

1 Aktueller Stand der Digitalethik

Alle kennen es: Man möchte einen neuen Ausweis beantragen oder den Wohnsitz ummelden und muss lange auf einen Termin warten. Dann muss man zum Bürgeramt und dort wieder warten. Das soll durch die Digitalisierung der Verwaltungsprozesse bald schneller gehen. Durch das Gesetz zur Verbesserung des Onlinezugangs zu Verwaltungsleistungen (Onlinezugangsgesetz – OZG) sind „Bund und Länder (...) verpflichtet, ihre Verwaltungsleistungen auch elektronisch über Verwaltungsportale anzubieten.“¹ Das heißt, dass es möglich sein soll, Verwaltungsleistungen online zu beantragen, ohne auf einen Termin beim Amt warten zu müssen. Dies wird durch die Digitalisierung ermöglicht. Im Englischen werden zwei Dimensionen der Digitalisierung unterschieden: Während „Digitization“ die technischen Möglichkeiten der Digitalisierung beschreibt und den Einsatz von digitalen Technologien zur Umwandlung von analogen Werten in digitale Formate bezeichnet², beschreibt „Digitalization“ den Prozess der Digitalisierung, also die Einführung und verstärkte Nutzung dieser technischen Möglichkeiten.³ Diese verstärkte Nutzung der technischen Möglichkeiten („Digitalization“) soll auch – auf der Grundlage des Onlinezugangsgesetzes (OZG) in die Verwaltung eingeführt und dort verstärkt genutzt werden. Dies fußt auf der „Digitization“, welche es ermöglicht, Daten digital zu speichern und schneller zu verarbeiten. So soll zum Beispiel das Beantragen von Wohngeld durch den Einsatz digitaler Technologien von Zuhause aus möglich sein.

Um solche digitalen Anträge möglich zu machen, müssen sogenannte Datenökosysteme gestaltet werden. Datenökosysteme können als komplexe sozio-technische Systeme definiert werden, welche

1 §1a, Abs. 1, Satz 1 OZG

2 Donges 2022:212f.; Görs et al. 2022:12; Klenk et al. 2020:5; Altintac 2022:3,6; Corsten, Roth 2022:10; Harwardt 2022:3; Traum et al. 2017:2

3 Görs et al. 2022:12; Traum et al. 2017:2f.

durch das Zusammenspiel zwischen menschlichen Akteuren⁴ und Technologien das Sammeln, Speichern und Verarbeiten von Daten sowie das Bereitstellen der Ergebnisse ermöglichen.⁵ Im Folgenden wird ausführlich definiert, was ein Datenökosystem ist, und es wird ein Modell von der Struktur eines Datenökosystems erstellt, in welchem der Datenfluss zwischen vier Akteuren abgebildet wird: Datenquellen, Datengenerierenden, Datenverarbeitenden und Datennutzenden. Teil eines Datenökosystems zu sein kann viele Vorteile haben, aber auch Risiken bergen. In dieser Arbeit soll der Umgang mit digitalen Daten in Datenökosystemen systematisch beleuchtet werden.

Die Vorteile und Risiken von Datenökosystemen lassen sich anhand von Beispielen aus der öffentlichen Verwaltung einführend aufzeigen: Bei der digitalen Gestaltung des Wohngeldantrages in Form eines Datenökosystems, geben Antragstellende ihre Daten online an. Um einen Antrag auf Wohngeld stellen zu können, müssen viele persönliche Informationen preisgegeben werden, wie zum Beispiel Name, Anschrift, Einkommensnachweise, gegebenenfalls der Nachweis einer Schwerbehinderung sowie Bescheide über Kindergeld, Kinderzuschlag oder Elterngeld. Die Wohngeldbehörde speichert und verarbeitet diese Daten, um feststellen zu können, ob die Antragstellenden Anspruch haben auf Wohngeld und der Staat verpflichtet ist, ihnen dies auszuzahlen. Nach der Prüfung bekommen die Antragstellenden Bescheid, ob sie Anspruch auf Wohngeld haben oder nicht. Gleichzeitig werden die Daten anonymisiert (ohne Namen und Anschrift) an das Statistische Landesamt, an das Statistische Bundesamt sowie an das Bundesministerium des Innern, für Bau und Heimat und an das Bundesamt für Bauwesen und Raumordnung weitergegeben, damit diese die Daten im Rahmen einer Wohngeldstatistik verarbeiten können.⁶

Der Vorteil der digitalen Gestaltung des Wohngeldantrages in einem Datenökosystem liegt auf der Hand: ein geringerer Arbeitsaufwand für die Antragstellenden und die Wohngeldbehörde. Für die Antragstellenden ist es weniger Aufwand, ihre Daten online

4 Mit dem Begriff „Akteur“ sind hier und im Folgenden Personen aller Geschlechter gemeint.

5 Oliveira et al. 2019:590

6 §§ 34–36 WoGG; Siehe zum Beispiel der Wohngeldantrag des Landes Bremen: https://www.wohngeld.org/wp-content/uploads/AntragMZ_LZ_2021.pdf

anzugeben als auf einen Termin in der Behörde zu warten und dort hinzufahren. Die Mitarbeitenden der Wohngeldbehörde haben weniger Aufwand, weil sie die Dateneingabe einsparen, alle notwendigen Dokumente digital abrufen können und bei der Auswertung durch einen Algorithmus unterstützt werden. Ein solches digitales Datenökosystem kann allerdings auch Risiken bergen. Zum einen werden diese personenbezogenen Daten digital gespeichert und dadurch dem Risiko der Zweckentfremdung ausgesetzt. Zum anderen erzeugt der Einsatz von Algorithmen bei der Anspruchsbewertung das Risiko, dass die Entstehung der Ergebnisse für Antragstellende schwer oder gar nicht nachvollziehbar ist. Es besteht die Gefahr, dass Antragstellende nicht nachvollziehen können, warum ihr Antrag abgelehnt wurde und nicht mehr überblicken, für welche Zwecke ihre Daten ansonsten verarbeitet werden.

Diese Risiken lassen sich auch in Projekten des Bundes identifizieren. So soll zum Beispiel das vom Bundesministerium für Wirtschaft und Klimaschutz (BMWK) geförderte Projekt MERLOT (MarkEtplace foR LifelOng educaTional dataspaceS and smart service provisioning) Bildungsdatenräume und -dienste in Form von einem Datenökosystem erschaffen.⁷ Dabei geht es unter anderem darum, Lernende bei der Bildungs- und Berufsorientierung zu unterstützen. Hierfür sollten ab 2022 digitale Karriereorientierungsassistenten, ein Weiterbildungsassistent und ein Bildungsprofilprognoseassistent entwickelt werden. Zu dem Zweck sollten Daten über reale Bildungswege erhoben werden und schon vorhandene Bildungsprofildaten aus unterschiedlichen Quellen zugänglich gemacht werden. Diese Bildungsdaten sollten mit Hilfe von Algorithmen verarbeitet werden, mit dem Anspruch, den Datenschutz, die Datensouveränität⁸ der Datenquellen nicht zu gefährden. Mit Hilfe des Einsatzes von Künstlicher Intelligenz (KI) sollten die kognitiven Positionen der Lernenden dem (Weiter-) Bildungsangebot zugeordnet werden. Das Ergebnis sollten Hinweise für Lernende sein, wel-

7 Mir liegt die Gesamtvorhabenbeschreibung aus dem Jahr 2021 vor auf Grund von internem Zugang durch Dataport (Lehmann et al. 2021). Der Abschlussbericht liegt mir jedoch nicht vor, weshalb die Schritte des Projektes als Vorhaben beschrieben werden.

