichern oder fördern. Dies führt jedoch zu der Frage, wer in der Lage sein sollte, sich mit normativen Fragen zu befassen, die durch technische Geräte verursacht werden oder sich auf diese beziehen. Es liegt auf der Hand, dass ethische Fragen auch aus der Perspektive der Entwickler behandelt werden müssen. Dies ist zum einen eine Frage der Curricula und der Ausbildung (Center for Engineering, Ethics, and Society & National Academy of Engineering, 2017; Shen et al., 2021; Tavani, 2013). Zum anderen ist diese Forderung jedoch ebenso relevant für die Ermöglichung bzw. Förderung normativen Denkens in der Arbeitswelt. Ethische Deliberation erfordert gerade wegen ihrer kontingenten Struktur, die sich aus dem freien Willen speist, eine hinreichend effektive strukturelle Integration.

3.2 Transparenz und Opazität als Reflexionsbegriffe

Die Schwierigkeiten, Informationstechnologie und Informationstechnik philosophisch zu analysieren und ethisch zu evaluieren, sind der Vielschichtigkeit des Untersuchungsobjekts zu verdanken. Diese wird sowohl in der Software selbst, den Entwicklungsprozessen als auch in der Anwendung bzw. im Gebrauch sichtbar. Software wird hierbei als Teil technischer Sachsysteme, wie z.B. des Verkehrs, als interaktive Mensch-Maschinen-Relation, aber auch als digitales Artefakt verstanden werden müssen. Dabei gilt es bereits an dieser Stelle zu betonen, dass Software somit in ihrer Relationalität, ihrer Kompositionalität, aber auch in ihrer Singularität einer normativen Betrachtung unterzogen werden

muss, möchte man ein fundiertes moralisches Urteil erzielen. Diese metatheoretische Perspektive ermöglicht zudem, die Auswirkungen des Status von Informationstechnologie und Informationstechnik auf menschliche Wahrnehmungshorizonte und Urteilsvermögen (bspw. *decision supportive technology*), zwischenmenschliche Beziehungsstrukturen (bspw. *social media*) und gegenseitige Erwartungshaltungen oder gesamtgesellschaftliche Einstellungen (bspw. *e-democracy*) zu thematisieren. Die Multidimensionalität, die es bei einer sorgfältigen Evaluation zu berücksichtigen gilt, verliert sich daher häufig in der Betrachtung eines bestimmten Anwendungsbereichs, wie bspw. der Kommunikation, eines bestimmten gesellschaftlichen Bereichs, wie bspw. der Medizin, oder einer bestimmten Informationstechnologie oder -technik, wie bspw. der Künstlichen Intelligenz. Vor allem seit der empirischen Wende der Technikphilosophie hat dies zu einer Fülle an Einzelstudien geführt, die sich je nach fachlicher Perspektive dieser Vermittlungsstruktur paradigmatisch annehmen (Brey, 2010). So diskutieren Friedman (1997) sowie Flanagan, Howe und Nissenbaum (2008), die einen Value-Sensitive-Design-Ansatz verfolgen, Software als Exklusivität, da sie eine moralische Bewertung von Software als eigenständiges Untersuchungsobjekt vornehmen. Sie markieren somit einen Verantwortungsraum für Softwareentwickler. Wiederum andere nehmen eine moralische Bewertung menschlicher Handlungen vor, weshalb Software als Betrachtungsgegenstand in den Hintergrund rückt und meist der Nutzer oder ein Nutzer-Kollektiv (Grimm, Keber & Zöllner, 2019), seltener der Entwickler, in einer Berufsethik (Bynum, 2004) evaluiert wird. Der *social constructivism* hingegen, wie ihn z.B. Feenberg (1991) entwickelte, legt vor allem den Determinismus von Sachsystemen dar. Die Wechselbeziehungen, die sich aus den verschiedenen Akzentuierungen ergeben, werden aufgrund ihrer Modellbildungsabsicht beiseitegeschoben, um effizient und zielorientiert Phänomene zu rekonstruieren, die durch den Einsatz von Informationstechnik zunehmend beeinflusst werden. Eine meta-theoretische Betrachtung hingegen sollte gerade diesen Charakter der informationstechnischen Vermittlung in seiner Komplexität abstrahierend thematisieren können, um eine beabsichtigte Verortung philosophischer Erörterung oder normativer Bewertung zu rechtfertigen. Hierfür bedarf es zunächst eines epistemischen Zugangs, um normative Tatsachen oder allgemein für eine normative Bewertung relevante Tatsachen in Anbetracht des Untersuchungsobjekts oder -szenarios zu konstatieren.¹⁵

¹⁵ Der metaethische Realismus moralischer Tatsachen und die Möglichkeit einer Erkenntnis sowie Einsicht kann hier nicht weiterdiskutiert werden (siehe hierzu Nida-Rümelin, 2019).

Folgende Tabelle soll hierfür eine erste Orientierung bieten, um einen normativen Erkenntnisgewinn zu erzielen.

Betrachtungsgegenstand (an-sich)		Gebrauch (für-sich)	Strukturelle normative Auswirkungen auf	Generische ethische Fra- gestellungen
Exklusivität: Informationstechnisches Artefakt (Softwaresystem)		Anwen- dungsbereich (z.B. Kommu- nikation)	Individuum: Wahrnehmungshorizont und Urteilsvermögen	z.B. Pri- vatsphäre, Si- cherheit, Dis- kriminierungs- freiheit, Inklus- sion, Fairness, Nachhaltig- keit...
Komposition: Informationstechnisches Sachsys- tem (system-of-systems)		Gesell- schaftsbe- reich (z.B. Medizin)	Intersubjektivität: Gegenseitige, interaktive Erwartungshaltungen	
Relationalität: Mensch-Maschinen- Schnittstellen (sozio-technische Systeme)	Expansion: Cyber-phy- kalische Systeme	Institutionali- sierung, Kon- ventionalisie- rung, kultu- relle Kontex- tualisierung	Kollektivität: Gesamtgesellschaftliche Erwartungshaltungen	

Tabelle 3 Erweiterte Zuordnung der normativ relevanten Domänen auf die idealtypischen technik-generischen oder strukturellen Analyseebenen

So können Forschungs- und Evaluationsformate oder technikphilosophische Theorien z.B. ein konkretes informationstechnisches Artefakt, bspw. ein Softwaresystem, das in der Gesichtserkennungstechnik eine Rolle spielt, in seiner Exklusivität betrachten. Ein Gesichtserkennungssystem kann dann als System-of-Systems (SoS) bewertet werden, insofern darunter die Komposition vieler einzelner Software- und Hardwarekomponenten verstanden wird. Ein SoS besteht wie ein System auch aus Einzelteilen, deren Verbindung zueinander sowie der Annahme, dass das Ganze, das finale digitale Artefakt, mehr als seine Teile ist. Ein SoS wird jedoch zusätzlich dadurch charakterisiert, dass es Komponenten auf eine andere Weise aufnimmt und somit in eine neuartige Verbindung zu- und miteinander bringt (*legacy*), was zudem in einem (zuvor) unvorhersehbaren Zweck des gesamten SoS als Singulär resultiert (*mystery*) (Boardman & Sauser, 2006, S. 121-122). Ein Gesichtserkennungssystem ist zudem ein sozio-technisches System, das in seiner Klärung und Ausgestaltung neben technischen Funktionalitäten wesentlich auch Menschen, Organisationen und gesellschaftliche Bezüge miteinbezieht (Baxter & Sommerville, 2011). Ropohl (2009) umreißt dementsprechend den Begriff des sozio-technischen Systems als eine Metaebene, die aus zwei Subsystemen – einem Sachsystem und einem menschlichen System – besteht, die zusammengetragen eine neue Arbeits- und Handlungsform konstituieren. Ist ein Softwaresystem eingebettet in zumeist industrielle Produktionsabläufe, die weniger auf die Interaktion mit Menschen ausgerichtet sind, sondern eine bestimmte Aufgabe wiederholend substituieren, spricht man dagegen von einem cyber-physikalischen System (Wolf, 2009). Gesichtserkennungssysteme können sodann hinsichtlich ihrer technischen Funktionalitäten, aber auch weiterer normativer Bewertungsmerkmale innerhalb eines Anwendungsbereichs diskutiert werden, z.B. als Diagnoseinstrument seltener Krankheiten in einem gesellschaftlichen Bereich, hier bspw. in der Medizin. Die Analyse dieses digitalen Artefakts innerhalb eines Systems ließe sich wiederum hinsichtlich der strukturellen Beeinflussung des jeweils individuellen Wahrnehmungshorizonts sowohl des Arztes als auch des Patienten (hier beide als Nutzer) erweitern (Teil I, Kapitel 1.1 und 1.2). Des Weiteren sollen dann diese Auswirkungen auf ihre ethische Zumutbarkeit hin bewertet werden: Soll der Arzt seine Diagnosen auf digitale Entscheidungsinstrumente stützen? Die strukturelle Betrachtung muss somit auch den intersubjektiven Moment des Arzt-Patienten-Verhältnisses oder die gesamtgesellschaftliche Wertschätzung eben jener Beziehung einer ethischen Bewertung unterziehen: Wie verändert sich durch den Einsatz eines digitalen Diagnoseinstruments das Verhältnis zwischen Arzt und Patient? Ist diese Transformation zumutbar oder ist eine Veränderung gar wünschenswert? Sollen neue Formen der Vertrauensbeziehung erörtert werden?

Eine vollumfassende normative Erörterung bedarf zusätzlich einer generisch-ethischen Systematisierung, die sich aus der Beschaffenheit der Informationstechnik und -technologie selbst ergibt (Teil II, Kapitel 2.1), wie bspw. Aspekte der Datensicherheit oder bestimmte Formen der Diskriminierung. Die

Übersichtstabelle ermöglicht somit eine Zuordnung philosophisch und ethisch relevanter Sachverhalte, die einen je unterschiedlichen theoretisch begründeten Problem- und Lösungsraum begründen. Von daher können Einzelstudien oder gar Theorien niemals alle Themenfelder und deren Wechselwirkungen berücksichtigen, sondern müssen sich auf einige Beziehungselemente in ihrer Betrachtung und Bewertung beschränken. Von Bedeutung ist daher, hervorzuheben, welcher Schwerpunkt zu welchem Zeitpunkt besprochen werden sollte und welche theoretischen oder empirischen Lösungsansätze hierfür Orientierung bieten. Bspw. ist bei einer Analyse der gesellschaftlichen Erwartungshaltungen Nida-Rümelins *Theorie einer Strukturellen Rationalität* hilfreich, um normative Aspekte zu lokalisieren und angemessen adressieren zu können (Teil I, Kapitel 2), während value-sensitive Ansätze eine Übersetzungshilfe sind, die normative Sprache in informationstechnische sowie -technologische Termini, Methodik und Verfahren übertragen möchte (Teil II, Kapitel 7.2).

In der informationstechnisch-philosophisch geprägten Literatur wird diese mehrdeutige Verwobenheit, die den zu betrachtenden Gegenstand oder das zu betrachtende Phänomen nur schwer oder gar nicht greifbar werden lässt, mit den Begriffen der *Transparenz* und der *Opazität* diskutiert. Opazität bezeichnet Sachverhalte, in denen der Gegenstand oder das Phänomen nicht erkennbar oder nicht nachvollziehbar ist. Sie bleiben außerhalb des Kontrollierbaren. Dieses epistemische Defizit bezieht sich auf die Erkenntnis von theoretischem oder praktischem Faktenwissen. So sind solche Situationen oder Sachverhalte normativ opak oder transparent, die sich der praktischen Urteilskraft entziehen oder in denen nicht unmittelbar praktische Schlüsse gezogen werden können. Merdes (2020, S. 254) unterscheidet zudem noch zwischen „prinzipieller oder essentieller Opazität und lediglich akzidentieller epistemischer Intransparenz. Eine Methode ist *essentiell* [Herv. d. A.] opak, wenn sie prinzipiell opak ist, unabhängig von kontingenten Gegebenheiten [...]. Kann die Opazität dagegen grundsätzlich überwunden werden, ist sie lediglich akzidentiell oder kontingent.“ Normative Aspekte von Software, Softwaresystemen und Lebensweisen, in die diese eingebettet werden, müssen folglich dem praktischen Urteilsvermögen zugänglich gemacht werden können. Ein normatives Urteil kann gefällt werden, denn Einsicht in normative Tatsachen ist möglich. Dieses hoffnungsvolle Szenario verweist zugleich auf den Umstand, dass viele Gegenstände einer ethischen Deliberation zwar nicht essentiell opak sind, sich aufgrund ihres transparenten Charakters jedoch einer ethischen Bewertung entziehen können. Das Konzept der Transparenz umfasst somit die Problematik einer Analyse oder Evaluation aufgrund der Durchsichtigkeit des Gegenstands bzw. des Phänomens. Transparenz beeinträchtigt somit die unmittelbare Wahrnehmbarkeit eines Sachverhalts, da der Gegenstand, die Technologie oder das zugrunde gelegte Modell nicht wahrzunehmen sind oder gar sein sollen. Van Den Eede (2011, S. 154) definiert demgemäß, „(t)ransparent' here means that something is not perceived, that it is 'invisible', or escapes conscious attention—it still *is* there in some capacity,

but one sees ‘through’ it. On the contrary, ‘opaque’ is something that lies clearly in view or to which deliberate attention is paid, but in any case, importantly, ‘through’ which one cannot look.”

Transparenz und *Opazität* sollen daher als Reflexionsbegriffe verstanden werden, die als normative Gedankenstützen eine ethische Analyse leiten sollen. Sie fungieren somit als Zielgrößen, die als unbekannte Variablen den Blick darauf lenken sollen, was gerade nicht unmittelbar ersichtlich oder für den Akteur gar nicht erkennbar sein kann. Wichtig hierbei ist, wie Merdes (2020) betont, dass epistemische Transparenz und Opazität immer Akteur-relativ sind, d.h. in Abhängigkeit einer Perspektive gedacht werden müssen. Selbst wenn einige opake Sachverhalte zukünftig erkenntnistheoretisch überwindbar sein werden, muss dieser Sachverhalt auf den Ist-Zustand hin reflektiert werden. An dieser Stelle sei die *Explainability*-Problematik Künstlicher Intelligenz zu nennen, insofern Künstliche Intelligenz als Black-Box-Modell Anwendung findet. Hierunter fallen sogenannte Neuronale Netze, die in ihrem Vorgehen nicht erklärbar gemacht werden können (Fenech, Strukelj & Buston, 2018; Knight, 2017). Aber auch die Kombination verschiedener statistischer Modelle, sogenannter *conceptual models*, oder die Oberflächenausgestaltung von Homepages, die User-Verhalten mittels kognitiver Verhaltensmodelle strukturieren, sind aus Endnutzersicht oft nicht nachvollziehbar und entziehen sich somit einer vernünftigen Rechtfertigung, weshalb sie auch als *dark patterns* bezeichnet werden. Dies gilt für Laien meist noch mehr als für Experten. Aber auch hier muss erwähnt werden, dass die geforderte Sichtbarkeit nicht ausreicht, um theoretische oder praktische Urteile vollständig zu begründen, wie es bspw. Ananny und Crawford (2018) für die Thematik der Transparenzforderung hinsichtlich Künstlicher Intelligenz zeigen.

Transparenz ist dabei nicht nur der selektiven Wahrnehmung des Betrachters geschuldet, insofern darunter verstanden wird, dass die Aufmerksamkeit auf einen Sachverhalt gelenkt wird, während andere zwangsläufig in den Hintergrund treten müssen. Sondern es geht verstärkt darum, dass informationstechnische Artefakte im Gebrauch nicht bemerkbar sein sollen: Der Computer, auf dem dieser Beitrag geschrieben wird, soll verschwinden, sobald man arbeitet. Jedes Update, jede Unhandlichkeit, das oder die den Computer als technisches Artefakt wieder in den Mittelpunkt der Aufmerksamkeit rückt, erscheint als lästiges Arbeitsübel oder Hemmnis. Der Computer soll der Arbeitspraxis nicht im Wege stehen. Ebenso soll ein Gesichtserkennungssystem im Hintergrund operieren, wenn es optimal bspw. für die Nachverfolgung von Straftätern eingesetzt werden soll. Zudem kann es für die Betroffenen gar nicht als technisches System in seiner Gänze wahrgenommen werden. Es entzieht sich dem Wissensstand vieler Stakeholder: Kameras mögen evtl. noch erkennbar sein, aber die Übermittlung der Daten, der Speicherort, der Umgang mit diesen und die Analyse der gesammelten Daten bleiben (zunächst) transparent, da sie sich der unmittelbaren Wahrnehmung entziehen, und zu großen Teilen auch prinzipiell opak. Ganz zu schweigen von solchen strukturellen Veränderungen, die manche Systeme in gefestigten kulturellen Netzwerken nach sich ziehen. Für *praktisches Urteilen*

muss daher ausdrücklich erwähnt werden, ob computergestützte Systeme für die Stakeholder zu jenem Zeitpunkt der Beurteilung opak oder transparent sind, um sodann eine ethische Analyse des Umgangs mit eben jenen Systemmerkmalen zu ermöglichen (Merdes, 2020).

Die Vielschichtigkeit, die sich aus den mannigfaltigen Schnittstellen gepaart mit der inhärenten Transparenz und Opazität von Softwaresystemen ergibt, bedarf einer genauen Kenntnis des Untersuchungsobjekts und dessen Umwelt, um die normativen Belange zielgerichtet thematisieren zu können. Van Den Eede (2011) hebt zurecht hervor, dass Transparenz und Opazität häufig gleichzeitig auftreten: So mag die Transparenzforderung an Neuronale Netzwerke erfüllt werden, indem sie aufgezeigt werden, zugleich jedoch erweisen sie sich als essentiell opak (Ananny und Crawford, 2018). Die meisten technikphilosophischen Ansätze diskutierten allerdings entweder nur den Charakter der Transparenz, während andere sich der Opazität widmen, aber deren Gleichzeitigkeit wird häufig außer Acht gelassen. Dies liegt nach Van Den Eede darin begründet, dass sich philosophische Analysen den technisch medierten Zwischenraum meist nur aus einem Blickwinkel annehmen. Entweder setzen sie auf die Interaktion Mensch-Technik, wie bspw. die postphänomenologischen Theorien eines Peter Paul Verbeek oder eines Don Ihde, erörtern soziale Wechselbeziehungen, wie bspw. Andrew Feenberg, oder diskutieren technische Opazität vor allem im Zusammenhang mit Künstlicher Intelligenz (Angerer & Frank, 2020; Heil, 2021; Wright, 2021). Dabei wird häufig auf eine Verzahnung der unterschiedlichen Einzelaspekte sowie ein damit verbundenes Explizieren des technologisch medierten Miteinanderseins (*being-with-each-other*) abgezielt. Häufig wird jedoch das Ziel verfehlt (Van Den Eede, 2011). Mithilfe einer strukturellen Betrachtung lässt sich diese gegenseitige Bedingtheit einfangen (Teil II, Kapitel 2.2).

Die oft stark, teilweise mehrfach kontextualisierten Fallstudien können daher der Frage gar nicht gerecht werden, welche phänomenologischen oder gar ontologischen Transformationen die Technologie durch ihr wesentliches Dazwischenschieben hervorruft. Hinsichtlich einer ethischen Bewertung von Softwaresystemen, System-of-Systems, sozio-technischer oder cyber-physikalischer Systeme ist daher geboten, sich der beiden Reflexionsbegriffe anzunehmen und die Systeme auf jene beiden Kategorien hin zu durchdenken. Dabei werden sowohl essentielle Strukturmerkmale als auch generische normative Aspekte lokalisiert. Gerade hier zeigt sich bereits, wie Haltungen, Einstellungen, Werte sowie Normen durch das Design entweder direkt vermittelt, wie in Breys *causalist* Interpretation (Brey, 2000a, 2010), oder ausgedrückt werden, wie es Nissenbaum mit ihrer *expressionist* Interpretation vornimmt (Flanagan et al., 2014). Diese zunächst verborgenen, eingebauten Tendenzen (*built-in tendencies*) (Brey, 2010, S. 49) müssen durch eine aufdeckende und aufklärende Philosophie und Ethik artikuliert werden. Brey plädiert in seinem Aufsatz *Values in Technology and Disclosive Computer Ethics* (2010) für ein solches Vorgehen, indem er Opazität und Transparenz auch in direktem Bezug zu moralischen Praxen, als individuell akzeptierte, getragene vernünftige Lebensweisen (Nida-

Rümelin 2020; Teil I, Kapitel 2) denkt, die es in einer ethischen Evaluation offenzulegen gilt. Opazität und Transparenz betreffen hier somit wesentlich den strukturellen Aspekt normativer Überlegungen, weniger den generischen, da sie nicht direkt moralische Verhaltensweisen steuern, sondern aus der technischen Bedingtheit des Untersuchungsobjekts resultieren.

Allerdings konnotiert Brey den Begriff der Transparenz anders als es Van Den Eede oder Merdes tun: Er formuliert ihn geradezu als Vorwurf an die Computer-Community. Diese würde vor allem transparente moralische Praxen erörtern und diffizile Fälle ignorieren. Diese offenkundigen Praxen seien bereits normativ geklärt und gerade deshalb nicht mehr moralisch kontrovers zu diskutieren, wie bspw. der Fall der Piraterie patentierter Software. Dennoch, so Brey, würde sich die Computer-Ethik häufig dieser Fälle annehmen (Brey, 2010). Somit wird hier Transparenz mit der Idee einer moralischen Klarheit gleichgesetzt, die moralisch eindeutige(re) Fälle diskutiert. Hier seien aber die moralischen Werte und Normen bereits einsichtig und bedürften keiner fundamentalen moralischen Klärung mehr. Brey betont jedoch, dass gerade viele computerunterstützte Praxen in ihrer moralischen Wertigkeit unklar und unbekannt seien. Diese Praxen sind daher normativ epistemisch opak, weil unzureichende Kenntnis über den Status ihrer Moralität herrsche. Dies begründet Brey damit, dass diese Praxis entweder nicht einmal von Spezialisten eingesehen werden kann, also in Merdes' Terminologie tatsächlich eine essentielle Opazität vorherrsche, oder fälschlicherweise als moralisch neutral bewertet würde. In Merdes' Terminologie wäre dieser Fall der akzidentiellen Opazität zuzuordnen (Brey, 2010; Merdes, 2020). Diese normativen Uneindeutigkeiten benötigen eine Analyse, die sachliche, kontextuelle und gebrauchsspezifische Intransparenz und Opazität aufdecken kann, um normative Tatsachen hervorzuheben. Eine Ethik für Softwareentwickler muss auf diesem aufklärenden, theoretischen Fundament aufbauend Handlungsorientierung geben können, um ein ethisch angemessenes Design von Software zu erreichen. Zugleich muss jedoch betont werden, dass Opazität und Transparenz erkenntnistheoretische Leitlinien sind und weder eine notwendig kluge oder richtige Entscheidung nach sich ziehen müssen, noch Handlungsgarantie bieten: Auf die Erkenntnis normativer Tatsachen kann nach wie vor eine unangemessene Handlung folgen, sei es aufgrund von Entscheidungs- oder Willensschwäche.

3.3 Die Methode der levels of abstraction

Eine fundierte Diskussion wird zudem nicht nur durch einen unterschiedlichen Gebrauch der Terminologie oder verschieden stark ausgeprägter Kommunikationsfähigkeiten erschwert, sondern auch aufgrund der verschiedenen Ebenen der Untersuchungsobjekte, die innerhalb einer normativen Ana-

lyse adressiert werden sollen. Diese kategorialen Unterschiede, die sich durch die Betrachtungsebenen begründen lassen, werden in der informationstechnologischen Literatur mit dem Konzept des *levels of abstraction* (LoA) diskutiert. In diesem Zusammenhang wird betont, dass gerade diese Misskommunikation auf einer unterschiedlichen Setzung von Systemgrenzen beruht (Baxter & Somerville, 2011). Luciano Floridi's *Methode der* LoA hat daher wesentlich einen epistemischen Charakter (Floridi, 2008). Sie soll eine vertikale und horizontale Taxonomie der Informationseinheiten ermöglichen. Die Schwierigkeit liegt darin, dass je nach LoA andere ethische Aspekte einer Erörterung bedürfen und somit auch praktisch anders thematisiert sowie technisch oder politisch gehandhabt werden müssen. Das je nach Entwicklungsstand oder gesellschaftlichem Bedarf passende LoA zu lokalisieren, ist wiederum aufgrund seiner transparenten oder opaken Natur kompliziert und heikel. Man kann Gesichtserkennungssysteme als sozio-technisches System diskutieren, das hinsichtlich des Gebrauchs durch Staaten strengere Governance und Regulierungslösungen sowie politische Kommunikationsprozesse benötigt als es in der Betrachtung von Gesichtserkennungssystemen als kontext-freies informationstechnisches Artefakt der Fall ist. Letzteres erfordert Überlegungen hinsichtlich technischer Zuverlässigkeit im Hinblick auf generische ethische Fragestellungen, wie bspw. die Garantie der Einhaltung von *general data protection guidelines*, weshalb zunächst verstärkt *compliance*-Richtlinien in Betracht gezogen werden. Obwohl sich die Perspektiven mit demselben Untersuchungsobjekt befassen und auch notwendigerweise auf Wissensüberschneidungen rekurrieren, sind der Problem- und Lösungsraum unterschiedlicher Natur. Tabelle 3 verweist somit schon auf eine Form der Orientierung von Problem- sowie Lösungsräumen, die jedoch in den Einzelkategorien wiederum verzweigt darstellbar ist.

Hierzu ist Luciano Floridi (2013) makroethischer Ansatz hilfreich, da er nicht nur einen bestimmten Aspekt einer der oben angeführten Kategorien ethisch artikulieren möchte, sondern eine umfassende Ethik entwickelt, die alle (Meso-) und (Mikro-)Ethiken einfängt (Teil I, Kapitel 6.1). Floridi stützt seine makroethische Vorgehensweise ontologisch ab. Die Interpretation des Seins als Informationseinheit teilen alle Gegenstände sowie Lebewesen als fundamentale kategoriale Bestimmung. Alles und jedes kann als Information verstanden sowie erklärt werden. Dies ermöglicht es Floridi (2008), einen *levelism* einzuführen, den er durch Abstraktionsgrade der zu betrachtenden Entitäten auffasst, die in vertikalen Relationen zueinanderstehen. Wieso Floridi eine ontologische Untersuchung und keine Betrachtung der Intentionalität einführt, erscheint nur dann nachvollziehbar, wenn jeder Informationseinheit aufgrund der Gleichwertigkeit dieselbe normative Aufmerksamkeit zukommen soll (Brenner, 2010). Floridi geht so weit, dass er aufgrund seiner ontologischen Betrachtungsebene jeder Entität des Universums aufgrund des Seins-Levels einen moralischen Wert zuschreibt, ungeachtet dessen, ob Mensch oder Maschine, Baum oder Tier (Floridi, 2013, S. 131 ff.).

Floridis LoA-Methode soll hier vor allem eine hinreichende Beschreibung von moralisch bedeutsamen Universen und deren kategorialer Gleichrangigkeit gewährleisten, um erkenntnistheoretisch normative Tatsachen identifizieren und lokalisieren zu können. Der von Floridi eingeführte informationstechnische Begriff der LoA bestimmt sich somit relational, d.h. LoA zeigen eine bestimmte Seinsdimension auf, allerdings ohne dadurch in einen Relativismus zu fallen. Somit ergänzen sich die Betrachtungen des Säugetiers auf dem nächsten LoA ‚Mensch und Hund‘, aber substituieren sich nicht, d.h. das LoA ‚Säugetier‘ kann aus der Sicht des Menschen beschrieben werden, ohne dabei rein humanzentrisch vorgehen zu müssen.

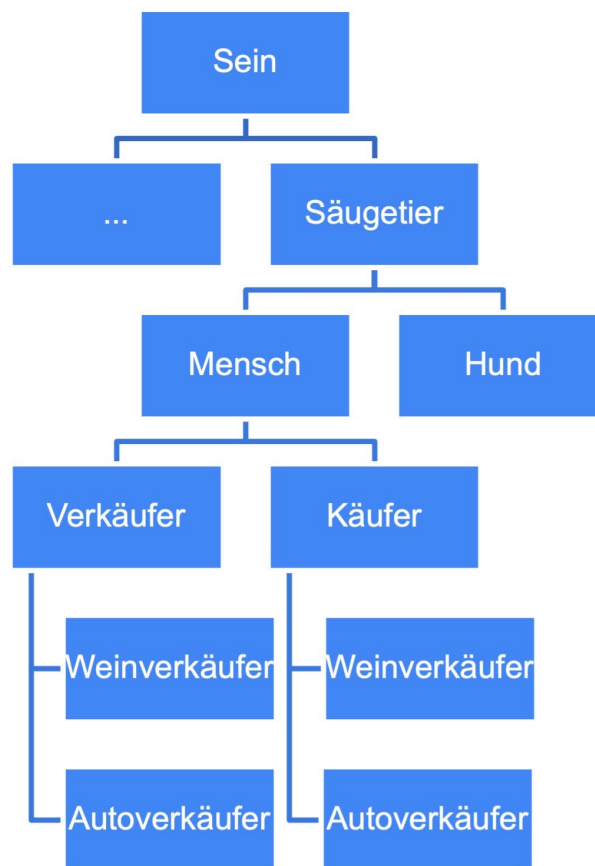


Abbildung 7 Levels of Abstraction

Zudem führt Floridi (2008) in seinem Artikel *The Methods of Levels of Abstraction* den Überbegriff des *gradients of abstraction* (GoA) ein, mit dem verschiedene LoA korrespondieren und der als Perspektive funktioniert, d.h. LoA werden in Abhängigkeit von GoA definiert (Floridi 2008, S. 311). Die unterschiedlichen GoA ergänzen sich, weshalb Floridi von disjunkten GoA im Unterschied zu *nested GoA* spricht. Ein GoA wäre demnach:

GoA: Weinkäufer und Weinverkäufer → Wein

GoA: Autokäufer und Autoverkäufer → Auto

GoA disjunkt: Qualität → Farbe und Geschmack

GoA nested: Qualität → Farbe und Wellenlänge

GoA hybrid: disjunkt und nested

Demnach ist ‚Qualität‘ ein GoA mit zwei LoA, Farbe und Geschmack, die disjunkt sind, weil sie sich nicht auf dasselbe *observable* – eine Wahrnehmungseinheit – beziehen und somit nicht aufeinander reduzierbar sind. Wenn ein GoA mit mindestens einem LoA wiederum selbst als LoA betrachtet wird, spricht Floridi von einem *nested GoA*, denn eine weitere Differenzierung führt zu einer präziseren, d.h. konkreteren, Bestimmung des vorangehenden LoA (Floridi, 2008). Wenn zum Beispiel über die Qualität des Weines gesprochen wird, muss nicht nur über den Geschmack rasoniert werden, sondern auch über die Farbe, die wiederum über die Wellenlänge des Lichts bestimmt werden kann. Allerdings kann eine ethische Analyse nicht die Wellenlänge mit der Blumigkeit des Geschmacks gemeinsam als Tatsachen bewerten, sondern die Farbe Rot und der blumige Geschmack sowie die Wellenlänge und das (Geschmacks-)Molekül wären jeweils die passenden *observables*, die ein LoA konstituierten. Ein LoA wird also darüber bestimmt, welches *observable* zugrunde liegt, d.h. welche Tatsachen auf derselben Ebene korrespondieren. Zudem kann ein *observable* diskret sein, was heißt, dass es eine endliche Anzahl an Werten gibt, die es annehmen kann. Ist dies nicht der Fall, handelt es sich um einen analogen Wert, d.h. das *observable* kann unendliche kontinuierliche Werte annehmen. Bei solchen LoA handelt es sich um eine theorieunterstützte Methode, um epistemische Klarheit hinsichtlich normativer Tatsachen zu gewinnen.

Als Beispiel soll hier zunächst ‚Wein‘ näher spezifiziert werden. Wein ist als Produkt ökonomisch bewertbar, kann aber dann weiter als kartesisches Produkt von Geschmack und Farbe näher beschrieben werden. Qualität ist sodann als *Type* (Farbe x Geschmack) beschreibbar, wonach „Rot“ demnach als *Token* des *Types* ‚Farbe‘ erkennbar wird. Qualität ließe sich demnach weder auf Farbe noch Geschmack hinreichend reduzieren, da beide vertikal zueinanderstehen sowie disjunkt sind, sich also auf unterschiedliche Seins-Atome beziehen.

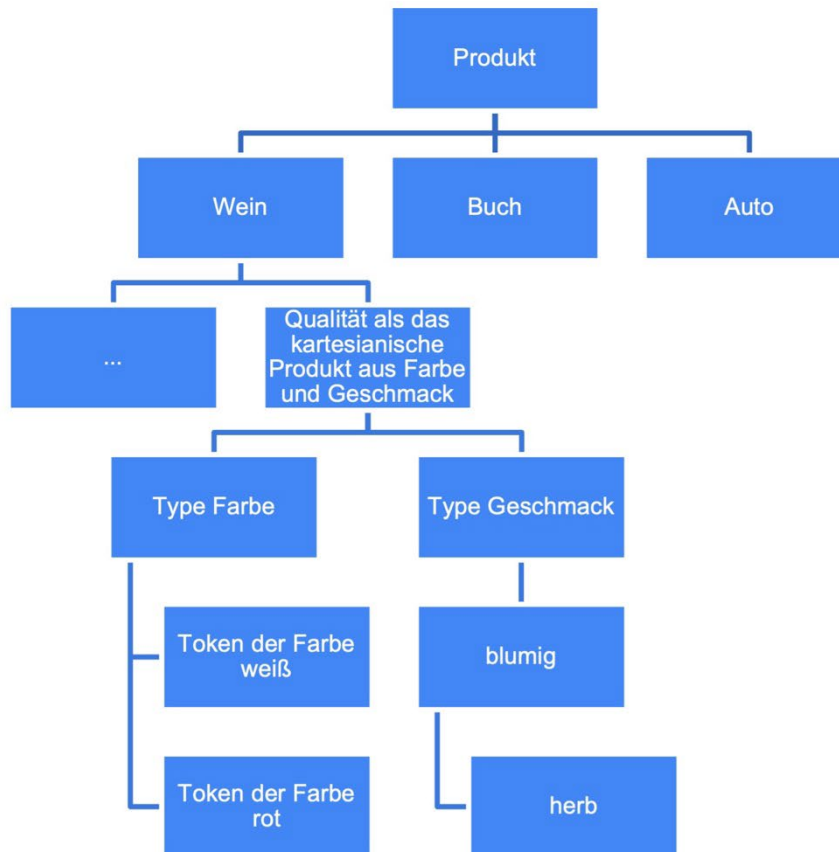


Abbildung 8 Levels of Abstraction und Gradient of Abstraction am Beispiel ‚Wein‘

Floridis Ansatz lässt sich dahingehend interpretieren, dass eine ethische Analyse von Information zuerst offenlegen muss, welche Perspektive (LoA) innerhalb welchen Universums (GoA) besprochen werden soll. Dies ermöglicht einen Erkenntnisgewinn, der als Grundlage einer gelungenen Artikulation und Evaluation dienen kann, wenn bspw. Autokäufer, Autoverkäufer und Produkt versehentlich analysiert würden anstelle von Autokäufer, Autoverkäufer und Auto. Die Abstraktion ‚Käufer, Verkäufer und Produkt‘ wäre wiederum als GoA denkbar und könnte konkretisiert in das LoA des Auto-GoA überführt werden. Die Zusammensetzung dieser GoA ist gerade deshalb schwierig, weil es sich um kein Subsumptionsverhältnis handelt, sondern um eine horizontale Strukturanalyse. Es bleibt unklar, weshalb Floridi 2013 in seinem Buch *The Ethics of Information* Abstand von den GoA nimmt, die ein relatives Universum beschreiben. Möglich ist, dass GoA und LoA ohnehin immer wieder wechseln und daher nur schwer zu bestimmen sind, und somit horizontale Erweiterungen mit vertikalen Abstraktionsniveaus unnötig vermischt würden. Ein Softwaresystem beschreibt Floridi (2013) demnach aus der Perspektive mehrerer LoA, die jeweils durch Modelle repräsentiert werden können. Diese Modelle basieren auf verschiedenen Variablen, die sich in *observables* zeigen. Hierfür bedarf es wiederum guter Theorien, die erläutern, welche Variablen ein GoA konstituieren und welche Merkmale

für dieses GoA konstitutiv sind. Dies ist der Grund, weshalb im ersten Teil dieser Arbeit drei philosophische Paradigmen eingeführt werden, die einer Lokalisierung geeigneter, d.h. in sich abgrenzbarer, LoA Orientierung bieten sollen.

Da Floridis ontologische Thematisierung keine Aussagen über die Normativität erlaubt, werden in den folgenden zwei Kapiteln die philosophisch-ethischen Grundlagen für die strukturelle Erfasstheit sowie deren deontischen Normen entwickelt, um diesen methodischen Ansatz auch als Handlungsorientierung einzusetzen zu können. Nida-Rümelins *Theorie einer Strukturellen Rationalität* (Teil I, Kapitel 2) ist aufgrund ihres zuvorderst deskriptiven Charakters hilfreich, um eine Lokalisierung und Identifikation normativer Tatsachen zu gewährleisten. Dies ist notwendig, um *observables*, also normative Tatsachen, zu klassifizieren und somit normative Tatsachen oder deren ungenaue relative Zuordnung zu identifizieren. Denn der Wert ‚Privatsphäre‘ korrespondiert je nach LoA mit einer unterschiedlichen Bewertungsgrundlage und bedarf folglich unterschiedlicher Handhabungen: So kann Privatsphäre in einer strukturellen Dimension als Schutzraum bestimmter individueller Verhaltensweisen oder in einer technik-generischen Dimension als Speicherortauswahl diskutiert werden. Die Ausprägungen des Werts ‚Privatsphäre‘ sind dementsprechend je nach LoA zu interpretieren und zu evaluieren. Nida-Rümelins *Theorie einer Strukturellen Rationalität* kann sodann auch als normative Theorie praktische Handlungsorientierung geben, wenn es um die rationale Begründung und Rechtfertigung handlungsleitender Gründe gehen soll. Im Anschluss daran wird der Frage nachgegangen, ob digitale Artefakte unabhängig von Nutzerhandlungen normativ zu bewerten sind. Dies wird im Rahmen der Einzigkeit und Besonderheit von Computertechnologie und -technik diskutiert werden, um die aus deren Natur resultierenden ethischen Fragestellungen erörtern zu können (Teil II, Kapitel 5, 6 und 7). Im Gegensatz zu den normativen Ethiken, die eine Handlung anleiten sollen, bestimmt und begrenzt die ontologische Topologisierung den ethisch zu bewertenden Gegenstandsbereich. Normative Ethiken hingegen sollen Antworten auf die Frage liefern, was über alle Felder und Bereiche hinweg getan werden soll, weshalb sie somit einen vollständig universalistischen Anspruch aufweisen und sich gerade nicht mit dem Gegenstand beschäftigen müssen (Henning, 2019).

Eine Technikethik beschränkt sich aber gerade schon deshalb, weil sie sich digitaler Artefakte und deren moralischer Handhabung annimmt und bleibt somit auf einen bestimmten, gesondert zu betrachtenden Bereich begrenzt (Teil I, Kapitel 6). Denn eine vertikale Ansammlung von LoA bezieht einzelne Ebenen ein, die es zu berücksichtigen gilt, wenn eine gründliche und sachliche normative Analyse durchgeführt werden soll, die sodann zu einer weiteren horizontalen und vertikalen Untergliederung führt. Durch die Zusammenführung aller LoA können Nutzer, Stakeholder und Entwickler jeweils in Abhängigkeit oder vermittelt durch die Informationstechnik gemeinsam gedacht werden und kann ihre strukturelle Einbindung Berücksichtigung finden. Zudem ist hierbei möglich, metaethische

Annahmen zu erörtern, um widerspruchsfreie Paradigmen oder Prämissen zu bestimmen sowie zu begründen.

Zur Veranschaulichung soll an dieser Stelle das Gebot der Diskriminierungsfreiheit von Gesichtserkennungssystemen diskutiert werden. Hier gilt es zum einen, den Gegenstandsbereich abzugrenzen, auf den diese Diskriminierung zurückzuführen ist: Darunter fallen vielfältige Erscheinungsformen, in denen und durch die sich Diskriminierung offenbaren kann. Die moralische Forderung der Gleichstellung aller Personen, gleichgültig ihres Geschlechts, ihrer Hautfarbe, ihrer Herkunft oder ihrer Religion, adressiert somit unterschiedliche Handlungskontexte, auf die diese moralische Forderung Anwendung finden muss. Um konkreten Handlungen normative Orientierung bieten zu können, müssen zunächst diese moralischen Bedeutungsstrukturen in ihrer Dimensionalität aufgezeigt werden können. Nur so wird klar, wer Adressat dieser moralischen Forderung sein kann und wie diese Person in ihrem Handlungskontext dieser Forderung nachkommen kann.

Erst die Lokalisation relevanter normativer Tatsachen in Abhängigkeit technischer und sozialer Phänomene ermöglicht eine grundlegende Bewertung technik-ethischer Anforderungen bezüglich wünschenswerter Lebensformen. Da jedoch, wie Vallor (2016) betont, die techno-soziale Opazität, aufgrund schneller technologischer, soziopolitischer und ökologischer Veränderungen, existentielle Risiken schwer kalkulierbar werden lässt, zeigt sich der dringende Bedarf einer erkenntnistheoretischen Unterstützung sowie einer praxis-geleiteten Umsetzungsmethodik. Letzteres verweist bereits auf die Notwendigkeit einer antizipatorischen, prozeduralen und diskursiven Technikethik. Die Methode der LoA bedarf einer soliden Theorie, die es ermöglicht, die jeweils relevanten Sachverhalte einer Ebene zuzuordnen. Diese wurde in Teil I der Arbeit ausführlich dargelegt. Hierfür ist jedoch auch notwendig, sich der Begriffe *Digitalisierung*, *digitale Transformation* sowie *digitale Artefakte* anzunehmen und diese voneinander abzugrenzen, um die jeweils notwendige normative Deliberation hinsichtlich des technischen Machbarkeitsraums offenzulegen.

3.4 Zur Normativität der Digitalisierung

Ursprünglich verstand man unter Digitalisierung die Umwandlung analoger, kontinuierlicher Informationen in diskrete Datenpakete: „Digitalisierung [ist] die binäre Repräsentation von Texten, Bildern, Tönen, Filmen sowie Eigenschaften physischer Objekte in Form von aneinandergereihten Sequenzen aus „1“ und „0“, die von heutigen Computern mit extrem hoher Geschwindigkeit – Milliarden von Befehlen pro Sekunde – verarbeitet werden können.“ (Neugebauer, 2018a, S. 9) Der Digitalisierungsprozess, der bereits seit den 1970er Jahren im privatwirtschaftlichen und militärischen Sektor begann, ersetzte dort routinemäßige Arbeitsprozesse, die sich aufgrund ihrer wiederholenden Ablaufstruktur

durch simple algorithmische Steuerung einfach substituieren oder verstärken ließen (Hess, 2019a). Durch den Einsatz digitaler Technik wurde das Erstellen und Bearbeiten von Informationen z.B. durch Texterkennungsprogramme oder Datenbanken enorm erleichtert. Aber vor allem die Nachbildung von komplexen Systemen führte zu einer effizienten und effektiven Organisation von Arbeitsabläufen, wie es bspw. integrative Planungsprogramme im Unternehmenskontext sehr schön veranschaulichen. So erhofft man sich durch die Digitalisierung eine erfolgreiche Planbarkeit der Produktionsprozesse, weil Ressourcen aufgrund der kontinuierlich zur Verfügung stehenden Informationen und einer genauen computer-gestützten Berechnung zielgerichtet eingesetzt und eingespart werden können (Finger, 2012). Die weitere Ausgestaltung der PPS-Systeme, die dann ab den 1990er als ERP-Systeme (*efficient resource planing*) Einzug erhielten, integrierte weitere betriebswirtschaftliche Komponenten zusätzlich zur Materialbeschaffung, wie bspw. die Personal- und Finanzplanung. Ebenso wurde die Entwicklung durch die immer höheren Speicher- und Rechenkapazitäten angetrieben. Das Moor'sche Gesetz postuliert gar eine Verdopplung der Rechenleistung alle zwei Jahre (Mattern, 2005). Dies versetzt Unternehmen in die Lage, in Echtzeit ihre (globalen) Lager zu inventarisieren und diese Informationen auch zeitgleich in ihrem Finanzmanagement abzubilden.

Normative Aspekte der Automatisierung von Prozessen schlugen sich bspw. in der Diskussion um die Substitution von Arbeitsplätzen nieder, seien dies nun die Tätigkeiten der Berufsfahrer oder Fachkräfte in Steuer- und Rechtsanwaltskanzleien. Die Sorge vor steigender Arbeitslosigkeit, die im Zuge der Industrialisierung durch den verstärkten Einsatz von Maschinen in der Industrie bereits apokalyptische Interpretationen hervorrief (Fritz Langs *Metropolis* von 1927 und Charlie Chaplins *Modern Times* von 1937), sowie die Verelendung der Arbeiter durch einfache und monotone Arbeit an Maschinen und Fließbändern führten in eine marxistisch geprägte Technik-Feindlichkeit, die in einem Klassenkampf Mensch gegen Maschine enden sollte (Karafyllis, 2019). Ebenso war die Frage zentral, ob alles, was technisch machbar ist, auch wirklich ausgeführt werden sollte. Die Formulierung eines *can-implies-ought* wurde in der Technikphilosophie bis in die 1970er Jahre verstärkt diskutiert (Lenk & Ropohl, 1993), gerade auch wegen des Zweiten Weltkriegs, in dem es zu einem vermehrten Einsatz von Technik und Forschung gekommen war. Die schiere Grenzenlosigkeit technischer Anwendungen und die damit einhergehenden Risiken, wie sie die Kernenergie oder Wasserstoffbomben heraufbeschworen, verschärften den Dialog um die moralische Angemessenheit von Technik und Technologie. Die vielfach gravierenden negativen Folgen, die sich in der Umweltbelastung oder auch in der Gesundheitsgefährdung der Menschen zeigten, machten es notwendig, über eine Ethik bzw. eine Evaluationsmethodik zu reflektieren, um v.a. im Zuge der Demokratisierung und des Wertpluralismus technische Aufklärung und gemeinsame Entscheidungsfindungsprozesse für einen aufgeklärten Technikeinsatz zu erörtern. Die ethische Frage nach den guten Gründen, die für eine Handlung aus

der Menge aller möglichen Handlungen sprechen sollen, wird so auch für technische bzw. technologische Handlungen von der Forschung bis hin zur Entwicklung zentral. Die Disruptivität der Digitalisierung hält sich jedoch in Grenzen, wenn darunter die bahnbrechenden Innovationen verstanden werden, die sich in einer Verdrängung tradierter Kulturen manifestieren. Denn rein durch die Substitution analoger durch digitale Verfahren bzw. eine Umwandlung analoger Kulturgüter in digitale, die zu keinem Mehr an Wissen oder neuen Handlungsräumen führen, verbleiben Handlungsstrukturen zunächst unangetastet. Digitalisierung eröffnet somit erst einmal keine Fragen bezüglich inhärenter, normativer Bedingungen von Handlungszusammenhängen, sondern wirkt sozusagen von außen auf die Handlungsstruktur, indem durch den Einsatz von Maschinen Arbeitsplätze abgeschafft werden oder die Telefonie nicht mehr mittels des Standtelefons, sondern mittels eines Handys erfolgt. Die Erwünschtheit dieser neuen, digitalen Technologie und deren Einsatz wird in politischen Debatten öffentlich diskutiert.

Die zunehmende Vernetzung von Objekten und deren Repräsentation in einer virtuellen Struktur, für die Ashton 1999 das Konzept des „internet of things“ prägte (Ashton, 2009, S. 1), sowie die zu jederzeit und an jedem Ort möglich gewordene Verbundenheit mit dem Internet lässt die in der Digitalisierung innewohnende, umwälzende Kraft erahnen. ‚Ubiquitous computing‘ oder ‚ambient intelligence‘ bezeichnen denselben Sachverhalt: Eine immerwährende Verbindung mit dem Internet und eine damit einhergehende Auflösung separater, analoger und virtueller Räume. Heutzutage umgreift die Digitalisierung nun nicht mehr nur die Arbeitswelt und substituiert oder unterstützt einfache Prozesse, sondern aufgrund der Entwicklung von selbstlernenden Systemen, sogenannter Künstlicher Intelligenz¹⁶, wurde auch die Automatisierung weniger strukturierter Prozesse möglich. Unsystematische Abläufe finden sich vor allem im alltäglichen Leben der Menschen: in der Art der Kommunikation, bei der Suche nach den schnellsten Fahrzeiten oder bei der Bewertung bestimmter Ware. Diese Abläufe

¹⁶ Künstliche Intelligenz ist jener Teilbereich der Informatik, der Softwaresysteme zu entwickeln sucht, die gerade wenig strukturierte Prozesse zu automatisieren vermögen. Damit weniger strukturierte Prozesse automatisierbar, d.h. mathematisch abbildbar sind, kommen sogenannte intelligente Systeme zum Einsatz, die sich in der „Unordnung der Daten selber zurechtfinden müssen“, um zu einem zufriedenstellenden Resultat zu gelangen. Damit dies gelingen kann, benötigen künstlich intelligente Systeme große Mengen an Daten, anhand derer das System „trainiert“ wird, damit das System „erlernen kann“, wie es die Datenmengen nach einem gewünschten Output „zu sortieren vermag“. Die Klassifikation erfolgt mittels Algorithmen, wobei ein Algorithmus ein Rechenvorgang ist, der ein bestimmtes Problem löst. Bspw. löst der Google-Algorithmus das Problem des Webite-rankings, also welche Website an welcher Stelle angezeigt werden soll. Dieser Rechenvorgang wiederholt ein bestimmtes Schema, wobei alle Einzelschritte bereits vorgegeben sind. Allerdings bleibt unklar, nach welchem Muster die Daten wie gewünscht angeordnet werden. Dieses Vorgehen wird als Maschinelles Lernen bezeichnet, d.h. die Systeme erscheinen „intelligent“, denn sie „lernen“ die Datenmengen zu ordnen, ohne dass ihnen die wesentlichen Merkmale, wonach der gewünschte Output klassifiziert werden soll, vorgegeben werden. Diese Kriterien „erlernt“ das System von selbst. Die normativen Probleme, die auf diesen Vorgängen beruhen, werden in der Ethik der KI unter den Schlagwörtern *algorithm bias*, *big data* und *black box* diskutiert (Arnold & Scheutz, 2018; Floridi et al., 2018; Morley, Floridi, Kinsey & Elhalal, 2020; Winfield & Jirotko, 2017).

sind von nun an individualisiert mittels Software beschreibbar und ausführbar. Dies verdeutlicht, weshalb heute schon weit mehr unter dem Begriff *Digitalisierung* verstanden wird als die reine Übertragung analoger in digitale Daten. Es werden vielmehr die weitreichenden Veränderungen der gesamten Lebenswelt adressiert, die aufgrund der digitalen Technologisierung aller Lebensbereiche stattfinden.

3.5 Zur Normativität digitaler Transformation

Der breite Einsatz digitaler Anwendungen in allen Bereichen des menschlichen Lebens wurde jedoch erst durch leistungsfähigere Hardware ermöglicht. Das Smartphone hat seit seiner Einführung vor gut zehn Jahren die Kommunikation in Echtzeit rund um den Globus ermöglicht und somit das Verhalten und die Lebenswelt beeinflusst. Die immer vorhandene Möglichkeit, überall und zu jeder Zeit mit dem Internet verbunden zu sein, verändert nicht nur „die Art, wie wir leben“, sondern sie verändert den Menschen, weshalb Markowetz (2015, S. 20) den Menschen als „homo digitalis“ bezeichnete. Das Smartphone, der kleine tragbare Computer, ermöglicht immerzu Fahrplanweisungen abzurufen, Waren zu kaufen, Restaurantbewertungen einzusehen oder einfach jederzeit auf vielfältige Weise kommunizieren zu können. In vielen dieser epistemischen Tätigkeiten unterstützt das Smartphone anstelle unhandlicher Straßenkarten, voller Einkaufszentren oder lästiger Erkundigungen. Der Einsatz digitaler Anwendungen hat das Verhalten der Menschen maßgeblich verändert, denn durch die ständige Vernetzung mit anderen und mit der Welt nimmt der *homo digitalis* vermittelt durch digitale Instrumente an der Welt teil. Der *homo digitalis* bestimmt seine Beziehungsgeflechte mit Hilfe digitaler Technik, richtet sein Verhalten nach neuartigen Anforderungserwartungen aus und verändert zudem Erfahrungs- und Wahrnehmungsräume. Der *homo digitalis* verändert sich nicht nur individuell, sondern verändert auch intersubjektive und kollektive Beziehungsmuster.

Die digitale Partizipation re-strukturiert und steuert Sinnzusammenhänge und transformiert dadurch normative Ausrichtungen, denn von nun an bestimmt der Imperativ der Optimierung aller Lebensbereiche. Analog zum Unternehmenskontext wurde die Alltagswelt digital transformiert, da auch hier ein Anpassungsdruck auf lebensweltliche Verhaltensweisen ausgeübt wurde, der sich nicht mehr nur auf Marktbedingungen oder Wertschöpfungsprozesse beschränkt, sondern bestimmte Umgangsformen dahingehend erfordert, wie mit diesen Folgen umzugehen ist. Dies zeigt sich im Unternehmenszusammenhang bspw. in der Etablierung neuer Managementstrukturen (Hess, 2019b). Für die gesamte Lebenswelt bedeutet dies, dass die Digitalisierung, der Prozess, eine Tätigkeit zu digitalisieren, zugleich den Kontext transformiert, in den sich diese Tätigkeit einfügen lässt, sprich andere Formen des Verhaltens erfordert, gleichgültig ob in der Arbeit oder in privater Kommunikation: Nicht nur die Ab-

hängigkeit von ständigen *rewards* und ständiger Ablenkung durch die „Unendlichkeit in meiner Tasche“ (Markowetz, 2015) verändert das Verhalten der Menschen, sondern auch die Kommunikationsarten verändern ihre normativen Bedingungen, z.B. wann und wie schnell in welcher Ausführlichkeit eine Nachricht beantwortet werden sollte, welcher Tonfall passend, welches Profil angemessen ist. Diese Umdeutungen verschieben die Wertigkeit der Beziehungsgeflechte, weshalb der Vorgang gerne als disruptiv bezeichnet wird.

Es ist somit klar, dass nicht jede digitale Abbildung bereits bestehender Prozesse als digitaler Wandel bezeichnet werden kann. Aber das Phänomen des digitalen Wandels ist folgenreich für jeden gesellschaftlichen Bereich: Es ist ein kultureller Wandel (Mohr, 2020). Das bedeutet nicht, dass Digitalisierung im Sinne der Prozessautomatisierung gering zu schätzen ist. Vorhandene Prozesse zu digitalisieren, verlangt jedoch zunächst kein wesentliches Umdenken, kein Zurechtfinden in einer immer stärker vernetzten Welt, die auch normative Konstituenzen verschiebt und in Frage stellt. Während Digitalisierung bspw. zwar den Alltag der beteiligten Mitarbeiter verändert und in der Regel Ressourcen einspart, sind für eine ethische Analyse die technischen Systeme mit ihren (nicht) vorhersehbaren Folgen und Risiken relevant, weisen jedoch noch keine beständige, gegenseitige Interaktivität auf. In den letzten Jahren rückt jedoch die Schnittstelle Mensch-Maschinen-Interaktion aufgrund der dialektischen Struktur der Informationstechnik immer mehr in den Fokus: Gerade die Veränderlichkeit digitaler Artefakte durch die Nutzer sowie die strukturellen Verhaltensverschiebungen des Nutzers durch diese Technik in ihren epistemischen, rationalen und normativen Dimensionen verweisen auf die Einzigartigkeit dieser Informationstechnik.

Auch wenn die Aktualität der Forderung im Zuge der Künstlichen Intelligenz verstärkt wurde, ist doch das gesamte Feld der Informatik zu adressieren, wenn von digitaler Datenverarbeitung gesprochen wird, denn jede (digitale) Technik trägt bestimmte normative Anpassungsanforderungen an ihre Nutzer und ihre Umwelt heran und (re-)strukturiert somit Handlungen und Handlungszusammenhänge (Grunwald, 2018a). Es geht gerade bei der Digitalisierung aber nicht nur um Technik, sondern „um die Suche nach digital realisierbaren Möglichkeiten, Geschäftsprozesse und Geschäftsmodelle zu optimieren oder neu zu definieren“ (Mohr, 2020, S. 16). Dadurch werden Kulturwandel und eine Veränderung von Sinnzusammenhängen sowie deren normative Ausrichtung ermöglicht. Diese Umwälzung tradierter und etablierter Zusammenhänge kann auch als Grund angeführt werden, weshalb digitalen Technologien ein disruptiver Charakter zugeschrieben wird. Sie verändern oder modifizieren Kultur, also vornehmlich „jene Prozesse [...], in denen soziale Bedeutung, also die normative Dimension der Existenz, durch singuläre und kollektive Handlungen explizit oder implizit verhandelt wird“ (Stalder, 2016, S. 16). Technologien erweitern oder durchbrechen wechselseitige Abhängigkeiten und formen neue (auch rationale) Bedeutungsketten, die wiederum das Handeln der Personen strukturieren. Die strukturellen Veränderungen, die durch die Umwandlung hervorgerufen werden, werden

als digitale Transformation bezeichnet: „Sie [die digitale Transformation N.Z.] schafft neue, schnelle Möglichkeiten der wirtschaftlichen Verwertung und der Selbstverwirklichung, aber auch Gefahren sozialer Art, mit denen umzugehen wir in vielen Fällen noch lernen müssen.“ (Neugebauer, 2018, S. 4) Folglich ist eher die Rede von digitaler Transformation, wenn in den letzten Jahren in der Politik, in der Forschung und in der Presse von Digitalisierung gesprochen wurde. Diese digitale Transformation bezeichnet zusätzlich zur Digitalisierung einen Wandel der Kultur, einen Wandel normativer Ansprüche an die gegenseitigen Beziehungen und führt aufgrund der Permanenz und Ubiquität digitaler Technik zu einer Anpassungsforderung an die Nutzer. Die entscheidende Frage lautet also nicht, wie Mohr (2020, S. 9) schön im Sinne eines Digitalen Humanismus formuliert, „Wie kann ich digitalisieren?“, sondern „Was muss ich tun, um handlungsfähiger zu werden?“

3.5.1 Zur Normativität digitaler Technik und Technologie

Die Stimmen, die einen humanen Einsatz von Informationstechnologien fordern, kommen sowohl aus der Zivilgesellschaft, aus den Reihen der Konsumenten, den Medien, von Regierungen und supranationalen Institutionen als auch aus der Informatik-Gemeinschaft selbst. Dabei setzen sie freilich unterschiedliche Nuancen in ihren Forderungen nach einem verantwortungsvollen, menschlichen und vernünftigen Umgang mit Softwareprogrammen (Floridi & Cowls, 2019; Morley, Floridi, Kinsey & Elhalal, 2020). Meist sprechen sie Werte an, deren Einhaltung wünschenswert erscheint, in dem Sinne, dass dem Gemeinwohl, der Gemeinschaft, kein Schaden zugefügt werden dürfe, sondern „Produkte, Systeme und Infrastrukturen so zu realisieren [seien NZ], dass sie immer und ausnahmslos im Interesse der Menschen funktionieren und agieren, [und dies NZ] zu einem zentralen Ziel der technischen Entwicklung [würde NZ]“ (Neugebauer, 2018b). Bereits heute unterstützen Softwareapplikationen den Menschen in seinen unterschiedlichen Rollen, indem sie ein Zurechtfinden in einer immer komplexer werdenden Lebenswelt unter knappen Ressourcen wie Wissen oder Zeit erleichtern. So ermöglicht Software eine automatisierte Ordnung digitaler Fotoalben oder die Entsperrung von Geräten ohne lästige PINs, sondern per Fingerabdruck oder Gesichtserkennung. Die Möglichkeiten der *e-democracy* oder *civic technology* – also jener digitalen Technik, die es ermöglicht, sich als Bürger direkt und unkompliziert an staatliche Institutionen zu wenden oder durch eine aktive Partizipation an einer digitalen Zivilkultur im Sinne von Wikis oder Informationsplattformen teilzunehmen – können Demokratien bzw. kooperative Lebensräume stärken (Pfaff & Hasan, 2011). Digitale Arbeitsprozesse, wie komplexe ERP-Systeme, erschaffen effizientere Arbeitsorganisationen, weil sie die unterschiedlichen betriebswirtschaftlichen Bereiche sogar in Echtzeit vernetzt darzustellen vermögen. Der Einsatz dieser Technik in allen Lebensbereichen verändert die Beziehungen menschlicher Kooperation, z.B. in der Zusammenarbeit von Arbeitnehmer und Arbeitgeber, aber auch menschliche Vertrauensbeziehungen, wie z.B. familiäre oder freundschaftliche Kommunikationskulturen.

Diese Arbeitserleichterungen haben zusätzlich zu ihrer strukturell transformativen Kraft auch ihren Preis, denn durch die Nutzung digitaler Applikationen hinterlassen Nutzer Datenspuren und decken dadurch wesentliche Merkmale ihrer Identität(en) auf, die das Individuum in seiner Intimität einer Öffentlichkeit ausliefern. Informationen in Form von Bildern, Beiträgen bis hin zur Auswertung von Meta-Daten bleiben zudem auch noch nach Jahren in virtuellen Welten weiter existent. Das Recht auf informationelle Selbstbestimmung, das sich auf Artikel 8 der Europäischen Menschenrechtskonvention stützt, in dem das Recht auf Privat- und Familienleben festgestellt wird, „sichert dem Einzelnen die Befugnis, grundsätzlich „selbst über die Preisgabe und Verwendung seiner persönlichen Daten zu bestimmen“ und „zu entscheiden, wann und innerhalb welcher Grenzen persönliche Lebenssachverhalte offenbart werden“ (Franzius, 2015, S. 259). Eine Umsetzung dieses Rechts benötigt jedoch zum einen das Wissen auf Seiten der Nutzer, welche Daten in welchem Umfang verwendbar werden, man denke hier nur an die Kombinationsmöglichkeiten von Meta-Daten und die in jener Kombinatorik verborgene Aussagekraft. Zum anderen führen einseitig gesammelte Daten zu Schwierigkeiten all jener Technik, die auf Big Data beruht, denn Verzerrungen aufgrund nicht repräsentativer Daten kann man dann nur schwer begegnen, d.h. ein Datensatz umfasst dann evt. nur Personendaten einer bestimmten Gruppe, die bereitwillig oder unwissentlich ihre Daten zur Verfügung stellte (Žliobaitė & Custers, 2016).

Die Verewigung im Netz führte auch zur Einführung eines *Rechts auf Vergessenwerden* (Vollmer, 2021). Dieses Recht ist nicht nur juristisch schwierig umsetzbar, sondern auch die technischen Grundlagen eines Einsatzes müssen auf das Internet skalierbar sein. Zwar kann man mehr oder weniger problemlos mit Hilfe von Suchmaschinen ungewollte Einträge detektieren oder aber gleich aufgrund eines vom Nutzer eingetragenen Verfallsdatums gespeicherte Daten löschen (z.B. das 2011 hochgepriesene X-pire!). Allerdings ist eine automatisierte Detektion von Informationen aufgrund des semantischen Inhalts für Software nicht kategorisierbar und der Bedarf an Personen zur Lektorierung nach wie vor gegeben: Ein zielführendes *Labelling* irrelevanter und relevanter Information muss somit vom Menschen vorgenommen werden und ist daher zeit- und kostspielig. Ein Verfallsdatum wiederum selbst zu bestimmen, löst das Problem der von fremden Personen hochgeladenen und archivierten Beiträge nicht. Zudem weist der Wikipedia-Gründer Jimmy Wales das Vergessenwerden-Können-Sollen als unmoralisch aus, da es gegen das Recht auf Wissen und Informationsbeschaffung sowie gegen das Presserecht der Meinungsfreiheit verstoße (Lomas, 2014). Besonders aber betont er, dass dieses Recht zu Wissenslücken führen könne, da Einträge gelöscht würden, die eventuell doch von öffentlichem Interesse seien. Dieser Rechtsgüterstreit und seine angemessene Abwägung sind dann besonders schwierig, wenn es keine wünschenswerten und verlässlichen Verhaltensweisen gibt, die als etablierte Ethos-Normen Orientierung bieten können.

Die Vielzahl der skizzierten unterschiedlichen Szenarien, in denen digitale Artefakte zur Anwendung kommen, verdeutlicht die normative Ohnmacht, die Vallor (2016) so treffend analysiert hat. Der Einsatz digitaler Artefakte in allen Bereichen des täglichen Lebens führt zur kritischen Betrachtung der Bereitstellung von Infrastruktur durch die Tech-Industrie und wirft dabei auch datenökonomische Fragestellungen sowie unterschiedliche Möglichkeiten der Etablierung neuer Arbeitsnormen auf: Denn nicht mehr nur monotone Arbeitsabläufe werden maschinell ersetzt, damit der Mensch sich qualitativ höherwertiger Arbeit hingeben kann, sondern, wie Agar (2019, Introduction und Kapitel 2) betont, die Übernahme von *mind work* wird beschleunigt werden (Brynjolfsson & McAfee, 2017, Teil 1, Kapitel 2; McAfee & Brynjolfsson, 2017). Datenökonomische Probleme betreffen den Umgang mit der Non-Exklusivität von Wissen oder Information, d.h. gehen der Frage nach, ob ein freier Umgang mit Daten, im Sinne der Datensammlung, -auswertung und -preisgabe, gerechtfertigt sei und welche Auswirkungen ein freier Umgang auf die Wirtschaft sowie auf gesellschaftliche Systeme haben kann. Im Zentrum steht dabei die Ausgestaltung einer (digitalen) Wirtschaftsform (Agar, 2019; McAfee & Brynjolfsson, 2017), die sich um die ordnungspolitischen und wirtschaftsethischen Fragestellungen rankt, die sich aus den Slogans „Daten sind das neue Öl“ und „Algorithmen sind die neuen Raffinerien“ ergeben (Agar, 2019, S. 68 ff.). Aber nicht nur die Wirtschaftsethik oder der ordnungspolitische Rahmen wird durch die Mächtigkeit der Daten herausgefordert. Auch das Design der digitalen Artefakte – wie bspw. durch die notwendige Integration einer informierten Einwilligung (*informed consent*) oder die Berücksichtigung der Bedienerfreundlichkeit der Software, die als *usability* bezeichnet wird und später um die *user experience* erweitert wurde – verdeutlicht den Bedarf einer Kohärenz vieler auf den ersten Blick unabhängig erscheinender Tätigkeitsfelder (Hassan & Galal-Edeen, 2017). Das Bedürfnis nach Integrität der unterschiedlichen Handlungsdomänen ist damit zu erklären, da es gerade darum geht: „to create a new social norm of information sharing“ (Agar, 2019, S. 68). An diesem sozialen Informationsaustausch sind alle direkt oder indirekt Beteiligten in ihren verschiedenen Rollen in Bezug auf das digitale Artefakt zu berücksichtigen. Folglich wird durch den deutlich vermehrten Einsatz digitaler Technologie und Technik die menschliche Sozialität, d.h. menschliche Beziehungsgeflechte, neu strukturiert und neue Umgangsformen werden etabliert. Diese Umwälzung normativer Strukturmerkmale muss als Transformation tradierter Lebenswelten verstanden werden. Noch eindrücklicher wird diese disruptive Natur der Informationstechnik, wenn das menschliche Urteilsvermögen zunehmend durch Software ersetzt bzw. unterstützt werden soll. Genau dann, wenn Informationstechnik und -technologie den Kern des Humanen treffen, nämlich die Sozialität, die Emotionalität sowie die Vernunftfähigkeit, und somit die autonome Entscheidungsfähigkeit und Handlungsbefugnis berühren, drängen sich normative Fragestellungen auf und klare normative Haltungen werden unumgänglich. Es muss daher betont werden, dass Software vor allem dann in der Kritik steht, insofern Entschei-

dungen an Maschinen delegiert werden, welche die Emotionalität, Sozialität oder die kollektive Gemeinschaft tangieren. Dementsprechend fordert Agar (2019), dass Bereiche zu schützen sind, die vor allem für ein kooperatives Miteinander, menschliche Anteilnahme und gegenseitiges Verständnis konstitutiv sind (*social mind work*). Nicht selten führen Softwaresysteme nicht nur zu einer gesteigerten Flexibilität, Mobilität oder gar zu einer verstärkten demokratischen Partizipation, sondern häufig verschärfen sie persönliche Verletzlichkeiten oder fördern Einschränkungen und Misstrauen in demokratische Strukturen (Coeckelbergh, 2013; Eubanks, 2018).

Floridi (2013) hat diejenigen strukturellen Lebensbeziehungen, die sich durch Informationen beschreiben und erklären lassen, als *Infosphere* bezeichnet. Dieser *terminus technicus* umfasst daher mehr als nur den Cyberspace und beinhaltet auch analoge, nicht virtuelle Informationsdomänen, selbst wenn sich analoge Lebensräume immer mehr digitalisieren lassen. Dieses informationsgestützte Habitat wird durch den Einsatz von Informationstechnik und -technologie stark transformiert, aber auch erweitert, denn immer mehr Lebensbereiche werden durch diese Technik in eine Informationsform gebracht und somit Teil der *Infosphere*. In diesem Zusammenhang spricht Floridi von *re-ontologizing*, worunter er die kraftvolle Transformation von Lebensformen bzw. Lebensräumen versteht, die durch die Anwendung digitaler Technik und Technologien verursacht werden. Floridi (2010, S. 19) versteht diese Verschiebung als eine, die „fundamentally transforms its [infosphere N.Z.] intrinsic nature, that is, its ontology or essence“. Die Seinsstruktur der Lebensbereiche wird durch die Formalisierbarkeit der Information und die Einführung digitaler Technik verändert, denn Informationen sind jene Entitäten, die sich perfekt und ohne großen Aufwand kopieren lassen (Wallach & Allen, 2009): Dies führt zu einem Vervielfältigungspotenzial, das zunächst den Zugang zu Wissen und Informationen erweitert und an der Exklusivität von Eliten rüttelt. Zudem sind die Nutzungsrechte in Bezug auf Informationen genauso zentral wie die Eigentumsrechte, sodass nicht nur wesentlich für Informationen ist, wem sie gehören, sondern ebenso zu berücksichtigen gilt, wer sie in welchem Sinne nutzen darf oder gar muss, damit sie eine Wertigkeit oder Bedeutung erhalten. Aus den Eigentumsrechten lassen sich nicht notwendigerweise Nutzungsrechte folgern: Eine Onlineshoppingplattform mag ein Eigentum über die Daten ihrer Kunden haben, insofern sie diese schützen können muss, aber sie darf diese nicht nach Belieben weiterverwerten, verändern oder die Nutzer aus dem Gebrauch ihrer Daten ausschließen. Allerdings führen Eigentumsrechte ohne Nutzungsrechte zur Handlungsunfähigkeit: Ein Haus, das jemandem gehört, der es jedoch weder verändern, bewirtschaften noch bewohnen darf, ist nicht mehr ohne Weiteres als Eigentum zu bezeichnen.

Zentral für eine ethische Auseinandersetzung jenseits konsequentialistischer Zustandsbewertungen ist Floridis ontologischer Beitrag: Das Sein der *Infosphere* adressiert nicht mehr die Beharrlichkeit des zu betrachtenden Gegenstandes oder die Wahrnehmungsfähigkeit des Menschen, sondern betrifft

die Interaktivität, insofern sich die Information durch alle Beteiligten und die Beteiligten durch die Information beeinflussen oder verändern lassen (Floridi, 2010). Das Sein der Information, das sich gerade durch seine Veränderlichkeit auszeichnet, wird durch einen gegenseitigen Austausch und die jeweils notwendige Rückinterpretation der in Form gebrachten Informationen bedingt, wobei damit nicht gemeint sein soll, dass die Bedeutung der Information willkürlich und konstruktivistisch gedacht werden darf, sondern vielmehr, dass die Bedeutung der Information niemals unabhängig von sozialem Austausch begriffen werden kann. Diese Interaktivität, die zwischen Information, Absender und Adressat vorherrscht, folglich zwischen der Intention des Absenders und der Reaktion des Adressaten, verändert das Sein der Information. Information ist somit konstitutiv für Interaktionen, aber zugleich auch abhängig von der Intentionalität des Absenders und der als angemessen empfundenen Reaktion des Empfängers. Information zeigt sich in der Äußerung einer Absicht des Absenders, bspw. in der Frage nach dem Weg, hin zur Beantwortung der Frage durch den Empfänger in Form einer Geste, die in eine bestimmte Richtung verweist. Diese Bedeutungstransmission ist somit nur in ihrer Sozialität greifbar und sollte der geteilte, normative Bedeutungshorizont fehlen, bildet sich dies in der Schwierigkeit ab, wünschenswerte strukturelle Bedeutungsabsichten und -ausrichtungen auszumachen. Dadurch werden Situationen schwer bewertbar und erwecken den Anschein einer ihnen innewohnenden Relativität. Wesentlich ist hierbei, dass der Umgang mit Informationen das Habitat normativ umstrukturieren kann. Eben dies wird durch den vermehrten Einsatz digitaler Artefakte hervorgerufen. Folglich handelt es sich um eine dialektische Interaktivität, in der sich das jeweils andere Bedingende zugleich durch diese Interaktion verändern kann.

Bevor sich der Frage genähert werden kann, wie man mit dieser transformativen *re-ontologizing* Kraft umgehen kann und vor allem, welche Bedeutung diese Erkenntnis für die Ausgestaltung digitaler Artefakte hat, ist ausschlaggebend, den Untersuchungsgegenstand und den Adressaten definitorisch zu umreißen. Von welchen Systemen, von welchen digitalen Artefakten oder Produkten ist im Zusammenhang mit einer Ethik für Softwareentwickler die Rede? Und welche Handlungen führen Softwareentwickler aus, für die sie sich entscheiden und für die sie somit auch moralische Verantwortung übernehmen können müssen?

Erst eine Begriffssortierung ermöglicht, diesen ethischen Fragen anzugehen, die sich mit den technischen Handlungsmöglichkeiten und Begründungen von Handlungsregeln auch im Hinblick auf soziotechnisches Handeln ergeben (Lenk & Ropohl, 1987). Dadurch kann auch der spannenden Frage nachgegangen werden, ob Informationstechnik und -technologie eine spezifische Technikethik benötigen. Grundlegend ist hierfür, ob durch oder in Informationstechnologien und -techniken neue ethische Konzepte entstehen (Maner, 1999) oder ob klassische ethische Werte in einem neuen Gewand erscheinen (Johnson, 1985). Diese Diskussion ist auch als Maner-Johnson oder *uniqueness*-Debatte in die Geschichte der Computereethik eingegangen (Bynum, 2001).

3.5.2 Moral Machines

Digitale Technologie und Technik begründen dann eine genuin spezifische Ethik, wenn Maschinen selbstständig moralisch sein können, d.h. wenn sie selbst moralisch angemessene Resultate erzielen können. Ihre Einzigartigkeit und Spezifikation wäre darauf fußend, dass Maschinen selbst moralische Akteure sein würden (Bendel, 2019) und nicht nur menschliche Moral bzw. Normativität integrierten. Dieser Ansatz firmiert unter dem Namen der Maschinenethik und wird auch als *Moral Machines* bezeichnet. Er wird in seiner philosophisch-technischen Umsetzungsmöglichkeit kontrovers diskutiert. Wallach und Allen (2009) erörtern in ihrem Buch *Moral Machines - Teaching Robots Right from Wrong* die Möglichkeit, eine Maschine so zu konstruieren, dass diese sich selbstständig ethisch verhalten kann, wobei in dieser Monographie ethisch und moralisch synonym verwendet werden. Die Autoren unterscheiden dabei zwischen autonomen moralischen Systemen (*autonomous moral agents*) und Systemen, die eine Wertesensitivität aufweisen (Wallach & Allen 2009). Während an dem einen Ende wertsensitive Systeme durch die verantwortungsvolle Sorgfalt des Entwicklers entstehen und der Entwickler die Kontrolle über die Technik behält, finden sich am anderen Ende des Spektrums Maschinen, die selbstständig moralische Konfliktsituationen bewältigen können: Diese Maschinen wären in der Lage, autonom eine moralische Entscheidung zu fällen und sollen sich in ihren Ausführungen danach ausrichten können. Hier geht es dann nicht mehr um einen metaphorischen Gebrauch der Begriffe „denken, entscheiden, wissen, fühlen oder handeln“, sondern Maschinen unterscheiden sich gerade nicht mehr kategorial von Menschen in eben diesen Fähigkeiten. Dabei gilt zu verstehen, dass Maschinen, so die These der Autoren, moralische Fragestellungen immer dann von digitalen Akteuren auszuführen wäre, wenn eine selbstinteressierte Optimierung zu gesellschaftlich unerwünschten Resultaten führt. Hierunter fällt der Sachverhalt, dass die moralische Forderung sodann darin bestünde, eine Handlung zu verbieten, die Leid verursacht, und eine Handlung zu gebieten, die eben jenes minimiert. Da Softwaresysteme auch in unsere Lebenswelt eingebunden sind, müssen sie diese Sachlage selbstständig mitberücksichtigen (Wallach & Allen, 2009, v.a. Kapitel 2).

Diese Position entspricht der Annahme einer starken Künstlichen Intelligenz, unter der man nach John Searle versteht, dass ein Computerprogramm mit einem Bewusstsein (*mind*) gleichzusetzen ist (Allen & Wallach 2009). Searle widerlegt diese These in seinem berühmten Beitrag *The Chinese Room*, in dem er zeigt, dass ein Computer nicht bewusst verstehen können muss, wenn Zeichen nach einer bestimmten Vorgabe richtig zugeordnet werden (Searle, 1980). In seinem Gedankenexperiment werden durch einen Spalt in der Wand Fragen und Geschichten auf Chinesisch durchgeschoben. In diesem Zimmer hält sich eine Person auf, die des Chinesischen nicht mächtig ist und die Antworten auf die Fragen in Chinesisch schreiben soll. Sie erhält noch ein Handbuch in ihrer Muttersprache, das ihr dazu verhilft, alle Symbole der Geschichte und Fragen zu übertragen. Die Antworten,

die in chinesischer Zeichensprache verfasst wurden, werden wiederum durch den Spalt in der Wand nach draußen geschoben und von einer Person, die des Chinesischen mächtig ist, gelesen und verstanden. Searle zeigt damit, dass eine Übertragung möglich ist, ohne dass die Person im Zimmer selbst Chinesisch verstehen können muss, um diese Aufgabe erfolgreich zu bewältigen. Die Person muss keine mentalen Zustände und Erwartungshaltungen, Intentionen und referentielle Beziehungsmomente aufweisen, um die Arbeit auszuführen. Ein Sprachverständnis wird folglich nur simuliert. Aber Searle geht noch einen Schritt weiter, indem er zeigt, dass Computer nicht nur keine semantischen Maschinen sein müssen, sondern dass Computer nicht einmal als syntaktische Maschinen verstanden werden können (Searle, 1990). Denn, so das Argument, Zeichen und Symbole haben nur Bedeutung für Menschen, sind also betrachtungsrelativ. Die Bedeutung wird durch den Menschen zugeschrieben, indem das Zeichen referentiell verwendet wird, es in bestimmten Kontexten bestimmte Sinngebungen aufweist. Dies impliziert, dass Zeichen immer erst durch praktische Sozialität entstehen, in Form von Konventionen oder innerhalb kooperativer Strukturen. Computer sind aber in erster Linie nur materielle Gegenstände, die sich mit den Begriffen der Physik vollständig beschreiben und erklären lassen. Da jedoch Syntax kein Teil der Physik ist, weil innerhalb der Physik keine grammatischen Regeln, keine logischen Schlüsse, keine Algorithmen beschrieben werden, können Computer nicht als syntaktisch begriffen werden (Nida-Rümelin, 2022). Somit kann man nicht einmal von einer schwachen KI im Sinne der Simulation menschlicher Bereichsfähigkeiten sprechen, denn der Computer simuliert gar keine menschlichen Fähigkeiten, sondern ist auf physikalische Gesetzmäßigkeiten reduzierbar. Es kann hier also keine Rede davon sein, dass der Computer über menschliche Fähigkeiten im Sinne einer eigenständigen Bedeutungsbegabung und eigenständigen Anwendung von Regelsystemen aufgrund reflexiver Wissensstrukturen verfügt. Und noch mehr: Computer müssen nicht einmal menschliche Fähigkeiten simulieren, um funktionieren zu können. Die Prädikate *denken*, *meinen*, *fühlen* etc. sind daher nicht auf den Computer anzuwenden, sondern bleiben wesentlich dem Menschen oder anderen Lebewesen vorbehalten. Dies zeigt, dass Computer nicht einfach als autonome Akteure gehandhabt werden können. Inwiefern das menschliche Bewusstsein und v.a. das Denken simuliert oder gar in toto nachgebildet werden können, um sogar unabhängige moralische artifizielle Akteure entwickeln zu können, bleibt hier zunächst unbeantwortet. Aber ein funktionstüchtiger Computer muss weder über die Fähigkeit verfügen, noch muss die Fähigkeit simuliert werden. Wichtig dabei ist, dass Denken sich gerade nicht auf Logik reduzieren lässt bzw. dass – folgt man dem Unabhängigkeitstheorem Kurt Gödels (1931) – die Beweiskraft der Logik sehr wohl an ihre Grenzen kommt, die dem menschlichen Denkvermögen so nicht vorgegeben sind.