8 Datensouveränität bedeutet in dem Fall, dass diejenigen, über die die Daten eine Aussage machen, der Verwendung ihrer Daten zustimmen oder diese ablehnen können und somit die „Souveränität“ über ihre Daten haben.

che Bildungsangebote oder Berufswege ihnen zur Verfügung stehen. Anschließend sollten neue KI-Systeme mit Hilfe der erhobenen Bildungsdaten trainiert werden.⁹

Der Schutz der Bildungsdaten wird in der Vorhabenbeschreibung garantiert, aber gleichzeitig werden Bildungsdatenräume geschaffen, damit die Bildungsdaten zu verschiedenen Zwecken verwendet werden können – zum Beispiel zum Trainieren von KI-Systemen oder zum Zuordnen der kognitiven Positionen der Lernenden zu dem (Weiter-)Bildungsangebot. Damit die Lernenden, deren Daten erhoben wurden, ihre Datensouveränität erhalten können, sollte ihre Zustimmung zu der Verarbeitung ihrer Daten zweckbezogen eingeholt werden und widerrufbar sein. Auf diese Weise wird das Risiko der Zweckentfremdung aktiv thematisiert und minimiert. Was jedoch nicht thematisiert wird, ist das Risiko der fehlenden Nachvollziehbarkeit der Ergebnisse. Denn wenn eine Person einen der digitalen Assistenten nutzt und eine Empfehlung für eine Weiterbildung oder einen Berufsweg bekommt, sollte es für sie nachvollziehbar sein, warum sie *diese* Empfehlung bekommt und nicht eine andere. Da die Empfehlungen durch die Verarbeitung von Bildungsdaten anderer Personen mit Hilfe von KI-Systemen generiert werden, muss man sich fragen, welche Aussagen durch die Analyse von Bildungsdaten anderer überhaupt gemacht werden können. Angedeutet werden soll an dieser Stelle schon mal, dass dieses Vorgehen das Risiko birgt, Bestehendes zu replizieren, ohne dass dies durchsichtig wird. Diese Frage wird im Folgenden ausführlich erörtert (Kapitel 7). Hier soll aber schon deutlich werden, dass die Gestaltung von Datenökosystemen in der öffentlichen Verwaltung nicht nur Vorteile mit sich bringt, sondern auch Risiken birgt – ethische Risiken.

Wie genau ethische Risiken zu definieren sind, wird im Verlauf dieser Arbeit geklärt. An dieser Stelle soll nun aber begründet werden, warum das Digitalisieren der Datenökosysteme in der Verwaltung so einer kritischen Betrachtung unterzogen werden muss. Denn man könnte argumentieren, dass beim Beantragen von Verwaltungsleistungen auch jetzt schon analoge Datenökosysteme bestehen. Die Antragstellenden müssen vor Ort beim Amt die gleichen sensiblen Daten angeben, wie in einem Online-Antrag. Die Wohngeldbehörde muss die Daten auch bearbeiten und dann abheften, um bei

9 Lehmann et al. 2021

Beschwerden oder Fehlern eventuell darauf zurückgreifen zu können. Zum Schluss wird den Antragstellenden ebenfalls mitgeteilt, ob sie ein Anrecht auf Wohngeld haben oder nicht. Der Unterschied bei einem digitalen Datenökosystem ist aber, dass die Daten nicht analog abgeheftet werden und dementsprechend in einem ganz bestimmten Aktenschränk in einer konkreten Wohngeldbehörde zu finden sind, sondern sie werden digital gespeichert und sind dementsprechend prinzipiell von überall aus zugänglich. Natürlich sollen die Daten nicht von überall für jeden zugänglich sein, weshalb sie geschützt werden – leider reicht aber der Schutz der Daten oft nicht aus.

Auf Grund menschlicher Fehler und der Komplexität der Software gibt es oft inhärente Schwachstellen in einem System oder seinem Design, die es Unbefugten ermöglichen, Anweisungen auszuführen, sich ohne entsprechende Berechtigung Zugang zu Daten zu verschaffen und Angriffe durchzuführen. Angriffe bezeichnen vorsätzliche Handlungen, die darauf abzielen, einem System Schaden zuzufügen oder seinen regulären Betrieb zu unterbrechen, indem Schwachstellen durch verschiedene Taktiken und Werkzeuge ausgenutzt werden. So kann zum Beispiel Schadsoftware in ein System gebracht werden, indem interne Mitarbeitende zum Betrachten von Phishing Mails und dem Öffnen deren Anhänge bewegt werden. Dadurch können sich die Angreifenden Zugang zum System und den darin enthaltenen Daten verschaffen. Alternativ kann durch massenhaftes Aufrufen eines Systems der Zugang zum System blockiert werden (Denial-of-service) oder das System wird anderweitig gestört bzw. zerstört. Angreifende können einzelne Personen mit Fachkenntnissen im Bereich Hacking sein oder aber Banden des organisierten Verbrechens, und auch staatliche Akteure können hinter Angriffen stecken. Die Angreifenden können unterschiedliche Motive und Ziele haben, wie zum Beispiel finanziellen Gewinn, Beeinflussung der öffentlichen Meinung, oder Spionage. Sobald sie aber Zugang haben zu personenbezogenen Daten, werden auch Identitätsdiebstahl oder Erpressung möglich.¹⁰

Leider gibt es dafür viele aktuelle Beispiele mit fatalen Auswirkungen: Im Jahr 2017 wurde Equifax, eines der weltweit größten Finanzdienstleistungsunternehmen, gehackt. Dabei wurden die persönli-

10 Abomhara, Koien 2015:71–74; Sasi et al. 2024:457

chen Daten von fast 143 Millionen Kund:innen erbeutet, darunter Sozialversicherungsnummern, Wohnadressen und Geburtsdaten.¹¹ Die Britische Post (Royal Mail) hatte Anfang 2023 für 6 Wochen Probleme mit dem Versand von Briefen und Paketen nach Übersee auf Grund eines Cyberangriffs auf die Computersysteme, welche die Britische Post für den Versand von Post ins Ausland nutzt.¹²

Vastaamo war Finnlands größtes Psychotherapieunternehmen mit jährlich weit über 10.000 Kunden. Es wurde zunächst als Telegesundheitsdienst entwickelt und bot später auch psychotherapeutische Behandlung über eine Plattform an. Im Oktober 2020 wurde bekannt, dass Vastaamo von einer Datenpanne betroffen war, bei der es dem Angreifer gelang, persönliche Informationen über mehr als 33.000 seiner Patienten und deren psychotherapeutische Behandlungen zu stehlen – darunter auch vertrauliche Therapieunterlagen, wie zum Beispiel Notizen aus den Therapiesitzungen. Der Hacker kontaktierte den CEO von Vastaamo erstmals am 28. September 2020 mit einer Lösegeldforderung in Höhe von 40 Bitcoin (damals entsprach das einem Wert von rund 450.000 Euro). Er behauptete, dass er jeden Tag 100 Datensätze durchsickern lassen würde, bis das Lösegeld eingegangen sei. Da Vastaamo den Lösegeldforderungen nicht nachkam, verlagerte der Angreifer seine Bemühungen auf die Erpressung der Patient:innen, deren Daten gestohlen wurden. Es wurden ca. 30.000 der Patient:innen angeschrieben und ein Lösegeld von 200 bis 500 Euro gefordert. Um seine Drohung zu verstärken, veröffentlichte der Angreifer Patientenakten und Notizen von Sitzungen von mehr als 2.000 Vastaamo-Kunden im Darknet.¹³ Laut der finnischen Tageszeitung Helsingin Sanomat hat ein Rechtsvertreter, der etwa 1.500 Opfer vertritt, erklärt, dass sich mehrere der Vastaamo Kund:innen, deren Therapienotizen veröffentlicht wurden, daraufhin das Leben genommen haben.¹⁴ Hier wird deutlich, wie fatal die Zweckentfremdung von personenbezogenen Daten sein kann. Zusätzlich zu dem Schmerz, der durch die Weitergabe hochsensibler Informationen verursacht wurde, sind die Vastaamo Kund:innen auch anfälliger für mögliche Identitätsdiebstähle in der

11 Durnell et al. 2020:1836

12 Espiner, Tidy 2023; Sweney 2023

13 Looi et al. 2025:107f.; Ghanbari, Koskinen 2024:1f., 4ff.

14 Helsingin Sanomat 2024; Ghanbari, Koskinen 2024:2

Zukunft, da ihre persönlichen Informationen im Darknet verfügbar waren.¹⁵

Solche und ähnliche Datenpannen und Cyberattacken werden im DSGVO-Portal gesammelt¹⁶ – es sind oft auch Einrichtungen der öffentlichen Verwaltung betroffen. So wurde zum Beispiel am 25. April 2025 die digitale Infrastruktur der Stadt Berlin lahmgelegt. Es gab einen Denial-of-service Angriff auf die IT-Systeme der Berliner Verwaltung, indem eine Flut künstlich generierter Anfragen die Onlineportale der Stadt für Nutzende unerreichbar machten. In diesem Fall wurden keine Daten gestohlen, doch auch das kommt regelmäßig vor.¹⁷ Mitte April 2025 hat die Safepay-Ransomware-Gruppe behauptet, 85 GB Daten aus den Verwaltungssystemen der Stadt Heilbronn erbeutet zu haben und drohte diese zu veröffentlichen.¹⁸ Am 28. Januar 2025 hat die Ransomware-Gruppierung „Cloak“ tatsächlich einen Datensatz veröffentlicht, den sie nach eigener Angabe von der Webseite der Gemeinde Kaisersbach in Baden-Württemberg gestohlen hatten.¹⁹ Das bedeutet, wenn mehr Daten im Verwaltungskontext digital gespeichert werden, sind bei Städten und Gemeinden noch mehr sensible Daten zu holen, und die Angriffe werden noch attraktiver für die Angreifenden.