Wären Maschinen bereits autonome moralische Akteure, würde sich jedoch eine eigenständige Ethik der Informationstechnik unmittelbar durch den Akteur-Status begründen lassen. Eine wichtige ethische Frage wäre dann, ob solche moralischen Maschinen wünschenswert wären, also ob aus einem

technischen Können auch ein Sollen folgen sollte. Diese Überlegungen würden in den Verantwortungsbereich einer *research ethics* fallen. Diese Debatte wird im Rahmen dieser Arbeit weder nachgezeichnet, noch geführt, denn hier werden Systeme betrachtet, die keine moralische Eigenständigkeit im Sinne von Interaktivität, Autonomie und Adaptabilität (Floridi & Sanders, 2004) aufweisen.¹⁷

Graduell mögen einige der Aspekte technisch erfüllbar sein, allerdings gibt es derzeit kein System, das all diese Merkmale aufweist. Selbst wenn einige Forschende dieses Ziel vor Augen haben mögen, spricht viel dafür, sich ethisch mit solchen Maschinen zu beschäftigen, die bereits existieren und die normativ auf soziale Strukturen wirken. Der IT-Dienstleistungs- und Softwaremarkt in Deutschland wuchs 2019 um 5 % (Pütter, 2019). Auch steigt die Anzahl an Personen, die in diesem Segment tätig sind, seit 2008 kontinuierlich an (Statistisches Bundesamt, 2019). Selbst die Umsatzeinbußen, welche die Corona-Pandemie verursachte, sind verhältnismäßig gering und es wird in Deutschland mit einer Umsatzsteigerung von knapp 75 % bis 2025 gerechnet. Der derzeitige Anteil von 4,2 % am Bruttoinlandsprodukt soll bis 2025 sogar um 2,8 Prozentpunkte ansteigen (Gandorfer, 2020). Diese nur für Deutschland genannten ökonomischen Fakten belegen deutlich die wirtschaftliche Bedeutungszunahme der Softwareentwicklungsbranche sowie des Einsatzes von Software und IT-Dienstleistungen über alle Branchen hinweg. Allein das Marktwachstum lässt die durch, in und mit IT-Technik und Technologie aufgeworfenen Fragestellungen sehr viel drängender und somit unausweichlich werden. Es ist daher zentral, eine Ethik für Softwareentwickler zu begründen, die bereits während des Prozesses der Softwareentwicklung praktische, normative Orientierung bietet.

Wenn nun die moral agency von Softwaresystemen als Grundlage einer speziellen Ethik für Softwareentwickler ausscheidet, so bedarf es einer genauen Betrachtung, ob anderen Merkmalen dieser argumentative Status zukommen kann. Hierfür sollen im folgenden Unterkapitel *Embedded Values Approach* die Grundsteine gelegt werden, um sodann im Kapitel *Die Einzigartigkeit der Informations-technologie und -technik* vertieft auf zentrale Eigenschaften digitaler Technologie und Technik eingehen zu können.

¹⁷ Unter Interaktivität fallen Systemeigenschaften oder Fähigkeiten, die es ermöglichen, auf die Umwelt reagieren zu können. Als Autonomie wird die Möglichkeit verstanden, den eigenen Zustand ohne externen Stimulus, ohne direkte, externe Responsivität verändern zu können. Adaptabilität bezeichnet wiederum die Fähigkeit, Transitionsregeln zu ändern, die den Übergang von einem in einen anderen Zustand bestimmen. Wenn eine Maschine alle jene Aspekte, so die Autoren, erfüllt, kann man von einem autonomen moralischen artifiziellen Akteur sprechen.

3.5.3 Embedded Values Approach

Als Fundament einer eigenständigen Ethik für Computer schied die *moral agency* digitaler Artefakte aus. Sollte die Werteinhärenz als ein Alleinstellungsmerkmal digitaler Technologie und Technik gelten, kann sie als Argument einer eigenständigen Ethik fungieren. Diese Thematik wird in diesem Kapitel mit dem Ansatz der *embedded values* diskutiert.

Gegenstand dieser Betrachtung sind folglich – um sich der Terminologie John Moors (2006) zu bedienen – *ethical impact agents* und *ethical implicit agents*: Softwaresysteme, die entweder externe, negative moralische Effekte verursachen, z.B. AirBnB, oder solche, in denen bereits (verzerrte) Moralausdrücke implementiert wurden, z.B. die Diskussion um *bias* im *machine learning*. Nicht Teil dieser Analyse sind die *explicit ethical agents* und *full ethical agents* – also artifizielle Akteure, die selbstständig deontische Abwägungen in einem bestimmten Bereich oder über viele Situationen hinweg durchführen können sollen, d.h. die moralisch deliberieren und ihre Entscheidung begründet rechtfertigen können (Moor, 2006). Ebenso wenig wird hier von wertinhärenten Maschinen ausgegangen, die selbst über Intentionen verfügen oder Werte aus sich heraus generieren. Unter der Werteinhärenz wird vielmehr die oft unbewusste Inkorporation der Haltungen bzw. Einstellungen der Entwickler in technischen Objekten verstanden (Friedman & Kahn, 2002).

Wenn nun autonome, moralische Maschinen als ethische Begründung ausscheiden, muss man der Fragestellung, ob eine eigene Bereichsethik der Informationstechnik als Subdisziplin der Technikethik gerechtfertigt ist, auf eine andere Weise begegnen. Hierfür muss eine Ethik der Informationstechnik und -technologie unabhängig der Techniknutzung begründet werden, denn wenn ethische Anforderungen an den Umgang mit oder der Nutzung von Informationstechnik gestellt werden, reichen klassische ethische Prinzipien und Theorien aus, um diese ethischen Fragestellungen angemessen zu erörtern. So könnte man jenes Nutzerverhalten beurteilen, indem man die angemessene moralische Verpflichtung aufgrund eines Universalisierungstests ermittelt, also im klassischsten Sinne Kants kategorischem Imperativ Folge leistet. Demgegenüber kann die Frage nach dem Sollen einer konsequentialistischen Ethiktradition folgen wie bspw. dem Utilitarismus. Hier gilt es dann, diejenige Handlung als moralisch geboten zu betrachten, die den Weltzustand optimiert, wobei dieser bspw. als eine Zunahme der Nutzensumme aller Menschen interpretiert wird. Wenn die Frage nach dem moralisch gebotenen Handlungsgrund tugendethisch begründet wird, rücken der gute Charakter und die Ausbildung wünschenswerter Tugenden in den Mittelpunkt der Betrachtung. Die moralisch wertvolle Handlung begründet sich dann in der Ausführung wohlbedachter Tugenden, die das Gute fördern sollen. Eine Anwendung dieser klassischen Ansätze wäre komplex und schwierig, aber die ethischen Konzepte wären angemessen und würden einem solchen Evaluationsprozess genügen. Der tugendethische Ansatz hebt sich von den beiden anderen Deliberationsweisen ab. Die ersten beiden Modi

der ethischen Deliberation benötigen keine Kenntnis des informationstechnischen Produkts, denn sie beurteilen die Handlung nach klassischem, rationalistischem, ethischem Vorgehen und nach Prinzipien. Sie bedienen sich ethischer Prinzipien und wenden diese auf Einzelfälle, d.h. auf individuelle oder kollektive Handlungen, an. Sie fragen nicht, was ein wünschenswertes digitales Artefakt sein soll oder was ein solches ausmache, sondern welche Handlungen in der Nutzung oder im Umgang mit demselben moralisch geboten sein sollen. Daher gilt es hier auch nicht explizit nach den Handlungen der Entwickler zu fragen, denn die Betrachtung fällt zunächst nicht auf das Informationstechnische und somit auch nicht auf den Herstellungsprozess. Ein tugendethischer Ansatz hingegen verweist nicht nur auf eine Art und Weise der Deliberation, sondern auch auf die zu entwickelnden Tugenden und Praxen, derer es bedarf, um mit lebensweltlichen Situationen bestmöglich, im Sinne von moralisch angemessen, verfahren zu können. Es geht hierbei darum, Haltungen zu festigen, zu denen auch das praktische Urteilsvermögen zählt, also um eine Bereitschaft, moralische Deliberationen durchzuführen, wenn diese notwendig sind (Vallor, 2016). Diese Deliberation ist kontextsensitiv, d.h. es ist unmöglich, anhand der Anwendung eines Prinzips die moralisch gebotene Handlung zu ermitteln, sondern es bedarf Erfahrung und Wissen, um zu beurteilen, welche tugendhafte Handlung in diesem Gegenstandsbereich dem Guten dient (Vallor, 2016).

So hat es den Anschein, dass normative, ethische Ansätze zunächst keinen Rekurs auf die Informationstechnologie oder -technik nehmen müssen, da sie sich, unabhängig vom Gegenstand, auf die richtigen oder guten Handlungen beziehen. Jedoch verweist die tugendethische Herangehensweise bereits auf die Öffnung der Handlungskontexte und bezieht diese mit in die Überlegungen zur Charaktergüte ein: Denn erst in den lebensweltlichen Situationen beweist sich die wahre Güte des Charakters. Dieser starke Bezug zur erfahrbaren Lebenswelt mag auch erklären, weshalb gerade in der Informationstechnikethik die Tugendethik eine große Renaissance erlebt. Ebenso ermöglicht diese Tugendethik, das Individuum in Rückbezug zu einer Gemeinschaft, zu einem Ethos, zu denken, durch die und in der es seine rationalen Lebensentwürfe erst moralisch ausleben kann. Obgleich diese kulturelle Einbindung auch als Kritikpunkt gelten mag, sind die strukturellen Momente tradierter moralischer Rationalität essentiell, wenn eine hermeneutische Verortung und moralische Thematisierung der Informationstechnologie und -technik gelingen soll (Teil I, Kapitel 2.2).

Die Frage nach einer IT-spezifischen Ethik muss daher in der Besonderheit der Informationstechnologie und -technik begründet sein (Brey, 2010). Daraus folgt allerdings, dass Informationstechnik nicht mehr nur als reines Mittel zum Zweck betrachtet werden kann, denn dieser Mittelcharakter der Technik übersieht zum einen die Wertinhärenz der Informationstechnik und zum anderen die mögliche Transformation des Wahrnehmungs- und Erfahrungsraums der Nutzer. Zudem kann über die Medialität der digitalen Technik die strukturelle Einbettung des autonomen Nutzers und Entwicklers in ein Netz gegenseitiger Verpflichtungen eingefangen werden, sodass die transformative Kraft, die von

mancher informationstechnischen Errungenschaft ausgeht, für eine ethische Analyse fruchtbar gemacht werden kann. Rohbeck bettete 1993 die moderne Ablehnung eines strengen Technikdeterminismus in seiner Einleitung in *Technologische Urteilskraft – Zu einer Ethik technischen Handelns* in einen philosophiehistorischen Kontext ein: So verdeutlicht Rohbeck, dass sich die *technologische Kränkung* aus dem Ohnmachtsgefühl speist, trotz des herstellenden Schöpfungsaktes nicht mehr Herr über die eigenen geschaffenen Werkzeuge zu sein (Rohbeck, 1993). Dieser Kränkung stellt Rohbeck die philosophischen Begriffe der Subjektivität und Freiheit gegenüber, die den modernen Handlungsbegriff tragen, insofern ein autonomes Subjekt sich selbst Ziele zu setzen und sich durch die Erreichung eben jener selbst zu verwirklichen vermag (Rohbeck, 1993). Dieser philosophischen Haltung entspringt eine vehemente Ablehnung jeglicher Ansätze, welche die autonome Zwecksetzungsvernuft in ihre Schranken verweisen, weil sie diese dadurch entmachten. Allerdings führt der traditionell-moderne Handlungsbegriff zugleich in eine groteske Darstellung des Mittel-Zweck-Verhältnisses: Denn gerade in der Geschichte der Technik und insbesondere der Informationstechnik gibt es eine Fülle an Beispielen, die auf den Umstand verweisen, dass das Artefakt vormals für einen gänzlich anderen Zweck entwickelt wurde, wie bspw. die SMS, sodass gerade hier die Zwecke die Mittel nicht vollständig bestimmen. Ein solcher Handlungsbegriff würde wieder zurück zur Annahme führen, dass der Zweck die Mittel bestimme und die Mittel diesem vollständig unterlägen (Rohbeck, 1993). Der moderne Handlungsbegriff, der von autonomen Individuen belebt wird, führt in die Aporie der Zweck-Mittel-Beziehung: Zum einen postuliert er die Eigenständigkeit des Subjekts und somit der Zwecke. Zum anderen führt er dadurch zurück zu einem klassischen Mittelbegriff, der für Informationstechnologie und -technik nicht haltbar ist.

Die traditionelle Sicht der Wertfreiheit der Technik (Simon, 2016; Van den Hoven, Vermaas & Van de Poel, 2015) besagt, dass die Güte oder die Minderwertigkeit eines technischen Produkts oder einer Fertigkeit erst durch die Bedienung oder Ausübung des Nutzers entsteht. Erst durch einen spezifischen Umgang mit dieser Technik entstünde die Wertfrage der Technik, denn Werkzeuge werden nur durch ihren Gebrauch oder Missbrauch, durch ihre Zweckentfremdung, evaluativ zugänglich und entbehren als Werkzeuge zunächst jeglicher moralischen Bewertungsgrundlage. Der Nutzer verleiht durch seinen Umgang mit einem Werkzeug, z.B. einem Hammer, diesem einen Wert, da er den Hammer nutzen kann, um Nägel einzuschlagen oder um jemanden zu erschlagen. Ein Hammer wird erst durch den Einsatz des Nutzers zur Waffe, zu einem Tatwerkzeug, wenn der Nutzer ihn als solchen nutzt. Die Wertigkeit kommt dem Hammer exogen, sozusagen über die Handlung des Nutzers, zu. Der Umgang mit dem Hammer als Werkzeug oder als Waffe entscheidet darüber, ob eine moralisch wünschenswerte Handlung vorliegt oder nicht (Ihde, 1990). Die Intention, den Hammer als Waffe zu missbrauchen, wird nicht durch den Hammer selbst erzeugt – selbst wenn der Hammer als solcher offen zugänglich ist und jemand ihn im Affekt als Waffe zweckentfremdet, kann man nicht davon

sprechen, dass dem Hammer jene ‚Waffenheit‘ qua Design intrinsisch angelegt sei und man von daher einen Hammer bspw. nur noch aus Schaumstoff anfertigen sollte. Dem Täter kann man jedoch den strafrechtlichen und moralischen Vorwurf machen, absolut unangemessen gehandelt zu haben. Ethisch relevant werden in diesem Zusammenhang also Fragestellungen, die einen guten, fairen, gerechten Umgang miteinander adressieren und dies bedeutet, dass moralisch wünschenswerte Handlungen im Zentrum der Analyse stehen. Eine solche Betrachtung befreit den Gestaltungsprozess, normativen und vor allem moralischen Geboten Folge leisten zu müssen, denn weder der Designer, der Gestalter noch der Entwickler können etwas zur Werthaftigkeit des Produkts beitragen. Die moralische Verantwortung liegt hier ganz beim Hammer-Nutzer. Diese Annahme greift jedoch in vielen Fällen der Informationstechnik zu kurz.

Der dieser Hypothese zugrundeliegende Mittel-Zweck-Charakter der Technik, der Werkzeugen eigen ist, trifft jedoch nicht ohne weiteres auf Informationstechnik zu, denn Informationstechnik und -technologie umfasst voreingestellte Produkte, über die der Nutzer keine vollständige Kontrolle verfügen kann.‘ (Grunwald, 2018a) Man denke hier bspw. an Waschmaschinen, Suchmaschinen oder auch Assistenzsysteme in der Robotik. In ihnen stecken bereits viele Annahmen über den Wasserverbrauch, den Stromverbrauch, über die Platzierungsverfahren der Suchergebnisse bis hin zur äußeren Gestaltung eines Roboters, die anthropomorph oder zoomorph sein kann und so ein Gegenüber vorgaukelt. All diese Merkmale sind gerade nicht wertfrei. Die Art und Weise, wie der Nutzer mit dieser digitalen Technik umgehen kann, ist bereits zum Teil durch die Technik selbst vorgegeben und die Handlungsfreiheit der Nutzer somit eingeschränkt. Dies mag teils auch auf klassische Technik übertragbar sein, insofern jede Technik bestimmte Verhaltensweisen einfordert, weil ein bestimmtes reaktives Verhalten meist erforderlich ist, damit Technik funktioniert. Aber im Falle von Informationstechnik und -technologie weisen viele vorgeformte Eigenschaften und Infrastrukturen bereits bestimmte Wertungen auf, die nicht unmittelbar einsichtig sind und die von außen nicht (oder nur eingeschränkt) beeinflussbar sind.

Zusätzlich spricht die Möglichkeit bestimmter digitaler Technik und Technologien, auf den Weltbezug der Nutzer einzuwirken, auch gegen einen reinen Mittelcharakter (Hubig, 2006; Teil I, Kapitel 3). Die durch den Einsatz von Technik sich verändernden Relationalitäten führen zu neuen Wissens- oder Wahrnehmungshorizonten (Ihde, 1990), die nicht mehr durch die instrumentelle Natur des Technischen erklärbar sind, sondern wesentlich einen Medialitätscharakter aufweisen (Hubig, 2006). Technik ist dabei „kein spezielles Medium, sondern macht eine bestimmte Art und Weise aus, in der die realen, intellektualen und sozialen Möglichkeitsräume strukturiert sind“ (Hubig, 2006, S. 155). Dadurch wird bereits ersichtlich, worin die transformative Kraft begründet liegt, die Kulturen beeinflussen und neue Welten eröffnen kann. Sowohl der Aspekt der Modalitätsveränderung des Weltzu-

gangs als auch der Aspekt der Strukturierung von Handlungskontexten verdeutlicht die informations-technikimmanente Wertausrichtung. Beide Gesichtspunkte werden durch den Medialitätscharakter der Informationstechnik konzeptualisierbar und veranschaulichen dadurch zugleich die immanente Wertigkeit der Informationstechnik und -technologie. Dies eröffnet nun auch Raum für den Gedanken, dass Technik normativ gut gestaltet werden kann, indem man diese technischen Aspekte, die Werte perpetuieren, im Entwicklungsprozess bereits berücksichtigt. Der Gestalter, Designer oder Entwickler ist somit maßgeblich an der Entstehung und Weitergabe von Werten beteiligt.

Diese Wertinhärenz stellt somit die Möglichkeit dar, durch Technik wesentlich individuelle, intersubjektive – hierunter fallen bestimmte relationale Beziehungsgeflechte – und gesellschaftliche Phänomene zu beeinflussen. So kann eine digitale Kommunikationsplattform jemanden individuell ausgrenzen (Diskriminierung), auf die Beziehungen zwischen Individuen einwirken (Kommunikationskultur) oder gar kollektive Auswirkungen haben (z.B. auf die Demokratie). Diese Gestaltungsmacht betont auch Brey (2000a) in seinem Aufsatz *Disclosive Computer Ethics* als Grundbedingung der *values in design*-Bewegung oder Friedman (1997) in *Human Values in the Design of Information Systems and Technology*. Diese Bewegung entstand in den 1990er Jahren vornehmlich aus der Softwareentwicklung heraus und stellte die ethische Forderung auf, dass bereits zu Beginn und fortlaufend moralische Werte und generell normative Haltungen Berücksichtigung finden müssen, um moralisch gelungene informationstechnische Objekte zu entwickeln: „Value Sensitive Design is a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process.“ (Friedman, Kahn & Borning, 2006, S. 350)

Somit ist die gänzliche Abwälzung moralischer Verantwortung ausschließlich an die Nutzer selbst moralisch verwerflich, denn IT-Technik und IT-Technologie können nicht als neutraler und transparenter Vermittler menschlicher Intentionalität verstanden werden (Simon, 2016). Eine ethische Bewertung kann somit auch unabhängig der menschlichen Nutzung oder des Umgangs mit dieser Technologie erfolgen. Brey (2000a, 2010) versteht unter den in Technik und Technologien implementierten Werten (*embedded values*) Handlungsausrichtungen, die, sobald das System implementiert oder genutzt wird, die Lebenswirklichkeit moralisch beeinflussen. Somit bemisst sich die Güte des Systems aus den tatsächlich entstandenen sozio-technischen Realitäten (Brey, 2000b, 2010). Da die in der Informationstechnik enthaltenen oder implizierten Werte tatsächliche moralische Auswirkungen in der Lebenswelt haben, muss eben jene Beziehung im Entwicklungsprozess und im Verwaltungsprozess informationstechnischer Produkte miteinbezogen werden. Brey (2000a, S. 12) hebt vier Werte hervor, die durch den Einsatz digitaler Technologie und Technik verstärkt unter Druck geraten: Gerechtigkeit, Autonomie, Privatsphäre und Demokratie. Darauf aufbauend entwickelte Brey zur Ergänzung der *mainstream computer ethics* einen normativen Zugang zur Computertechnologie, den er als *disclo-*

sive ethics bezeichnet. Dieser Ansatz fokussiert sich zunächst nicht darauf, moralisch-qualitative Kriterien, sondern Klarheit in der Relationalität von Software und Lebenswelt zu gewinnen. Unter *mainstream computer ethics* versteht Brey normative Analysen, die sich auf bekannte moralische Konfliktfälle beziehen, wie bspw. Hacking, Internetpornographie oder Phishing. In der ethischen Deliberation über jene Konfliktfälle werden sodann die klassischen, wissenschaftlich-theoretischen Ethiken angewendet, um im Anschluss Regulierungsanforderungen formulieren zu können. Dieses Vorgehen hält Brey jedoch für verkürzt, denn es werden gerade nur offensichtlich problematische Fälle berücksichtigt. Hingegen können moralisch wertvolle Einflüsse keine Berücksichtigung erhalten, denn es werden traditionell nur diese moralisch kontroversen Fälle im Hinblick auf individuelle, intersubjektive oder kollektive Handlungen diskutiert. Somit steht nur der Nutzer und niemals das Design informationstechnischer Artefakte im Mittelpunkt einer ethischen Analyse (Brey, 2000). Brey verweist jedoch darauf, dass die Umsetzung von Werten und die Beeinflussung sozialer Strukturen durch den Einsatz digitaler Technik einer eigenständigen Analyse und normativen Evaluation bedürfen. So stellt er den Begriff der Opazität ins Zentrum einer ethischen Aufklärung, denn: „[c]omputer-related practices may be morally opaque for two reasons: because they are unknown or because they have a false appearance of moral neutrality“ (Brey, 2000, S. 11). Es gibt folglich für Brey zwei gravierende Probleme, die aufgrund der *mainstream approach* auftreten: der Fokus auf gängige, bekannte moralische Probleme sowie die Ignoranz gegenüber der Annahme, dass Computertechnologie selbst Quelle moralischer Werte und deren Beibehaltung, Aufrechterhaltung oder Verzerrung sein kann. Somit beginnt jede Computereethik zunächst mit der Aufdeckung eingebetteter Normativität in Computersystemen, Anwendungen und Praxen, sodass die *disclosive ethics* vor allem einen deskriptiven Charakter aufweist, da die Computertechnik und die damit zusammenhängende Praxis analysiert werden, um deren moralische Bedeutung herauszuarbeiten, ohne diese zunächst zu bewerten (Brey, 2000; Teil II, Kapitel 3).

Die für die Computertechnik einzigartigen Aspekte, durch welche die bestimmten normativ relevanten Sachverhalte differenziert werden können, werden im nachfolgenden Teil der vorliegenden Arbeit behandelt. Sie grenzen den Bereich der normativen Deliberation auf die für die Computertechnik wesentlichen Momente ein.

4 Die Einzigartigkeit der Informationstechnologie und -technik

Die als *uniqueness*-These bekannte Forderung nach der Begründung einer eigenständigen Ethik der Informationstechnik verlangt nach einer Argumentation, weshalb klassische, ethische Theorien nicht ausreichen, um sich normativ angemessen mit der Thematik der Digitalisierung, der digitalen Transformation sowie der digitalen Artefakte zu befassen. Vielmehr gilt es dem Anspruch nachzugehen, eine Bereichsethik einzuführen, die sich am Sachgegenstand orientiert, da im Gegensatz zu anderen Bereichsethiken, wie bspw. der Wirtschaftsethik, kein eindeutiges gesellschaftliches Subsystem mehr adressiert werden kann. Dies liegt darin begründet, dass Digitalisierung oder digitale Transformation in allen menschlichen Lebensbereichen stattfinden und somit verschiedene gesellschaftliche Systeme tangieren. Die Wertinhärenz digitaler Technik, die These der *embedded values* verdeutlicht, dass Software und Softwaresysteme selbst Gegenstand einer normativen Beurteilung werden müssen, da ethisch relevante Situationen nicht erst im Umgang mit oder in der Nutzung digitaler Artefakte entstehen. Eine Ethik für Softwareentwickler muss sich folglich über einen eigenständigen Gegenstandsbereich oder ein idiosynkratisches Vorgehen abgrenzen, die jeweils eine separate ethische Deliberation rechtfertigen. Es gilt somit, diejenigen Situationen einzugrenzen, die dezidiert aufgrund des informationstechnischen oder -technologischen Charakters auftreten. So müssen diejenigen Auswirkungen betrachtet werden, die durch einen Einsatz von Informationstechnik und -technologie in der lebensweltlichen Realität entstehen, die direkt auf die technischen Bedingtheiten zu reduzieren sind bzw. muss thematisiert werden, wie auf diese normativen Unzulänglichkeiten technisch eingewirkt werden kann (Maner, 1996). Die Frage, welche normativen Deliberationen vonnöten sind, und wie sich diese von klassischen, ethischen Deliberationen unterscheiden, gilt es in den letzten Kapiteln dieser Arbeit nochmals hervorzuheben.

Eine selbstverantwortliche Ethik für Softwareentwickler bedarf zunächst der Herausarbeitung jener Aspekte, die als Unterscheidungsmerkmale dienen und dadurch den Charakter der Einzigartigkeit digitaler Technologien und Techniken hervorheben:

- der mediale Charakter digitaler Technologie und Technik
- das Merkmal einer *emerging technology*
- Vernetzung durch Transparenz und Allgegenwart
- die Formbarkeit digitaler Technologie und Technik
- die Machtposition digitaler Technologie und Technik
- Verantwortungsdiffusion

Diese sechs Erfassungsbereiche sollten bei jeder normativen Betrachtung graduell fokussiert werden, d.h. je nach Kontext und digitaler Technologie und Technik mehr oder weniger.

4.1 Der mediale Charakter der Informationstechnologie und -technik

Impact und *implicit moral agents* wurden als Kategorien moralischer Implikationen von Informationstechnik herausgestellt. Der kategoriale Unterschied beruht auf der je unterschiedlichen Rückführbarkeit normativer Implikationen: So wurden *impact moral agents* negative, externe, moralische Effekte zugeordnet, die zwar durch den Einsatz eines digitalen Artefakts entstehen oder maßgeblich durch ihren Einsatz verstärkt werden. Diese Effekte sind jedoch von solchen zu unterscheiden, die durch *implicit moral agents* hervorgerufen werden. Letztere entstehen gerade durch inhärente Werte, weshalb von einem kategorialen Unterschied der moralisch relevanten Ursachen gesprochen werden muss.

Impact moral agents → negative externe moralische (Neben-)Effekte

Implicit moral agents → moralische Effekte aufgrund technikhärender Werte

Die verschiedenen Ursachen moralischer Güte lassen die Debatte um den instrumentalen oder medialen Charakter von Informationstechnologie und -technik in den Mittelpunkt der Betrachtung rücken: Um diese Unterscheidung in den Zusammenhang um *impact* und *implicit* zu stellen, ist die kategoriale Verschiedenheit von Werkzeugen, Maschinen und Automaten einzuführen. Diese verhilft dazu, den moralischen Status eingrenzen zu können sowie die Besonderheit der digitalen Technologien und Techniken im Gegensatz zu anderen technologischen und technischen Verfahren hervorzuheben.

Digitale Artefakte können nicht mehr als reine Werkzeuge verstanden werden, sondern sind auch als Maschinen oder Automaten zu verstehen. Dies begründet Joseph Weizenbaum mit der Tatsache, dass einige digitale Artefakte zunehmend auf menschliche Kontrolle verzichten können (Kastendiek, 2003), d.h. eigenständig Prozesse in Gang halten oder gar in Gang bringen können. Diese Ersetzung menschlicher Fähigkeiten ist zunächst vor allem auf die menschliche, physische Arbeitskraft beschränkt gewesen: Mit Garn und Nadel lässt es sich besser Stoffe fest und über einen längeren Zeitraum zusammenhalten, mit einer Nähmaschine können sodann komplexere Abfolgen im Eigenantrieb erfolgen, wie z.B. mit einer Nähmaschine ohne Fußbrett. Die automatisierte Nähmaschine hingegen verfügt über noch mehr Eigenleistung: So kann eine digitalisierte Nähmaschine in Abhängigkeit der Stoffqualität eigenständig zwischen Programmen und Stichmustern variieren.



Abbildung 9 Die Entwicklung vom Werkzeug zum Automaten hinsichtlich der Unterstützung der physis hin zur phronesis. Eigene Abbildung.

Unter Informationstechnik versteht man zudem nicht nur diejenigen autonomen Maschinen, die in bestimmten Arbeitskontexten menschliche physische Arbeitsfähigkeit unterstützen oder gar menschlich-körperliche Fähigkeiten durch maschinelle Ausführungen ersetzen. Vielmehr sollen Software-gestützte Artefakte, die als *decision supportive technology* (Shim et al., 2002) begriffen werden, *mind work* (Agar, 2019) verrichten: Sie sollen folglich die *phronesis* unterstützen, weshalb sie einer besonderen ethischen Überlegung bedürfen, die nicht nur negative Externalitäten betrachtet, sondern auch wesentlich die normative Ausgerichtetheit der Welt mitberücksichtigt. In dieser Funktion sind digitale Artefakte nicht ausschließlich instrumentell zu verstehen, sondern transzendieren einen strengen Mittelbegriff. Digitale Artefakte sind somit auch als Medium zu begreifen (Hubig, 2006), gerade weil vielen digitalen Artefakten die Möglichkeit innewohnt, Wissen oder den Wissenszugang auf eine bestimmte, selektive Art und Weise aufzubereiten. Technik als Medium betont gerade, dass die Dringlichkeit normativer Deliberation über den Einsatz und Gebrauch digitaler Technik, aufgrund ihres transformativen Charakters, den Weltbezug des Menschen stark beeinflussen kann. Dies liegt darin begründet, dass digitale Technik sich mehr und mehr im Urteilsvermögen des Menschen verankert und das Urteilsvermögen unterstützt oder gar ersetzt, wenn Entscheidungen oder der Entscheidungsfindungsprozess, d.h. die Deliberation, gänzlich an Automaten delegiert werden. So können digitale Artefakte nicht mehr nur als ein Instrument verstanden werden, um Ressourcen optimal einzusetzen,

sondern müssen in erster Linie als Medium begriffen werden, durch das sich Handlungs- und Erfahrungsspielräume konstituieren, strukturieren und transformieren lassen (ebd.).

Dies ermöglicht es, eine weitere Unterteilung hinsichtlich der moralisch bedeutsamen Spielräume vorzunehmen, die technisch relevant sind: *Impact moral agents* können in Form von Werkzeugen und Maschinen als Mittel verstanden werden, die positive oder negative Externalitäten zur Folge haben können. Diese Folgen entstehen durch den Einsatz digitaler Technik in bereits existenten mikro-, meso- oder makro-sozialen Kontexten, z.B. in der verschlüsselten peer-to-peer-Kommunikation, die staatliche Überwachung umgehen soll und sich somit auf das Bürger-Staats-Verhältnis auswirkt. *Implicit moral agents* hingegen erzeugen moralische (un-)wünschenswerte Realitäten aufgrund technikinhärenter Werthaltungen. Selbstredend können digitale Artefakte moralisch bedeutsam sowohl hinsichtlich ihres *impacts* als auch hinsichtlich ihrer *implicit*-Folgen sein: Die kategoriale Unterscheidung ist somit keineswegs exklusiv zu verstehen, sondern dient der Orientierung moralisch bedeutsamer Momente und einer adäquaten praktischen Adressierung.

Impact moral agents → Werkzeug/Maschine → negative, externe, moralische Effekte (Nebeneffekte)
Implicit moral agents → Maschine/Automat → moralische Effekte aufgrund von technikhärenten Werten

Interessant ist, dass der Übergang von Mittel zu Medium sowie von Werkzeug zu Maschine jeweils einen anderen normativen Handlungsspielraum konstituiert: *Ethical impact agents* zeichnen sich vor allem dadurch aus, dass ihre ethischen Probleme nicht genuin technischer Natur sind. Die normative Analyse und Evaluation befassen sich demnach mit denjenigen Artefakten, deren ethisch unerwünschten oder wünschenswerten Konsequenzen externer Natur sind, d.h. weil ihr Einsatz in einem lebensweltlichen Kontext dort zu Verzerrungen führt, wie es bspw. AirBnB auf dem Wohnungsmarkt gezeigt hat. Allerdings können diese negativen oder positiven ethischen Einflüsse nicht rein technisch erklärt oder aufgefangen werden, sondern bedürfen einer wirtschaftsethischen, unternehmensethischen oder politischen Diskussion. Daraus lässt sich folgern, dass diese normativen Aspekte kollektiv und somit kooperativ adressiert werden. Individuell – in unserem Kontext bedeutet dies auf der Ebene des Softwareentwicklers oder des Softwareentwicklungsteams – kann hier nur schwierig agiert werden, ohne den einzelnen Softwareentwickler unter *moral distress* zu stellen. Softwareentwickler allein können zwar ihren Handlungsspielraum eigenständig eingrenzen, aber dies zu fordern käme einer supererogatorische Forderung gleich. Diese externen ethischen Konsequenzen sollten zwar einer Analyse unterworfen werden, sobald ein Artefakt entworfen und entwickelt wird, aber viele der Rahmenanforderungen sind nicht Teil des zu diskutierenden Optionen- und Lösungsraums. Ob die Ungleichheiten, die durch bspw. AirBnB auftreten, wünschenswert sind oder nicht, lässt sich nicht einzig durch die Diskussion der technischen Begebenheiten klären. Zwar muss betont werden, dass sich

die negativen Auswirkungen, die AirBnB aufgrund der extremen Reichweite und einfachen Teilnahme verursacht, auch informationstechnisch erklären lassen. Allerdings lässt sich diese normative Schiefelage nicht einzig durch das Design der Technik klären. Selbst wenn sich AirBnB selbst begrenzen und bspw. nur fünf Wohnungen pro km² ausweisen würde, ist dies eine eigenständige, wirtschaftliche Regulierung, die nicht als Sollensanspruch an das Unternehmen allein gestellt werden kann.

Diskriminierungsaspekte, die aufgrund eines implementierten Algorithmus entstehen, sind hingegen individuell auf der Ebene der Softwareentwickler technisch lösbar. Diese moralische Aufforderung, Diskriminierungen zu unterlassen, kann somit an Softwareingenieure delegiert werden, denn hier gilt es bereits, die deontologische Forderung der Gleichbehandlung aller – gleichgültig der Hautfarbe, des Geschlechts, der Religion oder der Herkunft – zu berücksichtigen.

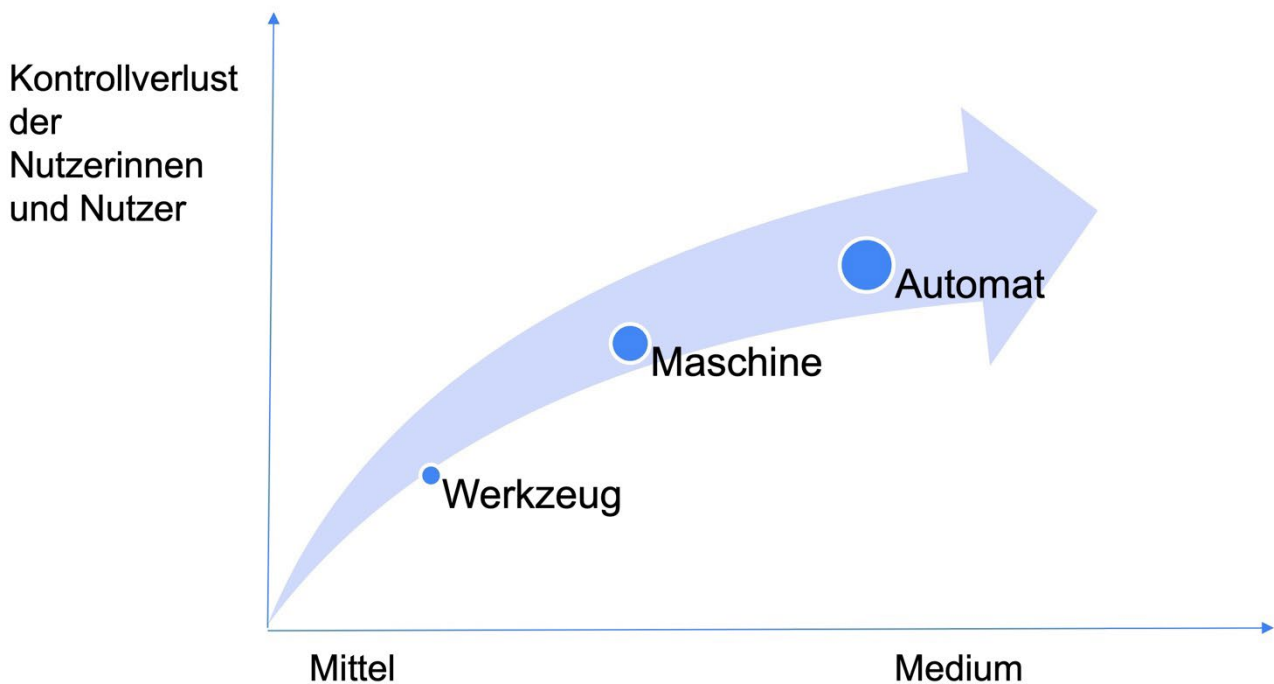


Abbildung 10 Werkzeuge als Mittel sowie digitale Technologie und Technik als Medium. Eigene Abbildung.

Der Werkzeugcharakter ist demnach als ein qualitatives Merkmal zu verstehen, das den *impact* von *implicit* unterscheidet. Dies ist relevant, insofern gezeigt werden kann, welche Aspekte sich als genuine informationstechnische oder informationstechnologische Probleme herausstellen lassen, an denen Softwareingenieure qua ihrer Professionalität arbeiten können.

4.2 Das Merkmal einer emerging technology

Oft gibt es (noch) keine lebensweltlich eingebundene, für moralisch angemessen befundene Orientierung, d.h. es fehlt an für moralisch gut befundenen Lebenspraxen sowie der Förderung moralisch wünschenswerter Tugenden. Dies kann man zum einen auf die Dynamik, die digitaler Technologie und Technik eigen ist (Teil III, Kapitel 4), zum anderen auch durch deren Zielsetzung begründen. Letzteres kann auch auf andere Technologien zutreffen, allerdings zielen digitale Artefakte gerade darauf ab, dem Menschen nicht nur als Werkzeug oder Maschine dienlich zu sein, sondern unterstützen oder substituieren ihn in seiner Urteilskraft, indem sie als Medium fungieren (Hubig, 2006). Softwaresysteme veranlassen die Endnutzer dazu, in bestimmte Richtungen zu denken oder zu handeln, sodass selbst das Urteilen und Entscheiden an Softwaresysteme delegiert wird. Dies geht häufig mit einem Verlust der Autonomie einher, denn so werden Personen in ihrer Vernunftfähigkeit untergraben und verlieren zunehmend Vertrauen in ihre eigenen erlernten Fähigkeiten und Kompetenzen (Frischmann & Selinger, 2018; Harris, 2016). Je mehr ein Softwaresystem folglich nicht nur Werkzeug oder Maschine, sondern ein medialer Automat geworden ist, desto mehr nehmen die ethischen Aspekte zu. Die ethisch zu überblickenden Sachverhalte betreffen den Kern der Humanität, weshalb nach sorgsam entwickelten und vertrauenswürdigen digitalen Artefakten verlangt wird: Es geht von nun an verstärkt darum, die Ausübungen der Vernunft technisch zu unterstützen. *Decision support systems* erweitern nicht nur Handlungsräume, sondern wirken auf Gedanken und Urteile ein. Im Gegensatz dazu ersetzen Maschinen menschliche Arbeitskraft.

Der pragmatistischen Tradition folgend führen moralische Überzeugungen genau dann nicht mehr zu angemessenen, zumutbaren Resultaten, sondern laufen ins Leere, wenn sie wegen einer fehlenden oder mangelhaft gewordenen Praxis Konflikte erzeugen, die eine philosophische und ethische Deliberationen notwendig werden lassen (Dewey, 1922; Mead, 1934; Nida-Rümelin, 2008). Tradierte moralische Kriterien können keine wünschenswerten Handlungen mehr garantieren, sondern neue Handlungsoptionen fordern diese moralischen Kriterien heraus. Ethik wird hier als die zur Deliberation gebrachten, strukturellen Konflikte verstanden. Diese herausgeforderten Strukturmerkmale gilt es nun zu überdenken, weil die tradierten moralischen Überzeugungen nicht ohne Weiteres auf einen neuen Lebenskontext übertragen werden können. Folglich muss die Handlungsebene verlassen und eine ethische Reflektion über die moralischen Überzeugungen begonnen werden, die das gesellschaftliche System nicht mehr unvermittelt tragen. Moralische Überzeugungen verbergen sich in Routinen und stabilisieren so lebensweltliche Handlungsketten, die dadurch Erwartungshaltungen konstituieren und darauf gerichtete Handlungen vorhersagbar machen (Teil I, Kapitel 2.1). Aufgrund ihres in alltäglichen Strukturen verborgenen Charakters sind moralische Überzeugungen und Haltungen implizit und offenbaren sich nur dann, wenn sie ihrer Gewandtheit verlustig oder wenn sie durch eine

ethische Deliberation explizit gemacht werden. Als rationale, normative Konzepte sind sie wiederum erlernbar und so tradierbar, bleiben deshalb aber auch modellierbar. Moral ist als solche unproblematisch und die meisten moralischen Überzeugungen sind weitreichend akzeptiert. Ethische Deliberationen explizieren diese Überzeugungen oder Haltungen und können sie so öffentlich zur Debatte stellen. In diesen Deliberationen wird normativ abgewogen, welches das moralisch wünschenswerte Szenario sein soll und wie Strukturen sowie Handlungen dementsprechend ausfallen sollen. Es ist daher wenig verwunderlich, dass durch den Einsatz neuer Technologien oder Technik moralische Überzeugungen herausgefordert werden können, gerade weil sie feste Strukturen transformieren und somit die unsichtbaren Muster akzeptierter Moral offenbaren, die sodann zugleich zur Disposition stehen (Swierstra & Rip, 2007). Diese Transformationen oder moralischen Überwerfungen führen zunächst in eine Ratlosigkeit, die auch Vallor (2016) vortrefflich belegt, indem sie hervorhebt, dass selbst in Handbüchern zur Angewandten Ethik normative Erörterungen informationstechnologischer Situationen eine Ausnahme darstellen. Ethische Abwägungen betreffen dagegen bspw. den Fleischkonsum, Abtreibung oder Folter, da in diesen abgeschlossenen Szenarien menschliche Handlungskontexte vertraut sind und mit lebensweltlichen Praxen korrespondieren. Die ethische Frage kann sodann, so Vallor (2016, S. 27), hinreichend präzise als „[i]s x (where x is an act or practice from the above list) right or wrong?“ gestellt werden. Die konkreten Fragen sind so vielfältig wie der Einsatz der digitalen Technik selbst. Diese Frage kann aufgrund der fehlenden Kenntnis um strukturelle Bedingtheiten wünschenswerter Praxis im Zusammenhang mit Informationstechnik so nicht formuliert werden. Vallor spinnt das Beispiel weiter und beginnt mit der Frage „Is twitter right or wrong?“ Dabei gelingt es ihr aufzuzeigen, dass die moralische Güte von Twitter als Technik-Objekt nicht erörtert werden kann, weil die Bedeutung aufgrund einer mangelnden Klarheit über die strukturelle Einbettung nicht verständlich gemacht werden kann. Dasselbe gilt auch für die Umformulierung der Frage in Bezug auf die moralische Qualität einer Handlung, hier als „Is tweeting right or wrong?“ (Vallor, 2016, S. 26-27). Somit macht sie deutlich, dass nicht Technik oder Technologie als wesentliche Differenzierungsmerkmale auftreten, denn weder digitale Artefakte noch Handlungen oder der Umgang mit dem digitalen Artefakt werden durch die Frage ethisch angemessen adressiert. Doch welche Fragen müssen formuliert werden, damit auch normative Urteile hinsichtlich digitaler Technik gefällt werden können, die nicht jeglicher Sinnhaftigkeit entbehren? Denn Twitter als Kommunikationsmöglichkeit ist moralisch nicht unangemessen, sondern fördert zunächst das menschliche Bedürfnis nach einem gemeinsamen Austausch. Schließlich bleibt unklar, worauf sich die Normativität beziehen soll: Sowohl Twitter als technisches Objekt als auch die Handlung, einen 200-Zeichen-Kommentar zu kommunizieren, können ohne Einbettung in soziale Strukturen, ohne Kenntnis über entsprechende Erwartungshaltungen und deren moralisch angemessene Beantwortung gar nicht oder bestenfalls nur

unzureichend beantwortet werden. Zudem betont Vallor, dass auch in anderen Erörterungen Angewandter Ethik Technik eine zentrale Rolle spielt, sei es in Abtreibungsfragen oder in der ökologischen Ethik. Die Unklarheit, die normative Orientierungslosigkeit beruht also nicht darauf, dass Informationstechnik eine Technik ist, sondern vielmehr darauf, dass es sich um eine *emerging technology* handelt. Es handelt sich hierbei also um Technik, die:

„Not yet sufficiently developed to be assignable to specific practices with clear consequences for definite stakeholders. They present open developmental *possibilities* [Herv. d. A.] for human culture as a whole, rather than fixed options from which to choose. The kind of deliberation they require, then, is entirely different from the kind of deliberation involved in former sets of problems.“ (Vallor, 2016, S. 28)

Dies ist ausschlaggebend, denn ohne moralische Überzeugungen, was als moralisch wünschenswert gilt, ohne einen *phronimos*, der als Vorbild dienen kann, muss zunächst ein normativer Bedeutungshorizont begründet werden. Somit ist deutlich geworden, dass auch die moralische Güte von Informationstechnik und Technologien, wie z.B. soziale Medien, Gesichtserkennungstechnologie oder Pflegeroboter, nur in Abhängigkeit von Handlungszusammenhängen und Praxen, in die sie implementiert werden, ethisch erörtert werden kann.

Zudem wird zugleich ersichtlich, dass für einen vernünftigen Umgang mit Informationstechnik, allein aufgrund der alltäglichen Nutzung in nahezu allen lebensweltlichen Bereichen, sowohl die Tätigkeiten der Nutzer als auch diejenigen der Entwickler betrachtet werden müssen. Folglich ist auf der einen Seite zu verstehen, welche lebensweltlichen Handlungszusammenhänge durch den Einsatz von Informationstechnik transformiert werden und ob diese Transformation zu denkbar moralisch wünschenswerten Praxen führt. Das heißt bspw., ob die Informationstechnik bestimmte Nutzer-Haltungen gegenüber dieser Technik fördert, die moralisch wünschenswerte Praxen stabilisieren oder gar fördern, diese aber nicht unterminieren. Auf der anderen Seite kann durch eine solche Kenntnis ein Berufsethos, eine *professional ethics*, artikuliert werden, die sich dezidiert auf die Etablierung ingenieursrelevanter Handlungen und Haltungen fokussiert und somit den Herstellungsprozess praktisch-moralisch anreichert. Ethisch brisant werden Informationstechnologien insbesondere dann, wenn sie als sozio-technische Systeme beschreibbar werden, d.h. wenn durch einen Einsatz sowie durch den Umgang mit Informationstechnik normative Verpflichtungsgefüge transformiert werden und dadurch der Bezug zur Welt und den Mitmenschen transformiert wird.

4.3 Vernetzung durch Transparenz und Allgegenwärtigkeit

Ein weiteres Merkmal, das sich in dieser Arbeit bereits mehrfach zeigte und sich auch grundlegend für die Reflexionsbegriffe der Transparenz und Opazität erwiesen hat, ist die allgegenwärtige Präsenz

von Software, die sich jedoch oft der direkten Wahrnehmung entzieht. Im Gegensatz zu einem Desktopcomputer ist in diesem Zusammenhang die Rede von mobilen Geräten wie Smartphones oder Tablets sowie dem kontinuierlichen Informationsaustausch mit anderen Geräten wie Autos und Kühlschränken. Hier bedarf es bestimmter Middleware, Sensorik oder Prozesse, damit dieser Informationsfluss zustande kommen kann. Dieses Phänomen wird unter dem Begriff des *ubiquitous computing* thematisiert:

Die zentralen Charakteristika des *Ubiquitous Computing* [Herv. d. A.] sind das weitgehende Verschwinden von Hardwarekomponenten und der Mensch-System-Schnittstelle, die Adaptivität und Smartness des Systems, die Selbstorganisiertheit und Kontextwahrnehmung – d. h. die Fähigkeit, Handlungssituationen zu erkennen – des Systems, die informatische Aufladung der Mesosphäre, die ubiquitäre Nutzbarkeit sowie die handlungsrelevante Verknüpfung von lokalen und globalen Informationen.“ (Wiegerling, 2013, S. 375)

So bedient eine Person ihr Auto, ohne dabei wissen zu müssen, welche Computer sich in diesem befinden oder mit wie vielen anderen Systemen diese wiederum im Austausch stehen. Die Softwaresysteme, um optimal wirksam werden zu können, gehen in ihrer Umwelt auf. Aber nicht nur die Transparenz der Softwaresysteme, sondern auch ihre Funktionalität erschließt sich dem Endnutzer nicht mehr ohne Weiteres: Weder muss man über das Wissen verfügen, welche Algorithmen zur Lösung bestimmter Probleme eingesetzt werden, noch bedarf es eines Verständnisses darüber, welche Nutzerverhalten das System steuern oder von diesem gesteuert werden. Die Zielfunktionen bleiben verborgen: So sind sie entweder transparent oder gar opak. Im Hintergrund operieren Systeme, die gewisse, dem Endnutzer unbekannt Ziele verfolgen und auf mathematischen oder kognitionspsychologischen Modellen beruhen, die sich dessen Kenntnis entziehen.

„Ubiquitous computing“ umfasst daher sowohl das „persuasive computing“, vor allem wenn es um die Verhaltenssteuerung in ökonomischen Kontexten geht, als auch den Begriff der „ambient intelligence“, der meist in Europäischen Forschungsprojekten Verwendung findet. Ebenso ist in diesem Zusammenhang oft vom Internet der Dinge die Rede. Allen Konzepten ist der kontinuierliche Informationsaustausch vieler einzelner Geräte gemein, die im Verborgenen operieren. Unter dem Begriff des „persuasive computing“ wird verstärkt der Einfluss auf das Verhalten und die Handlungen des Nutzers thematisiert. Es handelt sich hierbei um ein strategisches Konzept, das erwünschte Verhaltensweisen bewusst mittels des Einsatzes digitaler Technologie und Technik zu erzielen versucht. Aufgrund dieses Merkmals wird sich in dieser Arbeit verstärkt auf das „persuasive computing“ fokussiert, denn vor allem die Annahme der gezielten Verhaltensänderung ist auch für wünschenswerte normativ gestaltete digitale Artefakte bedeutsam: Zum einen kann man diese kognitionspsychologischen Modelle auf positive Resultate lenken, zum anderen muss man zugleich diskutieren, ob ein solcher willentlicher Eingriff, selbst bei noch so wünschenswerten Resultaten, moralisch zumutbar ist. Denn beim „persu-

asive computing‘ handelt sich nicht um die automatische Eingabe von Daten, sondern um eine Beziehung von Daten zu Handlungsintentionen des Nutzers. Der Begriff des ‚persuasive computing‘, den B.J. Fogg 1997 einführte und weiterhin prägt, verdeutlicht die Bedeutung der intentionalen Steuerung des Verhaltens oder der Wahrnehmung des Endnutzers: „Simply put, a persuasive computer is an interactive technology that changes a person’s attitudes or behaviors.“ (Fogg, 1997, S. 225)

Als ‚persuasive technology‘ versteht man daher die Menge an digitaler Technik und Technologie, die psychologische Prinzipien der Überzeugung anwendet, wie Glaubwürdigkeit, Gegenseitigkeit oder Vertrauen, um Veränderungen im Verhalten und in den Einstellungen der Nutzer hervorzurufen (de Kort et al., 2007). Das beabsichtige Eingreifen in die menschlichen Entscheidungsstrukturen ist in diesem Zusammenhang ausschlaggebend, denn auch externe Ereignisse, wie z.B. ein Regen, können zu einer Verhaltensanpassung führen. Allerdings wird die durch einen Regenschauer verursachte Verhaltensänderung nicht angemessen mit dem Prädikat „der Regen überzeugte den Menschen, sich dementsprechend zu verhalten“ beschrieben. Wesentlich für die ‚persuasive technology‘ ist somit die beabsichtige Verhaltensänderung, die auch ohne subtilen Zwang oder subtile Wahrnehmungsverchiebungen zu teils auch gewollten Verhaltensveränderungen führen kann (Berdichevsky & Neuenchwander, 1999; Oinas-Kukkonen & Harjumaa, 2009).

Während der Nutzer im *persuasive computing* noch Anerkennung findet, heben Søraker und Brey (2007) bezüglich der *ambient intelligence* den wichtigen Aspekt hervor, dass gerade keine direkten Schnittstellen mehr zum Endnutzer notwendig sind, weshalb sein Verhalten, seine Handlungsgründe gar nicht mehr zentral sind. Vielmehr berechnet ein Algorithmus aus einer Menge an Daten, die Verhaltensspuren der Nutzer sind, Muster und triggert durch den berechneten Output wiederum ein bestimmtes Nutzerverhalten (ebd.). „Ambient Intelligence differs from traditional IT in the sense that we no longer consider what our desires are and interact with the device (behave) accordingly. Instead, we leave it up to the device itself to infer – what we really want on the basis of our natural behaviour.“ (ebd, S. 9)

Ambient intelligence beinhaltet somit auch die Annahme, dass Software die Wünsche und Vorlieben einer Person schneller und besser aus ihren Datenspuren herauslesen kann, als es den Personen häufig selbst gelingt. Die Software greift dadurch dem Selbsterkenntnisprozess der Person voraus, indem vergangenes Verhalten auf zukünftige Empfehlungen hin abgebildet wird, z.B. durch *predictive algorithms*. Auch hier gilt, dass dem Endnutzer die berechnenden Prozesse nicht ersichtlich oder erklärbar sind. Das Stattfinden dieses Austausches sowie die Art und Weise des Austausches entziehen sich der Wahrnehmung: Zum einen, weil die Technologie und Technik nicht jedem Endnutzer epistemisch zugänglich gemacht werden kann und zum anderen, weil die offensichtliche Schnittstelle

zum Menschen zunehmend verschwindet. Systeme und Geräte sind so für die Endnutzer nicht sichtbar, d.h. sie sind transparent. Die Systeme steuern sich dabei häufig selbst, d.h. aus den Daten des Nutzerverhaltens werden weitere Schritte in Gang gesetzt. Dabei steuert das System wiederum das Verhalten des Nutzers, ohne diesen aktiv in die Entscheidungssituation einzubeziehen. Solche Softwaresysteme befinden sich zunehmend überall und vernetzen unterschiedliche gesellschaftliche Bereiche oder unterschiedliche individuelle Rollen miteinander. Die Systeme bleiben aufgrund ihrer Komplexität in ihrer technischen Funktionalität für die Endnutzer opak. Diese Charakterisierung trifft nur auf Automaten zu und kann nicht von Geräten oder Maschinen erfüllt werden, d.h. es handelt sich um genuine, Software-spezifische Merkmale.

Tristan Harris, ein ehemaliger Mitarbeiter von Google und 2015 Mit-Begründer des *Center for Humane Technology* – einer Bewegung, die sich um die Belange der IT-Konsumenten bemüht und als Auslöser des *Time Well Spent Movement* gilt – verknüpft den Aspekt des manipulativen Designs von Technik mit den Zielen von Wirtschaftsinteressen auf Seiten von Unternehmen. Unternehmen würden, so Harris (2016), die Autonomie der Nutzer absichtlich unterminieren, indem sie gewollt durch technische Applikationen oder Features bestimmte Nutzerverhaltensweisen hervorrufen, um den Nutzer als Kunden bestmöglich an ihr Produkt zu binden. Dies gelingt, indem das Design des technischen Artefakts so ausgestaltet wird, dass es den Nutzer dahingehend verführt, bspw. das Produkt häufig zu verwenden. Dienlich sind hierfür z.B. kontinuierliche Benachrichtigungen oder ein systematischer Einsatz normativ aufgeladener Motivation, wie das Anerkennungsbedürfnis durch *likes* in Sozialen Medien sehr eindrücklich zeigt, sowie das Phänomen, dass *dislikes* vermehrt zu schädlichen, unkooperativen Verhaltensformen führen können. YouTube gab 2021 daher die Abschaffung der Dislike-Funktion bekannt (Garcia, Abisheva & Schweitzer, 2017; Khan, 2017). Dieser Eingriff in die menschlichen Handlungsstrukturen hinterlässt Spuren und verändert die lebensweltlichen Bedingungen menschlichen Miteinanders, weil auf normative Konstituenzen eingewirkt wird. Gerade dieser subtile Druck, den digitale Artefakte auf diese Eckpfeiler gemeinsamer Verhaltensweisen ausüben, führt auch dazu, dass „[e]ntgegen unserer persönlichen Wahrnehmung [...] unser exzessives Nutzungsverhalten nämlich in den wenigsten Fällen eine absichtliche Entscheidung [ist NZ]. Die Strukturen der Interaktion und die Algorithmen, auf denen die meisten Apps basieren, sprechen bei uns unterbewusste Automatismen an, die uns instinktiv zu einer Handlung verleiten“ (Markowitz, 2015, S. 25). Allerdings sei an dieser Stelle im Gegensatz zu Markowitz betont, dass es nicht nur um die Auslösung unterbewusster Reize-Reaktion-Mechanismen geht, sondern dass gerade die Verschiebung normativer Orientierungscluster vormals rationale, unbewusste, aber normativ angemessene Verhaltensreaktionen zu normativ fragwürdigem Verhalten transformieren kann.

Eine beabsichtigte Manipulation von Nutzerverhalten, die den Gedanken der Wertneutralität technischer Artefakte ohnehin aufhebt, verdeutlicht zugleich, dass sich Normativität hinsichtlich digitaler Artefakte nicht in einer reinen Güterabwägung erschöpft, sondern dass Softwaresysteme aus der Perspektive psychologisch-individueller Bedürfnisse und Heuristiken gedacht werden, um vollfunktionsfähige Applikationen, d.h. nicht nur technisch funktionale Software, entwickeln zu können. Dabei werden normative Anforderungen zwischenmenschlicher Beziehungsgeflechte auf die Interaktion von Menschen und Maschinen übertragen (Fogg, 2003). Fogg, Gründer und Leiter des *Behavior Design Lab* der Universität Stanford (<https://captology.stanford.edu/>), greift dabei dezidiert die Frage nach der Ausgestaltung von Softwaresystemen und Hardwarekomponenten auf, um sich der Fragestellung einer Verhaltensänderung und somit Verhaltenssteuerung anzunähern. Er stellt die Gestaltbarkeit von Computersystemen aus Nutzersicht und deren Reaktionen auf bestimmte Features in den Mittelpunkt seiner Betrachtung. Die Verhaltensanpassung wird dadurch zu einer steuerbaren Verhaltensänderung. Hierbei dreht sich bei der Formierung von Technik alles um die *usability* von Mensch-Maschinen-Interaktionen, die meist mit Hilfe der *5-Es* gemessen wird: *efficient, effective, easy to learn, engaging, error tolerant* (Albers & Mazur, 2014). Jedoch stellt sich bereits bei der Messung von *usability* die Frage nach der Metrik, denn so gilt es, das Nutzererlebnis eher nach Qualitäten zu unterscheiden, wie bspw. nach mehr Freude oder nach mehr gelebter Gemeinschaft. Letzteres betont sogleich die Werteausgerichtetheit menschlichen Daseins und überführt eine ökonomisch-soziologische Betrachtung von *usability* in eine Ethik, die auch Vallor (2016) techno-moralisch einzufangen versucht. Um wünschenswerte Tugenden zu formulieren, verweist sie zunächst auf vier Charaktereigenschaften, die sie für das technische Zeitalter als unangemessen deklariert:

Consider the person who: 1) is characteristically incapable of empathizing with or giving moral consideration to others beyond their local circle; 2) is unable to communicate or deliberate well with others, especially with those holding different metaphysical and value commitments; 3) reasons unusually poorly in circumstances involving great uncertainty and risk; and 4) has concerns for the good life that rarely if ever extend beyond maximizing gains in the present and immediately foreseeable future. (Vallor, 2016, S. 32)

Eine Person, die diese vier Wesenszüge aufweist, kann in der vernetzten Welt keine positiven Akzente setzen, denn gerade hier treffen Personen aus den verschiedensten Ländern, mit den unterschiedlichsten Interessen und divergierenden Werten aufeinander. Auch ist der Einsatz der digitalen Technik aufgrund ihres jungen Alters nicht immer vorhersehbar, weshalb sie auch in langfristiges Denken eingebunden werden muss, wenn die Bewertung nicht nur nach technischer Funktionalität und ökonomischem Nutzen erfolgen soll. Wenn diejenigen Charaktermerkmale – oder klassisch ausgedrückt, wenn diejenigen Tugenden ausfindig gemacht werden können, die zu wünschenswerten Handlungsweisen mit und Ausgestaltungsweisen von digitaler Technik führen, dann können diese Artefakte auf eine normative, wünschenswerte Weise konstruiert werden sowie positiv auf jene Ver-

haltensweisen einwirken, die zu einem selbstbestimmten Leben beitragen. Wenn dieser tugendethischen Technikethik Gehör geschenkt wird und ihr Kraft verliehen werden soll, dann müsste Fogs kognitionspsychologischer Ansatz mit der Idee moralisch wünschenswerter Praxen verbunden werden. Sodann müssen Technik und Technologie nicht nur in einer individualistisch geprägten Autonomieforderung gedacht werden, sondern zudem Sozialität und Kollektivität mit in unsere Betrachtung einbeziehen: Der Mensch ist nicht allein frei.

Die Idee, Technik so zu entwickeln, dass sie das Verhalten der Nutzer positiv ändern kann, stützt sich auch auf die ökonomische *Nudging-Theorie*, wie sie von Thaler und Sunstein eingeführt wurde: Diese betont, dass Fehlverhalten – verstanden als ein suboptimales oder irrationales Verhalten – wieder auf die richtige, optimale, da ökonomisch-rationale Spur gebracht werden kann. Dies geschieht nicht notwendigerweise durch eine rationale Reflektion, die den Akteuren aufgrund des Verlassens der Handlungsebene oft mühsame Anstrengungen abverlangt, sondern bereits durch kleine Anreize, die auf der Handlungsebene integriert werden und zu großen Verhaltensänderungen führen können (Thaler & Sunstein, 2009). Durch die Einführung bestimmter Entscheidungsarchitekturen können Menschen so beeinflusst werden, dass sie ohne große Mühen, folglich ohne die Anstrengungen des *reasoning*, die rational-optimale Entscheidung treffen. Strenggenommen kann hier nicht von rationaler Entscheidung gesprochen werden, denn das Individuum wird in seinem Verhalten von außen gelenkt – es reagiert auf die äußeren Umstände. Die vom Individuum nicht wahrnehmbare Manipulation führt in ethische Probleme des Paternalismus und zur Frage danach, inwiefern es zumutbar ist, Menschen zu ihrem Glück zu zwingen, für sie Entscheidungen zu fällen, sie in eine bestimmte Richtung zu lenken (Dittmeyer, 2020; Fischer, 2017). Hier stellen sich Fragen nach der Überbetonung kollektiver versus individueller Autonomie, die hier nur angerissen werden können.

Aufgrund dieser manipulativen Macht und gezielten strategischen Ausrichtung digitaler Technologie und Technik räumt Fogg (2003, S. 2012) ethischen Beurteilungen seiner Überredungsmaschinerie großen Raum ein und betont mehrfach: „The key for those who design, study, or use persuasive technologies is to become sensitive to the range of ethical issues involved.“

Durch die Aufnahme der gezielten Verhaltensänderung mittels bewusst gesetzter Design-Akzente kann Technik nicht mehr unabhängig vom Nutzer gedacht werden und das Forschungsprogramm muss um eine Human-Computer-Interaction, d.h. ein menschlich-zentriertes Design, erweitert werden. Der Fokus auf *human-centred Design*, der aus dem Ansatz der Mensch-Maschinen-Interaktion resultiert, kann aber nicht ohne Weiteres mit der humanistischen Forderung gleichgesetzt werden. Die humanistische Forderung umfasst gerade, dass diese genuin menschliche Fähigkeit zur Autonomie des Einzelnen innerhalb einer Gemeinschaft in ihren spezifischen Rollen zu gewährleisten und zu fördern sei, aber nicht unterminiert werden dürfe. Somit werden aus einer humanistischen Sicht

sogleich ethische Gesichtspunkte relevant, die ein Nutzer-basiertes Design nicht notwendigerweise beinhalten muss. Die Fähigkeit, das eigene Leben eigenständig zu gestalten, gründet auf der Fähigkeit des Gründegebens und Gründenehmens und somit auf einer vernünftigen Ausrichtung der eigenen Lebenspläne, die wiederum als eigenständige Zwecksetzungsrationalität zu verstehen sind (Nida-Rümelin, 2020). Diese Autonomiefähigkeit, ein selbstbestimmtes Leben zu gestalten, kann das Individuum nicht allein erreichen, sondern sie ist gestützt durch ein Miteinander, das soziale Autonomien und kollektive Freiheiten zu berücksichtigen gedenkt. Eine radikale Autonomie, die nur die individuellen Bedürfnisse in den Mittelpunkt stellt, scheitert demnach schon an der Lebensrealität der Akteure, die jedoch die Grundlage einer soliden ethischen Reflexion bildet.

Die Allgegenwärtigkeit, die Transparenz und die mangelnde Nachvollziehbarkeit digitaler Artefakte führen, wie es König (2013) in seinem Artikel *VDI-Richtlinie in der Technikbewertung* beschreibt, zu einem Kontrollverlust, da die Widerspenstigkeit der Realität eliminiert wurde. Dem Nutzer bleibt unklar, in welchem Szenario er sich gerade aufhält. Der Kontrollverlust, der sich bereits durch die Werthärenz der *built in tendencies* der Softwaresysteme bemerkbar machte, wird durch diese Form der Technologie verstärkt. Um diesem Kontrollverlust entgegenzuwirken, entstand das Konzept der Parallelkommunikation, das als Versuch verstanden werden muss, Probleme wie die Entmündigung des Nutzers und die Unkontrollierbarkeit der Technik abzumildern (Wiegerling, 2011). Diese Parallelkommunikation soll dem Nutzer ad hoc durch im System implementierte Kanäle über das Zustandekommen der Information sowie der Metainformation Auskunft geben: Dem Nutzer soll eine Möglichkeit der Systemkontrolle nach Plausibilitätskriterien und eine Eingriffsmöglichkeit in Systemkonfigurationen gewährt werden. Parallelkommunikation ist für Laien entwickelt worden und stellt die Entlastungsfunktion der digitalen Technologie nicht infrage. Es geht darum, die im Hintergrund agierenden Systeme in ihrer Funktions- und Wirkungsweise wahrnehmbar zu machen sowie einen Beitrag zur Kompetenzwahrung bzw. -wiedererlangung zu leisten. Im Zentrum steht die Autonomiewahrung des Nutzers gegenüber dem System durch die Aufweisung und Gewährung von Handlungsoptionen sowie Ausstiegsmöglichkeiten aus der Systembegleitung.

Fallmann (2007) spitzt die Frage nach der Qualität des normativen Moments jedoch zu, indem er die Normativität sogleich in Zusammenhang mit einem offenen Entscheidungsspielraum bringt. Nicht mehr die Offenlegung aller Momente, die sich der Wahrnehmung und Erkenntnis entziehen, ist Gegenstand der Überlegung, sondern vielmehr gilt es, die Gütekriterien hinsichtlich ihrer Ausrichtung zu überdenken. Denn welches qualitative Moment ist ausschlaggebend, damit ein Computersystem als moralisch zumutbar beurteilt werden kann: „One that persuades one more [Herv. d. A.] or one that persuades one into something good [Herv. d. A.] ? Who is to decide?“ (Fallman, 2007, S. 296)

4.4 Die Formbarkeit digitaler Artefakte: Der Überschuss an Zwecken

Die Vorhersehbarkeit aller positiven oder negativen Folgen, die aus einem Einsatz resultieren, sowie eine ausreichende Konzeptualisierung struktureller Lebensformen, die eine mögliche Transformation durchlaufen, sobald die Technik oder Technologie eingebunden wurde, ist allein aufgrund des inhärent dynamischen Charakters digitaler Technologie und Technik unmöglich. Bereits die schiere Anzahl an Möglichkeiten, für die diese digitalen Technologien und Technik eingesetzt werden können, erschwert den Versuch, normativ gezielt einzugreifen. Wie bereits in Teil III, Kapitel 2 dargelegt, führt die Verselbständigung der Mittel zu einer nicht einzufangenden Dynamik, die jedoch für den innovativen Antrieb dieser Technologien und Technik wesentlich ist. Diese allgemeine Zweckdienlichkeit verdeutlicht sich bspw. an der Verselbständigung der Idee, die SMS als Fehlermeldung intern zu entwickeln. Erst nach und nach etablierte sich die SMS als Kommunikation zwischen Privatpersonen (Wietlisbach, 2017). Die normativen Dimensionen der Informationstechnik sind aufgrund der vielfältigen Zweckdienlichkeiten nur schwer greifbar, was Moor (1985, S. 27) in seinem Aufsatz *What is Computer Ethics?* mit dem Label *malleable* beschreibt:

„Because computers are logically malleable, they will continue to be applied in unpredictable ways, generating numerous policy vacuums for the foreseeable future. Moreover, because computerized situations often become informationally enriched, we will continue to find ourselves in conceptual muddles about how precisely to understand these situations.“

Allerdings betont Moor vor allem, dass nicht nur die Einsatzmöglichkeiten unbekannt seien, sondern auch die damit verbundenen ethischen und legalen Kategorien (ebd.). Er diskutiert diese unbeabsichtigte Dynamik anhand einiger Konzepte, wie der Privatsphäre oder des Stehlens. So zeigt er auf, dass man im Digitalen nicht mehr ohne Weiteres von einem klassischen Diebstahl sprechen kann, da die Originaldatei nicht entwendet, sondern kopiert werden müsste. Somit werden solche Situationen *informationally enriched*: Hierunter fällt die Veränderung ethischer und legaler Konzepte, die durch die Einführung binärer Informationen verursacht wird. Ethische Imperative – wie, dass man nicht stehlen möge – werden im Zusammenhang mit dem Kopieren von Dateien herausgefordert, da hier strenggenommen nichts gestohlen, also im Traditionellen nichts weggenommen wird. Sie dienen keiner hinreichenden Handlungsorientierung mehr und bedürfen einer anderen ethischen und legalen Handhabung. Moor (1985) verweist darauf, dass es bei digitaler Technologie und Technik verstärkt darum geht, dass es an normativer Orientierung im Umgang mit diesen Transformationen mangelt. Er bezeichnet dies als „policy vacuum“ (ebd. S. 26). Allerdings, wie in dieser Arbeit bereits erörtert, gilt es nicht nur, eine politische, externe, normative Orientierungslosigkeit zu adressieren und Regulierungsbestrebungen einzufordern. Vielmehr müssen sich Praxen, verstanden als etablierte, wünschenswerte Lebensformen, ausbilden. Dies kann auch beabsichtigt erfolgen und benötigt ausreichend Zeit.

Die fehlende Orientierung, die sich aus dem dynamischen Charakter digitaler Technologien und Technik ergibt, beruht folglich auf nicht etablierten Verhaltensweisen, die sich aufgrund der Schnelllebigkeit nur schwer oder zu langsam ausbilden können. Informationstechnik als Gebrauchsgegenstand wird gerade aufgrund der Ahnungslosigkeit über die für Informationstechnik relevanten, moralisch wünschenswerten Strukturen zu einer Quelle lebensweltlicher Orientierungslosigkeit.

Die Mehrzweckdienlichkeit, das Unwissen über zukünftige Zustände sowie die Ungenauigkeit moralischer Begrifflichkeiten verstärken den bereits aufgerührten Charakter einer *emerging technology*. Dieser Überschuss an Zwecken begründet jedoch auch gerade die *economies of scope*: Digitale Produkte, die für einen bestimmten Zweck entwickelt wurden, lassen sich meist schnell und einfach für andere Zwecke einsetzen. Diese Schnelllebigkeit, die sich im vielfältigen Einsatz zeigt, ist keiner anderen Technologie oder Technik eigen und zeigt sich auch in der Schwierigkeit, digitale Technologie und Technik einer gezielten Bereichsethik zuzuordnen.

If I had told you 25 years ago that, in a quarter century's time, one-third of the human race would be communicating with one another in huge global networks of hundreds of millions of people—exchanging audio, video, and text—and that the combined knowledge of the world would be accessible from a cellphone, that any single individual could post a new idea, introduce a product, or pass a thought to a billion people simultaneously, and that the cost of doing so would be nearly free, you would have shaken your head in disbelief. All are now reality. (Rifkin, 2014, S. 134)

Digitale Technologie und Technik sind zudem einzigartig, weil sie wie kein anderes technisches Instrument nicht nur *economies of scope*, sondern auch *economies of scale* ermöglichen. Sobald ein digitales Artefakt technisch funktioniert, lässt es sich aufgrund seiner diskreten Natur billig vervielfachen, auf den Markt bringen und sogar relativ schnell wieder ändern, wenn dies notwendig sein sollte. Dieses Phänomen beschreibt Rifkin in seiner Theorie einer digitalen Ökonomie als eine Null-Grenzkosten-Gesellschaft: Es gibt Fixkosten, aber nur geringe variable Kosten. Allerdings bleibt die Herstellung des ersten Produkts aufgrund der forschungsintensiven Phase kostspielig (Agar, 2019). Nach dem Prototyping ist das Produkt jedoch einfach skalierbar: So mag die Entwicklung der ersten Version eines Textverarbeitungsprogramms noch aufwendig sein, wenn es sich jedoch als technisch funktional erwiesen hat, lässt es sich vielfach kopieren. Die Massenproduktion, die aufgrund der Kopierbarkeit einfach erzielt werden kann, ist unbestritten ein weiteres Alleinstellungsmerkmal digitaler Technologie und Technik.

4.5 Die Machtposition digitaler Technologie und Technik

Digitale Artefakte sind vielfältig: Man denke an Textverarbeitungsprogramme, Steuerungssysteme, Tamagochis, Games, Webseiten bis hin zu Plattformen. Digitalen Artefakten war zunächst gemein, dass sie sowohl aufgrund ihres Überschusses an Zwecken als auch aufgrund ihrer diskreten Natur

eine gewisse kontinuierliche Dynamik sowohl in ihrer Erscheinung als auch in ihrer Anwendung aufwiesen. Die schnelle Vervielfältigung, die ihre binäre Charakterisierung ermöglicht, führte zudem in eine ungeheure Nutzerreichweite: Ein Textverarbeitungsprogramm lässt sich weltweit auf unzähligen Rechnern installieren und nutzen. Auf eine Plattform kann von vielen simultan zugegriffen werden. Ebenso ist es möglich, sich zugleich in Echtzeit mit anderen Nutzern global auszutauschen. Diese Reichweite und die unermüdlichen Einsatzmöglichkeiten verleihen digitalen Artefakten eine unglaubliche Macht. Kurz: Sie sind ein Massenprodukt. Aber nicht nur die Eigenschaft der Massenproduktion lässt digitale Artefakte tonangebend werden, sondern auch ihr vermehrter Einsatz in sozialen und emotionalen Bereichen oder in Gemeinsinn stiftenden Praktiken etabliert eine durch Technologie und Technik nie dagewesene Machtstruktur. Gerade weil digitale Artefakte nicht nur lebensweltliche kollektive Praktiken nachhaltig transformieren können, sondern auch gänzlich neue Handlungs- und Denkräume eröffnen, üben sie Druck auf die normativen Konstituenzen der Lebensformen ihrer Nutzer aus. Die in dieser Technologie innewohnenden Möglichkeiten verleihen digitalen Produkten, Herstellungspraktiken und dem dazu notwendigen Wissen eine zuvor nicht dagewesene Macht, die sich am eindrucklichsten am Beispiel von Plattformen demonstrieren lässt.