Bei der digitalen Gestaltung von Verwaltungsleistungen in Form eines Datenökosystems, werden sensible Daten digital gespeichert und sind dadurch dem Risiko einer Zweckentfremdung ausgeliefert. Hinzu kommt, dass diese Daten mit Hilfe von Algorithmen verarbeitet werden. Dadurch können im Verwaltungskontext mehr Anträge schneller bearbeitet werden, sodass die Antragstellenden früher ihren Bescheid bekommen – ein klarer Vorteil. Im Fall eines Wohngeldantrages sind die Bedingungen für eine Bewilligung des Geldes gesetzlich eindeutig vorgegeben – hier spielen Faktoren wie die Höhe des Einkommens und die Größe der Wohnung eine Rolle. Deshalb würden in diesem Kontext nur Wenn-Dann-Algorithmen

15 Ghanbari, Koskinen 2024:6

16 <https://www.dsgvo-portal.de/sicherheitsvorfall-datenbank/>

17 https://www.dsgvo-portal.de/sicherheitsvorfaelle/ddos_angriff_stadt-berlin-9278.php

18 https://www.dsgvo-portal.de/sicherheitsvorfaelle/ransomware_angriff_stadt-heilbronn-9177.php

19 https://www.dsgvo-portal.de/sicherheitsvorfaelle/ransomware_angriff_gemeinde-kaisersbach-8211.php

eingesetzt, welche diese klaren Vorgaben in den Anträgen überprüfen. Es gibt aber auch viele andere Beispiele, bei denen die Bedingungen nicht so deutlich festgelegt oder definierbar sind. Wenn, wie bei MERLOT, (Weiter-) Bildungsangebote aus Bildungswegdaten generiert werden sollen, gibt es neben Faktoren wie Ausbildung und Berufserfahrung einen gewissen Ermessensspielraum. Um trotzdem Empfehlungen aussprechen zu können, werden selbstlernende KI-Algorithmen eingesetzt. Dies kann jedoch zu systematischen Fehlern führen, die zunächst schwer zu identifizieren sind.

So hat zum Beispiel das Technologieunternehmen Amazon 2014 für den Zeitraum von einem Jahr bei seinem Einstellungsverfahren ein KI-System eingesetzt. Das KI-System wurde anhand von zuvor eingereichten und bewerteten Lebensläufen trainiert. Auf diese Weise lernte der Algorithmus, welche Bewerbungen in der Vergangenheit erfolgreich gewesen waren, und er identifizierte bestimmte Merkmale dieses Erfolgs, diese wurden dann zu seinen Entscheidungskriterien. In der Vergangenheit waren es vor allem Männer, die im Bewerbungsverfahren erfolgreich waren. Dies hatte zur Folge, dass Bewerbungen von Frauen vom Algorithmus abgelehnt wurden – sogar wenn das Geschlecht nicht direkt genannt wurde, sondern nur durch zum Beispiel eher Frauen zugeordnete Hobbys erkennbar war.²⁰

In Österreich wurden im Jahr 2020 die Chancen auf dem Arbeitsmarkt von Jobsuchenden mit Hilfe eines Algorithmus bewertet. Danach wurden die Jobsuchenden in Gruppen eingeteilt, um Fördermaßnahmen gezielt an diejenigen mit mittelguten Chancen zu vergeben. Dabei wurden Frauen, unabhängig von ihrer Ausbildung, geringere Chancen am Arbeitsmarkt zugesprochen als Männern. Zudem wurden Frauen schlechtere Chancen ausgerechnet, wieder Arbeit zu finden, wenn sie Kinder zu betreuen hatten, während die Anzahl von Kindern auf die Chancen von Männern am Arbeitsmarkt keine Auswirkungen hatte. Die Folge dieser ungleichen Einteilung war, dass Frauen auf Grund ihres Geschlechts von Fördermaßnahmen ausgeschlossen wurden. Der Einsatz dieses Algorithmus ist noch im Jahr 2020 gerichtlich verboten worden und wird weiterhin geprüft.²¹

20 Stahl et al. 2023:10f.

21 Wimmer 2019; Köver 2019; Pflügl 2024

In Bezug auf die Digitalisierung von Daten und deren algorithmische Verarbeitung gibt es in der wissenschaftlichen Literatur ausführliche ethische Debatten. Die digitale Ethik ist der Zweig der Ethik, der sich mit der Frage befasst, was wir tun sollen angesichts der Konfrontation mit Digitalisierung und ihren Entwicklungen.²² Der praktische Philosoph und Wirtschaftsinformatiker Rafael Capurro formuliert das so: „Digital ethics undertakes a critical reflection about ourselves in a world shaped by digital technology.”²³ Diese Reflexion über die Menschen im Kontext von technologischen Entwicklungen gibt es laut Capurro schon seit den 1940er Jahren und nannte sich zuerst Computerethik.²⁴ Dieses Feld wurde zwischendurch auch als Informationsethik bezeichnet und nennt sich heute meist Digitaletik. Der Begriff der Digitaletik soll die Aufmerksamkeit weg von den technologischen Entwicklungen selber und hin zu ihrer Nutzung lenken. Hierunter fallen auch konkretere Bereiche wie zum Beispiel Datenethik oder KI-Ethik.²⁵

Die Technologiefilosoph:innen und Digitaletiker:innen Luciano Floridi und Mariarosaria Taddeo unterteilen die ethischen Herausforderungen der Digitalisierung in drei Achsen: Datenethik, Algorithmus-Ethik und Ethik der Praktiken. Demnach liegt der Schwerpunkt der Datenethik auf ethischen Problemen bei der Sammlung und Analyse großer Datenmengen – hierdurch werden philosophische Themen wie Privatsphäre, Vertrauen und Transparenz relevant. Bei der Algorithmus-Ethik geht es vorrangig um die Frage, wer für Fehler, wie zum Beispiel Diskriminierung durch algorithmische Ergebnisse, zur Rechenschaft gezogen werden kann, sodass hier der philosophische Begriff der Verantwortung eine große Rolle spielt. Die Ethik der Praktiken beschäftigt sich damit, wie Datenprozesse gestaltet werden sollten.²⁶ In dieser Arbeit sollen diese drei Bereiche eine Rolle spielen, um Datenökosysteme umfassend ethisch bewerten zu können.

In der wissenschaftlichen Literatur werden verschiedene Zugänge gesucht zu einer systematischen Digitaletik. Der Medientheoretiker

22 Öhman, Watson 2019:2

23 Capurro 2017:277

24 Ebd.

25 Müller 2023:3

26 Floridi, Taddeo 2016:3

Lev Manovich identifiziert drei Gruppen von Menschen im Kontext von Datenverarbeitung: diejenigen, die Daten erzeugen (sowohl bewusst als auch durch das Hinterlassen digitaler Fußabdrücke), diejenigen, die über die Mittel verfügen, sie zu sammeln, und diejenigen, die über das Fachwissen verfügen, sie zu analysieren.²⁷ Die Gruppe derjenigen, die die Daten analysieren, ist die kleinste und die privilegierteste, da sie die Regeln für die Nutzung der Daten und die Teilnahme daran bestimmen können.²⁸ Den Digitaethikern Carl Öhman und David Watson zufolge hängen die ethischen Auswirkungen der Digitalisierung von den Zielen ihrer Entwickler:innen und Nutzenden ab. Digitaethik ermöglicht die Betrachtung von widersprüchlichen Werten und das Ermitteln von Kompromissen. Dadurch kann die digitale Ethik theoretische Analysen von grundlegenden ethischen Fragen – nach zum Beispiel Autonomie, Privatsphäre und Gerechtigkeit – in praktikable Leitlinien für die Gestaltung und Nutzung digitaler Technologien umsetzen.²⁹

Es gibt bereits viele solcher ethischen Leitlinien, sowohl von staatlichen als auch privaten Akteuren. Die Digitaethiker:innen Anna Jobin und Marcello Ienca sowie die Bioethikerin Effy Vayena haben im Jahr 2019 84 bestehende KI-Leitlinien analysiert und große inhaltliche Übereinstimmung, jedoch Unterschiede bei den Lösungsansätzen festgestellt. In mehr als der Hälfte der analysierten Leitlinien wird Bezug genommen auf die ethischen Prinzipien Transparenz, Gerechtigkeit, Unparteilichkeit, Verantwortung und Privatsphäre – hier herrscht Einigkeit, dass diese Werte beim Einsatz von KI relevant sein sollten.³⁰ Es geht in den Leitlinien hauptsächlich darum, Schaden zu verhindern und Werte, wie Privatsphäre, Menschenwürde und Autonomie, zu wahren, anstatt sie aktiv zu fördern. In dieser Arbeit soll es deshalb neben der Bewahrung auch bewusst um die Förderung der relevanten Prinzipien gehen.

27 Manovich 2012:470

28 Boyd, Crawford 2012:675

29 Öhman, Watson 2019:1, 3

30 Werte sind orientierungsbietende Ideale, die als moralisch gut und dementsprechend als wünschens- und erstrebenswert wahrgenommen werden (Friedman et al. 2008:70f.). In Kapitel 6 und 7 dieser Arbeit werden Privatsphäre und Transparenz als solche Werte im Kontext von Datenökosystemen etabliert, indem begründet wird, warum sie wünschens- und erstrebenswert sowie moralisch gut sind.

Ein großes Problem bei der Umsetzung von Leitlinien der digitalen Ethik sind Prinzipien, die miteinander im Konflikt stehen. Jobin et al. haben in ihrer Studie festgestellt, dass in der Praxis Unsicherheit bezüglich der Wahl der Prinzipien besteht, insbesondere wenn es widersprüchliche Prinzipien gibt.³¹ Prinzipien können im Konflikt stehen, wenn eines zugunsten der Erfüllung des anderen aufgegeben werden muss. Wenn zum Beispiel eine Wohngeldstatistik erstellt werden soll, um an den bedürftigsten Orten am schnellsten bezahlbaren Wohnraum schaffen zu können, dann braucht man die Zustimmung der Wohngeldempfänger:innen, ihre Daten zu diesem Zweck zu verwenden. Wird die Zustimmung vorausgesetzt und zur Bedingung für die Antragstellung gemacht, so wird die Autonomie der Antragstellenden eingeschränkt. Ist die Zustimmung jedoch freiwillig, kann mit hoher Wahrscheinlichkeit keine repräsentative Wohngeldstatistik erstellt werden und somit können die Baugenehmigungen für bezahlbare Wohnungen in Deutschland nicht gerecht verteilt werden. Hier steht die Autonomie der Wohngeldempfänger:innen im Konflikt mit dem Grundsatz der Verteilungsgerechtigkeit.

Den Technologie- und Wissenschaftsphilosoph:innen Jess Whittlestone, Rune Nyrop, Anna Alexandrova und Stephen Cave zufolge können solche Wertekonflikte nur dadurch gelöst werden, dass einem der Werte Vorrang gewährt wird. Diese Entscheidung lässt sich nur dann begründen, wenn die konfligierenden Werte klar definiert sind, wenn klar ist, was technisch möglich ist, um die Werte zu schützen und wenn die Bedürfnisse der betroffenen Gemeinschaften bekannt sind.³²

In der Realität aber bleibt es oft der Interpretation überlassen, was die Prinzipien besagen, insbesondere weil ein Prinzip unterschiedliche Bedeutungen haben kann. So enthielten zwar 73 der von Jobin et al. analysierten Leitlinien den Grundsatz „Transparenz“, aber dieser hatte verschiedene Bedeutungen von Erklärbarkeit über Explizierbarkeit, Verständlichkeit, Interpretierbarkeit, Kommunikation bis zu Offenlegung.³³ Die Lösungsansätze für die Wahrung von Transparenz beziehen sich in vielen der analysierten Leitlinien auf

31 Jobin et al. 2019:396

32 Whittlestone et al. 2019:198f.

33 Jobin et al. 2019:391, 395

eine verstärkte Offenlegung von Informationen durch diejenigen, die KI-Systeme entwickeln oder einsetzen. Aber die Anforderungen, welche Informationen offengelegt werden sollten, unterscheiden sich stark. Manche fordern Informationen über die Verwendung der KI, andere wollen Zugriff auf den Quellcode, um Transparenz zu ermöglichen. Auch die Offenlegung der Verwendung von Daten wird gefordert oder eine Evidenzbasis für die Verwendung von KI, sowie eine Antwort auf die Frage, wer Verantwortung trägt für die KI.³⁴ In dieser Arbeit werden die relevanten Prinzipien und Forderungen nach ihrer Umsetzung deshalb ausführlich definiert, um Missverständnissen vorzubeugen.

An dieser Stelle fragt sich, woher man weiß, welche Prinzipien relevant sind. Ethisch relevante Prinzipien werden normalerweise aus ethischen Theorien abgeleitet – die drei meist genutzten ethischen Theorien sind der Konsequentialismus, die Deontologie und die Tugendethik. Ein Beispiel von hieraus abgeleiteten Prinzipien sind die Prinzipien der biomedizinischen Ethik der Philosophen Tom Beauchamp und James Childress: Achtung der Autonomie, Wohltätigkeit, Nicht-Schädigung und Gerechtigkeit.³⁵ Diese Prinzipien wurden sowohl aus einem konsequentialistischen als auch aus einem deontologischen Ansatz abgeleitet und sollten ursprünglich im bioethischen Diskurs und in der bioethischen Praxis eine Grundlage für ethische Urteile bilden.³⁶ Heutzutage werden sie auch im Kontext der digitalen Ethik verwendet.³⁷ Dem Datenethiker Brent Mittelstadt zufolge lassen sich die Grundsätze der biomedizinischen Ethik jedoch nicht einfach auf die digitale Ethik übertragen.³⁸ Diese Debatte ist zwar interessant, doch in dieser Arbeit werden die im Datenökosystem relevanten Prinzipien aus einer der ethischen Theorien selber abgeleitet, sodass die Prinzipien der biomedizinischen Ethik weiter keine Rolle spielen.

Die KI-Ethikerinnen Larisa Bolte und Aimee van Wynsberghe teilen die Debatten zur KI-Ethik in drei Wellen ein: Während sich die erste Welle auf abstrakte und nicht-existierende Technologie konzen-

34 Jobin et al. 2019:391

35 Beauchamp, Childress 1979

36 Beauchamp, Childress 2019:9

37 Floridi, Cowls 2021; Jobin et al. 2019; Ryan, Stahl 2021

38 Mittelstadt 2019; Diese Debatte wird vertieft behandelt in Kempkens 2025.

trierte, befasse sich die zweite Welle mit den Risiken, die von der derzeit bestehenden KI-Technologie ausgehe – dies sei eine Verlagerung vom Spekulativen zum Konkreten. Die Autorinnen verstehen die zweite Welle der KI-Ethik als Ausdruck einer grundlegend anderen Einstellung zur Frage, wo die relevanten Probleme liegen. Und sie identifizieren ein weiteres Problem, das aus ihrer Sicht durch die zweite Welle der KI-Ethik nicht abgedeckt wird: Nachhaltigkeit. Ihrer Meinung nach ist eine neue, dritte Welle der KI-Ethik erforderlich, um die mit der Entwicklung von KI-Systemen verbundenen Umweltkosten zu berücksichtigen. Während das Ziel der zweiten Welle die konkrete Analyse bestimmter KI-Systeme ist, geht es bei der dritten Welle darum, alle ethischen Bedenken zu berücksichtigen, die sich aus einer strukturellen Perspektive auf die Technologie ergeben – hier findet eine Verlagerung von einem isolationistischen zu einem strukturellen Ansatz statt.³⁹