Plantin und Punathambekar (2016) verdeutlichen anhand ihrer Fallstudien von Google und Facebook die Herrschaftsstellung, die sich aus einer plattformig ausgerichteten Infrastruktur ergibt. Zwar wurde digitale Technologie zu Beginn vermehrt als öffentliches Gut behandelt, was sich in den hohen staatlichen Investitionen zwischen 1960-1980 zeigte, die – vom öffentlichen Interesse gestützt – getätigt wurden (Plantin & Punathambekar 2019, S. 11 ff.). Erst die Privatisierungsbestrebungen im Zuge des Neoliberalismus führten ab den 1980er Jahren dazu, dass nicht nur staatliche oder staatlich überwachte, private Monopole wie die Post oder die Telekommunikation, sondern auch andere, kleinere Anbieter konkurrenzfähige (digitale) Infrastrukturgüter zur Verfügung stellen konnten. Insbesondere sollte so einer Explosion der Konsumentenpreise entgegengewirkt werden. Selbstredend blieben staatliche Regulierungsbestrebungen hinsichtlich der Bereitstellung oder Herstellung digitaler Güter bestehen. Dieser Trend der Entstaatlichung, so die Autoren, führte schlussendlich dazu, dass sich auch die Plattformökonomie immer mehr mit der Terminologie öffentlicher Güter einfangen ließ. „[T]he rise of digital technologies, in a neoliberal, political, and economic climate, has simultaneously facilitated a ‚platformization‘ of infrastructures and an ‚infrastructuralization‘ of platforms.“ (Plantin et al., 2018, S. 301)

Infolge dieser Feststellung analysieren Plantin und Punathambekar (2018) digitale, netzförmige Artefakte nicht mehr nur ausschließlich mit Theorien der Plattformökonomie oder der Infrastrukturtheorie, sondern verknüpfen die wesentlichen Aspekte beider Theorien, um den Charakter dieser digitalen Artefakte vollständig erklärbar zu machen. Google, Amazon oder Facebook verfolgten bei ihrer Gründung noch bestimmte Zwecke, wie Internetsuche, Einkauf oder Kommunikation, weshalb sie deutlich

als Plattformen beschrieben werden können. Plattformen weisen die Merkmale der Programmierbarkeit und Erschwinglichkeit, die Verbindung heterogener Akteure sowie die Zugänglichkeit von Daten und Logik durch Anwendungsprogrammierschnittstellen (API) auf (Platin & Punathambekar, 2019). Plattformen haben somit einen festen Software-Kern, der sich meist modular erweitern lässt und dabei auf ökonomische Gewinne und Nutzerbefindlichkeiten ausgerichtet bleibt. Ihre Standardisierung bestimmen die Plattformen häufig selbst, weshalb sie sich meist gegenseitig ausschließen, d.h. ein Umzug von einer Plattform zu einer anderen ist nicht ohne Opportunitätskosten möglich. Die Opportunitätskosten steigen mit der Zeit meist an, denn auch der Nutzer profitiert von der Vielzahl der Mitnutzer: Hier findet er mehr potentielle Kunden, mehr Austausch, mehr Informationen oder eine verbesserte Kartographie. Plattformen können klein sein, also nur wenige Personen verbinden, aber können über die Zeit hinweg eine große Reichweite erzielen. So nutzten 2021 1,9 Millionen Personen Facebook täglich, was einem Wachstum von 6,8 % gegenüber dem Vorjahr entsprach (Dean, 2022). Der Zugang wird über ein dem Nutzer zur Verfügung stehendes Gerät sowie dessen finanzielle Zusage mittels Online-Abonnements, *pay-per-use*-Funktionen oder über geschaltete Werbeeinnahmen reguliert. Häufig ist es dem Endnutzer möglich, sich durch eigene Kreationen an der Plattform zu beteiligen, sei es in Form von Kurzbeiträgen, Suchaufträgen oder durch das Betätigen eines Einkaufs. Die Plattformen wirken sich somit auf die Ausdrucksmöglichkeiten ihrer Nutzer aus, aber auch die Plattform selbst wird durch die Bedienung, die ökonomischen Anforderungen sowie durch technologische Sachzwänge geformt.

Allerdings weisen Plattformen zunehmend Merkmale globaler Infrastruktur auf (Plantin & Punathambekar, 2019). Im Gegensatz zu Plattformen versteht man unter einer Infrastruktur zunächst Systeme, wie die Telekommunikation oder die gesamtgesellschaftliche Versorgung mit Elektrizität. Diese werden mit den Merkmalen der Allgegenwärtigkeit, Zuverlässigkeit, Transparenz, Zugangsmöglichkeit, Langfristigkeit und, wenn eine Versorgung ausbleibt, mit einem Zusammenbruch identifiziert (ebd.). Im Gegensatz zu Plattformen zielen Infrastrukturen auf ein gesamtgesellschaftliches Ziel ab, das gemeinsam verwaltet wird, weshalb sie als öffentliches Gut bezeichnet werden können. Meist kann man mit einem einmaligen Kauf eines Tickets an der Infrastruktur teilnehmen, bspw. am öffentlichen Nahverkehr, oder aber Teilnehmende müssen explizit austreten, wenn sie das Gut nicht konsumieren wollen. Bei einigen öffentlichen Gütern wird ein Austritt nicht gewährt. Bspw. ist dies in Deutschland im Falle der Bildung durch die Schulpflicht bis zu einem gewissen Alter nicht gestattet. Infrastrukturen müssen nachhaltig geplant werden, da sie wesentlich kultur- und austauschstiftend sind und sich ein Ausfall gravierend auf das gesellschaftliche Leben auswirken würde.

Im Zuge der Zeit haben sich diese Plattformen zu Infrastrukturen ausgebildet, die das soziale Leben bedingen, indem sie den Zugang zu bestimmten Austauschformen oder Kulturgütern regeln. Sie üben dadurch auch Druck auf die Ökonomie selbst aus, denn auch hier gilt fortan, dass Firmen über digitale

Plattformen präsent sein müssen. Somit fungieren Plattformen als digitale Marktstrukturen, aber ihre gesellschaftliche Funktionalität lässt sich nicht auf diese Marktfunktion reduzieren: Sie kanalisieren, sie organisieren und strukturieren zwischenmenschliche Beziehungsgeflechte in allen gesellschaftlichen Bereichen. Plattformen ermöglichen auf einen Schlag, Millionen Personen zu geringen Kosten zu kontaktieren und weisen somit für jeden eine nie dagewesene Möglichkeit auf, seine Ideen oder Produkte global mit nur geringem Aufwand und wenig Wissen feilzubieten oder kundzugeben zu können. Diese Merkmale verstärken jedoch den Einfluss von Plattformen und lassen sie Charakteristika von Infrastrukturen annehmen: Wer über diese Herrschaftsstruktur verfügt, hat potentiellen Zugang zu vielen Nutzern, kann die Nutzerdaten für weitere Zwecke verwenden und übt auf die Teilnehmenden mittels *dark pattern*, *nudging* oder aufgrund der Wahl oder Bereitstellung von bestimmtem *content* Einfluss aus. Zudem bleibt der Nutzer aus Sicht der Plattformbetreibenden meist Konsument und wird nicht in seiner Funktion als Bürger betrachtet, obgleich er gerade auf diese bürgerlichen Werten und Pflichten angesprochen wird, wie es bspw. der Skandal um *Cambridge Analytica* aufzeigt.

Infrastrukturen werden jedoch erst dann verwirklicht, wenn sie in alltägliche Lebenspraktiken eingebunden sind, d.h. wenn sie wesentlich zum Erhalt dieser Lebensform beitragen, in diesen aufgehen oder gar selbst eine neue Praxis etablieren (Plantin & Punathambekar, 2019). Ein Ausschluss ist dann als moralisch unzumutbar auszuweisen, wenn dadurch das Mitwirken an gesellschaftlich relevanten Praxen nicht mehr möglich ist, wie es auch die Debatte um das Internet als Menschenrecht zeigt (Klimke, 2020; Voigt, 2014).

Alle vormals skizzierten Kriterien, die digitale Technologie und Technik als eine einzigartige Verfahrensweise mit und Zugangsmöglichkeit zu der Welt aufzeigen, konstituieren einen mehrdimensionalen Herrschaftsraum. Allerdings kann über die Verortung der Optionen- und Lösungsräume, in welchen die Medialität, der Überschuss an Zwecken ebenso wie die potentiell ubiquitäre, transparente Präsenz sowie die Wertinhärenz digitaler Artefakte verstärkt zum Vorschein treten, eine Verantwortungszuschreibung ermöglicht werden. Hierfür kann wiederum die Methode der *levels of abstraction* hilfreich sein, indem die verschiedenen normativ relevanten Sachverhalte hinsichtlich ihrer Ausgestaltungsmöglichkeit und Wertigkeit lokalisiert und in ihrer Relationalität adäquat thematisiert werden. Dies kann durch die Zuordnung dieser Phänomene zu den jeweiligen Handlungsbereichen gelingen. Die unterschiedlichen Handlungskompetenzen können wiederum einer Verantwortungszuschreibung dienen, um mit der herausfordernden Mächtigkeit, die den für Informationstechnologie und -technik typischen Phänomenen entspringt, normativ zielgerichtet umzugehen. Die Opazität, die vielen digitalen Artefakten und Technologien inhärent ist oder durch deren strukturelle Einbettung auftritt, erschwert die Offenlegung von Verantwortungsstrukturen und bestärkt dadurch Machtverhältnisse. Deshalb wurde auch in Teil II, Kapitel 3 Opazität als Reflexionsbegriff eingeführt, der den Deliberationsprozess kontinuierlich anleiten soll.

Normativ relevante Sachverhalte, die aus der technologischen Beschaffenheit sowie deren Einsatz resultieren, bedürfen nicht nur mühsamer, epistemischer Anstrengungen, sondern fordern praktisch, evaluativ Stellung zu beziehen. Es gilt hier, den ethisch schwierigen und komplexen Umständen mit Entscheidungskraft entgegenzutreten. Dies führt zu Fragen der Verantwortungsübernahme dahingehend, wer im Rahmen digitaler Technologien und Techniken für welche normativen Handlungsräume verantwortlich gemacht werden kann.

4.6 Verantwortungsdiffusion

Die Vielfalt der informationstechnischen Artefakte zeigt sich darin, dass hierunter riesige Stand-alone-Computer, private Desktoprechner oder Laptops, Netzwerke zum Austausch von Informationen zwischen diese Computer und Roboter fallen. Zudem zählen hierzu alle Gegenstände, die sich binär erfassen lassen, wie Software, Programme, Anwendungen oder kurz: alles, was Maschinen programmierbar macht. Allen diesen Artefakten ist gemein, dass sie eine bestimmte Interaktivität mit der Umwelt aufweisen, also nicht mehr ausschließlich durch den Menschen gesteuert werden müssen, sondern über Schnittstellen, wie z.B. Sensoren, zur Umwelt verfügen, und dadurch mit der Außenwelt interagieren. Ihr Einsatz in menschlichen, interaktiven Handlungsketten wirkt demnach auf die Handlungsmotivation und bestimmt somit auch Handlungsabfolgen. Indem sich digitale Artefakte zwischen den Menschen und seine Mitwelt schieben, strukturieren sie Handlungsabläufe und gestalten Lebensräume mit. Digitale Technik mischt sich in zwischenmenschliche Interaktionsgefüge oder individuelle Deliberations- und Dezisionsverfahren ein, indem durch eben jene Technik bestimmte Ausdrucksweisen geboten und andere nicht ermöglicht werden oder Empfehlungen sowie prognostische Resultate die Entscheidungsfindung beeinflussen. Diese Fragestellung nach der Zumutbarkeit solcher medialer Vermittlung darf nicht mit der Frage nach der Möglichkeit autonom agierender Systeme verwechselt werden, solange diesen mentale Zustände, also eine maschinelle Intentionalität oder Emotionalität, zugeschrieben werden. Die Erhaltung oder Förderung der Nutzerautonomie muss auch bei Systemen, die nicht auf Künstlicher Intelligenz – mit dieser oft einhergehenden hypothetischen Zustandszuschreibung – beruhen, sondern bei jeder programmierbaren Maschine kritisch hinterfragt und praktisch gewährleistet werden. Ob ein Nutzer nämlich noch autonom oder in seinen Entscheidungen und Handlungen bereits durch teils gar nicht sichtbare Systeme manipuliert wird, ist eine drängende ethische Frage, die sich bei der Gestaltung und Entwicklung digitaler Artefakte unmittelbar stellen sollte.

Unvermittelte Handlungen, die auf schlechten Entscheidungen autonomer Personen beruhen und zu unzumutbaren Ereignissen führen, werden handelnden Personen zugeschrieben, sofern diesen AI-

alternativen zur Verfügung standen oder sie die Handlung auch hätten unterlassen können. Diese Personen werden für diese Handlungen zur Verantwortung gezogen: Sie werden legal belangt oder moralisch verachtet. Was geschieht jedoch, wenn diese Entscheidung Software-gestützt getroffen wurde und sich der Nutzer maßgeblich durch Software beeinflusst sieht, z.B., weil er sich auf das Softwaresystem beim Einparken verließ und dabei ein Kind übersah oder wenn ein Arzt aufgrund von prognostischen Diagnosetools eine falsche Therapie einleitet?

Auf kollektiver Ebene lässt sich durch den Einsatz digitaler Systeme eine Unterhöhlung ganzer Handlungsräume erkennen, wie bspw. die Auswirkungen von Social Media auf demokratische Grundprinzipien, die den demokratischen Lebensraum tragen, gezeigt haben (Teil I, Kapitel 2.3 und Teil II, 7.2). Sollten die Produzenten dieser digitalen Artefakte für die größtenteils unerwünschten oder gar unzumutbaren Handlungsverzerrungen verantwortlich gemacht werden? Sollte ein Unternehmen wie Facebook die Verantwortung übernehmen, weil die Bereitstellung dieser Plattform gerade Verhaltensweisen ihrer Konsumenten förderte, die zu gesellschaftlichem Aufruhr und in eine gesellschaftliche Zerspaltung führten? Oder gilt es, die Nutzer zu adressieren, die auf den Plattformen zuweilen nicht wissen, wie sie sich angemessen zu verhalten haben? Hier spielt sicherlich eine Rolle, ob diese Artefakte als Medien, Plattformen oder als öffentliche Infrastruktur behandelt werden (Plantin & Punathambekar, 2019).

Die angerissenen Beispiele, die allesamt aufgrund ihrer Komplexität lohnend nachzuzeichnen sind, verdeutlichen die Schwierigkeit einer Verantwortungszuschreibung im Zusammenhang mit digitaler Technologie und Technik. Wie viel Verantwortung der Gesellschaft, dem Nutzer, dem Unternehmen und dem Entwickler zugemutet werden kann, ist womöglich nur kasuistisch und kontextuell, d.h. im Einzelfall oder für bestimmte Typen digitaler Artefakte (Gesichtserkennungssysteme im polizeilichen Einsatz, Kommunikationsplattformen, Fitnesstracker...), bestimmbar. Denn es geht nicht mehr nur um die Vermeidung technischer Fehler, die zu unerwünschten lebensweltlichen Resultaten führen, sondern um ein normativ angemessenes Design, welches die Vermittlung positiver Werthaltungen fördert und genuin menschliche Fähigkeiten nicht beeinträchtigt. Dabei ist zunächst gleichgültig, ob die Artefakte beabsichtigt normativ fragwürdig konstruiert wurden, d.h. ob bspw. eine Nutzermanipulation tatsächlich beabsichtigt wurde, oder ob keine böswillige Intentionalität seitens des Unternehmens oder der Entwickler aufzufinden ist.

Die Einführung digitaler Technologie und Technik in zwischenmenschliche Handlungsketten führt daher sowohl hinsichtlich ihrer Herstellung als auch hinsichtlich ihrer Nutzung in drängende Verantwortungsfragen. Gerade aufgrund der hohen Anzahl an unterschiedlichen Artefakt-Typen sowie deren Verbindungsmöglichkeiten mit anderen Technologien und deren direkten Auswirkungen auf das

menschliche Urteilsvermögen, auf menschliche Emotionalität oder Sozialität, muss die Verantwortungsthematik auf zwei Ebenen angesprochen werden: Zum einem ist der Herstellungsprozess der zu entwickelnden, digitalen Artefakte zu durchdenken (Nissenbaum, 1996) und zum anderen darf die *agency*, die Handlungsfähigkeit, des Nutzers nicht unberücksichtigt bleiben, wenn über gerechtfertigte Verantwortungszuschreibung gesprochen werden soll (Nida-Rümelin, Spiegel & Tiedemann, 2017).

Die Komplexität wird zusätzlich erhöht, wenn man sich vor Augen führt, dass selbst kleine Inkremente, die während des Entwicklungsprozesses programmiert werden, nicht auf individuelle Akteure zurückführbar sind: So entwickelt meist ein wechselndes Team Software und nutzt dabei bereits vorhandene, teils auch zugekaufte Softwarestücke. Zugleich führt Teamarbeit jedoch auch in notwendige gegenseitige Deliberations- und Kontrollmechanismen, die gerade bei normativ schwierigen Fragestellungen, die nicht von einer Person überblickt und entschieden werden können, individuelle Erleichterung herbeiführen können. Hinzu kommt, dass Software meist in Hardware implementiert wird und sich aus der Mischform weitere normative Tücken ergeben: So verstärken sich normative Unklarheiten, denn je mehr unterschiedliche Teams mit unterschiedlichen Schwerpunkten an einem Artefakt arbeiten, desto schwieriger wird die Lokalisation der Schnittstellen, die es bei einer normativen Betrachtung zu bewerten gilt. Ebenso wird die Tendenz erhöht, Verantwortung an andere Personen oder gar an andere Unternehmen zu delegieren. Qualitätssichernde Maßnahmen müssen daher zu verschiedenen Zeiten während des Entwicklungsprozesses implementiert werden, denn gerade die Qualität sozio-technischer Systeme zeichnet sich nicht nur durch technische Zuverlässigkeit aus, sondern muss auch die Interaktion der Beteiligten, indirekter Beteiligter, weiterer Systeme sowie kollektiver Lebensformen berücksichtigen. Auch staatliche Aufsichtsbehörden benötigen einen gewissen Überblick über Handlungsmöglichkeiten, um sich dieser Schnittstellenüberprüfung souverän annehmen zu können.

Als prominentes, wenn auch schon älteres Beispiel kann hier der Fall der Therac-25-Produktion genannt werden (Turner, 1993). Der Therac-25 war ein Elektronenlinearbeschleuniger, der in der Krebstherapie in den USA und Kanada zwischen 1984 und 1987 eingesetzt wurde. Aufgrund von Überdosierungen verstarben drei Personen und mindestens drei erlitten schwere Verletzungen. Eine Vielzahl sich gegenseitig verstärkender Sicherheitsmängel führten zu diesen verheerenden Ereignissen.

Die Herstellerfirmen, Energy Canada Limited (AECL) und das französische Unternehmen CGR, erweiterten den Therac-25 ab 1984 um ein Computersystem. Mit Hilfe der automatischen Einstellung der Behandlungstischhöhe sowie der Bestrahlungsart konnte man nun mit einem Therac-25 sowohl eine oberflächliche Bestrahlung mittels leichterer Elektronenbestrahlung vornehmen als auch bei ei-

ner notwendigen Tiefenbestrahlung auf Röntgenphotonen wechseln. Die Einführung des Computersystems führte dazu, dass das Sicherheitssystem allein durch die Software kontrolliert wurde, da die Messwerte von einem Computer ausgewertet und daraufhin weitere Aktionen automatisiert eingeleitet wurden. Dies sollte die Bedienung der Maschine erleichtern und dadurch Kosten einsparen. Die Ersetzung der mechanischen Sicherheitssysteme – meist Verriegelungen, die bspw. bei einer Übertretung eines Strahlenrichtwertes einrasteten und die Maschine stilllegten – durch Softwaresysteme ermöglichte zwar eine vereinfachte Bedienung, führte jedoch zugleich zu den drei tödlichen Ereignissen: Zugunsten der Bedienerfreundlichkeit des Systems wurden Sicherheitsaspekte auf verschiedenen Ebenen vernachlässigt. Bspw. wurde keine erneute Durchsicht bei schnellen Eingaben oder fehlerhaften Befehlen verlangt. Zudem wurden die Fehlermeldungen allesamt mit abstrakten Bezeichnungen eingeblendet, wodurch die Fehlergrade nicht mehr einfach unterschieden werden konnten, um nur einige wenige Beispiele zu nennen.

Die Untersuchung ergab zusätzlich, dass die gesamte Software von einer Person entwickelt wurde, wobei die Entwicklung nur rudimentär dokumentiert und das System nicht ausreichend getestet worden war. Die Forderung, dass unabhängige Parteien kritische Systeme kontrollieren müssen, bevor diese vermarktet werden dürfen, wurde im Rahmen dieser Katastrophe vehement diskutiert (Turner, 1993). Der heutige politische Diskurs über einen Einsatz von Gesichtserkennungssystemen in der EU kann als Versuch verstanden werden, eben jener mangelnden Regulierung Herr zu werden. Der Therac-25-Vorfall zeigt bereits eine notwendige Verzahnung unterschiedlicher Verantwortungsbereiche. Nicht alle Fehlerquellen sind rein technischer Natur und bedürfen zusätzlich zu einer Software-techno-ethischen Erörterung auch einer wirtschaftsethischen.

Ebenso wird hier ersichtlich, dass zwischen technischen Fehlern, die zu ungewollten Konsequenzen führen, und normativ ausgerichtetem Design unterschieden werden muss: So kann man zwischen Designfeatures unterscheiden, die aufgrund ihrer technischen Unzuverlässigkeit zu nicht wünschenswerten Resultaten führten, wie bspw. den Softwarefehlern, der erst die tragischen Ergebnisse implizierten, als kein mechanischer Mechanismus mehr als Sicherheitssystem implementiert war. Normative Designanforderungen spiegeln sich aber auch darin wider, ob die Bedienungsfreundlichkeit zu nicht wünschenswerten Verhaltensweisen der Nutzer führt und man daher bereits im Design auf diese Reaktionen hätte aufmerksam werden sollen. Letztere normativen Aspekte lassen sich gerade nicht durch eine Überprüfung technischer Zuverlässigkeit erreichen und berücksichtigen, sondern bedürfen Kenntnisse über soziale Bereiche und deren normativen Konstituenzen.

Nissenbaum (1996) nahm sich der durch die Therac-25 diskutierten Phänomene an und analysierte daraufhin den Entwicklungsprozess hinsichtlich der Möglichkeit, sich ethischen Überlegungen und der dementsprechenden Verantwortungsübernahme zu entziehen. Ungeachtet der organisatorischen

Grundstruktur liefe die Entwicklerkultur Gefahr, normative Fragestellungen vorschnell abzutun, so Nissenbaum. Sie identifizierte vier typische Haltungen, die sich systematisch in der Arbeitskultur der Softwareentwicklung finden: die Bugkultur, das *Many-Hands-Problem*, das Eigentum sowie die Sündenbock-Falle (ebd., 1996).

The *Many-Hands-Problem* umfasst die Schwierigkeit, dass viele Personen an der Entwicklung von Software beteiligt sind und Entscheidungen oder Fehler nicht mehr ohne Weiteres auf eine Person zurückzuführen sind. Unterschiedliche Kompetenzen der Unternehmer, Manager bis hin zu technischem Personal mit unterschiedlichen Schwerpunkten, wie Grafiker, Qualitätssicherung, Sicherheit, sowie der Nutzer führen in eine Verantwortungsdiffusion. Das Therac-25 verdeutlicht dieses Phänomen: Nicht nur die Softwareentwickler, auch die Mechaniker und Techniker in den Krankenhäusern vor Ort, schließlich die Unternehmenskultur und die staatliche Behörde hatten sich allesamt normativ suboptimal verhalten. Eine solche Diffusion führt nicht nur ex post zu Schwierigkeiten, Verantwortlichkeiten zuzuschreiben, sondern bietet bereits ex ante Nährboden für eine Haltung der Verantwortungsdelegation oder Gleichgültigkeit.

Unter Bugkultur wird eine Fehlerkultur verstanden, die sich in einer indifferenten Haltung gegenüber Softwarefehlern zeigt. Das Auftreten fehlerhafter Software sei ohnehin nicht zu vermeiden, da Software niemals perfekt sein könne. Dies führt, so Nissenbaum, zu einer Kultur nicht ernstzunehmender Qualitätssicherung, zu der auch normative Belange zu zählen sind. Diese indifferente Einstellung wird zusätzlich durch die Komplexität der Systeme verstärkt, da die Überprüfung der Software keinem Individuum allein gelingen kann, weshalb es auch zu einer moralischen Überforderung kommen mag. Dies hat zur Folge, dass unerwünschte Resultate weder technisch noch arbeitsorganisatorisch auf eine bestimmte Entscheidung oder Handlung zurückzuführen sind. Entscheidungsträger umfassen in diesem Kontext Politik, Unternehmen, Entwicklung sowie Nutzung.

Zudem zeigt sich auf der unternehmerischen Ebene kaum eine verpflichtende Verantwortungsübernahme: Die Unternehmen produzierten Güter, fühlten sich für deren Entgleisungen jedoch nicht verantwortlich oder delegierten diese an die Konsumenten, denen sie einen missbräuchlichen oder fehlerhaften Gebrauch unterstellten.

Der Computer fungiert somit verstärkt als Sündenbock: Der Computer ist schuld, nicht der Unternehmer, nicht der Entwickler und auch nicht der Nutzer. Für Fehler und unerwünschte Resultate in Form von Output und Outcomes sind Computer selbst verantwortlich zu machen, schließlich vermitteln diese Wahrnehmungskontexte und Handlungsstrukturen.

Diese für Informationstechnologie und -technik einzigartige Verantwortungsproblematik verlangt nach einem Modell der technischen Rückführbarkeit und Dokumentation sowie Prozessen, die eine ar-

beitsorganisatorische Konzeption der Verantwortungszuschreibung erlauben, die sich über Kompetenzbereiche definieren lässt. Dies bedarf Arbeitnehmern, die sich als autonom begreifen, damit sie sich ihrer Entscheidungsmacht bewusstwerden können (Zuber et al., 2021).

4.7 Ausblick

In diesem Kapitel wird die Einzigartigkeit digitaler Technologie und Technik diskutiert, die eine besondere Ethik zu begründen vermag. Maner (1996) folgend muss eine Ethik für Softwareentwickler auf Fälle fokussiert sein, in denen der Computer eine wesentliche und nicht nur eine zufällige Rolle spielt, d.h. das Computersystem muss als Quelle jener normativen Ereignisse verstanden werden. Dabei reicht es nicht mehr aus, Computertechnik nur als vermittelndes Gerät aufzufassen, sondern auch die performative, mediale Natur muss in einer normativen Analyse berücksichtigt werden. Es geht nicht mehr ausschließlich um zufällig eingetretene, unzumutbare Effekte, die sich technisch fassen lassen, sondern auch um die normative Ausprägung digitaler Artefakte und deren Herstellungsverfahren. Technik muss hier in einem doppelten Sinn verstanden werden: Auf der einen Seite steht das zu entwickelnde Produkt, auf der anderen Seite steht der Arbeitsprozess und stehen somit die Verfahrensweisen, die notwendig sind, um dergleichen Produkte zu entwickeln.

Jedes der in diesem Kapitel angeführten Unterkapitel kann eine gesamte Dissertation füllen. Hier wurden sie lediglich als Merkmale der Einzigartigkeit digitaler Technologien und Technik eingeführt, die bestimmte normativ relevante Sachverhalte konstituieren. Für eine eigenständige Ethik für Softwareentwickler ist bedeutsam, dass diese Merkmale auch eine eigenständige Ethik begründen. Eine Ethik für Softwareentwickler rechtfertigt sich über eine analytische Verknüpfung des Sachgegenstands mit dem Einsatzbereich: Hier können die qualitativen Kriterien herausgearbeitet werden. Zudem wird eine Ethik für Softwareentwickler wesentlich über die herstellende Tätigkeit erarbeitet, da es sich hier um eine Berufsethik handelt, die im Gegensatz zur Medizin oder Ökonomie nicht nur einer kollektiven Praxis gerecht werden muss.

Folglich erfordert eine Ethik für Softwareentwickler eine professionelle sachgegenständlich orientierte Berufsethik. Da aufgrund der oben angeführten Merkmale weder universell festgelegte Kriterien, die maßgeblich für alle digitalen Artefakte vollumfänglich gelten können, noch Maßnahmen, wie diese umgesetzt werden können, genannt werden können, ist ein prozedurales Vorgehen maßgeblich. Hier müssen alle ethischen Gründe sowie normativ relevanten Bereiche gemeinsam während des Entwicklungsprozesses deliberiert werden. Ein solcher ethischer Ansatz wird im nachfolgenden Kapitel betrachtet.

Die hier angeführten Momente der Einzigartigkeit digitaler Technologien und Techniken sind graduell zu verstehen. Sie treten häufig gemeinsam auf und verstärken ihre Wirkung gegenseitig. Die Komplexität einer normativen Analyse wird dadurch klar ersichtlich. Die vielfältigen Merkmale, die jenen digitalen Artefakten eine Machtposition zuteilwerden lassen, sind in dieser Weise anderen Technologien und Techniken nicht zuzusprechen. Die Schwierigkeit, auf diese epistemischen Umstände auch noch praktisch angemessen zu agieren, führt Nissenbaum (1994) zu der ernstzunehmenden Frage, ob ein privates Eigentum bei dieser Technologie überhaupt moralisch gerechtfertigt sei.

5 Schluss: Für eine prozedurale, proaktive und diskursive Technikethik

Der Versuch, dieser praktischen Ohnmacht entgegenzuwirken, die durch die epistemische Unterbestimmtheit verstärkt wird, soll mittels einer prozeduralen, proaktiven und diskursiven Technikethik gelingen, die zusammen mit deskriptiv-normativem Orientierungswissen eine praktische Anleitung bietet und ein Ethos der Verantwortung etabliert: Humanistisch wünschenswerte digitale Artefakte benötigen eine entsprechende Entwicklermentalität. Eine Zusammenführung aller Teilethiken, wie sie in Teil I, Kapitel 6.1, diskutiert wurde, kann nur dann gelingen, wenn bereits während des Herstellungsprozesses das digitale Artefakt in seinen relationalen Wechselwirkungen normativ evaluiert wird. Damit einer Forderung nach normativ wünschenswerten, digitalen Artefakten nachgekommen werden kann, müssen die normativen Anforderungen, die sich aus der ethischen Deliberation ergeben, zugleich auf ihre technische Umsetzbarkeit hin überprüft werden. Dies bedeutet jedoch nicht, dass diese Normativität ausschließlich technisch lösbar sein muss. Damit ethische Deliberationen in der Softwareentwicklung zielführend durchgeführt werden, ist es notwendig, den Handlungsspielraum einzugrenzen. Hierdurch erfahren Entwickler das Ausmaß ihrer Handlungsmacht und Steuerungshoheit, was eine Verantwortungsübernahme rechtfertigt. Nicht alle normativen Zusammenhänge können durch die technische Gestaltung beeinflusst werden. Von daher gilt es, die für die Berufsgruppe aussichtsreichen Lösungsräume zu diskutieren.

Die Komplexität kann dabei nur durch eine Lokalisation der Fragestellung innerhalb einer Ethik reduziert werden, die sich über eine bestimmte Ebene des Sachgegenstands begründet, d.h. der bewussten Auswahl eines sich zu diesem Zeitpunkt aufdrängenden Handlungsraums. Dies bedarf einer Verbindung von technischem Wissen und technikphilosophischen Grundlagen mit einer normativen Ethik: Erst wenn die normativen Sachverhalte erkannt, Wissen über deren technische Ursächlichkeit und Machbarkeit vorhanden ist, kann ein Lösungsraum umrissen und eine begründete Entscheidung gefordert werden. Nida-Rümelin (1999, S. 245) formuliert diesen Sachverhalt, der die Forderung eines verstärkt informierten, deskriptiven Urteils mit der normativen Handlungsdimension verbindet, folgendermaßen: „Die ethische Dimension kommt dabei zum einen unter einer empirischen und zum anderen unter einer normativen Perspektive in den Blick: Die empirische Perspektive bezieht sich auf die Analyse der Konstituenten moralischer Grundorientierung, ihre Verbreitung in der Gesellschaft, ihre mögliche Gruppierung [...], die Bildung von moralischen Idealtypen und damit auch die Akzeptanzdimension einzelner technologischer Optionen [...] bzw. technisch-wirtschaftlich-gesellschaftlicher Szenarien.“ Die zweite Rolle der Ethik ist nicht empirisch, sondern normativ: „Es geht um die wohlbegründete, d.h. möglichst rational fundierte normative Stellungnahme.“ (ebd.) Ethik ist daher

keine bloße Bestandsaufnahme empirischer Tatsachen, sondern vor allem eine gut begründete Entscheidung, die einer rationalen Deliberation folgt. Die Frage nach der guten Entscheidung lässt sich somit nicht auf die Frage der auffindbaren normativen Überzeugungen reduzieren, sondern führt selbst zu einem normativen Urteil, beruhend auf einer wohlbegründeten Entscheidung, die am Ende einer Deliberation steht.

„Es ist kein Novum der Technikethik, sich mit Sollensansprüchen auseinanderzusetzen, denn der Mensch sieht sich alltäglich mit der Frage konfrontiert, was er tun soll: was eine richtige, angemessene, vernünftige oder gute Handlung sei. Mit zunehmender Verbreitung von Maschinen und Algorithmen, die Aktionen ohne menschliche Aufsicht ausführen, stellt sich diese Frage insbesondere auch bei der Entwicklung und dem Einsatz solcher Systeme, gerade weil diese einen direkten Einfluss auf das individuelle Leben und die Gesellschaft haben.

Viele Alltagssituationen bedürfen keiner normativen Deliberation, da das erwünschte Verhalten oder die angemessene Handlung, zumindest für den jetzigen Zeitpunkt, eindeutig bestimmbar ist: viele Rechenregeln, Verkehrsregeln oder Höflichkeitsregeln fallen hierunter. Personen wissen, ohne darüber nachdenken zu müssen, was getan werden soll, welches Verhalten erstrebenswert ist. Diese dispositionellen Verhaltensweisen sind akzeptiert, internalisiert und bedürfen in dem Moment keiner weiteren Reflexion – was jedoch nicht ausschließt, dass sie zu einem späteren Zeitpunkt überdacht und verändert werden können. Lebensweltlicher Komplexität wird somit durch angeeignete und gefestigte Dispositionen, man mag hier auch an Tugenden denken, begegnet (Nida-Rümelin, 2019; Vallor, 2016). Anders verhält es sich bei Handlungskonflikten, also in Situationen, in denen unklar ist, welche Handlung, welcher Zustand oder welches Ziel erstrebenswert ist.

Aufgabe der normativen Ethik ist es somit, zu reflektieren und zu begründen, welcher Handlungsgrund Orientierung bieten soll. Diese Begründung ist Resultat einer ethischen Deliberation, die normative Argumente systematisiert und zuletzt selbst ein normatives Urteil fällt: So gilt es, die Zulässigkeit oder Zumutbarkeit von Entscheidungen zu er- und begründen. In der ethischen Deliberation werden Handlungsgründe gewichtet und unterschiedliche Handlungsoptionen abgewägt. Letztendlich muss die Entscheidung bzw. die Handlungsintention wohlbegründet sein, um als vernünftige Handlungsorientierung gelten zu können. Die Informatik kann diese normativen Entscheidungen dann in der Entwicklung technischer Artefakte berücksichtigen und praktisch in ihrem Design umsetzen.

Dabei lassen sich normative Urteile nicht auf epistemische Tatsachen reduzieren. Ein Mehr an Wissen führt nicht notwendigerweise zu mehr Klarheit in der Entscheidungsfindung bzw. zu einem besseren Urteil. Anders ausgedrückt: Gewissheit lässt sich nicht allein in Daten finden, weshalb alle Arten von Algorithmen zu keiner normativen Lösung führen können. U.a. gilt das für die weitverbreiteten Machine-Learning-Algorithmen, die Korrelationen in Daten ausfindig machen oder die Dimensionalität von Daten reduzieren können. Dies ist schon allein der Unterbestimmtheit normativer Handlungsgründe geschuldet, also dass gute Gründe nicht immer eindeutig ermittelbar oder gar berechenbar sind. Zwar sind Fälle denkbar, in denen der Handlungsgrund unmittelbar erkenntlich scheint, beispielsweise in der Pflicht, Verletzten zu helfen. Aber selbst hier müssen situativ-evaluative Perspektiven in der Abwägung der jeweiligen Einzelsituation Berücksichtigung finden, wie z.B. Zeitdruck oder Gefährdung des Hilfeleistenden. Entscheidungen bleiben somit kontingent. Es lässt sich eben nicht bei allen Tatsachen eindeutig, d.h. notwendig, bestimmen, welcher Handlungsgrund gelten soll. Die Gründe entziehen sich somit einer Berechenbarkeit, weshalb wir beim Entstehungsprozess technischer Objekte ansetzen müssen. Zudem lassen sich normativ uneindeutige Situationen nicht nur durch Wissenserwerb lösen, denn sie fordern zusätzlich praktische Kompetenzen heraus, wie Entscheidungsstärke oder andere Charaktermerkmale.

Die Handlungsanleitung, die Normativität, erschöpft sich nicht in juristischen Gesetzen oder sozialen Normen. Daher ist es nicht weiter verwunderlich, dass im Fokus des aktuellen, öffentlichen Diskurses gerade Algorithmen stehen, die durch ihre enorme Reichweite viele Menschen betreffen und beeinflussen können und so eine enorme normative Kraft entwickeln. Menschen stehen alltäglich in Situationen, in denen sie entscheiden müssen und in denen sie auch richtig, angemessen und gut entscheiden wollen. Hierbei kommt auch Werten bzw. Werthaltungen eine herausragende Rolle zu, denn sie können konative Einstellungen konstituieren, d.h. sie können Intentionen hervorbringen, die ausdrücken, wie sich Menschen in dieser oder jener Handlungssituation entscheiden *wollen*. Individuell für gut befundene und im Laufe eines Reifeprozesses internalisierte Werte können also Handlungsgründe als besonders wohlbegründet erscheinen lassen, sodass diese Gründe einen zwanglosen Zwang auf die Ausführung der Handlung ausüben. Folglich kann eine ethische Deliberation als ein Denkmodus, ein Abwägungsprozess verstanden werden, der zu jenem normativ wirksamen Handlungsgrund führen soll.

Normative Gründe müssen keine moralischen Gründe sein, aber alle, auch moralische, Gründe sind normativ, d.h. sie können mit mehr oder weniger Berechtigung zur Realisierung drängen. Ökonomische und juristische Tatsachen oder soziale Regeln können ebenso normative Gründe konstituieren und Handlungen Orientierung bieten. Konflikte entstehen, wenn sich Gründe gegenseitig ausschließen. So kann ein ökonomischer Grund einen moralischen Grund ausschließen oder zwei ökonomische Gründe weisen in verschiedene Richtungen. Von moralischen Konflikten kann man dann sprechen, wenn sie die Handlungsfreiheit tangieren – wenn also eine Handlungsausführung Freiheit(en) beschneidet (Nida-Rümelin, 2016).“ (Zuber et al., 2020).

Der autonome Nutzer muss sowohl in seiner individuellen Weltwahrnehmung als auch in seiner strukturellen Einbettung in der Entwicklungsarbeit maßgeblich berücksichtigt werden (Teil I, Kapitel 1.2; Teil I, Kapitel 3). Dies geht über eine Betrachtung des Nutzers als Konsument hinaus, wie es Usability-Ansätze vornehmen. Doch können digitale Technologien oder technische Artefakte die Autonomie-sichernde Sphäre der Nutzer auf vielfältige Weise berühren, weshalb sowohl technik-generische als auch strukturelle Werte auf diese Lebensentwürfe hin zu überprüfen sind (Teil II, Kapitel 8). Da Autonomie jedoch ohne Vernunft, die auch menschliche Emotionalität und Sozialität umfasst, und Freiheit, also die Möglichkeit, sich von der Vernunft auch leiten lassen zu wollen, nicht denkbar ist, sind sowohl die menschliche Vernunftfähigkeit als auch die Möglichkeit, vernünftig handeln zu können, als Gütekriterien digitaler Artefakte zu nennen. Darüber hinaus muss auch die Herstellungspraxis jenen qualitativen Maßstäben genügen. Nur so bietet die Entwicklungsarbeit die Möglichkeit, sich dem digitalen Gegenstand aus unterschiedlichen Perspektiven nähern zu können.

Die Schwierigkeit, die konstituierenden Momente dieser humanistisch interpretierten ethischen Paradigmen in Anbetracht digitaler Technologie und Technik zu identifizieren und ihre Bewertung vorzunehmen, liegt in der dynamisch medialen und performativen Natur dieser Technologie begründet (Teil III). Digitale Artefakte sind keine Geräte, sondern fordern die Nutzer kontinuierlich dazu auf, sich Technologie-konform zu verhalten, indem sie subtil Denk- und Handlungsanleitungen geben oder individuelle sowie kollektive Wahrnehmungsinhalte und -formen verändern (Teil I, Kapitel 2; Teil III, Kapitel 7). Auch ein Hammer verlangt eine bestimmte Abfolge physischer Handlungen, sozusagen ein Hammer-konformes Verhalten der Nutzer. Im Gegensatz zu digitalen Artefakten kann hier jedoch

nicht von einem Kontrollverlust gesprochen werden. Digitale Artefakte zeichnen sich aufgrund ihrer transparenten und opaken Natur als manipulativ aus, da der Nutzer die vielfältigen Handlungsanweisungen nicht unmittelbar erkennen oder gar begreifen kann. Dies führt zu einem Kontrollverlust der eigenen Handlungssteuerung (Teil III, Kapitel 9).

Die Multidimensionalität des Sachgegenstands sowie die vielen, sich im Umlauf befindenden Konzepte machen eine systematische Begriffsanalyse erforderlich, damit eine eindeutige gemeinsame Kommunikation gelingen kann. Sofern durch den Einsatz von Informationstechnologien und -techniken neue Formen menschlicher Handlungsräume erschaffen werden, die „conceptual muddles“ (Moor, 1985, S. 266) verursachen, muss zunächst eine begriffliche Ordnung geschaffen werden. Erst eine Begriffsklärung ermöglicht präzise Formulierungen von Regulierungsbestreben, Entwicklungsanforderungen oder von konkreten Erwartungen im Hinblick auf eine ethisch wünschenswerte Nutzung. Die Begriffsanalyse erlaubt zudem eine differenzierte Erkenntnis der normativ relevanten Sachverhalte: Phänomene können so voneinander abgegrenzt und in Bezug zueinander gesetzt werden. Dieses Vorgehen ist unabdingbar, wenn auf die idiosynkratischen Merkmale digitaler Technologie und Technik, wie sie in Teil III dargelegt wurden, nicht nur reaktiv, sondern auch proaktiv eingewirkt werden soll. Eine Ethik, die dieser Forderung Folge leisten will, bedarf ausreichend deskriptiven Wissens über diese Technik und Technologie sowie zugleich normativen Orientierungswissens, das die normative Deliberation systematisierend erleichtert (Moor, 2007, S. 111).

In diesem Sinne ist es auch von großer Bedeutung, Technik und Technologie nicht nur durch naturwissenschaftliche Terminologie, sondern auch als Kulturleistung zugänglich zu machen (Hartmann & Janich, 1998). Diese als kulturalistische Technikphilosophie bezeichnete Herangehensweise betont die sprachliche Aufbereitung technischer Prozesse, die durch handelnde Subjekte getragen werden (Janich, 2013, S. 103). Erst durch den performativen Charakter wird die Rationalität der Herstellung ersichtlich, denn das digitale Objekt wird nicht mehr rein isoliert verstanden, als Objekt der Natur und somit durch physikalische oder chemische Gesetzmäßigkeiten bestimmt, sondern ebenso als Kulturgut, das aufgrund der menschlichen Intentionalität erst erschaffen und genutzt werden kann. Diese durch die Sprache methodisierte Abfolge von Handlungen der Herstellungspraxis sowie des Gebrauchs lässt sowohl die Entwicklung als auch den Umgang mit Technologie und Technik als praktisch vernünftig denken, d.h. umfasst nicht ausschließlich Anforderungen der instrumentellen Rationalität. Denn die einzelnen Herstellungsschritte werden zu vernünftigen Handlungen, indem diese notwendig in vernetzte Lebenspraxen integriert werden. Versteht man unter Handlung jenen Akt, der dem Menschen vom Menschen als Leistung oder Versäumnis zugeschrieben wird, dann erhalten technische Ausführungsprozesse ein theoretisches und ethisches Moment, das sich mit Hilfe von

Kulturbegriffen erfassen lässt (Janich, 2013, S. 105). Sprache ermöglicht so, den Herstellungsprozess und das Herstellungsgut in eine grundlegende rationale Methodik zu überführen und Verantwortungskomplexe zuzuschreiben.

Winter und Butler (2022) betonen in ihrem Aufsatz *Responsible Technology Design: Conversations for Success*, dass die Verantwortungszuschreibung paradox anmutet. Es wird der Eindruck vermittelt, dass Softwareingenieure die alleinige Verantwortung über die wünschenswerten oder negativen Auswirkungen digitaler Artefakte tragen. Dies impliziert, dass Softwareingenieure bei Fehlentwicklungen allein zur Verantwortung gezogen werden, obgleich nicht alle ethischen Probleme technisch steuerbar sind. Im Rahmen dieser Arbeit wurde bereits betont, dass sich normativ wünschenswertes Design nicht ausschließlich über nicht vorhandene negative externe Effekte definieren lässt, sondern maßgeblich auch positive Verhaltensweisen und Haltungen fördern können muss. Normativ wünschenswertes Design ist somit immer ethisch proaktiv positives Design: Ein Design, das Werten folgt, die an sich erstrebenswert sind. Diese Werte steuern wünschenswerte Praxen. Bspw. kann hier der Wert der Inklusivität genannt werden, der sich technisch als Kriterium der Barrierefreiheit zeigt. Die Verantwortung reduziert sich nicht nur auf die Vermeidung negativer Folgen, sondern muss als positive Affirmation erstrebenswerter Werthaltungen verstanden werden. Die Autoren sprechen in diesem Zusammenhang von der Notwendigkeit einer *superhuman power*, über die Softwareingenieure verfügen müssten, wenn sie alle normativ relevanten Sachverhalte eigenständig zu erkennen und zu beheben hätten. Allerdings entscheiden Menschen jeden Tag und benötigen keine *superhuman power*, um angemessene Handlungen auszuführen oder zu unterlassen. Das Argument trägt folglich zur Debatte nicht sonderlich viel bei, sondern verstärkt das Problem der Verantwortungsübernahme, wie es Nissenbaum (1996) diskutiert.

Eine moralische Überforderung beruht auf der Unmöglichkeit, Situationen immer moralisch eindeutig zu erkennen oder moralisch angemessen handeln zu können. Dies zeigt sich in beruflichen Situationen vermehrt, indem

- 1) moralische Überzeugungen mit institutionellen Strukturen konfliktieren können oder
- 2) die moralischen Konflikte erlebt, aber deren normativen Ursachen nicht wahrgenommen werden, weshalb in diesen Situationen nur moralisch unzureichend gehandelt werden kann (Gräb-Schmidt & Stritzelberger, 2019).

Das Phänomen der moralischen Überforderung, das häufig als *moral distress* bezeichnet wird, ist im Pflegebereich besonders gut dokumentiert und diskutiert worden (Gräb-Schmidt, 2019, v.a. Kapitel 2): Es zeigt sich hier, dass das Pflegepersonal oft intrapersonell überfordert ist, weil die individuelle Pflegekraft entweder die ethische Schiefelage nicht erkennt oder aber gegen ihre eigenen moralischen Überzeugungen handeln muss. Ersteres deutet auf ein epistemisches Defizit hin, während letzteres

den interpersonellen, moralischen Druck hervorhebt, der moralische Konflikte verursacht. Interpersonelle Konflikte entstehen durch die Interaktion mit Institutionen, Kollegen oder anderen Berufsgruppen. Die moralische Belastung behindert zum einen die moralische Deliberation selbst und zum anderen wird die Äußerung moralischer Überzeugungen erschwert. Allerdings kann hieraus nicht geschlossen werden, dass Entwicklern keine Verantwortung übertragen werden kann, denn viele alltägliche Situationen können ebenso als normativ unsicher klassifiziert werden. Daher fordern die Autoren, dass die Verantwortungsübernahme strukturell implementiert und kollektiv getragen werden muss, sowohl auf Unternehmens- als auch auf gesellschaftlicher Ebene.¹⁸ Dies kann durch die Einführung diskursiver Elemente in der Softwareentwicklung gelingen, denn durch die Konversationen mit anderen können verschiedene Perspektiven miteinbezogen werden, die gemeinsam ein Bild möglicher unzumutbarer, zukünftiger Szenarien ergeben. Auf dieser Grundlage soll dann gemeinsam eine ganzheitliche Entscheidung getroffen werden:

Rather, responsible design involves expecting the technologies to be designed in ways that provide for active, meaningful, ongoing conversations between the developer and the technology, between the user and the technology, and between the user and the developer – and expecting that designers and users will commit to engaging in those conversations. It is well within our ability to create systems and technologies that provide the affordances for iterative designer-tech, user-tech, and designer-user conversations. [...] Instead of defining success as developers creating a responsible design, we must expect that they engage in the never-ending process of responsible design. Responsible design implies responsible engagement with the larger socio-technical system and the processes by which meaning, purpose, and values emerge. These conversations can occur at multiple levels and in diverse forms. (Winter & Butler, 2022, S. 273)

Die Forderung nach einem partizipativen und strukturierten Diskurs ist nicht neu, sondern bildet den Kern jeder Diskursethik (Habermas, 1998; Krings et al., 2011; Toulmin, 2003).¹⁹ Der Diskurstyp muss sich jedoch an die jeweils verschiedenen Kontexte adaptieren. Diese Anpassung erfolgt durch unterschiedliche Konzepte und Methoden für diskursive und partizipative Verfahren, die in verschiedenen gesellschaftlichen Systemen zur Anwendung kommen (Skorupinski & Ott, 2000, Kap. VI.5). Im Bereich des Designs digitaler Artefakte sind diese partizipativen und diskursiven Elemente vor allem im Ansatz des *participatory designs* berücksichtigt worden:

It has to do with participation, with how stakeholders – especially users, developers and planners – cooperatively make or adjust systems, technologies and artefacts in ways which fit more appropriately to the needs of those who are going to use them. Participation can be approached as an ideology, and also clearly refers to questions of ethics, politics, democracy and empowerment. (Simonsen & Robertson, 2012, S. 42)

¹⁸ Rophol legt 1994 in seiner Grafik der „Morphologischen Matrix der Verantwortungstypen“ eine Verantwortungszuschreibungspraxis dar, welche die moralische Belastung des individuellen Ingenieurs verringern soll.

¹⁹ Für eine gute Übersicht über diskursethische Ansätze in der Technikethik siehe Ott (2013).

Dieser Wunsch eines methodologischen Vorgehens betrifft nicht nur die Softwareentwicklung, sondern generell den Bereich neuer Technik und Technologie, weshalb für die Entwicklung einer strukturierten, ethischen Deliberation für die Softwareentwicklung durchaus ein Blick auf bereits bestehende Methoden und Konzepte, z.B. in der Nanotechnologie (Swierstra & Rip, 2007) oder der Biotechnologie (Beekman et al., 2002), geworfen werden sollte. Die Ausarbeitung einer Methodik, die Softwareentwickler in normativen Deliberationen unterstützt, bedarf einer interdisziplinären Kollaboration.

Es wird also nicht nur begriffliche Klarheit gefordert, sondern vor allem auch ein methodisches Vorgehen einer proaktiven Ethik, die bereits antizipativ ethische Fragen lokalisieren und beantworten können soll. Dies kann folglich nur im engen Austausch mit anderen Disziplinen gelingen, denn das ethische Urteil benötigt technisches, ökonomisches, soziologisches, psychologisches und moralisches Orientierungswissen, wollen wir ein informiertes Urteil fällen. Tatsächlich muss eine solcher Austausch interner und integrativer Bestandteil des Entwicklungsprozesses sein. Ein solches Vorgehen erweist sich nur dann als zielführend, wenn die gesellschaftlichen Kosten durch eine in die Entwicklungsarbeit integrierte normative Deliberation reduziert werden (Van der Burg & Swierstra, 2013). Die Durchführung einer normativen Deliberation bedarf einer ethischen Expertise, die „sich nicht auf das Element der moralischen Intuition [bezieht] und ebenso wenig auf die empirischen Implikationen der Realisierung bestimmter normativer Überzeugungen, sondern ausschließlich auf die begriffliche Präzisierung, logische Relationen und die Kenntnis unterschiedlicher ethischer Theorien“ (Nida-Rümelin, 1999, S. 265). Vornehmlich findet diese Form der techno-ethischen Abwägung in einem Diskurs statt, der – sofern er vorurteilsfrei, offen und sorgfältig durchgeführt wird – unterschiedliche Expertisen und Interessen aufzunehmen vermag, um letztendlich ein normatives Kriterium im Hinblick auf den zu betrachtenden Gegenstand zu formulieren (ebd.).

Eine proaktive Haltung bedarf einer gemeinsamen Vermittlung, d.h. einer gemeinschaftlich ausgeführten diskursiven Bestandsaufnahme normativ relevanter Tatsachen, die bereits während der Entwicklungsarbeit stattfindet (Nida-Rümelin, 1999; Ott, 2005). Diese normativ relevanten Sachverhalte werden zugleich evaluiert und in technisches Machbarkeitsdenken überführt. Seit den 1970er Jahren begann jedoch mit der Einführung einer empirischen Technikfolgenabschätzung, die vor allem auf ökonomischen und soziologischen Auswertungen beruhte und eine empirisch informierte Handlungsorientierung für die Politik geben sollte (van der Burg & Swierstra 2013, S. 3 ff.), eine Abkehr von technikethischen Deliberationen. In der Technikethik führte dies zu einem Bruch mit metaphysischen Erörterungen mitsamt ihrer oft dualistischen Auffassung einer Technologie gegen die Menschheit, wie sie bspw. noch von Heidegger bis hin zu Marcuse thematisiert wurde (Vallor, 2016, S. 20-34). Die Hinwendung zur Empirie bezeichnet Vallor als den „empirical turn“ (ebd., S. 31), der dazu führte, dass nun eine Ansammlung an Einzelfällen sozialwissenschaftlich ausgewertet wurde, was jedoch – ohne

in einen größeren theoretischen Rahmen gestellt zu werden – zu einem endlosen Katalog vieler auf den ersten Blick unzusammenhängender Phänomene führte (Vallor, 2016). Vallor (2016, S. 3) betont deshalb, dass die Empirie zwar die praktische Vernunft mit Tatsachen „füttert“ und somit auch neue Deliberationsgrundlagen schafft, aber zugleich kann der Empirismus die Einzelfälle nicht priorisieren, folglich auch nicht in eine Rangfolge erstrebenswerter Ereignisse überführen.

Die empirische Wende muss, um diesen wichtigen Aspekt, den Vallor noch zaghaft mit einem Können auffasst, von der praktischen Vernunft – im Sinne einer vernünftigen Systematisierung und Aufbereitung von Gemeinsamkeiten und Differenzen – zu einem in sich stimmigen Gefüge geordnet werden, sollen die Einzelfälle Handlungsorientierung geben können. Letzteres verweist bereits auf die unabdingbare menschliche Fähigkeit der Urteilskraft. Eben dieses humanistische Ideal der Verwirklichung menschlicher Vermögen, die sich in einer selbstbestimmten Ausübung eigener Vernünftigkeit zuspitzen, betont zugleich die Bedingungen der Möglichkeit, sich als Person auch ausleben zu dürfen. Dies verdeutlicht erneut, dass das Anforderungsmanagement den Nutzer in seiner Lebenswelt berücksichtigen können muss, will man die humanistischen Werte nicht durch den Einsatz digitaler Technologie und Technik gefährden. Für die Softwareentwicklung ist vor allem die Zuschreibung des Kompetenzbereichs relevant: So benötigen Softwareentwickler zusätzlich zu ihrem technischen Wissen auch geübtes Machbarkeitsdenken hinsichtlich normativer Anforderungen, wie es Ansätze des *value sensitive design*, *participatory design* oder *responsible design* aufzeigen (Van den Hoven, Vermaas & Van de Poel, 2015). Welche Methoden für welche normativen Fragestellungen zielführend sind, sprengt den Rahmen dieser Arbeit, soll hier doch zunächst die Frage beantwortet werden, welche normativen Sachverhalte eine normative Deliberation einfangen können muss.

Technikethik und ihre Nachbarsethiken, wie die *Ethics of New and Emerging Technology and Science* und die *Ethics of Responsible Research* sehen sich mit der Schwierigkeit konfrontiert, eine real-ethische Analyse mit vielen ungewissen Variablen bewerkstelligen zu müssen: Denn es gilt hier eine ethische Bewertung eines Szenarios vorzunehmen, das vermehrt auf Unsicherheiten gründet. Auf der einen Seite beruhen viele Unklarheiten auf nicht vorhersehbaren, zukünftigen Konsequenzen, die durch einen Einsatz digitaler Technologie und Technik verursacht werden. Auf der anderen Seite gilt es, sich Ungewissheiten zu stellen, die nicht rein epistemischer Natur sind, d.h. es handelt sich hierbei um kein Defizit, das durch ein Mehr an Wissen und Information zu beheben ist (Sollie, 2007), sondern das gerade Haltungen und etablierte Praxen zur normativen Orientierung benötigt. Das normative Urteil, das in der praktischen Philosophie durch eine Entscheidung zu einer begründeten Handlung führen können soll, lässt sich nicht vollständig theoretisch lösen, d.h. stützt sich nicht vollständig auf Rationalitätsmaßstäbe, Logik oder empirische Fallbeispiele. Es mag vielmehr eine Kompetenz notwendig sein, die das *know-how* mit dem *know-that* und *know-why* zu verbinden vermag.

Diese Umstände führen dazu, dass Technikethiken in prozedural ausgerichtet oder in inhaltlich-substantielle unterschieden werden können: in solche, die sich um die Bedingungen der Möglichkeit bemühen, wünschenswerte technische Artefakte zu gestalten, und in solche, die inhaltlich normative Begründungen bzw. inhaltliche Richtlinien ethisch wertvollen Handelns zu liefern imstande sind. Erstere versuchen, normative Entscheidungsmethoden zu entwickeln, während letztere im Sinne einer Offenlegung Unterstützung in der Erkenntnisgewinnung bieten können sollen. Eine Technikethik, die beides gewährleisten möchte, muss sowohl in der Lokalisierung und Identifizierung normativ relevanter Aspekte als auch in der Entscheidungsfindung und Durchführung Orientierung bieten können.

Sollie (2007) bezeichnet Entscheidungen hinsichtlich neuer Technologien als Entscheidungen, die trotz ontologischer Unsicherheiten getroffen werden müssen: Allein die schiere Anzahl an möglichen Konsequenzen kann nicht gänzlich erkannt und berücksichtigt werden. Viele Einsatzmöglichkeiten sind zu diesem Zeitpunkt nicht greifbar und können somit in der ethischen Evaluation nicht erwogen werden. Aber auch ein Mangel an normativen Praxen verschärft die Problematik normativen Urteilens (Teil II, Kapitel 6). Zudem erschweren unvorhersehbare Naturereignisse, wie mögliche Katastrophen oder Krankheiten, Veränderungen der institutionellen Rahmenbedingungen oder der gesellschaftlichen Arbeitsorganisation den Entscheidungsprozess (Sollie, 2007). Die Unsicherheit rechtfertigt jedoch keinen zögerlichen Umgang mit ethischen Fragstellungen, sondern fordert geradezu dazu auf, eine wohlbegründete Haltung gegenüber normativ strittigen digitalen Technologien und Techniken sowie deren Implementierung einzunehmen. Dies gilt sowohl für die Forschung und Entwicklung als auch für die Endnutzung digitaler Technik und Technologien.

Floridi führt hier seine Kritik an einer mangelnden Beachtung einer *Ethik der Poesis* an. Mit Hilfe einer *Ethik der Poesis* kann die Herstellungstätigkeit selbst einer moralischen Bewertung zugänglich gemacht werden. Indem Floridi den Menschen als wesentlich herstellend, als *homo poeticus*, herausarbeitet, der sich durch seine Weltbezogenheit eine Welt herstellend erschafft (Floridi, 2011), wertet er die herstellenden Tätigkeiten auf. Floridis *constructivism* ist sowohl als ein metaphilosophisches Konzept als auch als eine Methode der Welterschließung zu verstehen (ebd.). Floridi schenkt so dem herstellenden Aspekt hinreichend Aufmerksamkeit, ohne dabei der Welt der Dinge ihre Objektivität absprechen zu müssen (Russo, 2012b, 2016). Eine konstruktivistische Ethik kann Herstellungstätigkeiten somit nicht als passive Reaktionen auf eine vorgegebene Welt verstehen, sondern muss die Herstellungstätigkeit selbst als Handlungsakt begreifen, sich die Welt der Dinge praktisch zu erschließen. Eine normative Bewertung der hergestellten Werke führt aufgrund des inhärent dynamischen Charakters jedoch zu keiner zufriedenstellenden Lösung: Es bleibt unmöglich, einen konkreten Anforderungskatalog zu erstellen. Somit muss die normative Güte der Werke über den Herstellungsprozess thematisiert werden. Dies bedeutet, dass der Herstellungsprozess ethisch wünschenswerte

Handlungsstrukturen aufweisen muss (Manders-Huits & Zimmer, 2009), indem bspw. ethische Deliberationsräume implementiert werden, die ethische Reflexionen während der Entwicklungsarbeit in bestimmten Zeitintervallen forcieren. Durch eine wünschenswerte Herstellungspraxis, ein Ethos der Softwareentwicklung, kann die normative Unbestimmtheit digitaler Artefakte eingefangen werden. Diese Herangehensweise ist einer prozeduralen Technikethik eigen, welche die Bedingungen der Möglichkeit einer ethischen Ausgestaltung digitaler Produkte berücksichtigt. Eine prozedurale, emanzipatorische Technikethik bereichert den technisch-funktionalen Herstellungsprozess um normative Deliberationen, wie es Floridi und Sanders (2005, S. 4) ausführen:

„But ethics is not only a question of dealing morally well with a given world. It is also a question of constructing the world, improving its nature and shaping its development in the right way. This proactive approach treats the agent as a world owner, a game designer or referee, a producer of moral goods and evils, a provider, a host, or a creator. The agent is supposed to be able to plan and initiate action responsibly, in anticipation of future events, in order to (try to) control their course by making something happen, or by preventing something from happening rather than waiting to respond (react) to a situation, once something has happened, or merely hoping that something positive will happen.“

Eine prozedurale Technikethik ist folglich wesentlich antizipatorisch und bedarf somit einer ethischen Methodik, die sachbezogen bereits im Entwicklungsprozess eingebunden werden kann (Brey, 2012a; Palm & Hansson, 2006; Stahl, Timmermans & Mittelstadt, 2016; Wright, 2011). Proaktiv ist sie, indem sie von Anbeginn an ethische Überlegungen in den Forschungs- und Entwicklungsablauf integriert, um der Forderung gerecht werden zu können, Technologie und Technik einer stetigen normativen Überprüfung unterzogen zu haben. Sie ist deswegen zugleich auch antizipatorisch, da eine Technikethik in diesem Sinne Zukünftiges und noch nicht Existentes erahnt und diese Vermutung auch in der Gestaltung berücksichtigt. Dies kann nach Brey (2011) auf spekulative oder aber auch konservative Art und Weise erfolgen. Eine Haltung, die Brey als *spekulativ* bezeichnet, versucht auf die Unvorhersagbarkeit zukünftiger Szenarien mit Hilfe von Konzepten und Methoden, wie z.B. der *Delphi-Methode* (Häder & Häder, 2000), zu reagieren. Als Zielvorgabe wird die Formulierung möglichst vieler Zukunftsszenarien artikuliert, die dann auf ihre wünschenswerten Konsequenzen hin untersucht werden. Demgegenüber definiert Brey (2011) eine *konservative* Haltung, die sich mit dem beabsichtigten Szenario und dessen Folgen auseinandersetzt, d.h. nicht auf Fremdnutzungsmöglichkeiten eingeht, da diese gerade unbestimmbar bleiben.

Die Notwendigkeit einer proaktiven Haltung wird von Moor (2005) mit seiner dreistufigen Einteilung der Technologieentwicklung belegt, indem die Dringlichkeit ethisch relevanter Aspekte von der Einführungsphase über die Durchbruchphase bis hin zur Power-Phase kontinuierlich zunimmt. Das Moor'sche Gesetz besagt: „As technological revolutions increase their social impact, ethical problems

increase“ (ebd., S. 117). Diese Hypothese mag nicht verwundern, entstehen oder verstärken sich doch normative Schwierigkeiten durch den Einsatz digitaler Technik in lebensweltlichen Strukturen. Ähnlich einer unbehandelten Angststörung, die sich schleichend in allen Lebensbereichen manifestiert, verstärken und verschieben sich unberücksichtigte, wenn auch unbeabsichtigte, negative ethische Auswirkungen von der Forschung bis hin zum finalen (Massen-)Konsumprodukt. Die zunehmenden sozialen Kosten beruhen auf Umwelteinwirkungen oder gesellschaftlichen Missständen, die ein Massenprodukt aufgrund seiner Reichweite hervorruft. Der mediale Charakter digitaler Artefakte restrukturiert oder erschafft neue Handlungsräume (Teil II Kapitel 7.2; Teil III), was zu einer normativen Orientierungslosigkeit beiträgt. Die Wertinhärenz digitaler Technologie und Technik verschärft zudem die Dringlichkeit einer ethischen Evaluation (Teil II, Kapitel 7), da die Nutzer in ihrer Autonomie eingeschränkt werden. Somit ist auch ökonomisch ratsam, ethische Deliberationen frühzeitig in den Softwareentwicklungsprozess einzubinden: Ignoranz führt zu steigenden Kosten (Moor, 2005).²⁰

Aufgrund der zentralen Aufgabe, die der Urteilskraft zukommt, ist es wenig verwunderlich, dass eine Renaissance der Tugendethik oder Moralphysikologie in der Ethik der Informationstechnologie und -technik stattfindet (Floridi 2002; Vallor 2016; Reijers & Coeckelbergh 2020). Dieser tugendethische Ansatz ermöglicht es, die Universalisierbarkeitsansprüche praktischer Vernunft in Bezug auf Informationstechnologien zu artikulieren, da die Idee des guten Lebens mit der Notwendigkeit sozial eingeübter, rationaler Haltungen zusammengebracht wird. Die Urteilskraft ist gerade dasjenige menschliche Vermögen, das Universalansprüche und Partikularinteressen, Emotionen und Körperlichkeiten zusammenführend denken können soll. Diese Urteilskraft ist somit ein ständiges Schwanken zwischen konkreten Situationen und universellen Ansprüchen, zwischen technischem *Know-How* und ethischen Haltungen, zwischen instrumenteller und normativer Rationalität.

Die Unmöglichkeit, eine Ethik der Digitalisierung, der digitalen Transformation oder der digitalen Artefakte ausschließlich über einen ausgewiesenen gesellschaftlichen Bereich zu rechtfertigen, bedarf einer ethischen Strukturanalyse, die sich der Kontextspezifität annehmen kann. Eine Strukturanalyse der Lebensform ermöglicht es, normativ relevante Aspekte zu lokalisieren, um qualitative Kriterien für eine techno-ethische Evaluation zu formulieren. Jedoch können nicht alle ethischen Fragestellungen, die in diesen Zusammenhängen auftreten, auf eine strukturelle, normative Analyse reduziert werden: Wie in der Diskussion um die Mini- und Mesoethiken bereits ersichtlich wurde, gibt es ethisch relevante Bereichsethiken, die sich über den Sachgegenstand, nicht über den gesellschaftlichen Bereich

²⁰ Das Gesetz von James Moor ist von *Moor's Law* zu unterscheiden, welches die Verdopplung der Transistoren und die damit verbundene Rechenleistung alle zwei Jahre vorhersagt.

rechtfertigen lassen. Diese technik-generischen, ethischen Fragestellungen explizieren für den technischen Sachgegenstand relevante Werte, die als qualitative Entscheidungskriterien einer normativen Bewertung dienen (Teil II, Kapitel 2.1).

Die diffuse Ausgangslage unterstreicht die Notwendigkeit, sich von den Reflexionsbegriffen *Opazität* und *Transparenz* leiten zu lassen, um die normativen Kriterien herauszuarbeiten, die mit dem jeweiligen Betrachtungsraum verknüpft sind (Teil II, Kapitel 3). So wurde bereits ersichtlich, dass die Einbettung digitaler Artefakte in alle Lebensbereiche, im Sinne des *ubiquitous* oder *ambient computing*, und deren subtile Verstrickung mit den wesentlichen Merkmalen menschlicher Lebensführung – Vernunft, Verantwortung und Freiheit – einer zunächst dezidiert deskriptiven und erkenntnisgewinnenden *disclosive ethics* bedarf (Teil III, Kapitel 3). In diesem Schritt können mit Hilfe der Reflexionsbegriffe *Opazität* und *Transparenz* normativ schwierig erkennbare Betrachtungseinheiten lokalisiert werden. Diese Domänen umfassen sowohl Werte als auch lebensweltliche Strukturen sowie deren relationale Wechselwirkung: Hierfür ist es hilfreich, sich der Methode der *levels of abstraction* anzunehmen, um das (sozio-)technische System in die für die normative Analyse relevanten Ebenen zu differenzieren (Teil II, Kapitel 4). Dies ist unerlässlich, sollen kategoriale Fehlinterpretationen oder Kommunikationsschwierigkeiten vermieden werden. Dieses zunächst vor allem erkenntnistheoretische Vorgehen soll Betrachtungseinheiten abgrenzen und dadurch notwendige Bewertungsräume offenbaren. Die Bewertung kann, wie bereits betont, nach Brey spekulativ oder konservativ erfolgen: Je nachdem, wie mit der Unbestimmtheit des zukünftigen Gebrauchsszenarios theoretisch umgegangen werden soll, wird eine dieser Haltungen gegenüber dem Untersuchungsobjekt eingenommen.

Eine *disclosive ethics* bedarf jedoch einer weiteren philosophischen Grundlage, will sie kein Entscheidungsvakuum generieren: Eine Ethik für Softwareentwickler benötigt eine Theorie, welche die lebensweltliche Normativität, in der sich Menschen immer schon bewegen, erklärbar macht und zugleich ein rationales Gerüst zur Verfügung stellt, das die alltäglichen gegenseitigen Verpflichtungsansprüche als Grundlage normativer Orientierung und praktisch normativen Wissens einführt (Teil I, Kapitel 2). Dies gelingt mit Hilfe von Nida-Rümelins *Theorie einer Strukturellen Rationalität*, die diese kontextuelle Einbettung in gegenseitige Erwartungshaltungen als eine notwendige Bedingung für Sollensanforderungen herausarbeitet. Somit wird Vernunft in ihren individuellen und kollektiven Bezügen denkbar. Die Transformation solcher teils tradiert rationaler Beziehungsgeflechte erfordert eine kritische ethische Analyse hinsichtlich der Erwünschtheit ihrer Zielsetzungen. Die durch den Einsatz digitaler Technik und Technologie hervorgerufenen lebensweltlichen Transformationen der gegenseitigen Erwartungshaltungen der, wenn man so sagen möchte, normativen Grundpfeiler gesellschaftlicher (Sub-)Systeme verweisen auch auf eine tugendethische Interpretation der normativen Analyse innerhalb eines deontologischen Rahmens: die Frage nach dem guten Leben innerhalb nicht verhandel-

barer Normen. Eine Ausrichtung digitaler Artefakte hinsichtlich deren Einbindung in strukturelle Rationalitäten führt weiter als ein Design, das *usability* oder die *user experience* berücksichtigt, und kann dadurch auch zu moralisch wertvollen digitalen Artefakten führen, da nicht nur individuelle Rationalitäten Berücksichtigung finden können.

Die vielfältigen theoretischen Paradigmen sollen gemeinsam zu einer gezielten ethischen Deliberation führen, die sich nicht ausschließlich durch das Subsystem einer Berufsethik rechtfertigt, sondern sich auch durch den technischen Sachgegenstand bestimmen lässt. Eine Ethik für Softwareentwickler ist daher vonnöten, welche die Güte des Entwicklungsgegenstandes nicht mehr rein über die technische Funktionalität formuliert, d.h. eine gute Software ist nicht nur eine, die funktioniert, die ausreichend getestet ist, die effizient läuft, die einfach gepflegt werden kann oder sich durch eine angenehme Oberflächengestaltung auszeichnet. Vielmehr soll normativ wünschenswerte Software die humanistischen Paradigmen fördern und nicht verletzen.

Es gilt somit, jene Werte zu lokalisieren, die ein normativ angemessenes Design ermöglichen, und diejenigen Praktiken aufzuzeigen, die ethische Deliberationen für wertvoll erachten. Maner (1999) betont deshalb in seinem Aufsatz *Unique Ethical Problems in Information Technology*, dass eine einseitige Betrachtung reiner *high-risk*-Fälle bloß zu einer Ansammlung an Fallstudien führt, die aufgrund ihres negativen Vorbildcharakters gar keine Orientierung mehr bieten. Diese *high-risk*-Fälle sind jedoch vornehmlich Ausreißer bzw. Härtefälle, weshalb ihnen weder eine Leitfunktion zugesprochen werden kann, noch sie das Alltagsgeschäft der Entwicklungsarbeit abbilden (Maner, 1999, S. 140). Verstärkt gilt es, sich derjenigen Fälle anzunehmen, die innerhalb der deontologischen Demarkationslinie liegen und keine extremen Spannungsverhältnisse aufzeigen, wie Abbildung 12 zeigt, um normative Orientierung erlangen zu können.

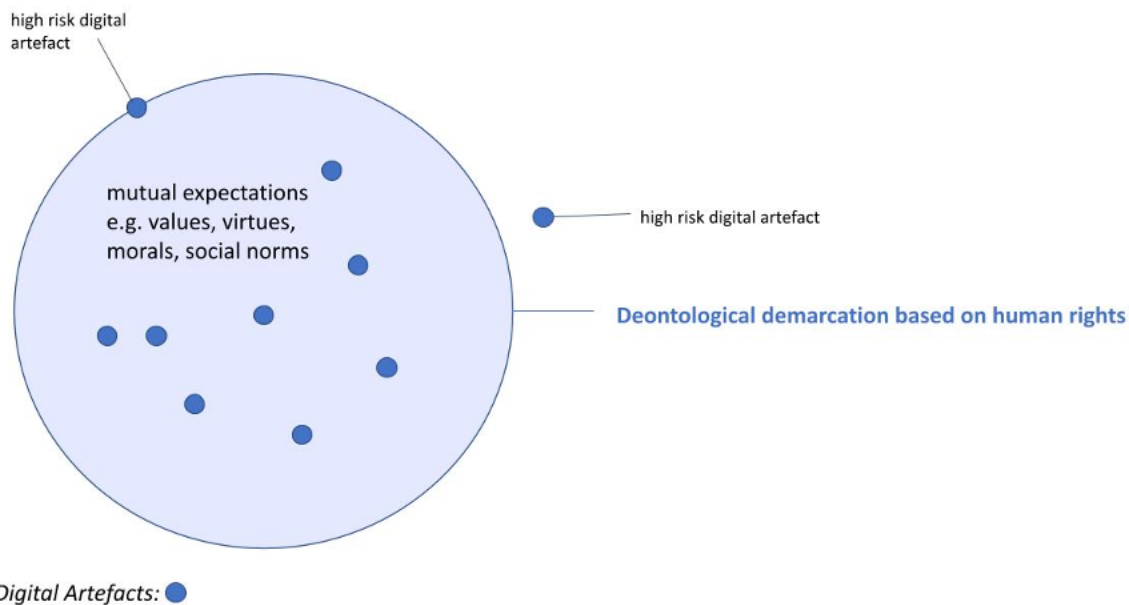


Abbildung 11 Deontologische Demarkationslinie und digitale Artefakte. Eigene Abbildung.

Das Netzwerk gegenseitiger Erwartungshaltungen, das von Werten, Gesetzen und sozialen Normen getragen wird, konstituiert Lebensformen. Digitale Artefakte, die sich auf oder außerhalb dieser deontologischen Demarkationslinie befinden, können als *high-risk*-Artefakte ausgewiesen werden. Lebensformen können jedoch nur eine begrenzte Anzahl an *high-risk*-Fällen ertragen: Beispielsweise nutzt Facebook die Wertesysteme ‚Demokratie‘ und ‚Freundschaft‘, unterminiert sie aber aufgrund der Plattformausgestaltung, wie bspw. ‚dem Daumen nach unten‘.

Somit muss sich ein philosophisches Verständnis mit dem normativen Druck befassen, den ein digitales Artefakt oder technisches Sachsystem auf individuelle Handlungen, intersubjektive Beziehungsgeflechte und kollektive Zielsetzungen ausübt. Eine philosophische Untersuchung verhilft dadurch dazu, normative Urteile über Technologie und Technik zu fällen, wenn durch eben jene Betrachtung diejenigen normativen Konstituenzen, die ein wünschenswertes, gemeinschaftliches Leben tragen, sichtbar gemacht werden können. Technologie und Technik kann sodann vor dem Hintergrund eines empirisch-philosophischen Horizonts normativ hinterfragt werden: Denn der Einsatz dieser Technologie oder Technik kann hinsichtlich der Förderung oder Behinderung derjenigen normativen Bedingungen evaluiert werden, die für ein gemeinschaftlich getragenes Zusammenleben oder für ein selbstbestimmtes Leben innerhalb der Gemeinschaft vonnöten sind. Dies verweist auf eine humanistische Ethik, denn das technische Artefakt wird an der Möglichkeit gemessen, ob es wünschenswerten menschlichen Bedürfnissen oder Fähigkeiten dient, indem nach deren lebensweltlichen Bedin-

gungen gefragt wird. Dadurch kann nicht nur Negatives vermieden, sondern vor allem auch der Versuch unternommen werden, Anforderungen an ein gutes normatives Design zu formulieren und – viel bedeutsamer – diese dann auch zu entwickeln.

Ein technisches Artefakt in seiner gesellschaftlichen Einbettung zu verstehen, bedeutet folglich die Erwartungshaltungen hervorzuheben, die Technik in eben jenem gesellschaftlichen Zusammenhang repräsentiert. Dafür ist wesentlich, die Strukturen hervorzuheben, in die das technische Objekt integriert ist. Eine philosophische Analyse wirkt sich somit direkt auf die normativen Überlegungen aus, denn ein kontextualisierter Technikbegriff verbindet Technik und Technologie mit normativen Bedingungen menschlicher Daseinsformen oder, wenn man will, mit der Idee des guten Lebens. Die normativen Beziehungselemente, welche diese zwischenmenschliche Ebene der Kommunikation und Kooperationen hervorbringen, können jedoch auch missachtet werden: So kann eine gelungene Kommunikation nicht gänzlich auf Lügen basieren, wenn man z.B. eine weitere Fortführung einer Freundschaft nicht gefährden möchte (Nida-Rümelin, 2011b). Das kooperative Miteinander, in diesem Falle der Freundschaftsstrukturen, wird durch das kontinuierliche Lügen gestört, denn Gemeinschaft erfordert eine Kommunikationskultur, die nicht durch verborgene Gründe getragen werden kann, sondern auf authentischen Äußerungen basiert. Z.B. gilt es als tugendhaft oder eben moralisch geboten, dem Freund nicht aus Eigeninteresse zu helfen, sondern der Freundschaft wegen. Punktuell können kooperative Strukturen Trittbrettfahrer-Handlungen verkraften. Kommt es jedoch zu einer flächendeckenden Erosion geteilter Strukturen, weil Trittbrettfahren einen systematischen Charakter annimmt, d.h. wenn das Verhalten zur Regel wird, verlieren diese Lebensformen ihren normativen Anspruch. Wenn Strukturen, die zur Aufrechterhaltung einer gewünschten gemeinsam-geteilten Lebenswelt benötigt werden, durch unbeabsichtigte oder gar beabsichtigte kontinuierliche Missachtung beschädigt werden, werden diese Lebensformen ihrer gelebten normativen Performativität verlustig: Freundschaften, die bspw. durch dauerhafte Lügen gekennzeichnet sind, bezeichnet man nicht mehr als Freundschaft, sondern das Beziehungsgeflecht mutiert zu einem einseitigen Abhängigkeitsverhältnis, das nur aus der Perspektive des Eigeninteresses zugänglich wird. Die gemeinsam geteilte Struktur der Freundschaft bedarf jedoch einer sie wertschätzenden Kommunikationskultur. Eine solche Kommunikation beruht auf den normativen Bedingungen der Wahrheit, Wahrhaftigkeit und Verlässlichkeit (Nida-Rümelin, 2011b, Kapitel I), die es überhaupt erst ermöglichen, dass Sprechakte nicht nur einer reinen strategischen Rationalität dienen, sondern auch kollektiv zu Gemeinsamkeiten vordringen können.