Bolte und van Wynsberghe werfen dem Ansatz der zweiten Welle Tech-Solutionismus (techno-solutionism) vor, weil die Autor:innen der zweiten Welle jede KI-Anwendung isoliert betrachten und annehmen, jedes Problem könne innerhalb des technischen Systems, das dieses Problem verursacht hat, gelöst werden. Die Erwartung ist dabei, dass die Gestaltung der KI-Systeme einen großen Einfluss auf ihre moralischen Eigenschaften und Wirkungen hat.⁴⁰ Doch laut Bolte und van Wynsberghe dient diese Kombination einer isolationistischen Sichtweise auf KI und einer techno-solutionistischen Ideologie nur den Interessen mächtiger Akteure.⁴¹ Die dritte Welle hingegen betrachtet nicht nur Probleme einzelner KI-Anwendungen, sondern strukturelle Probleme, die durch den Einsatz von KI entstehen oder weiter bestehen.⁴² Diese strukturellen Probleme werden hierbei den Autorinnen zufolge nicht auf technologische Rätsel reduziert: „The second wave approach relies on better design to bring about systemic or structural change. The third wave approach relies on systemic or structural change to bring about better design.“⁴³ Strukturelle Probleme erscheinen in Ansätzen der zweiten Welle als

39 Bolte, van Wynsberghe 2024:1735f.

40 Ebd.:1739

41 Ebd.:1738

42 Ebd.:1740

43 Ebd.:1738

solche, die in der Technologie und durch technologisches Design korrigiert werden müssen. In den Ansätzen der dritten Welle hingegen werden Technologie und Technologiedesign als selbst in Strukturen und Systeme eingebettet konzipiert.⁴⁴

Die vorliegende Arbeit bewegt sich zwischen der von Bolte und van Wynsberghe sogenannten zweiten und dritten Welle der KI-Ethik. Denn einerseits befasse ich mich konkret mit den Risiken, die derzeit von der bestehenden KI-Technologie ausgehen (zweite Welle), aber andererseits identifiziere ich diese Risiken mit Hilfe eines strukturellen Ansatzes (dritte Welle). Ich gehe nicht davon aus, dass die strukturellen ethischen Probleme innerhalb von einzelnen KI-Anwendungen gelöst werden können (Techno-solutionismus; zweite Welle), aber ich nehme auch nicht an, dass sich erst die Strukturen ändern müssen, bevor Technologie ethischer gestaltet werden kann (dritte Welle). Ich verfolge einen strukturellen Ansatz, der ethische Fragen auf einer systemischen Ebene aufdeckt (dritte Welle) und setze auf besseres Design, um strukturelle Veränderungen herbeizuführen (zweite Welle).

In Bezug auf den Einsatz von KI ist die größte Sorge in der Debatte der Digitaethik, dass Ungerechtigkeiten reproduziert werden.⁴⁵ Durch das Analysieren von Daten können Regelmäßigkeiten erkannt werden. Dies kann zum Beispiel dann nützlich sein, wenn man herausfinden will, wie viele der Mitarbeitenden bei Amazon weiblich sind. Dieses Beispiel zeigt, dass Daten nicht neutral sind – sie spiegeln gesellschaftliche Vorurteile und auch gefestigte Ungerechtigkeiten wider. Denn die Daten wurden in der Gesellschaft gesammelt, und in dem Maße, in dem die Gesellschaft Ungleichheit, Ausgrenzung oder andere Spuren von Diskriminierung aufweist, werden dies auch die Daten tun.⁴⁶ Durch eine reine Analyse können die Verhältnisse beschrieben und Ungerechtigkeiten reflektiert werden. Wenn die Daten aber als Trainingsdaten für Algorithmen verwendet werden, dann werden auch die Ungerechtigkeiten zur Entscheidungsgrundlage. Deshalb finden sich solche Vorurteile und Ungerechtigkeiten den Technologiephilosoph:innen Solon Barocas

44 Bolte, van Wynsberghe 2024:1739

45 Barocas, Selbst 2016:671; Goodman, Flaxman 2017:53; Whittlestone et al. 2019:198

46 Goodman, Flaxman 2017:53

und Helen Nissenbaum zufolge auch in algorithmischen Ergebnissen wieder: „(...) an algorithm is only as good as the data it works with.“⁴⁷ Das Problem ist, dass algorithmische Ergebnisse oft als neutral wahrgenommen werden – das sind sie aber nicht. Denn ein Algorithmus kann nur wiedergeben, was in den Trainingsdaten steckt.

In der digitalethischen Debatte geht es also in Bezug auf den Einsatz von KI meist darum, ob KI in bestimmten Kontexten eingesetzt werden sollte und insbesondere darum, wer dann die Verantwortung dafür trägt. Denn Algorithmen des maschinellen Lernens, ein Teilbereich der KI, können erstaunliche Ergebnisse liefern und zum Beispiel, wie im Fall von ChatGPT, Aufsätze oder Code schreiben. Dabei sind ihre Ergebnisse jedoch unvorhersehbar.⁴⁸ Dadurch, dass solchen komplexen Algorithmen nicht die Entscheidungskriterien vorgegeben werden, sondern die KI sich diese selber sucht auf Basis sehr vieler bereits analysierter Datensätze, ist es für Menschen sehr schwer, ihre Ergebnisse, aber auch ihre Fehler zu verstehen, zu erklären oder vorherzusagen. Dem Rechtsberater Andrew Tutt zufolge macht das die KI sowohl wertvoll, als auch besonders gefährlich.⁴⁹ Denn die überraschenden Ergebnisse entstehen nur, weil der Algorithmus aus den ihm zur Verfügung gestellten Daten lernt, doch die Ergebnisse können falsch sein und enthalten sicher aus den Daten übernommene Vorurteile. Wer aber steht dafür gerade? In der Literatur ist die Rede von einer Verantwortungslücke⁵⁰, welche in Kapitel 7 genauer betrachtet wird.

Es gibt also viele ethische Probleme im Kontext der Digitalethik. Dem Technologiephilosoph Vincent C. Müller zufolge ist es wichtig diese früh zu erkennen: „The ethics of the design and use of computers is clearly an area of very high societal importance and we would do well to catch problems early on (...).“⁵¹ Sobald Menschen nicht wissen, welche Handlung richtig wäre, gibt es ein ethisches Problem, das Aufmerksamkeit verdient. Oft handelt es sich dabei um in der Disziplin der Philosophie schon bekannte Probleme wie

47 Barocas, Selbst 2016:671

48 Tutt 2017:87

49 Ebd.:90

50 Matthias 2004; Sparrow 2007; Nyholm 2018; Himmelreich 2019; Kiener 2022:579; Oimann, Tollon 2024; Nyholm 2024

51 Müller 2023:15

zum Beispiel die Verletzung der Privatsphäre.⁵² Diese hat auch schon vor der Digitalisierung aller Lebensbereiche eine Rolle gespielt und philosophische Konflikte hervorgerufen, bekommt aber durch die Digitalisierung eine verstärkte Bedeutung, weil das Offenbaren von persönlichen Informationen nun viel weitreichender ist. So sind zwar die Phänomene, mit denen sich die Digitalethik befasst, neu, aber die Fragen, die sie aufwerfen, sind alt und womöglich zeitlos. Laut Öhman und Watson ist es die Aufgabe der Philosophie, ihre grundlegenden Konzepte, die immer wieder zur Entschlüsselung neuer Phänomene beitragen, im Lichte neuer Herausforderungen, wie der Digitalisierung, neu zu definieren.⁵³

Die Digitalethiker:innen Andréa Belliger und David Krieger versuchen nicht nur Konzepte neu zu definieren, sondern die digitale Ethik in ein neues Licht zu rücken und fordern eine Revolution der Ethik, die mit der Revolution der Technologie vergleichbar ist. Sie sind davon überzeugt, dass die Aufgabe der Ethik nicht darin besteht, Werte und Normen abzuleiten aus religiösen Lehren, Grundrechten, Erklärungen von Ethikkommissionen oder staatlichen Vorschriften und Richtlinien.⁵⁴ Solch ein Vorgehen ist im Kontext der Digitalethikdebatte jedoch tatsächlich zu beobachten. So haben zum Beispiel die Technologiephilosophen und Digitalethiker Luciano Floridi und Josh Cowls eine vergleichende Analyse von sechs Veröffentlichungen über Künstliche Intelligenz angestellt, um aus den Übereinstimmungen einen einheitlichen Rahmen zusammenzustellen. Die Veröffentlichungen stammen von einem Institut aus Belgien und den USA, einer kanadischen Universität, der Europäischen Kommission, dem Oberhaus des britischen Parlaments, einer Multistakeholder-Organisation in den USA und einer globalen Abhandlung nach dem Crowd-Sourcing-Prinzip mit Beiträgen von 250 globalen Vordenker:innen. Alle sechs Veröffentlichungen seien sehr renommiert und unmittelbar relevant für KI und ihre Auswirkungen auf die Gesellschaft als Ganzes. Floridi und Cowls haben darin insgesamt 47 Prinzipien identifiziert und diese in fünf Prinzipien zusammengefasst: Die vier biomedizinischen Prinzipien Achtung der Autonomie, Wohltätigkeit, Nichtschadensverursachung und Ge-

52 Müller 2023:16

53 Öhman, Watson 2019:5

54 Belliger, Krieger 2023:41

rechtheit sowie das Prinzip der Erklärbarkeit.⁵⁵ Dieses Vorgehen offenbart zwar einen erstaunlichen Konsens unter vornehmlich politischen und akademischen Autor:innen aus westlichen liberalen Demokratien, eine ethische Begründung für die Auswahl der Prinzipien gibt es jedoch nicht.

Digitalethik ist somit einerseits nicht dazu da, Werte und Normen aus Richtlinien abzuleiten, und andererseits ist sie auch keine philosophische Rechtfertigung für von oben verordnete Richtlinien, Gesetze und Regierungsvorschriften.⁵⁶ Laut Belliger und Krieger besteht digitale Ethik vielmehr aus konkreten Governance-Verfahren. Dabei bezeichnet Governance eine dezentrale, von unten nach oben gerichtete Regulierung von Netzwerken, welche als Designprozess geplant und umgesetzt wird.⁵⁷ Diese Netzwerke definieren sich, wie Datenökosysteme, über die Verbindung zwischen verschiedenen Akteuren mit der Fähigkeit, Informationen zu verarbeiten. Ihre Lösungsidee ist, dass jedes Netzwerk einen eigenen Governance-Rahmen entwickelt, in dem alle von dem Netzwerk Betroffenen (Stakeholder) sich an der Gestaltung beteiligen und sicherstellen, dass der Rahmen den Werten und Normen des Netzwerks entspricht. Die Werte und Normen des Netzes seien dabei Konnektivität, Fluss, Beteiligung, Transparenz, Authentizität und Flexibilität. Diese sechs Normen haben Krieger und Belliger in einer Veröffentlichung von 2014 aus Studien über sogenannte „neue Medien“ (gemeint sind digitale Medien) abgeleitet, welche zeigen sollten, wie Netzwerken nach der digitalen Medienrevolution funktioniert. Die Autor:innen sind der Meinung, dass sie mit den sechs Normen die Netzwerk-normen identifiziert haben, welche das kommunikative Handeln bestimmen.⁵⁸ „Since networks are dynamic processes, they follow certain inherent norms that describe what ‘good’ networking is. These values do not belong to a realm of what ‘ought’ to be as opposed to what ‘is’. They describe, not prescribe.“⁵⁹ Belliger und Krieger meinen, die sechs Normen leiten sich von den Möglichkeiten digitaler Technologien ab und beschreiben, was Netzwerke von

55 Floridi, Cowsls 2021:8–9

56 Belliger, Krieger 2023:41, 52

57 Ebd.:40, 53

58 Krieger, Belliger 2014:10

59 Belliger, Krieger 2023:42

Natur aus für wünschenswert, angemessen und nützlich halten und was für ein gutes Funktionieren erforderlich ist.⁶⁰ Ihnen geht es hier nicht darum vorzuschreiben, wie digitale Netzwerke sein sollten, sondern nur darum zu beschreiben, wie sie sind. Die Werte und Normen dieser neuen digitalen Ethik sind also nicht mehr inhaltlich (substantive), sondern verfahrensbezogen (procedural): „(...) it is no longer *what* is held to be good by some person or group, which guides and legitimates action, but *how* anything at all is done well that becomes the foundation of ethics.“⁶¹ Die Grundlage der Digital-ethik soll nach Belliger und Krieger demnach sein, wie ein Netzwerk gut funktioniert.

Hier liegt jedoch ein Missverständnis der Rolle der Ethik vor. Laut dem Philosophen Immanuel Kant beantwortet die Ethik, als Teilbereich der Philosophie, die Frage „Was soll ich tun?“⁶² Die Antwort darauf ist nicht „Du solltest das Netzwerk so gestalten, dass es gut funktioniert“. An dieser Stelle fangen die ethischen Fragen erst an: Für wen funktioniert das Netzwerk gut? Funktioniert es für alle Beteiligten gleichermaßen gut oder besser für diejenigen Akteure mit mehr Kontrolle? Die Aufgabe der Ethik ist es, zu hinterfragen was „gut“ generell und in diesem Kontext bedeutet. Die reine Beschreibung eines Netzwerks und seiner Funktionsweisen findet auch in der Disziplin der Soziologie einen Platz. Aber die Ethik geht darüber hinaus und fragt, ob es das Netzwerk überhaupt geben sollte oder wie die einzelnen Beziehungen zwischen Akteuren innerhalb des Netzwerks gestaltet sein sollten, sodass es als ethisch bezeichnet werden kann. Wie man in dieser Arbeit sehen wird, gibt es unterschiedliche Ansätze bezüglich der Frage, wann etwas als ethisch bezeichnet werden kann. Konkret werden der Utilitarismus, die Tugendethik und Kants Deontologie dazu herangezogen.

Die Ethik spielt als reflektierende und hinterfragende Disziplin eine wichtige Rolle in der Debatte zur Digitalisierung – sie kann nicht einfach ersetzt werden durch soziologische Beschreibungen von Netzwerken oder durch rechtliche Festschreibungen. Das soll nicht bedeuten, dass die rechtlichen Entwicklungen in diesem Bereich nicht hilfreich oder notwendig sind. Im Gegenteil: Insbesondere

60 Belliger, Krieger 2023:43

61 Ebd.

62 Kant 1956:727f.

re die neusten europäischen Verordnungen, die Datenschutzgrundverordnung (DSGVO) und die Verordnung zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (AI Act), sind Meilensteine in der Regulierung des Umgangs mit Datenverarbeitung. Die DSGVO reguliert einerseits den „Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten“⁶³ und andererseits den „freien Verkehr solcher Daten“⁶⁴. Die Verarbeitung personenbezogener Daten⁶⁵ ist nach DSGVO dann rechtmäßig, wenn die betroffene Person ihre Einwilligung für einen oder mehrere bestimmte Zwecke gibt; aber auch ein Vertrag, eine rechtliche Verpflichtung zum Schutz lebenswichtiger Interessen von natürlichen Personen, öffentliches Interesse und die Wahrung berechtigter Interessen von natürlichen Personen können rechtmäßige Gründe für die Verarbeitung personenbezogener Daten sein.⁶⁶ Dieser Datenschutz gilt nicht für anonyme Informationen, das heißt für Daten, bei denen es nur mit einem hohen zeitlichen und finanziellen Aufwand möglich wäre, die betroffenen Personen zu identifizieren.⁶⁷

Der AI Act ist dazu da „(...) die Einführung einer auf den Menschen ausgerichteten und vertrauenswürdigen künstlichen Intelligenz (KI) zu fördern und gleichzeitig ein hohes Schutzniveau (...) vor schädlichen Auswirkungen von KI-Systemen in der Union zu gewährleisten (...).“⁶⁸ Zu diesem Zweck verbietet der AI Act bestimmte Praktiken, wie zum Beispiel die Nutzung eines KI-Systems zum unterschweligen Beeinflussen einer Person, zum Ausnutzen von Vulnerabilität oder Schutzbedürftigkeit von Personen, oder zum Bewerten von Personen mit dem Ergebnis ihrer Schlechterstellung oder Benachteiligung.⁶⁹ Im AI Act werden KI-Systeme anhand der von ihnen verursachten Risiken unterschieden. Sogenannte Hochrisiko-KI-Systeme dürfen nur dann eingesetzt werden, wenn es ein Risikomanagementsystem gibt, die Daten vorgeschriebenen Qualitätskriterien entsprechen, es eine technische Dokumentation gibt,

63 Art. 1, Abs. 1 DSGVO

64 Ebd.

65 Personenbezogene Daten sind laut DSGVO alle Daten, die eine Aussage über eine bestimmte Person machen (Art. 4, Abs. 1 DSGVO).