Technikphilosophische Überlegungen bieten somit die notwendige Orientierung für normative Fragestellungen hinsichtlich der Wünschbarkeit, Akzeptanz oder Zumutbarkeit technischer Anwendungen. Die Abgrenzung des zu betrachtenden Szenarios, eine konkrete Verortung der relevanten ethischen

Sachverhalte und deren Beurteilung ist gerade in Bezug auf einen Einsatz von informationstechnologischen Artefakten unentbehrlich. Aufgrund der rasanten informationstechnologischen Entwicklung und der daraus resultierenden Dynamik – die den Technologien den Anschein verleiht, beharrlich einen Charakter von *emerging technologies* aufzuweisen – wird eine philosophietheoretische Grundlage benötigt. Diese muss es ermöglichen, normative Fragestellungen, die sich in den und durch die Technologien und Technik verbergen oder entstehen, zu lokalisieren, zu systematisieren und somit bewertbar zu machen. Die ihnen innewohnende Wandelbarkeit erschwert eine ethische Interpretation und Evaluation im Hinblick auf ihre Implikationen auf individuelle Weltbezüge sowie gesellschaftliche Systeme. Es bleibt folglich unklar, welche normativen Kriterien in welchen Situationen ethisch relevant sind, sodass eine normative Ethik im klassischen, angewandten Sinn nicht erfolgen kann, wenn darunter verstanden wird, dass von einem allgemeinen ethischen Axiom auf konkrete situationsspezifische Aussagen geschlossen werden kann (Nida-Rümelin, 2005, S. 863-864). Ein solches Vorgehen bezeichnet Nida-Rümelin als rationalistisch (ebd.), weil gerade die Verwendung eines Prinzips unabhängig aller Besonderheiten eine rein kognitive Anstrengung erfordert, die jeglicher Erfahrung, Interessen oder Emotionalität entbehrt. Aber auch die Hinwendung zu einer rein subjektiven Bewertung führt in eine Sackgasse, da Ethik dann zu einer reinen Akzeptanzethik verkümmert, die weniger um einen rationalen Diskurs unter Berücksichtigung vielfältiger, lebensweltlicher Begründungswelten bemüht ist, sondern lediglich eine subjektive Zustimmung anstrebt.

Technikphilosophie ist somit einerseits eine Freilegung der normativen Strukturen, in die Technik eingebunden ist. Dadurch erhält sie den aufdeckenden oder entdeckenden Charakter, weshalb Brey (2000; 2012) den Ansatz als *disclosive ethics* bezeichnet. Andererseits ist sie eine Analyse der Auswirkungen der technischen Einbindung auf die gesellschaftlichen Formen des Zusammenlebens. Dieser doppelte Ansatz der strukturellen Ausrichtung mutet zunächst soziologisch an, geht aber über einen rein deskriptiven Ansatz hinaus, da auch die Möglichkeiten der normativen Bedingungen berücksichtigt werden. Somit werden die tiefen, sogenannten disruptiven Transformationen erklärbar, die durch digitale Artefakte hervorgerufen werden. Technische Artefakte sowie deren Entwicklung und Nutzung werden als Lebensweisen verstanden, die sich aufgrund des Einsatzes eines technischen Artefakts nachhaltig verändern können.

Die strukturelle technikphilosophische Betrachtung digitaler Artefakte hebt sogleich die Wichtigkeit einer Analyse des Maschinen-Mensch-Verhältnisses hervor, denn eine normative Evaluation muss die Erwünschtheit der lebensweltlichen Veränderungen bestimmen können, die ein Einsatz digitaler Technik hervorruft. Der Druck, den eine Implementierung digitaler Technologie und Technik auf das menschliche Dasein ausübt, muss sowohl individuell-phänomenologisch als auch intersubjektiv und kollektiv evaluiert werden (Teil I; Teil II, Kapitel 7). Mit Hilfe der Theorie der *Strukturellen Rationalität*,

die als deontische Struktur ein Geflecht wechselseitiger Verpflichtungen aufzeigt, können die unterschiedlichen normativen Ebenen vernünftig systematisiert werden (Teil I, Kapitel 2). Die *Strukturelle Rationalität*, die sich in der Lebensform durch geteilte Praktiken menschlichen Miteinanders zeigt, ist aufgrund ihrer rationalen Legitimation erfassbar und somit auch als normatives Wissen tradierbar (Nida-Rümelin, 2020, II §4). Die normativen Knotenpunkte, die gegenseitige Anspruchshaltungen repräsentieren, konstituieren die praktische Lebenswelt und werden durch performative Akte bekräftigt und übertragbar. Dadurch ermöglicht die *Theorie Struktureller Rationalität* die vielfältigen Rationalitäten, verstanden als Denkprinzipien oder Denkweisen, zu vereinen, was einem humanistischen Umgang mit Technik und Technologien zuträglich ist.

Das dialektische Verhältnis von Technologie und Technik sowie der gesellschaftlichen, normativen Einbettung wird von Soziologen und Philosophen, wie es Reijers und Cockelbergh (2020, S. 9) in ihrer Monografie *Narrative and Technology Ethics* diskutieren, auf verschiedene Weisen definiert: als *technology influencing, co-shaping* oder als *translating, composing and delegating* oder als *configuration of practices by technologies*. Die Konfiguration lebensweltlicher Praktiken durch den Einsatz von Technologie und Technik ist das Phänomen der digitalen Transformation. Ebenso lässt sich dadurch der Bedarf an digitaler Technologie und Technik erklären, da kollektive Handlungsketten strukturiert und bestenfalls optimiert werden sollen. Der Bedarf einer Technik lässt sich aus einer existenziell-menschlichen Verletzlichkeit erklären, die der Mensch u.a. mittels des Einsatzes digitaler Technik zu beheben versucht (Coeckelbergh, 2013). Zugleich aber verschiebt er dadurch nur seine Verletzlichkeit, transformiert diese und kann sich nicht befreien, auch wenn dies der sehnlichste Wunsch der Transhumanisten sein mag. Durch die Sehnsucht dahingehend, seine Schwächen technisch beseitigen zu können, erschafft der Mensch neue Abhängigkeits- und Machtverhältnisse, wie es zum Beispiel Eubanks (2018) in *Automating Inequality – How High-Tech Tool profile, police and punish the poor* soziologisch-empirisch herausarbeitet. Der Gedanke der Gestaltbarkeit von Technik und Technologie wird hier erneut deutlich: Schließlich wird das Ziel formuliert, Technik bewusst normativ auszurichten oder sich zumindest dessen bewusst zu werden, wodurch das Bedürfnis nach Technik entsteht, in welchen Zusammenhängen Technik kreiert wird und auf welche strukturellen Gegebenheiten wiederum eingewirkt wird.

Ein vernünftiger Umgang mit Informationstechnologien ist daher vonnöten, um den vielfältigen aufgezeigten normativen Belangen auf eine angemessene Weise zu begegnen. Dies ist als Anforderung sowohl an die *episteme* als auch an die *phronesis* zu begreifen. Diese Arbeit ist keine klassische Erörterung im Sinne der Angewandten Ethik, sondern zielt darauf ab, metatheoretisch eine ethische Praxis für die Softwareentwicklung zu erarbeiten. Eine theoriegeleitete Praxis soll die Arbeitsweise der Softwareentwickler um normative Deliberationen erweitern. Es sollen folglich wünschenswerte Praxen erschaffen werden, d.h. eine Arbeitskultur, die normative Deliberation auch lebensweltlich

ermöglicht: Softwareentwickler müssen emanzipiert werden, d.h. müssen über eine hinreichend normative Bildung verfügen, um ihre ethische Entscheidungskompetenz zu stärken. Zusätzlich bedarf es einer Unternehmenskultur, die ethische Deliberation wertschätzt und fördert. Diese normativen Praktiken umfassen folglich den Entwicklungsprozess, weshalb eine prozedurale Ethik auch um die Ausformulierung eines Berufsethos versucht ist.

Während sich Technikfolgenabschätzungen meist mit Risikoprognosen befassen, wurden ethische Technikfolgenabschätzungen erst seit Anfang der 2000er Jahre entwickelt, wie z.B. die Norm VDI 3780 (Verein Deutscher Ingenieure 2000: *VDI 3780 - Technikbewertung – Begriffe und Grundlage*, 2000) oder der eTA-Ansatz (*ethical technology assessment*) von Palm und Hansson (2006). Im Hinblick auf die ethische Gestaltung von Softwaresystemen konzentriert sich der IEEE 7000/ D3 Standard vornehmlich auf die Kompetenzen von Teammitgliedern, die notwendig sind, um ethische Fragen bereits in der Softwareentwicklung zu berücksichtigen und zu behandeln (Systems and Software Engineering Standards Committee, 2020). Dabei wird betont, dass ethische Theorien bei der Bewertung von Softwaresystemen unterstützend wirken können, indem sie ethische Werte identifizieren, die sich auf deren Konsequenzen, auf Praktiken oder auf Aspekte universalistischer Ansprüche beziehen (Systems and Software Engineering Standards Committee (2020), S. 13-22). McLennan et al. (2020) schlagen daher vor, einen Ethiker in die Entwicklungsteams aufzunehmen. Dies scheint gelegentlich notwendig zu sein, wenn Softwareentwickler selbst die Entscheidungen nicht treffen oder sich aufgrund der prozeduralen Eingebundenheit diesen Fragen nicht widmen *wollen*. Auch wenn es supererogatorisch klingen mag, dass sich Entwickler täglich um normative Belange kümmern *wollen*, sind sie doch Träger von notwendigem Sachwissen. Eine juristische oder ethische Expertise von außen mag jedoch unerlässlich sein.

Softwareingenieure müssen daher befähigt und ermutigt werden, normativ zu argumentieren; sie müssen in ihrer *phronesis* und ihren ethischen Kommunikationsfähigkeiten geschult werden. Dabei gilt es zu betonen, dass die Suche nach technischen Lösungen für normative Fragen nicht alle ethischen Fragen im Zusammenhang mit der Informationstechnologie lösen wird, wie im Kapitel über die negativen Externalitäten digitaler Technologie und Technik gezeigt wurde (Teil II, Kapitel 7.2). Eine verantwortungsvolle Ingenieursarbeit inkludiert jedoch ethische Reflexionen: Sie berücksichtigt strukturelle und technik-generische Werte und bewertet die Zumutbarkeit digitaler Artefakte im Hinblick auf individuelle, intersubjektive und kollektive Praxen. Normative Deliberationen müssen wesentlicher Bestandteil der Entwicklungsarbeit sein, was sich auch in der Dokumentationskultur und der Qualitätssicherung niederschlagen muss. Idealerweise verbindet ein proaktiver und diskursiver Ansatz normative Deliberationen mit dem gesamten Software-Lebenszyklus: zum Beispiel durch die Einbindung in DevOps-Methoden, die sowohl die Entwicklung als auch den Betrieb von Software umfassen (Ebert et al., 2016).

Die Debatte um die Berechtigung einer eigenständigen Ethik benötigt ein argumentatives Fundament. Eine Ethik der Informationstechnologie und -technik bedarf einer begründeten Argumentation, die sich nicht auf den moralischen Status der Maschine berufen kann, solange darunter verstanden wird, dass die Maschine moralisch entscheidend auftritt. Maschinen kann nicht derselbe moralische Status zugesprochen werden wie Menschen (Teil II, Kapitel 7.1): Sie richten sich nicht selbständig nach ethischen Paradigmen aus; sie können moralische Deliberationen nicht einmal simulieren (Teil I, Kapitel 6.2). Die Berücksichtigung dieser genuin menschlichen Vermögen zeichnet jedoch ein normativ gutes digitales Artefakt aus²¹: Die menschliche Vernunftfähigkeit, welche die Herausforderung beinhaltet, über die eigene Lebensgestaltung zu entscheiden, soll von digitalen Artefakten weder eingeschränkt noch substituiert werden. Dadurch wird der Nida-Rümelin'sche *Digitale Humanismus* zum Imperativ der Entwicklungsarbeit: Digitale Artefakte sollen den Menschen in seiner Humanität fördern und dürfen ihn nicht in der Ausübung seiner wesentlichen Vermögen behindern.

Digitale Artefakte lassen sich aufgrund ihres dynamischen Charakters nicht in absoluten technischen Funktionalitäten beschreiben, die ein wünschenswertes Design garantieren. Zudem gibt es oftmals keine normative Orientierung, die sich bereits lebensweltlich etabliert hätte. Diese Aspekte erwecken den Anschein, es handele sich bei digitaler Technologie und Technik um eine permanente *emerging technology*: Weder lassen sich normative Kriterien aus der Technologie und Technik ableiten, noch finden sich gelebte wünschenswerte Erwartungshaltungen im Umgang mit der Technologie und Technik. Dies führt in eine normative Orientierungslosigkeit, die es näher zu bestimmen gilt, will man diese überwinden (Teil II, v.a. Kapitel 7 und Teil III).

Diese Arbeit konzentrierte sich daher vornehmlich auf die Ausgestaltung normativ wünschenswerter digitaler Technologie und Technik und beschränkte sich auf den Bereich der Entwicklungsarbeit. Somit ist eine Ethik für Softwareentwickler eine Berufsethik, die in der Entwicklung normative Reflexionen zulassen können muss. Hierbei ist bedeutsam, den Zusammenhang zwischen der Entwicklungsphase, deren Akteuren sowie der Notwendigkeit einer frühzeitigen und kontinuierlichen normativen Deliberation herauszustellen: Diesbezüglich eignete sich der Begriff der Gestaltung, denn hier treffen technisches Wissen und intentionales Handeln aufeinander (Teil I, Kapitel 1.1.1. (a)). Skorupinski und Ott (2000, S. 22) definieren „Gestaltung [...] als bewusstes Eingreifen in Prozesse der Genese und Umsetzung von Technik. Weil dadurch in die Verhaltensspielräume anderer eingegriffen wird, bedarf gestalterisches und planerisches Handeln der Rechtfertigung. Die Begriffe der Gestaltung bzw. der Planung sind immer auf den Begriff der Verantwortung und damit auf Werte und Normen rückbezogen.“ Dies ist ausschlaggebend für eine Aufwertung der herstellenden Tätigkeit als Handlung. Hierfür

²¹ Inwiefern andere Lebewesen graduell über diese Fähigkeiten verfügen, wird hier nicht weiter erörtert.

erwies sich die Praxistheorie als wegweisend (Teil I): Erst wenn die herstellende Tätigkeit als Teil einer moralisch wertvollen Praxis verstanden wird, können moralische Anforderungen bereits im Entwicklungsprozess überdacht werden. Dabei stellt es sich als zentral heraus, dass sowohl der Entwicklungsprozess als auch die Einbindung digitaler Artefakte in gesellschaftliche Systeme jeweils als Teil einer Praxis zu begreifen sind.

So kann über den Prozess auf das digitale Artefakt Einfluss ausgeübt werden: Zunächst, weil die Entwicklungspraxis moralisch bestimmt wird und sodann, weil das digitale Artefakt im Hinblick auf wertvolle Praxen gedacht wird (Teil I, Kapitel 3). Die normativ wünschenswerte Einbindung digitaler Artefakte muss dabei als normatives Kriterium für die Ausgestaltung eben jener digitalen Artefakte gelten (Teil I, Kapitel 1.1.1. (a)): Es gilt, diejenigen strukturellen Werte mit zu berücksichtigen, die ein normativ wünschenswertes Verhalten der Endnutzer fördern. Allerdings darf gutes normatives Design auch technisch-generische Werte nicht unberücksichtigt lassen, weshalb eine Überprüfung des Artefakts auf typisch technisch-generische Merkmale unerlässlich ist (Teil II, Kapitel 2.1).

Dabei zeigt sich, dass sich eine normative Deliberation von den Begriffen der Transparenz und der Opazität leiten lassen sollte, insofern gerade normativ relevante Sachverhalte, gleichgültig, ob diese auf ihre technische Beschaffenheit oder ihre strukturelle Einbettung zurückzuführen sind, nicht unmittelbar erkennbar oder nachvollziehbar sind. Eine Ethik, die in der Softwareentwicklung zum Einsatz kommt, hat so immer den Charakter einer *disclosive ethics* (Brey, 2000b).

Somit rechtfertigt die mangelnde Kenntnis empirischer und normativer Tatsachen keine Unterlassung einer ethischen Analyse, sondern fordert zu mehr Besonnenheit und Klugheit auf, um der Verantwortung gerecht werden zu können. Das beabsichtigte Eingreifen in die Lebenspraxen anderer durch den Einsatz digitaler Technik bedarf einer moralischen Begründung, insofern Dritte durch die Entwicklung und den Einsatz digitaler Technik und Technologie in ihrer Autonomie eingeschränkt bzw. in ihrer Selbstbestimmung tangiert werden können. Für digitale Artefakte ist zudem die Asymmetrie zwischen Entwickler und Nutzer zu beachten, denn Nutzer können das Artefakt schnell und einfach anderweitig anwenden oder durch mangelnde Erfahrung und mangelndes Verständnis in der Nutzung risikoreicher werden lassen, weshalb Floridis Forderung einer Makroethik an Bedeutung gewinnt, die sich um die Wechselwirkungen einzelner Meso- und Miniethiken bemüht (Teil I, Kapitel 6.1). Die Rückkopplungseffekte zwischen Entwickler, Nutzer, Endnutzer, Maschinen sowie der Umwelt bedürfen daher einer stetigen wachsam normativen Überprüfung, da sich ethische Kriterien für die technische Ausgestaltung an dieser lebensweltlichen Dialektik bewähren müssen. Eine Ethik, die sich um die Normativität interaktiver Bestimmungsmomente bemüht und nicht reduktionistisch voranschreitet,

muss diese Perspektivenwechsel integrieren können. Hierzu sind die Methode der *levels of abstraction*, die Systematisierung der Mini-, Meso- und Makroethik sowie die theoretische Grundlage der *Strukturellen Rationalität* als sich gegenseitig ergänzend eingeführt worden (Teil I, Kapitel 2 und 3).

Resümierend wurde eine Proaktivität reklamiert, da bereits während der Entwicklung ethisch delibereert werden soll: Das Artefakt soll nicht nur der technischen und ökonomischen Rationalität Folge leisten, sondern sich auch anhand anderer normativer Gesichtspunkte evaluieren lassen. Da sich der antizipierende Charakter nicht mit einer Checkliste einfangen lässt (Gogoll et al. 2021), muss eine Ethik ebenso prozedural verstanden werden, um der Dynamik des Sachgegenstands Rechnung tragen zu können. Diese Ethik soll auch, will sie ihre ethischen Paradigmen leben, der Vernunftfähigkeit folgen, weshalb sie sich nicht nur auf ein ethisches Prinzip beschränken kann. Sie muss vielmehr alle Wechselwirkungen berücksichtigen und diese angemessen hinsichtlich ihrer lebensweltlichen Tauglichkeit abwägen. Deshalb rückt der Austausch von Gründen in den Mittelpunkt einer proaktiven und prozeduralen Technikethik. Dieser diskursive Moment kann die unterschiedlichen Perspektiven in einer zielführenden und kollektiven Deliberation vereinen. So kann für den jetzigen Zeitpunkt die normative Deliberation zu einem vernünftigen Abschluss kommen. Dabei folgte diese Arbeit keinem Konstruktivismus, sondern bleibt einem moralischen Realismus verbunden (Nida-Rümelin, 2018).

Um eine Ethik auch arbeitsweltlich praktisch werden zu lassen, müssen methodische Fragen gemeinsam mit dem Softwareengineering und der Managementtheorie beantwortet werden. Unerlässlich ist hierbei, wie diese proaktive und diskursive Ethik in den Entwicklungsprozess implementiert werden kann. Daher gilt es, diejenigen Handlungsbereiche zu lokalisieren, auf die der Softwareentwickler aufgrund seiner Profession Einfluss ausüben kann, um eine punktuelle Verantwortungsübernahme zu rechtfertigen. Hierfür ist das Konzept des *implicit moral agent* (Teil II, Kapitel 7.2) eingeführt worden, das den Sachverhalt umfasst, dass bestimmte Werthaltungen, die sich technisch manifestieren, vermittels des Artefakts transportiert werden können. Eine normative Betrachtung darf aber nicht auf technisch-generische Werte reduziert werden, da auch strukturelle Werte technisch manipuliert werden können (Teil I, Kapitel 7).

Dies bedeutet, dass hierfür das notwendige Wissen und die notwendigen Kompetenzen zu explizieren sind. Es bedarf nicht nur philosophischen, sondern auch informationstechnischen, soziologischen, juristischen und psychologischen Wissens, wenn ein normativ wohlbegründetes Urteil gefällt werden soll. Doch damit nicht genug: Das erkenntnisgestützte Urteil muss auch den Entwicklungsprozess leiten, soll letztendlich ein digitales Artefakt geschaffen werden, das nicht nur normativ unzureichende Folgen vermeidet, sondern auch normativ-affirmativ ausgerichtet ist. Dies verlangt einen

umsichtigen Entwicklungsansatz, der normativ erstrebenswerte Werte auf ihre technische Machbarkeit hin reflektiert. Diese Deliberation kann nicht nur an Hochschulen erfolgen, sondern muss auch lebensweltlich in der Entwicklungsarbeit verankert sein.

Aufgrund der Vielzahl an notwendigem Wissen und der Kompetenzen, die hierfür vonnöten sind, wurde eine diskursive Ethik ausgearbeitet. Der Unterbestimmtheit der guten Gründe geschuldet, bedarf es eines zielgerichteten Austauschs aller Stakeholder. Ein weiterer Schritt, der in dieser Arbeit nicht weiter ausgearbeitet ist, wäre demnach eine Methodendiskussion. Hierzu müssen diskursive Methoden, Kommunikationsweisen und strukturierte kollektive Denkprozesse auf ihre Erweiterung um normative Aspekte hin evaluiert werden. Diese sollten zudem in eine bereits vorhandene Entwicklungskultur etabliert werden können, die auch einer normativen Entscheidungsfindung als systematische Grundlage dienen kann. Die vorherigen Überlegungen verweisen allesamt auf eine Zertifizierung der Prozesse und nicht der digitalen Artefakte selbst, wie es bereits der Standard IEEE 7000 durchführt. Allerdings stellt sich im Anschluss die Frage, wer in der Lage sein soll, dieses komplexe Unterfangen zielgerichtet in Unternehmen umsetzen zu können.

Unbestritten bleibt, dass Entwickler für normative Belange sensibilisiert werden müssen. Die Fähigkeit, normativ zu deliberieren, darf dabei nicht unter die Rubrik *soft skills* fallen, sondern muss als wertvolle unternehmerische Ressource Beachtung finden. Dies würde sicherlich erleichtert, wenn gezeigt werden könnte, dass normativ erstrebenswerte digitale Artefakte auch zu mehr technischer Funktionalität oder zum ökonomischen Erfolg beitragen. Gutes normatives Design verhilft dann nicht nur dazu, Reputationsschäden des Unternehmens zu vermeiden, sondern zudem Kosten im Service oder der Maintenance zu reduzieren.

Eine Ethik für Softwareentwickler muss wirtschaftsethische und regulatorische Bemühungen ergänzen können. Wirtschaftsethische und regulatorische Bestreben treten von außen an die Entwicklungsarbeit heran, um normativ unerwünschte Folgen zu reduzieren. Legale Formulierungen werden meist als deontologischer Rahmen verstanden, innerhalb dessen es die eigenen Interessen zu optimieren gilt. Eine Ethik für Softwareentwickler ist jedoch um ein Ethos bemüht, das notwendige gute Entwickler-Haltungen prägt, die sich wiederum in einer vertrauenswürdigen, d.h. nicht rein der technischen und ökonomischen Rationalität gehorchenden Entwicklungsarbeit niederschlagen.

Eine Ethik für Softwareentwickler zielt verstärkt auf die intrinsische Motivation der Entwickler ab, damit die Entwickler erkennen, dass sie durch ihre Handlungen zu normativ besserer, d.h. legaler und ethisch angemessener, Technologie und Technik beitragen. Ein normativ erstrebenswertes digitales Artefakt ist ein solches, das wünschenswerte gesellschaftliche Systeme, die als Autonomie-sichernde Sphären verstanden werden, aufrechterhält und somit die individuelle Autonomie des Endnutzers

gewährleistet. Hierbei gilt es, die Autonomie des Einzelnen zu berücksichtigen, die ohne das kooperative Miteinander nicht denkbar ist, bereits in der Entwicklungsarbeit miteinzubeziehen.

Wichtig hierbei ist, dass jedes entwickelte Produkt durch eine normative Deliberation nachhaltig verbessert werden kann: Aber eine normative Deliberation digitaler Technologie und Technik muss sich auf bestimmte Reflexionsbegriffe stützen, um sich nicht in der Komplexität ihrer Materie zu verlieren. Diese Arbeit möchte Softwareentwicklern Mut machen, während ihrer Entwicklungstätigkeit bereits auf normativ wünschenswerte Aspekte zu achten, ohne sich in langen Wertekanons, in ethischen Theorien oder in einem Kulturrelativismus zu verlieren.

6 Literaturverzeichnis

- ACLU. (2020). *A Tale of Two Countries: Racially Targeted Arrests in the Era of Marijuana Reform*. Zugriff am 8.12.2021. Verfügbar unter: <https://www.aclu.org/report/tale-two-countries-racially-targeted-arrests-era-marijuana-reform>
- Agar, N. (2019). *How to be human in the digital economy*. MIT Press.
- Albers, M. J. & Mazur, M. B. (2014). *Content and complexity: information design in technical communication*. Routledge.
- Amrollahi, A. (2021). A Conceptual Tool to Eliminate Filter Bubbles in Social Networks. *Australasian Journal of Information Systems*, 25. Australian Computer Society.
- Anderson, M., Anderson, S. L. & Armen, C. (2004). Towards machine ethics. *AAAI-04 workshop on agent organizations: theory and practice*, San Jose, CA.
- Angerer, K. & Frank, D. (2020). Künstliche Intelligenz und Weltverstehen. *TATuP-Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis*, 29(3), 87–88.
- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20, 973 - 989.
- Arendt, H. (1960). *Vita activa oder Vom tätigen Leben*. Piper.
- Aristoteles. (NE) [2007]. *Die Nikomachische Ethik*. Die Nikomachische Ethik. De Gruyter (A).
- Arnold, T. & Scheutz, M. (2018). The “big red button” is too late: an alternative model for the ethical evaluation of AI systems. *Ethics and Information Technology*, 20(1), 59–69. Springer.
- Arrow, K. (1973). The theory of discrimination, discrimination in labor markets. *Achenfelter, A. Ress (eds.), Princeton–New Jersey*.
- Ashton, K. (2009). That „Internet of Things“ Thing. *RFiD Journal*, 1.
- Baxter, G. & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, 23(1), 4–17. <https://doi.org/10.1016/j.intcom.2010.07.003>
- Beekman, V., de Bakker, E., Baranzke, H., Baune, O., Deblonde, M., Forsberg, E.-M. et al. (2002). Ethical Bio-Technology Assessment Tools for Agriculture and Food Production, 60.
- Bendel, O. (Hrsg.). (2019). *Handbuch Maschinenethik*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-17483-5>
- Berdichevsky, D. & Neuenschwander, E. (1999). Toward an ethics of persuasive technology. *Communications of the ACM*, 42(5), 51–58. ACM New York, NY, USA.

- Bicchieri, C. (2005). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511616037>
- Blinden-und Sehbehindertenverein Westfalen. (2016). Seeing AI: Microsoft entwickelt App, mit der Blinde ihre Umgebung auf neue Art wahrnehmen können - www.bsvw.org. Blinden- und Sehbehindertenverein Westfalen e.V. Zugriff am 7.12.2021. Verfügbar unter: <https://www.bsvw.org/seeing-ai-microsoft-entwickelt-app-mit-der-blinde-ihre-umgebung-auf-neue-art-wahrnehmen-koennen.html>
- Boardman, J. & Sauser, B. (2006). System of Systems - the meaning of of. *2006 IEEE/SMC International Conference on System of Systems Engineering* (S. 6 pp.-). Gehalten auf der 2006 IEEE/SMC International Conference on System of Systems Engineering. <https://doi.org/10.1109/SYS-OSE.2006.1652284>
- Bobak, A. K., Hancock, P. J. & Bate, S. (2016). Super-recognisers in action: Evidence from face-matching and face memory tasks. *Applied Cognitive Psychology*, 30(1), 81–91. Wiley Online Library.
- Bostrom, N. & Yudkowsky, E. (2014). The ethics of artificial intelligence. *The Cambridge handbook of artificial intelligence*, 1, 316–334.
- Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77–91). PMLR.
- Brenner, J. E. (2010). A Logic of Ethical Information. *Knowledge, Technology & Policy*, 23(1–2), 109–133. <https://doi.org/10.1007/s12130-010-9099-3>
- Brey, P. (1998). The Politics of Computer Systems and the Ethics of Design. *Proceeding ACM/SIG-CAS Conference*, 64–75.
- Brey, P. (2000a). Method in computer ethics: Towards a multi-level interdisciplinary approach. *Ethics and Information Technology*, 2(2), 125–129. <https://doi.org/10.1023/A:1010076000182>
- Brey, P. (2000b). Disclosive computer ethics. *ACM Sigcas Computers and Society*, 30(4), 10–16. ACM New York, NY, USA.
- Brey, P. (2010). Values in technology and disclosive computer ethics. In L. Floridi (Hrsg.), *The Cambridge Handbook of Information and Computer Ethics* (S. 41–58). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511845239.004>
- Brey, P. (2011). Anticipatory technology ethics for emerging IT. *CEPE 2011: Crossing Boundaries*, 13, 868.
- Brey, P. (2012a). Anticipating ethical issues in emerging IT. *Ethics and information technology*, 4(14),

- 305–317. Springer. <https://doi.org/10.1007/s10676-012-9293-y>
- Brey, P. (2012b). Anticipatory Ethics for Emerging Technologies. *NanoEthics*, 6(1), 1–13. <https://doi.org/10.1007/s11569-012-0141-7>
- Brown, A. (2018). What is so special about online (as compared to offline) hate speech? *Ethnicities*, 18(3), 297–326. SAGE Publications Sage UK: London, England.
- Brynjolfsson, E. & McAfee, A. (2017). Machine, platform, crowd: Harnessing our digital future. *WW New York: Norton & Company*.
- Bynum, T. (2001, August 14). Computer and Information Ethics. Zugriff am 17.11.2021. Verfügbar unter: <https://stanford.library.sydney.edu.au/archives/sum2018/entries/ethics-computer/>
- Bynum, T. (2004). Ethical decision-making and case analysis in computer ethics. *Computer ethics and professional responsibility*, 60–87. Blackwell Publishing.
- Bynum, T. & Rogerson, S. (Hrsg.). (2004). *Computer ethics and professional responsibility*. Malden, MA: Blackwell Pub.
- Camus, A. (1953): Basierend auf: Interviews mit Camus im *Bulletin du RUA* vom 15. April 1953 und in *France-Football* vom 17. Dezember 1957.
- Capurro, R. (2009). Digital ethics. *GLOBAL FORUM ON CIVILIZATION AND PEACE* (S. 207–216).
- Capurro, R. (2017a). *Homo Digitalis: Beiträge zur Ontologie, Anthropologie und Ethik der digitalen Technik*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-17131-5>
- Capurro, R. (2017b). Digitization as an ethical challenge. *AI & SOCIETY*, 32(2), 277–283. <https://doi.org/10.1007/s00146-016-0686-z>
- Capurro, R. & Grimm, P. (2020). Digitale Ethik: Das Wesen des Digitalen im Narrativen finden? *Religiöse Bildung und Digitalisierung*, 61.
- Center for Engineering, Ethics, and Society & National Academy of Engineering. (2017). *Overcoming Challenges to Infusing Ethics into the Development of Engineers: Proceedings of a Workshop*. (C. Anderson, Hrsg.) (S. 24821). Washington, D.C.: National Academies Press. <https://doi.org/10.17226/24821>
- Chander, A. (2017), TheRacistAlgorithm?, 115 *MICH. L. REV.* 1023 (2017).
- Chaplin, C. (1937). *Modern Times*.
- Coeckelbergh, M. (2013). *Human Being @ Risk* (Philosophy of Engineering and Technology) (Band 12). Dordrecht: Springer Netherlands. <https://doi.org/10.1007/978-94-007-6025-7>

- Community Control Over Police Surveillance (CCOPS) Model Bill. (2021). *American Civil Liberties Union*. Zugriff am 5.8.2021. Verfügbar unter: <https://www.aclu.org/legal-document/community-control-over-police-surveillance-ccops-model-bill>
- Darwall, S. L. (1983). *Impartial reason*. Cornell University Press.
- Davis, M. (1998). Thinking like an engineer. *Studies in the Ethics of a Profession*, 143.
- Davis, J. P., Bretfelean, L. D., Belanova, E. & Thompson, T. (2020). Super-recognisers: Face recognition performance after variable delay intervals. *Applied Cognitive Psychology*. Wiley Online Library.
- Dean, B. (2022). "How Many People Use Facebook in 2022?" In: *Backlinko*. Zugriff am 06.01.2022. Verfügbar unter: <https://backlinko.com/facebook-users>
- Denkler, T. (2021, März 30). Der Tod von George Floyd in der Rekonstruktion. Zugriff am 5.8.2021. Verfügbar unter: <https://www.sueddeutsche.de/politik/george-floyd-tod-polizeigewalt-videos-rekonstruktion-1.4928047>
- Dewey, J. (1922) [200]. *Human nature and conduct*. Dover Publications.
- Dittmeyer, M. (2020). *Der programmierte Mensch*. mentis.
- Ebert, C., Gallardo, G., Hernantes, J., & Serrano, N. (2016). DevOps. *Ieee Software*, 33(3), 94-100.
- Edwards, F., Lee, H. & Esposito, M. (2019). Risk of being killed by police use of force in the United States by age, race–ethnicity, and sex. *Proceedings of the National Academy of Sciences*, 116(34), 16793–16798. National Academy of Sciences. <https://doi.org/10.1073/pnas.1821204116>
- Ess, C. (2013). *Digital media ethics*. Polity.
- Eubanks, V. (2012). *Digital dead end: Fighting for social justice in the information age*. MIT Press.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- European Union. (2019). EU High-Level Expert Group on Artificial Intelligence (2019) - Ethics Guidelines for Trustworthy AI I. Zugriff am 8.12.2021. Verfügbar unter: https://ec.europa.eu/info/publications/sustainable-finance-high-level-expert-group_en
- European Union. (2020). *Weissbuch Zur Künstlichen Intelligenz & ein europäisches Konzept für Exzellenz und Vertrauen, COM/2020/65 final*. Website. Publications Office of the European Union. Zugriff am 8.12.2021. Verfügbar unter: <http://op.europa.eu/de/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1>
- Fallman, D. (2007). Persuade into what? why human-computer interaction needs a philosophy of

- technology. *International Conference on Persuasive Technology* (S. 295–306). Springer.
- Feenberg, A. (1991). *Critical theory of technology* (Band 5). Oxford University Press New York.
- Fenech, M., Strukej, N. & Buston, O. (2018). Ethical, social, and political challenges of artificial intelligence in health. *London: Wellcome Trust Future Advocacy*, 12.
- Finger, J. (2012). *Erfolgreiche ERP-Projekte: Ein Rezeptbuch für Manager* (2. Auflage). Berlin Heidelberg: Springer-Verlag. <https://doi.org/10.1007/978-3-642-32829-9>
- Fischer, A. (2017). *Manipulation: zur Theorie und Ethik einer Form der Beeinflussung*. Suhrkamp Verlag.
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. & Srikumar, M. (2020). Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3518482>
- Flanagan, M., Howe, D. C. & Nissenbaum, H. (2005). Values at play: design tradeoffs in socially-oriented game design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 751–760). New York, NY, USA: Association for Computing Machinery. Zugriff am 17.8.2021. Verfügbar unter: <https://doi.org/10.1145/1054972.1055076>
- Flanagan, M., Howe, D. C. & Nissenbaum, H. (2008). Embodying values in technology: Theory and practice. *Information technology and moral philosophy*, 322. Citeseer.
- Flanagan, M., & Nissenbaum, H. (2014). *Values at play in digital games*. MIT Press.
- Floridi, L. (2002). What is the Philosophy of Information? *Metaphilosophy*, 33(1–2), 123–145. Wiley Online Library.
- Floridi, L. (2008). Foundations of information ethics. *The handbook of information and computer ethics*, 1. Wiley Hoboken.
- Floridi, L. (2010). Ethics after the information revolution. *The Cambridge handbook of information and computer ethics*, 3–19. Cambridge University Press.
- Floridi, L. (2011). A Defence of Constructionism: Philosophy as Conceptual Engineering. *Metaphilosophy*, 42(3), 282–304. <https://doi.org/10.1111/j.1467-9973.2011.01693.x>
- Floridi, L. (2013). *The ethics of information* (First edition.). Oxford: Oxford University Press.
- Floridi, L. (2016). Tolerant Paternalism: Pro-ethical Design as a Resolution of the Dilemma of Tolerant Paternalism. *Science and Engineering Ethics*, 22(6), 1669–1688. <https://doi.org/10.1007/s11948-015-9733-2>

- Floridi, L. & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1). PubPub. <https://doi.org/10.1162/99608f92.8cd550d1>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V. et al. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Floridi, L. & Sanders, J. (2005). Internet Ethics: the Constructionist Values of Homo Poieticus.
- Floridi, L. & Sanders, J. W. (2004). On the Morality of Artificial Agents. *Minds and Machines*, 14(3), 349–379. <https://doi.org/10.1023/B:MIND.0000035461.63578.9d>
- Floridi, L. (2008). Foundations of information ethics. *The handbook of information and computer ethics*, 1. Tivani and Himma (eds.)
- Fogg, B. J. (1997). *Charismatic computers: creating more likable and persuasive interactive technologies by leveraging principles from social psychology*. Zugriff am 12.12.2020. Verfügbar unter: <http://books.google.com/books?id=ZMnaAAAAMAAJ>
- Fogg, B. J. (2003). The ethics of persuasive technology. <https://doi.org/10.1016/B978-155860643-2/50011-1>
- Franzius, C. (2015). Das Recht auf informationelle Selbstbestimmung. *ZJS*, 3, 259–270.
- Friedman, B. (1997). *Human values and the design of computer technology* (Band 72). Cambridge University Press.
- Friedman, B. & Kahn Jr, P. H. (1992). Human agency and responsible computing: Implications for computer system design. *Journal of Systems and Software*, 17(1), 7–14. Elsevier.
- Friedman, B., Kahn, P., & Borning, A. (2002). Value sensitive design: Theory and methods. *University of Washington technical report*, 2, 12.
- Friedman, B., Kahn, P. & Borning, A. (2006). Value Sensitive Design and Information Systems. *From Human-Computer Interaction and Management Information Systems: Foundations Advances in Management Information Systems*, 5 (Advances in Management Information Systems), 348–372.
- Frischmann, B. & Selinger, E. (2018). *Re-Engineering Humanity* (1. Auflage). Cambridge University Press. <https://doi.org/10.1017/9781316544846>
- Fuest, B. (2019, August 19). Gegen Massenüberwachung - WELT. Zugriff am 5.8.2021. Verfügbar unter: https://www.welt.de/print/welt_kompakt/print_wirtschaft/article198759379/Gegen-Massenueberwachung.html
- Gandorfer, S. (2020, Juni 24). Die IT-Branche erlebt eine Dekade der Superlative.