66 Art. 6, Abs. 1 a-f) DSGVO

67 Erwägungsgrund 26 DSGVO

68 Art. 1, Abs. 1 EU AI Act

69 Art. 5, Abs. 1 a)-c) EU AI Act

Aufzeichnungs- sowie Transparenzpflichten erfüllt sind, menschliche Aufsicht möglich ist und die KI-Systeme ein angemessenes Maß an Genauigkeit, Robustheit und Cybersicherheit aufweisen.⁷⁰ Als Hochrisiko-KI-Systeme gelten die Systeme, die in den folgenden Bereichen eingesetzt werden: Biometrie, kritische Infrastruktur, Bildung, Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit, Zugänglichkeit und Inanspruchnahme grundlegender privater und grundlegender öffentlicher Dienste und Leistungen, Strafverfolgung, Migration, Asyl und Grenzkontrolle, sowie Rechtspflege und demokratische Prozesse.⁷¹

Aber auch hier gibt es Ausnahmen. So sollen zum Beispiel KI-Systeme für Verwaltungsverfahren in Steuer- und Zollbehörden sowie für die Bekämpfung von Geldwäsche nicht als Hochrisiko-KI-Systeme eingestuft werden.⁷² Denn der AI Act soll „(...) die Entwicklung und Verwendung innovativer Ansätze in der öffentlichen Verwaltung nicht behindern (...)“⁷³, solange das Risiko nicht zu hoch ist.

Solche rechtlichen Verordnungen, sind sehr wichtig, da man sich auf sie beziehen kann, wenn man Schutz verlangt bei der Datenverarbeitung und vor Schaden durch KI-Systeme. Aber die Ethik ist der Rechtsanwältin und Technologiephilosophin Mireille Hildebrandt zufolge vom Recht zu unterscheiden: „Ethics is both more and less than law: it is more because many ethical concerns are not addressed by the law and less because the outcome of ethical considerations are not necessarily transformed into legal norms and thus not enforceable by way of law.“⁷⁴ Der Ethik fehlt die Durchsetzungskraft des Rechts, aber es ist auch nicht die Aufgabe der Ethik, bestimmte Regeln durchzusetzen. Im Gegenteil: Die Ethik sollte nicht abschließend in Regeln gegossen werden, da es verschiedene und sogar widersprüchliche ethische Haltungen geben kann, die gleichermaßen berechtigt sind. Ohne Einigkeit über die Ethik erfüllt erst mal das Recht mit Hilfe von festgelegten Regeln eine Rolle, die für Verlässlichkeit und Sicherheit sorgt.⁷⁵ Während das Recht die soziale

70 Art. 8 – 15 EU AI Act

71 Anhang III EU AI Act

72 Nr. 59 der nahestehenden Gründe (Preliminaries) EU AI Act

73 Nr. 58 der nahestehenden Gründe (Preliminaries) EU AI Act

74 Hildebrandt 2020:297

75 Ebd.:296f.

Ordnung aufrechterhalten soll, ist das Ziel der Ethik, die Menschen bei der Entscheidung anzuleiten, was über die Gesetze hinaus falsch ist und wie sie sich in bestimmten Situationen verhalten sollen. So schreibt der Elektroingenieur und Technologiephilosoph Spyros G. Tzafestas: „Laws provide a minimum set of standards for obtaining good human behavior. Ethics often provides standards that exceed the legal minimum.“⁷⁶ Die Ethik geht über das Recht hinaus, weil viele ethische Bedenken durch Gesetze nicht berücksichtigt werden. Ethik kann an dieser Stelle Standards setzen, die nicht gesetzlich vorgeschrieben sind, und in Situationen, die nicht direkt vom Recht abgedeckt sind, eine Orientierungshilfe bieten.

So könnten laut der Technologiephilosophin und -ethikerin Esther Keymolen und der Professorin für internationale Datenverwaltung Linnet Taylor aus ethischen Gründen auch anonyme Daten als schützenswerte Daten behandelt werden – und zwar weil der Schutz der Privatsphäre für wichtig gehalten wird und nicht, weil man gesetzlich dazu verpflichtet ist: „In that case, they go beyond what the law expects from them and enter the realm of data ethics.“⁷⁷ Datenethik beantwortet also Fragen danach, was getan werden sollte im Kontext von Datenerhebung und -verarbeitung. So lassen sich prinzipiell wichtige Aspekte herausarbeiten – wie in dieser Arbeit der Schutz der Privatsphäre und die Bereitstellung von Transparenz. Mit Hilfe von Gesetzen können rechtlich verpflichtende Mindeststandards für gutes menschliches Verhalten festgelegt werden. Da aber nicht alle ethischen Aspekte in Recht gegossen und sie oft durch Ausnahmen verwässert werden, geht das, was moralisch verpflichtend ist, darüber hinaus. Doch während der Verstoß gegen das Recht geahndet und mit Freiheits- oder Geldstrafen belegt werden kann, wird der Verstoß gegen moralische Verpflichtungen nur durch soziale Ächtung bestraft – dennoch sind sie wichtig.

Denn insbesondere im Bereich der Digitalisierung könnte das Recht laut Keymolen und Taylor durch schnelle technologische Entwicklungen an seine Grenzen stoßen: „For instance, developments in AI may lead to complex questions concerning explainability, fairness, and accountability which require us to rethink and reinterpret

76 Tzafestas 2018:98

77 Keymolen, Taylor 2023:493

existing legal norms.”⁷⁸ Dieses Umdenken bestehender Rechtsnormen kann die Digitaethik leisten, weil sie viel beweglicher und flexibler ist als das Recht. Trotzdem kann die Ethik, auch laut Keymolen und Taylor, den Handlungsraum nicht alleine dominieren, da ihr die Durchsetzungskraft und Vorhersehbarkeit des Rechts fehlen. Denn wenn es keine rechtliche Regulierung gäbe, könnten Unternehmen ihre eigenen Spielregeln aufstellen, und andere Akteure müssten einfach mitspielen. Wenn allerdings Regeln aufgestellt und ohne Grundlage als „ethisch“ bezeichnet werden, handelt es sich laut Keymolen und Taylor um Missbrauch der Digitaethik: „(...) in the darkest scenario, data ethics in the hands of data-driven businesses becomes a vehicle for intentional, malicious conduct, when it is used to hamper and ditch legal regulation.“⁷⁹ Die Digitaethik ist nicht dafür da, kommerzielle Regeln zu rechtfertigen, sondern bietet die Möglichkeit zu durchdenken, was getan werden sollte. Digitaethik sollte also nicht dazu dienen, die Lücken zu füllen, welche das Recht nicht schnell genug füllen kann, sondern Digitaethik sollte im vom Gesetz vorgegebenen Handlungsspielraum dazu beitragen, einen verantwortungsvolleren Umgang mit der Digitalisierung zu entwickeln.⁸⁰ Auch Tzafestas plädiert für eine Balance: „(...) [F]or the best behavior, both law and ethics should be respected.“⁸¹ Denn im Rahmen des Rechts ist die Ethik dazu in der Lage, die Menschen auf unregulierte Probleme aufmerksam zu machen und Orientierung zu bieten.

Deshalb sollen in dieser Arbeit Datenökosysteme in der öffentlichen Verwaltung ethisch beleuchtet werden. Denn in Bezug auf die Digitalisierung von Verwaltungsdiensten kann die Ethik über das Recht hinaus Orientierung bieten. Das Onlinezugangsgesetz (OZG) legt fest, dass bei der Verbesserung des Onlinezugangs zu Verwaltungsleistungen Standards der IT-Sicherheit eingehalten werden sollen und Nutzerfreundlichkeit soll gegeben sein.⁸² Bei Fragen bezüglich der Verarbeitung von personenbezogenen Daten bezieht sich das

78 Keymolen, Taylor 2023:494

79 Ebd.:495

80 Ebd.:494f.

81 Tzafestas 2018:98

82 §5 und §7 OZG

OZG auf die DSGVO.⁸³ Das OZG schafft aber eine weitere Ausnahme für Artikel 9, Absatz 1 der DSGVO, da die Verarbeitung von besonders sensiblen Daten erlaubt wird, wenn dies für Verwaltungsverfahren erforderlich ist.⁸⁴ Gleichzeitig schreibt das OZG die Einrichtung eines „Datenschutzcockpits“ vor, mit dessen Hilfe sich die Bürger:innen über sie betreffende Datenübermittlungen zwischen öffentlichen Stellen informieren können.⁸⁵ Architekturvorgaben, Qualitätsanforderungen und Interoperabilitätsstandards müssen noch bis 2026 vom Bundesministerium des Innern und für Heimat zusammen mit dem IT-Planungsrat festgelegt werden.⁸⁶ Noch bevor rechtlich geregelt ist, wie genau der Onlinezugang zu Verwaltungsleistungen gestaltet wird, kann mit Hilfe der Ethik erörtert werden, wie diese Digitalisierung der öffentlichen Verwaltung, und damit die einzelnen Datenökosysteme gestaltet werden sollte. Im vom Gesetz vorgegebenen Handlungsspielraum kann die Ethik auf entscheidende Kriterien aufmerksam machen, die bei der Gestaltung von Datenökosystemen erfüllt sein sollten.

Aus diesem Grund möchte ich mit dieser Arbeit die digitaethische Debatte ergänzen um einen umfassenden Blick auf Datenökosysteme und wie diese auf Grundlage ihrer Risiken ethisch gestaltet werden können. Dabei ist es mir besonders wichtig, eine passende ethische Methode zu nutzen. In der Literatur herrscht jedoch Uneinigkeit darüber, welche ethische Methode im Kontext unbekannter, sich schnell entwickelnder Technologie verwendet werden sollte.⁸⁷ Um eine ethische Methode zu finden, die Orientierung bieten kann für die Gestaltung von Datenökosystemen, werde ich in dieser Arbeit zunächst einmal grundlegend analysieren, wie sich Datenökosysteme und die Risiken, die durch den Umgang mit digitalen Daten entstehen, ethisch bewerten lassen. Anschließend wird auf dieser Grundlage der ethischen Bewertung begründet, wie Datenökosysteme ethisch gestaltet werden können. Denn das *Argumentationsziel* dieser Arbeit ist es zu zeigen, wie der Umgang mit digitalen Daten in Datenökosystemen ethisch gestaltet werden kann. Zu diesem Zweck werden folgende Forschungsfragen erörtert:

83 §3a, Abs. 2 OZG; §8, Abs. 9 und 10 OZG

84 §8a, Abs. 1, Satz 2 OZG

85 §10, Abs. 1, Satz 1 OZG

86 §6, Abs. 1 OZG

87 Vallor 2016:7f.

1. Wie lassen sich Datenökosysteme ethisch bewerten?
2. Wie sollten Datenökosysteme dementsprechend gestaltet werden?

Um herausfinden zu können, wie sich Datenökosysteme ethisch bewerten lassen und somit die erste Forschungsfrage beantworten zu können, wird zuerst eine Konzeption von Datenökosystemen aufgestellt (Kapitel 2). Daraus wird in Kapitel 3 ersichtlich werden, dass im Datenökosystem zwischen den Akteuren das Wissen und die Möglichkeiten zur Kontrolle ungleich verteilt sind, und die Akteure mit mehr Kontrolle den Akteuren mit weniger Kontrolle zwei Risiken auferlegen: das Risiko der Verletzung der Privatsphäre und das Risiko der fehlenden Transparenz. Die Methode zur ethischen Bewertung von Datenökosystemen muss deshalb sowohl die ungleichen Wissensstände und Kontrollmöglichkeiten als auch die daraus erwachsende Auferlegung der Risiken bewerten können. Es wird deutlich, dass Utilitarismus und Tugendethik diese Bedingungen nicht erfüllen und die Deontologie die beste Grundlage für die ethische Bewertung von Datenökosystemen darstellt (Kapitel 4). Was ethisch bewertet werden muss, ist die Struktur von Datenökosystemen, sich durch ungleich verteiltes Wissen, ungleiche Kontrollmöglichkeiten der Akteure sowie dadurch mögliche Risikoauflegung auszeichnet.

Deshalb wird in Kapitel 5 die Frage beantwortet, unter welchen Bedingungen es zulässig ist, anderen ein Risiko aufzuerlegen. Zur Bestimmung dieser Bedingungen wird eine deontologische Risikoethik entwickelt, indem die Risikoethik um die dritte Formel des Kategorischen Imperativ von Kant ergänzt wird. Es wird deutlich, dass es nur dann zulässig ist, anderen Personen ein Risiko aufzuerlegen, wenn diese dabei nicht als bloßes Mittel gebraucht werden. Deshalb darf ein Risiko einer Person nur nach deren informierter Zustimmung auferlegt werden. Datenökosysteme lassen sich also ethisch bewerten mit der von mir auf der Grundlage von Kants Deontologie entwickelten deontologischen Risikoethik.

Um herausfinden zu können, wie Datenökosysteme dementsprechend gestaltet werden sollten, und damit die zweite Forschungsfrage beantworten zu können, werden in Kapitel 6 und 7 die Risiken genauer betrachtet. Denn es ist nicht nur die Risikoauflegung selbst ethisch relevant, sondern auch die weitestgehende Vermeidung der Risiken ist geboten. Dies wird gezeigt durch die Begründung einer Pflicht zum Schutz der Privatsphäre (6.3) sowie einer

Pflicht zur Bereitstellung transparenter Ergebnisse (7.3). Die beiden Risiken werden einer gründlichen Betrachtung unterzogen, um zu entwickeln, wie sich Privatsphäre und Transparenz bei der Datenverarbeitung in einem Datenökosystem schützen und fördern lassen (Kapitel 6 und 7). Zum Schluss werden mit Hilfe der deontologischen Risikoethik administrative Anforderungen abgeleitet, die bei ihrer Einhaltung den Schutz und die Förderung von Privatsphäre und Transparenz ermöglichen.

Daraufhin stellt sich in Kapitel 8 die Frage nach der Verantwortung, denn wenn verantwortungsvolles Handeln unterbleibt, ist es möglich, dass die Privatsphäre der Datenquellen oder die Transparenz für die Datennutzenden verletzt werden. Die Verantwortung ist der wahrscheinlich stärkste Durchsetzungsmechanismus der Ethik und kann denjenigen Akteuren zugeschrieben werden, die Kontrolle über das zu Verantwortende haben. Aus einer Position der Kontrolle über den Umgang mit Daten innerhalb eines Datenökosystems erwächst also Verantwortung für die Minimierung der Risiken und das ethische Gestalten.

Auf der Suche nach einer Umsetzungsmethode, mit der die ethischen Erkenntnisse wirksam in die Praxis überführt werden können, leite ich aus der bisherigen Argumentation Kriterien für eine ethische Gestaltung ab. Es liegen bereits verschiedene Vorschläge zur Umsetzung digitaler Ethik in die Praxis vor, die ich einteile in ethische Leitlinien, ethische Gestaltungsprozesse und Risikoanalysen. Jeder dieser Ansätze weist Lücken auf, wie ich in Kapitel 9 ausführlich aufzeige. Deshalb entwickle ich abschließend einen eigenen Verfahrensmaßstab zur Umsetzung digitaler Ethik in die Praxis, um die Vorteile von Datenökosystemen gefahrlos nutzen zu können. Als Antwort auf meine zweite Forschungsfrage stelle ich in Kapitel 10 meinen ethischen Verfahrensmaßstab vor, welcher expliziert, wie Datenökosysteme den Erkenntnissen dieser Arbeit zufolge gestaltet werden sollten. Folglich wird abschließend gezeigt, wie der Umgang mit digitalen Daten in Datenökosystemen ethisch gestaltet werden kann.

Beginnen wir das Vorhaben mit einer ausführlichen Betrachtung des zu bewertenden und gestaltenden Gegenstands: dem Umgang mit Daten im Datenökosystem.