- Garcia, D., Abisheva, A. & Schweitzer, F. (2017). Evaluative patterns and incentives in YouTube. *International Conference on Social Informatics* (S. 301–315). Springer.
- Gardner, J. & Warren, N. (2019). Learning from deep brain stimulation: the fallacy of techno-solutionism and the need for 'regimes of care'. *Medicine, Health Care and Philosophy*, 22(3), 363–374. <https://doi.org/10.1007/s11019-018-9858-6>
- George, J. & Offenhartz, J. (2020, August 14). NYPD Used Facial Recognition Technology In Siege Of Black Lives Matter Activist's Apartment. *Gothamist*. Zugriff am 5.8.2021. Verfügbar unter: <https://gothamist.com/news/nypd-used-facial-recognition-unit-in-siege-of-black-lives-matter-activists-apartment>
- Gigerenzer, G. & Selten, R. (2001). *Bounded rationality*. CogNet.
- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für mathematik und physik*, 38(1), 173–198. Springer.
- Gogoll, J. & Müller, J. F. (2017). Autonomous cars: in favor of a mandatory ethics setting. *Science and engineering ethics*, 23(3), 681–700. Springer.
- Gogoll, J., Zuber, N., Kacianka, S., Greger, T., Pretschner, A. & Nida-Rümelin, J. (2021). Ethics in the Software Development Process: from Codes of Conduct to Ethical Deliberation. *Philosophy & Technology*. <https://doi.org/10.1007/s13347-021-00451-w>
- Gosepath, S. (1999). Motive, Gründe, Zwecke. Theorien praktischer Rationalität [Motives, Reasons, Ends. Theories of Practical Rationality].
- Gotterbarn, D. W., Brinkman, B., Flick, C., Kirkpatrick, M. S., Miller, K., Vazansky, K. et al. (2018). ACM Code of Ethics and Professional Conduct. Association for Computing Machinery.
- Gräb-Schmidt, E. & Stritzelberger, C. P. (2018). Ethische Herausforderungen durch autonome Systeme und Robotik im Bereich der Pflege. *Zeitschrift für medizinische Ethik*, 64(4), 357–372.
- Gramlich, J. (2019). The gap between the number of blacks and whites in prison is shrinking. *Pew Research Center*. Zugriff am 5.8.2021. Verfügbar unter: <https://www.pewresearch.org/fact-tank/2019/04/30/shrinking-gap-between-number-of-blacks-and-whites-in-prison/>
- Grimm, P. (2018). Grundlagen für eine digitale Wertekultur. *Mehr als Algorithmen. Digitalisierung in Gesellschaft und Sozialer Arbeit. Sonderband TUP-Theorie und Praxis* https://www.awo.org/sites/default/files/2018-09/TUP-Sonderband_2018_Grimm.pdf.
- Grimm, P., Keber, T. O. & Zöllner, O. (2019). *Digitale Ethik: Leben in vernetzten Welten*. Reclam.

- Grunwald, A. (2018a). *Der unterlegene Mensch: Die Zukunft der Menschheit im Angesicht von Algorithmen, künstlicher Intelligenz und Robotern*. München: Riva.
- Grunwald, A. (2018b). *Technology assessment in practice and theory*. Routledge.
- Habermas, J. (1998). *On the pragmatics of communication*. MIT press.
- Häder, M. & Häder, S. (Hrsg.). (2000). *Die Delphi-Technik in den Sozialwissenschaften*. Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-663-09682-5>
- Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, 30(1), 99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Handelszeitung. (2020, Juni 12). Nach Amazon und IBM: Microsoft setzt Schranken bei der Gesichtserkennung | HZ. *Handelszeitung*. Zugriff am 5.8.2021. Verfügbar unter: <https://www.handelszeitung.ch/news/nach-amazon-und-ibm-microsoft-setzt-schranken-bei-der-gesichtserkennung>
- Harris, T. (2016). How technology hijacks people’s minds—from a magician and Google’s design ethicist. *Medium Magazine*.
- Hartmann, D. & Janich, P. (1998). Die kulturalistische Wende: zur Orientierung des philosophischen Selbstverständnisses.
- Hassan, H. M. & Galal-Edeen, G. H. (2017). From usability to user experience. *2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)* (S. 216–222). IEEE.
- Hegel, G. W. F. (2012). *Grundlinien der Philosophie des Rechts* (Band 13). Рипол Классик.
- Heil, R. (2021). Künstliche Intelligenz/Maschinelles Lernen. *Handbuch Technikethik* (S. 424–428). Springer.
- Henning, T. (2019). *Allgemeine Ethik*. utb GmbH.
- Hess, T. (2019a). Digitalisierung — Enzyklopaedie der Wirtschaftsinformatik. *Enzyklopaedie der Wirtschaftsinformatik*. Zugriff am 15.9.2021. Verfügbar unter: <https://enzyklopaedie-der-wirtschaftsinformatik.de/lexikon/technologien-methoden/Informatik--Grundlagen/digitalisierung>
- Hess, T. (2019b). *Digitale Transformation strategisch steuern: Vom Zufallstreffer zum systematischen Vorgehen*. Springer-Verlag.
- Hetzl, A. (2015). Praxis und praktische Vernunft. *Praxis denken* (S. 109–127). Springer.
- Hevelke, A. & Nida-Rümelin, J. (2015). Selbstfahrende Autos und Trolley-Probleme: Zum Aufrechnen von Menschenleben im Falle unausweichlicher Unfälle. *Jahrbuch für Wissenschaft und Ethik*, 19(1), 5–24. De Gruyter.

- Hill, K. (2020, Juni 24). Wrongfully Accused by an Algorithm. *The New York Times*.
- Hubig, C. (2006). *Die Kunst des Möglichen I: Grundlinien einer dialektischen Philosophie der Technik. Band 1: Technikphilosophie als Reflexion der Medialität*. transcript-Verlag. <https://doi.org/10.14361/9783839404317>
- Hubig, C. (2013). Technik als Medium. In A. Grunwald & M. Simonidis-Puschmann (Hrsg.), *Handbuch Technikethik* (S. 118–123). Stuttgart: J.B. Metzler. https://doi.org/10.1007/978-3-476-05333-6_22
- Hume, D. (1739). *A treatise of human nature*.
- Ihde, D. (1990). *Technology and the Lifeworld: From Garden to Earth*. Bloomington: Indiana University Press.
- Introna, L. D. (2005). Disclosive ethics and information technology: Disclosing facial recognition systems. *Ethics and Information Technology*, 7(2), 75–86. Springer.
- James, W. (1892). The stream of consciousness. *Psychology*, 151–175. World Cleveland & New York.
- Janich, P. (2001). *Logisch-pragmatische Propädeutik: ein Grundkurs im philosophischen Reflektieren*. Velbrück Wiss.
- Janich, P. (2003). Technik und Kulturhöhe. *Technikgestaltung zwischen Wunsch und Wirklichkeit* (S. 91–104). Springer.
- Janich, P. (2013). *Handbuch Technikethik*. (A. Grunwald & M. Simonidis-Puschmann, Hrsg.). Stuttgart: J.B. Metzler. <https://doi.org/10.1007/978-3-476-05333-6>
- Lomas, N (2014). “Jimmy Wales Blasts Europe’s “Right To Be Forgotten” Ruling As A “Terrible Danger”” in: *TechCrunch*. Zugriff am 17.11.2021. Verfügbar unter: <https://social.techcrunch.com/2014/06/07/wales-on-right-to-be-forgotten/>
- Joas, H. & Mead, G. H. (1985). *Das Problem der Intersubjektivität: Neuere Beiträge zum Werk George Herbert Meads*. Suhrkamp.
- Jobin, A., Ienca, M. & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Johnson, D. G. (1985). Computer ethics. *Englewood Cliffs (NJ)*.
- Kacianka, S. & Pretschner, A. (2021). Designing Accountable Systems (FAcCT '21). *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (S. 424–437). New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445905>

- Kaiser, L. (2019, Juli 5). : Gesichtserkennung in London hat miserable Trefferquote und kann Menschenrechte verletzen. Zugriff am 5.8.2021. Verfügbar unter: <https://netzpolitik.org/2019/gesichtserkennung-in-london-hat-miserable-trefferquote-und-kann-menschenrechte-verletzen/>
- Kang, S. K., DeCelles, K. A., Tilcsik, A. & Jun, S. (2016). Whitened résumés: Race and self-presentation in the labor market. *Administrative Science Quarterly*, 61(3), 469–502. Sage Publications Sage CA: Los Angeles, CA.
- Kaptein, M. & Schwartz, M. S. (2008). The effectiveness of business codes: A critical examination of existing studies and the development of an integrated research model. *Journal of Business Ethics*, 77(2), 111–127. Springer.
- Karafyllis, N. C. (2019). Soziotechnisches System. In K. Liggieri & O. Müller (Hrsg.), *Mensch-Maschine-Interaktion: Handbuch zu Geschichte – Kultur – Ethik* (S. 300–303). Stuttgart: J.B. Metzler. https://doi.org/10.1007/978-3-476-05604-7_56
- Kastendiek, A. (2003). *Computer und Ethik statt Computerethik.: Mit einem Geleitwort von Josef Weizenbaum*. LIT Verlag Münster.
- Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in human behavior*, 66, 236–247. Elsevier.
- Kiesewetter, B. (2017). *The Normativity of Rationality*. Oxford, New York: Oxford University Press.
- Klimke, R. (2020). Menschenrecht auf einen Internetzugang? Deutsche Nationalbibliothek.
- Knight, W. (2017). The dark secret at the heart of ai. *Technology Review*, 120(3), 54–61. TECHNOLOGICAL REV 1 MAIN ST, 13 FLR, CAMBRIDGE, MA 02142 USA.
- Kogge, W. (2008). Technologie des 21. Jahrhunderts. Perspektiven der Technikphilosophie. *Deutsche Zeitschrift für Philosophie*, 56(6). <https://doi.org/10.1524/dzph.2008.0073>
- König, W. (2013). VDI-Richtlinie zur Technikbewertung. *Handbuch Technikethik* (S. 406–410). Springer.
- Korsgaard, C. (1997). The normativity of instrumental reason. Cambridge University Press.
- de Kort, Y., IJsselstein, W., Midden, C., Eggen, B. & Fogg, B. J. (Hrsg.). (2007). *Persuasive Technology: Second International Conference on Persuasive Technology, PERSUASIVE 2007, Palo Alto, CA, USA, April 26-27, 2007, Revised Selected Papers* (Lecture Notes in Computer Science) (Band 4744). Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-77006-0>
- Krämer, S. (1998). Das Medium als Spur und als Apparat. *Medien, Computer, Realität. Wirklichkeitsvorstellungen und Neue Medien*, 1, 73–94.

- Kretschmar, D. (2020, Juli 1). Rassismus in Algorithmen: Gesichtserkennung für die Tonne. *Die Tageszeitung: taz*.
- Krings, H., Baumgartner, H. M., Wild, C., Kolmer, P. & Wildfeuer, A. G. (2011). *Neues Handbuch philosophischer Grundbegriffe*. Alber.
- Krügel, S., Ostermaier, A. & Uhl, M. (2021). Zombies in the Loop? People are Insensitive to the Transparency of AI-Powered Moral Advisors. *arXiv preprint arXiv:2106.16122*.
- Kühl, E. (2015, Juli 2). Gesichtserkennung: „Meine Freundin ist kein Gorilla“. *Die Zeit*.
- Kudina, O. (2019). The technological mediation of morality: value dynamism, and the complex interaction between ethics and technology. <https://doi.org/10.3990/1.9789036547444>
- van Lamsweerde, A. (2000). Goal-oriented requirements engineering: a guided tour. *Proceedings Fifth IEEE International Symposium on Requirements Engineering* (S. 249–262). Gehalten auf der Fifth IEEE International Symposium on Requirements Engineering, Toronto, Ont., Canada: IEEE Comput. Soc. <https://doi.org/10.1109/ISRE.2001.948567>
- Lang, F. (1927). *Metropolis*. Stummfilm.
- Langford, D. (2000). *Internet ethics*. Macmillan Basingstoke and London.
- Lehto, M., Huhtinen, A.-M. & Jantunen, S. (2011). The Open Definition of Cyber: Technology or a Social Construction? *International Journal of Cyber Warfare and Terrorism (IJCWT)*, 1(2), 1–9. IGI Global.
- Lenk, H. & Ropohl, G. (1987). *Technik und Ethik*. Reclam.
- Lenk, H. (2018). „Neue Realitäten–Herausforderung der Philosophie“. In *Neue Realitäten. Herausforderung der Philosophie* (pp. 15-25). Akademie Verlag.
- Lord, E. (2015). Acting for the Right Reasons, Abilities, and Obligation. In R. Shafer-Landau (Hrsg.), *Oxford Studies in Metaethics, Volume 10* (S. 26–52). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198738695.003.0002>
- Lord, E. (2018). *The Importance of Being Rational*. Oxford, New York: Oxford University Press.
- MacIntyre, A. C. (2007). *After virtue: a study in moral theory* (3rd ed.). Notre Dame, Ind: University of Notre Dame Press.
- Mainzer, K. (2019). *Künstliche Intelligenz – Wann übernehmen die Maschinen?* Technik im Fokus. Springer.
- Manders-Huits, N. & Zimmer, M. (2009). Values and Pragmatic Action: The Challenges of Introducing

Ethical Intelligence in Technical Design Communities, 10, 8.

Maner, W. (1996). Unique ethical problems in information technology. *Science and Engineering Ethics*, 2(2), 137–154. Springer.

Maner, W. (1999). Is computer ethics unique? EUT Edizioni Università di Trieste.

Marcuse, H. (1996). Kritische Theorie—Frankfurter Schule. *Technikkritik in der Moderne* (S. 79–116). Springer.

Markowetz, A. (2015). *Digitaler Burnout: warum unsere permanente Smartphone-Nutzung gefährlich ist*. Droemer eBook.

Materese, R. (2020, Juli 27). NIST Launches Studies into Masks' Effect on Face Recognition Software. *NIST*. text, Zugriff am 5.8.2021. Verfügbar unter: <https://www.nist.gov/news-events/news/2020/07/nist-launches-studies-masks-effect-face-recognition-software>

Mattern, F. (2005). Die technische Basis für das Internet der Dinge. In E. Fleisch & F. Mattern (Hrsg.), *Das Internet der Dinge: Ubiquitous Computing und RFID in der Praxis: Visionen, Technologien, Anwendungen, Handlungsanleitungen* (S. 39–66). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-28299-8_2

McAfee, A. & Brynjolfsson, E. (2017). *Machine, platform, crowd: Harnessing our digital future*. WW Norton & Company.

McLaughlin, M. & Castro, D. (2020). *The Critics Were Wrong: NIST Data Shows the Best Facial Recognition Algorithms Are Neither Racist Nor Sexist*. Information Technology and Innovation Foundation.

McLennan, S., Fiske, A., Celi, L. A., Müller, R., Harder, J., Ritt, K. et al. (2020). An embedded ethics approach for AI development. *Nature Machine Intelligence*, 2(9), 488–490. Nature Publishing Group.

McNamara, A., Smith, J. & Murphy-Hill, E. (2018). Does ACM's code of ethics change ethical decision making in software development? *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering - ESEC/FSE 2018* (S. 729–733). Gehalten auf der the 2018 26th ACM Joint Meeting, Lake Buena Vista, FL, USA: ACM Press. <https://doi.org/10.1145/3236024.3264833>

Mead, G. H. (1923) [1967]. Scientific method and the moral sciences. *The International Journal of Ethics*, 33(3), 229–247. Rumford Press.

Mead, G. H. (1934). *Mind, self and society* (Band 111). Chicago University of Chicago Press.

- Mead, G. H. (1938). *The Philosophy of the Act*. Edited, with Introd. By Charles W. Morris in Collaboration with John M. Brewster, Albert M. Dunham [and] David L. Miller.
- Merdes, C. (2020). Epistemische Intransparenz als Herausforderung für praktisches Urteilen. *Wissenschaftsreflexion*, 251–267. Brill mentis. https://doi.org/10.30965/9783957437372_011
- Metz. (2021). Facebook is shutting down its facial recognition software - CNN. Zugriff am 17.11.2021. Verfügbar unter: <https://edition.cnn.com/2021/11/02/tech/facebook-shuts-down-facial-recognition/index.html>
- Meyer, R. (2021, Mai 21). Gesichtserkennung mit meinen Selfies? - Wie man sich digital maskiert. *Deutschlandfunk Kultur*. Zugriff am 5.8.2021. Verfügbar unter: https://www.deutschlandfunkkultur.de/gesichtserkennung-mit-meinen-selfies-wie-man-sich-digital.2156.de.html?dram:article_id=497433
- Mill, J. S. (1859) [2015]. *On liberty and other essays*. Oxford University Press, USA.
- Miller, D. L. (1973). *George Herbert Mead: self, language, and the world*. Austin: University of Texas Press.
- Miller, D. L. (1982). The Meaning of Freedom from the Perspective of G.H.Mead's Theory of the Self. *The Southern Journal of Philosophy*, 20(4), 453–463. <https://doi.org/10.1111/j.2041-6962.1982.tb00312.x>
- Mitcham, C. (1994). *Thinking through technology: The path between engineering and philosophy*. University of Chicago Press.
- Mohr, T. (2020). *Der Digital Navigator: Ein Modell für die digitale Transformation*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-29704-6>
- Mondal, M., Silva, L. A., Correa, D. & Benevenuto, F. (2018). Characterizing usage of explicit hate expressions in social media. *New Review of Hypermedia and Multimedia*, 24(2), 110–130. Taylor & Francis.
- Moor, J. (1985). What is Computer Ethics? *Metaphilosophy*, 16(4), 266–275. Wiley.
- Moor, J. (1998). Reason, relativity, and responsibility in computer ethics. *Acm Sigcas Computers and Society*, 28(1), 14–21. ACM New York, NY, USA.
- Moor, J. (2003). *The Turing test: the elusive standard of artificial intelligence* (Band 30). Springer Science & Business Media.
- Moor, J. (2005). Why We Need Better Ethics for Emerging Technologies. *Ethics and Information Technology*, 7(3), 111–119. <https://doi.org/10.1007/s10676-006-0008-0>

- Moor, J. (2006). The Nature, Importance, and Difficulty of Machine Ethics. *IEEE Intelligent Systems*, 21, 18–21. <https://doi.org/10.1109/MIS.2006.80>
- Morley, J., Floridi, L., Kinsey, L. & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics*, 26(4), 2141–2168. <https://doi.org/10.1007/s11948-019-00165-5>
- Nechushtai, E. & Lewis, S. C. (2019). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, 90, 298–307. Elsevier.
- Neugebauer, R. (Hrsg.). (2018a). *Digitalisierung: Schlüsseltechnologien für Wirtschaft und Gesellschaft* (Fraunhofer-Forschungsfokus) (1. Auflage.). Berlin Heidelberg: Springer Vieweg.
- Neugebauer, R. (2018b). Digitale Information – der „genetische Code“ moderner Technik. In R. Neugebauer (Hrsg.), *Digitalisierung: Schlüsseltechnologien für Wirtschaft und Gesellschaft* (S. 1–7). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-55890-4_1
- Nida-Rümelin, J. (1993a). Begründung in der Ethik. *LOGOS Zeitschrift für systematische Philosophie*, 1(3), 306–321.
- Nida-Rümelin, J. (1993b). *Kritik des Konsequentialismus* (Scientia nova). München: R. Oldenbourg.
- Nida-Rümelin, J. (1996). *Angewandte Ethik: die Bereichsethiken und ihre theoretische Fundierung; ein Handbuch*. Kröner.
- Nida-Rümelin, J. (1997a). Structural Rationality and Coherentism. *Economic Rationality and Practical Reason* (S. 140–161). Springer.
- Nida-Rümelin, J. (1997b). Praktische Kohärenz. *Zeitschrift für philosophische Forschung*, (H. 2), 175–192. JSTOR.
- Nida-Rümelin, J. (1998). Ethische Begründungen: ein kritischer Überblick. *Ethica*, 6(1), 7–36.
- Nida-Rümelin, J. (1999). Zur Rolle ethischer Expertise in Projekten der Technikfolgenabschätzung. *Angewandte Ethik in der pluralistischen Gesellschaft, Freiburg/Schweiz*.
- Nida-Rümelin, J. (2001). *Strukturelle Rationalität: ein philosophischer Essay über praktische Vernunft*. Philipp Reclam jun.
- Nida-Rümelin, J. (2005a). *Über menschliche Freiheit*. Reclam.
- Nida-Rümelin, J. (2005b, Oktober 29). Persönliche Schuld und politischer Wahn. Gehalten auf der 8. Hannah Arendt Tage, Hannover.

- Nida-Rümelin, J. (2008). Moralische Urteilsbildung und die Theorie der „Angewandten Ethik“. *Praxis in der Ethik* (S. 17–46). De Gruyter.
- Nida-Rümelin, J. (2009). Philosophie und Lebensform. *Suhrkamp Verlag*. Zugriff am 5.8.2021. Verfügbar unter: <https://www.suhrkamp.de/buch/julian-nida-ruemelin-philosophie-und-lebensform-t-9783518295328>
- Nida-Rümelin, J. (2011a). *Verantwortung*. Reclam.
- Nida-Rümelin, J. (2011b). *Die Optimierungsfalle: Philosophie einer humanen Ökonomie*. Irisiana.
- Nida-Rümelin, J. (2016). *Humanistische Reflexionen*. Suhrkamp Verlag.
- Nida-Rümelin, J. (2018). *Unaufgeregter Realismus: Eine philosophische Streitschrift*. Brill mentis.
- Nida-Rümelin, J. & Weidenfeld, N. (2018). *Digitaler Humanismus: eine Ethik für das Zeitalter der künstlichen Intelligenz* (Originalausgabe.). München: Piper-rümelin,
- Nida-Rümelin, J. (2019). *Structural Rationality and Other Essays on Practical Reason* (Vol. 52). Springer.
- Nida-Rümelin, J. (2020). *Eine theorie praktischer vernunft* (1st ed.). Boston: DE GRUYTER.
- Nida-Rümelin, J. (2022). Digitaler Humanismus – philosophische Aspekte Künstlicher Intelligenz. In J. Chibanguza, C. Kuß & J. Steege (Hrsg.), *Künstliche Intelligenz. Recht und Praxis automatisierter und autonomer Systeme*. Baden-Baden: Nomos Verlag.
- Nida-Rümelin, J. & Schmidt, T. (2000). *Rationalität in der praktischen Philosophie: eine Einführung*. Akademie-Verlag.
- Nida-Rümelin, J., Spiegel, I. & Tiedemann, M. (2017). *Handbuch Philosophie und Ethik: Band 2: Disziplinen und Themen* (Band 2). UTB.
- Nissenbaum, H. (1994). Computing and accountability. *Communications of the ACM*, 37(1), 72–81. Association for Computing Machinery, Inc.
- Nissenbaum, H. (1996). Accountability in a computerized society. *Science and Engineering Ethics*, 2(1), 25–42. <https://doi.org/10.1007/BF02639315>
- Oelschlaeger, M. (1979). The Myth of the Technological Fix. *The Southwestern Journal of Philosophy*, 10(1), 43–53. University of Arkansas Press.
- Oinas-Kukkonen, H. & Harjumaa, M. (2009). Persuasive systems design: Key issues, process model, and system features. *Communications of the Association for Information Systems*, 24(1), 28.

- Orth, E. W. (1987). Zum Begriff der Technik bei Ernst Cassirer und Martin Heidegger. *Phänomenologische Forschungen*, 20, 91–122. JSTOR.
- Ott, K. (2005). Technikethik. In J. Nida-Rümelin (Hrsg.), *Angewandte Ethik: Die Bereichsethiken und ihre theoretische Fundierung. Ein Handbuch* (2. vollständig überarbeitete Auflage., S. 568–648).
- Ott, K. (2013). Diskursethik. *Handbuch Technikethik* (S. 163–168). Springer.
- Pakaluk, M. (2005). Aristotle's Nicomachean Ethics: An Introduction, 360.
- Palm, E. & Hansson, S. O. (2006). The case for ethical technology assessment (eTA). *Technological Forecasting and Social Change*, 73(5), 543–558. <https://doi.org/10.1016/j.techfore.2005.06.002>
- Pfaff, C. C. & Hasan, H. (2011). Wiki-based knowledge management systems for more democratic organizations. *Journal of Computer Information Systems*, 52(2), 73–82. Taylor & Francis.
- Plantin, J.-C., Lagoze, C., Edwards, P. N. & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293–310. SAGE Publications. <https://doi.org/10.1177/1461444816661553>
- Plantin, J.-C. & Punathambekar, A. (2019). Digital media infrastructures: pipes, platforms, and politics. *Media, Culture & Society*, 41(2), 163–174. SAGE Publications Ltd. <https://doi.org/10.1177/0163443718818376>
- Phillip (2020); „GPT-3 Bot Posed as a Human on AskReddit for a Week“. *kmeme*, <https://www.kmeme.com/2020/10/gpt-3-bot-went-undetected-askreddit-for.html>. Zuletzt darauf zugegriffen am 6.9.2022.
- Pütter, C. (2019). Prognose für 2019: Software- und IT-Services-Markt wächst um fünf Prozent.
- Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J. & Denton, E. (2020). Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing. *arXiv:2001.00964 [cs]*.
- Ramge, T. (2018). *Mensch und Maschine: wie künstliche Intelligenz und Roboter unser Leben verändern*. Reclam Verlag.
- Raymond, A. H. & Shackelford, S. J. (2013). Technology, ethics, and access to justice: Should an algorithm be deciding your case. *Mich. J. Int'l L.*, 35, 485. HeinOnline.
- Reijers, W. & Coeckelbergh, M. (2020). *Narrative and Technology Ethics*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-60272-7>
- Rese, F. (2013). Praxis und Poiesis in zeitlicher Perspektive. In W. Mesch (Hrsg.), *Glück — Tugend — Zeit: Aristoteles über die Zeitstruktur des guten Lebens* (S. 185–200). Stuttgart: J.B. Metzler. https://doi.org/10.1007/978-3-476-05343-5_11

Reviglio, U. (2017). Serendipity by design? How to turn from diversity exposure to diversity experience to face filter bubbles in social media. *International Conference on Internet Science* (S. 281–300). Springer.

Rifkin, J. (2014). *The zero marginal cost society: The internet of things, the collaborative commons, and the eclipse of capitalism*. St. Martin's Press.

Rip, A. & Kemp, R. (1998). „Technological change“. *Human choice and climate change* 2, Nr. 2 (1998): 327–99.

Robertson, D. J., Megreya, A. & Davis, J. P. (2020). Facial recognition: research reveals new abilities of 'super-recognisers'.

Rohbeck, J. (1993). *Technologische Urteilskraft: Zu einer Ethik technischen Handelns*. Suhrkamp.

Ropohl, G. (1994). Das Risiko Im Prinzip Verantwortung. *Ethik Und Sozialwissenschaften*, 5(1), 109–120. Lucius & Lucius.

Ropohl, G. (2009). *Allgemeine Technologie - Eine Systemtheorie der Technik*. universitätsverlag karlsruhe.

Russo, F. (2012a). The Homo Poieticus and the Bridge Between Physis and Techne (Philosophy of Engineering and Technology). In H. Demir (Hrsg.), *Luciano Floridi's Philosophy of Technology* (Band 8, S. 65–81). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-4292-5_4

Russo, F. (2012b). The homo poieticus and the bridge between physis and techne. https://doi.org/10.1007/978-94-007-4292-5_4

Russo, F. (2016). On the Poietic Character of Technology, (30), 28.

Saris, P. B., Breyer, C. R., Friedrich, D. L., Barkow, R. E., Pryor, W. H., Morales, M. et al. (2016). GUIDELINES MANUAL 2016, 628.

Scanlon, T. (2000). *What we owe to each other*. Belknap Press.

Scanlon, T. (2007). Structural irrationality. *Common minds: Themes from the philosophy of Philip Pettit*, 84–103. Oxford: Oxford University Press.

Schimmeck, T. (2017).. „Algorithmen im US-Justizsystem - Schicksalsmaschinen“. Hörspiel und Feature - Deutschlandfunkkultur, 2017. https://www.deutschlandfunkkultur.de/algorithmen-im-us-justizsystem-schicksalsmaschinen.3720.de.html?dram:article_id=385478. Zuletzt zugegriffen am 27.12.2021.

Schulz, S. (2012). Privacy by design. *Computer und Recht*, 28(3), 204–208. Verlag Dr. Otto Schmidt.

- Schwartz, H. A. (2020, April 14). How Accurate are Facial Recognition Systems – and Why Does It Matter? Zugriff am 5.8.2021. Verfügbar unter: <https://www.csis.org/blogs/technology-policy-blog/how-accurate-are-facial-recognition-systems-%E2%80%93-and-why-does-it-matter>
- Schwartz, M. (2001). The nature of the relationship between corporate codes of ethics and behaviour. *Journal of business ethics*, 32(3), 247–262. Springer.
- Searle, J. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, 3(3), 417–457.
- Searle, J. (1990). Is the Brain's Mind a Computer Program? *Scientific American*, 262(1), 25–31. Scientific American, a division of Nature America, Inc.
- Sen, A. K. (1977). Rational Fools: A Critique of the Behavioral Foundations of Economic Theory. *Philosophy & Public Affairs*, 6(4), 317–344. Wiley.
- Senges, M., Ryan, P. S. & Whitt, R. S. (2017). Composite Ethical Frameworks for IOT and Other Emerging Technologies. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3092362>
- Shen, H., Deng, W. H., Chattopadhyay, A., Wu, Z. S., Wang, X. & Zhu, H. (2021). Value Cards: An Educational Toolkit for Teaching Social Impacts of Machine Learning through Deliberation. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (S. 850–861). Gehalten auf der FAccT '21: 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event Canada: ACM. <https://doi.org/10.1145/3442188.3445971>
- Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R. & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision support systems*, 33(2), 111–126. Elsevier.
- Siep, L. (2013). *Vernunft und Tugend*. mentis.
- Simon, H. A. (1997). *Models of bounded rationality: Empirically grounded economic reason* (Band 3). MIT press.
- Simon, J. (2016). In J. Heesen (Hrsg.), *Handbuch Medien- und Informationsethik* (S. 357–366). Stuttgart: J.B. Metzler Verlag.
- Simonsen, J. & Robertson, T. (2012). *Routledge International Handbook of Participatory Design*. Routledge.
- Skorupinski, B. & Ott, K. (2000). *Technikfolgenabschätzung und Ethik: eine Verhältnisbestimmung in Theorie und Praxis*. Zürich: vdf, Hochschulverl. an der ETH.
- Smith, A. (1759) [2010]. *The theory of moral sentiments*. Penguin.
- Smith, M. (1987). The Humean Theory of Motivation. *Mind*, XCVI(381), 36–61.

<https://doi.org/10.1093/mind/XCVI.381.36>

Snow, J. (2018, Juli 26). Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots. *American Civil Liberties Union*. Zugriff am 5.8.2021. Verfügbar unter: <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>

Sollie, P. (2007). Ethics, technology development and uncertainty: an outline for any future ethics of technology. *Journal of Information, Communication and Ethics in Society*, 5(4), 293–306. <https://doi.org/10.1108/14779960710846155>

Søraker, J. H. & Brey, P. (2007). Ambient intelligence and problems with inferring desires from behaviour. *International Review of Information Ethics*, 8(1), 7–12.

Spiekermann, S. (2015). *Ethical IT Innovation: A Value-Based System Design Approach*. *Ethical IT Innovation: A Value-Based System Design Approach* (S. 248). <https://doi.org/10.1201/b19060>

Spiekermann, S. (2019). *Digitale Ethik: ein Wertesystem für das 21. Jahrhundert*. Droemer eBook.

Spiekermann, S. (2021). Value-based Engineering: Prinzipien und Motivation für bessere IT-Systeme. *Informatik Spektrum*, 44(4), 247–256. Springer.

Spiekermann, S. & Winkler, T. (2020). Value-based Engineering for Ethics by Design. *arXiv:2004.13676 [cs]*.

Stahl, B., Timmermans, J. & Mittelstadt, B. (2016). The Ethics of Computing. *ACM Computing Surveys*, 48, 1–38. <https://doi.org/10.1145/2871196>

Stalder, F. (2016). *Kultur der Digitalität*. Suhrkamp Verlag.

Statistisches Bundesamt (2019): *Branche - Unternehmensanzahl Deutschland bis 2019*. Statista. Zugriff am 17.11.2021. Verfügbar unter: <https://de.statista.com/statistik/daten/studie/189870/umfrage/anzahl-der-unternehmen-in-der-it-branche-in-deutschland/>

Stern, L. (1974). Freedom, blame, and moral community. *The Journal of Philosophy*, 71(3), 72–84.

Strawson, P. F. (1974) [2008]. *Freedom and resentment and other essays*. Routledge.

Süddeutsche Zeitung. (2019, Januar 24). Polizei - „Super-Recogniser“ helfen bei 200 Fällen. *Süddeutsche.de*. Zugriff am 5.8.2021. Verfügbar unter: <https://www.sueddeutsche.de/muenchen/super-recogniser-polizei-gesichtserkenner-1.4301590>

Sunstein, C. R. (2018). Is social media good or bad for democracy. *SUR-Int'l J. on Hum Rts.*, 27, 83. HeinOnline.

- Swierstra, T. & Rip, A. (2007). Nano-ethics as NEST-ethics: Patterns of Moral Argumentation About New and Emerging Science and Technology. *NanoEthics*, 1(1), 3–20. <https://doi.org/10.1007/s11569-007-0005-8>
- Van der Burg, S., & Swierstra, T. (Eds.). (2013). *Ethics on the laboratory floor*. Springer.
- Systems and Software Engineering Standards Committee. (2020). *IEEE 7000-2021*.
- Tavani, H. T. (1996). Cyberethics and the future of computing. *ACM SIGCAS Computers and Society*, 26(2), 22–29. <https://doi.org/10.1145/236394.236396>
- Tavani, H. T. (2013). *Ethics and technology: controversies, questions, and strategies for ethical computing* (Fourth edition.). Hoboken, NJ: Wiley.
- Thaler, R. H. & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- Thoreau, H. D. (1866) [2004]. *Civil disobedience, and other essays*. Courier Corporation.
- Toulmin, S. E. (2003). *The uses of argument*. Cambridge university press.
- Turilli, M. & Floridi, L. (2009). The ethics of information transparency. *Ethics and Information Technology*, 11(2), 105–112. <https://doi.org/10.1007/s10676-009-9187-9>
- Turner, C. S. (1993). An investigation of the therac-25 accidents. *Computer*, 18(9|62|93), 0700–001830300.
- Umbrello, S. & Gambelin, O. (2021). A Value Sensitive Design Toolkit for Agile Project Management. *A Value Sensitive Design Toolkit for Agile Project Management*.
- Vaidhyanathan, S. (2018). *Antisocial media: how facebook disconnects US and undermines democracy*. New York, NY, United States of America: Oxford University Press.
- Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.
- Van Den Eede, Y. (2011). In Between Us: On the Transparency and Opacity of Technological Mediation. *Foundations of Science*, 16(2–3), 139–159. <https://doi.org/10.1007/s10699-010-9190-y>
- Van den Hoven, J., Vermaas, P. E. & Van de Poel, I. (2015). Handbook of ethics, values and technological design. *Handbook of Ethics, Values, and Technological Design*.
- VDI (2000). *VDI 3780 - Technikbewertung - Begriffe und Grundlage*.
- Verbeek, P.-P. (2011). *Moralizing technology: understanding and designing the morality of things*. Chicago ; London: The University of Chicago Press.

- Verbeek, P.-P. (2016). *Materializing Morality: Design Ethics and Technological Mediation*. *Science, Technology, & Human Values*. Sage Publications/Sage CA: Thousand Oaks, CA. <https://doi.org/10.1177/0162243905285847>
- Vincent, J. (2020, August 18). NYPD used facial recognition to track down Black Lives Matter activist. *The Verge*. Zugriff am 5.8.2021. Verfügbar unter: <https://www.the-verge.com/2020/8/18/21373316/nypd-facial-recognition-black-lives-matter-activist-derrick-ingram>
- Voigt, H. C. (2014). Menschenrecht auf Internet? *Medienwandel kompakt 2011-2013* (S. 353–364). Springer.
- Vollmer, N. (2021, Juli 2). Article 17 EU General Data Protection Regulation (EU-GDPR). text, SecureDataService. Zugriff am 15.9.2021. Verfügbar unter: <https://www.privacy-regulation.eu/en/article-17-right-to-erasure-'right-to-be-forgotten'-GDPR.htm>
- Wallach, W. & Allen, C. (2008). *Moral machines: Teaching robots right from wrong*. Oxford University Press.
- Wallach, W. & Allen, C. (2009). *Moral machines: teaching robots right from wrong*. Oxford; New York: Oxford University Press.
- Walsh, T. (2018). *It's alive: Wie künstliche Intelligenz unser Leben verändern wird*. Edition Körber.
- Wedgwood, R. (2014). Rationality as a Virtue: Rationality as a Virtue. *Analytic Philosophy*, 55(4), 319–338. <https://doi.org/10.1111/phib.12055>
- Wedgwood, R. (2017). *The Value of Rationality*. Oxford, New York: Oxford University Press.
- Weinberger, O. (1979). Handeln und Schließen, Überlegungen zum Begriff der praktischen Inferenz. *Philosophica*, 23(1), 5–36.
- Werthner, H., Lee, E. A., Vardi, M., Hardman, L., Ghezzi, C. & Nowotny, H. (2019). Wiener Manifest für Digitalen Humanismus. Zugriff am 8.12.2021. Verfügbar unter: https://dighum.ec.tuwien.ac.at/wp-content/uploads/2019/07/Vienna_Manifesto_on_Digital_Humanism_DE.pdf
- Wewer, G. (2020). Digitale Ethik. *Handbuch Digitalisierung in Staat und Verwaltung* (S. 231–240). Springer.
- Wiegerling, K. (2011). *Philosophie intelligenter Welten*. Wilhelm Fink Verlag.
- Wiegerling, K. (2013). Ubiquitous Computing. *Handbuch Technikethik* (S. 374–378). Springer.
- Wiener, N. (1950). Cybernetics. *Bulletin of the American Academy of Arts and Sciences*, 3(7), 2–4. JSTOR.

- Wiener, N. (1988). *The human use of human beings: Cybernetics and society*. Da Capo Press.
- Wietlisbach, O. (2017). „13 Fakten zum SMS, das versehentlich Geschichte schrieb“. (2017). *watson.ch*. Zugriff am: 3.12.2021. Verfügbar unter: <https://www.watson.ch/!330480101>
- Winfield, A. F. & Jirotko, M. (2017). The case for an ethical black box. *Annual Conference Towards Autonomous Robotic Systems* (S. 262–273). Springer.
- Winter, S. J. & Butler, B. S. (2022). Responsible Technology Design: Conversations for Success. *Perspectives on Digital Humanism* (S. 271–275). Springer.
- Wolf, W. (2009). Cyber-physical systems. *Computer*, 42(03), 88–89. IEEE Computer Society.
- Wolfangel, E. (2021). Timnit Gebru: Wie viel Ethik verträgt Google? *Zeit Online*.
- Wolff, M-L. (2020). *Die Anbetung*. Über eine Superideologie namens Digitalisierung. Westend Verlag.
- Wright, D. (2011). A framework for the ethical impact assessment of information technology. *Ethics and Information Technology*, 13(3), 199–226. <https://doi.org/10.1007/s10676-010-9242-6>
- Wright, J. (2021). Suspect AI: Vibraimage, Emotion Recognition Technology and Algorithmic Opacity. *Science, Technology and Society*, 09717218211003411. SAGE Publications India. <https://doi.org/10.1177/09717218211003411>
- Xafis, V., Schaefer, G. O., Labude, M. K., Brassington, I., Ballantyne, A., Lim, H. Y. et al. (2019). An Ethics Framework for Big Data in Health and Research. *Asian Bioethics Review*, 11(3), 227–254. <https://doi.org/10.1007/s41649-019-00099-x>
- Zeng, Y., Lu, E. & Huangfu, C. (2018). Linking artificial intelligence principles. *arXiv preprint arXiv:1812.04814*.
- Zeng, Y., Lu, E., Sun, Y. & Tian, R. (2019). Responsible facial recognition and beyond. *arXiv preprint arXiv:1909.12935*.
- Zieglmeier, V. & Pretschner, A. (2021). Trustworthy Transparency by Design. *arXiv:2103.10769 [cs]*.
- Žliobaitė, I. & Custers, B. (2016). Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models. *Artificial Intelligence and Law*, 24(2), 183–201. <https://doi.org/10.1007/s10506-016-9182-5>
- Zuber, N., Kacianka, S., Gogoll, J., Pretschner, A. & Nida-Rümelin, J. (2021). Empowered and embedded: ethics and agile processes. *arXiv preprint arXiv:2107.07249*.
- Zuber, N., Kacianka, S., Pretschner, A. & Nida-Rümelin, J. (2020). *Ethische Deliberation für agile Softwareprozesse: EDAP-Schema* (Band Digitale Transformation und Ethik). Ecowin.

Zwitter, A. (2014). Big Data ethics. *Big Data & Society*, 1(2), 205395171455925.
<https://doi.org/10.1177/2053951714559253>