

Bachelorarbeit zur Erlangung des Hochschulgrades Bachelor of Arts (B. A.) in Sozialwissenschaften

Deutungsspielräume in Leitlinien für ethische KI

Eine qualitative Untersuchung der Differenzen der KI-Leitlinien auf deutscher Länderebene

Gutachterin:

Prof. Dr. Cordula Kropp
Institut für Sozialwissenschaften
Lehrstuhl für Soziologie mit dem
Schwerpunkt Risiko- und Technikforschung

vorgelegt von:

Emma Takacs

Inhaltsverzeichnis

Abkürzungsverzeichnis	3
1. Einleitung	4
2. Forschungsstand	5
2.1. Künstliche Intelligenz (KI)	6
2.2. Ethik-Prinzipien-Ansatz	6
2.3. KI-Strategien und ethische Leitlinien in Deutschland	9
2.4. Forschungslücke	12
3. Forschungsdesign	13
3.1. Scoping Study	13
3.2. Strukturierende qualitative Inhaltsanalyse nach Mayring	16
3.3. Kommunikationsmodell nach Mayring	18
4. Ergebnisse der qualitativen Inhaltsanalyse nach Mayring	19
4.1. Quantitative Auswertung	23
4.2. Qualitative Auswertung	25
5. Diskussion	37
5.1. Diskussion der Ergebnisse	38
5.2. Diskussion der Methodik	44
6. Fazit	45
Literaturverzeichnis	48
Leitlinienverzeichnis	50
Erklärung über die Eigenständigkeit	52

Abkürzungsverzeichnis

AGI/ ASI	Artificial General Intelligence/ Artificial Super Intelligence
AI	Artificial Intelligence
BStK	Bayrische Staatskanzlei
ELLIS	European Laboratory for Learning and Intelligent Systems
EM MV	Ministerium für Energie, Infrastruktur und Digitalisierung Mecklenburg-Vorpommern
KI	Künstliche Intelligenz
Landesregierung BW	Landesregierung Baden-Württemberg
MW LSA	Ministerium für Wirtschaft, Wissenschaft und Digitalisie- rung des Landes Sachsen-Anhalt
MW Niedersachsen	Niedersächsisches Ministerium für Wirtschaft, Arbeit, Ver- kehr und Digitalisierung
MWAE	Ministerium für Wirtschaft, Arbeit und Energie des Landes Brandenburg
MWK Niedersachsen	Niedersächsisches Ministerium für Wissenschaft und Kul- tur
SenWiEnBe	Senatsverwaltung für Wirtschaft, Energie und Betriebe Berlin
SK Sachsen	Staatskanzlei Sachsen
Stk Hessen	Hessische Staatskanzlei
Stk Saarland	Staatskanzlei des Saarlandes
Stk SH	Staatskanzlei Schleswig-Holstein
TMWWDG	Thüringer Ministerium für Wirtschaft, Wissenschaft und Digitale Gesellschaft

1. Einleitung

Mit den großen Fortschritten im Bereich der Künstlichen Intelligenz (KI) und deren immer breiter werdende Anwendung in vielen gesellschaftlichen Bereichen, wie unter anderem der Landwirtschaft, dem Transport und in Smart Cities, geht eine steigende Prominenz des Themas im öffentlichen Raum einher (vgl. Ryan und Stahl 2021, S. 63). Hierbei werden jedoch immer wieder Negativbeispiele sichtbar, wie die Benachteiligung von Frauen in KI-gestützten Recruitingprozessen bei Amazon, der gezielte Ausschluss von Immobilien-Anzeigen auf Facebook aufgrund von Hautfarbe, Geschlecht, Herkunft, Religion und Behinderungen oder die Verstärkung von Stereotypen durch die Unterrepräsentation von Frauen in Google-Suchergebnissen zu stereotypisch männlich dominierten Berufen (vgl. Orwat 2020, S. 30ff.). Aufgrund dieses starken Einflusses von KI auf die Gesellschaft und die Wirtschaft besteht ein großes Interesse der Öffentlichkeit an deren Regulierung, um sie möglichst ethisch zu gestalten (vgl. Ryan und Stahl 2021, S. 63).

Für die Gestaltung einer solchen ethischen KI wurde insbesondere ab 2016 international eine große Anzahl an Leitlinien durch den privaten und öffentlichen Sektor, die Wissenschaft, Zivilgesellschaft, staatliche sowie zwischenstaatliche Institutionen erarbeitet und veröffentlicht (vgl. Jobin et al. 2019, S. 391). In Deutschland war dies insbesondere nach der Veröffentlichung der KI-Strategie der Bundesregierung im November 2018 zu beobachten (vgl. Jobin et al. 2021, S. 8).

Für die tatsächliche Umsetzung einer ethischen KI wird vor allem in wissenschaftlichen Beiträgen betont, dass es hierfür einen übergreifenden Konsens braucht, sowohl in Bezug auf die Frage, was ethische KI überhaupt ist, als auch in Bezug auf die Frage nach geeigneten Maßnahmen für deren Realisierung (vgl. Jobin et al. 2019, S. 389). An diesem Punkt haben bisherige Forschungsarbeiten im internationalen Kontext bereits angesetzt, größtenteils, um die in den verschiedenen Leitlinien thematisierten Ethik-Prinzipien, die laut ihren Autor*innen ethische KI konstituieren, sowie Maßnahmen für deren Umsetzung zu identifizieren (vgl. bspw. Fjeld et al. (2020) oder Jobin et al. (2019)). Wenige Forschungsbeiträge gehen hierbei bislang über die Deskription hinaus und beschäftigen sich speziell mit der unterschiedlichen Konzeptualisierung ethischer KI in verschiedenen Leitlinien oder der Extraktion spezifischer normativer Maßnahmen für deren Umsetzung (vgl. bspw. Zeng et al. 2019 oder Ryan und Stahl 2020). Alle diese Beiträge halten jedoch fest, dass sich Deutungsspielräume in Bezug auf die Ethik-Prinzipien, und daher häufig auch klare Differenzen zwischen den Leitlinien, auf tun. Jobin et al. (2021) richteten den Blick diesbezüglich auf Deutschland

und stellten fest, dass vergleichbare Arbeiten auf subnationaler Ebene bislang komplett ausblieben (vgl. Jobin et al. 2021, S. 1).

An dieser Lücke soll die vorliegende Arbeit ansetzen und dabei folgende Fragen beleuchten: Welche Differenzen bestehen zwischen den in den Leitlinien der Bundesländer vorgeschlagenen Ethik-Prinzipien für KI? Und welche Differenzen bestehen zwischen den in den Leitlinien der Bundesländer vorgeschlagenen Maßnahmen zur Umsetzung der Ethik-Prinzipien?

Durch die Identifizierung der Differenzen der verschiedenen Leitlinien soll ein Ausgangspunkt für die Bildung eines Konsenses geschaffen werden, der die Voraussetzung für die effektive Umsetzung der Leitlinien, und damit für die Realisierung ethischer KI darstellt. Um dieses Vorhaben umzusetzen, wird eine qualitative Untersuchung der Leitlinien anhand der Methode der strukturierenden qualitativen Inhaltsanalyse nach Mayring (2000) durchgeführt. Diese ermöglicht es die Forschungsfragen durch quantitative und qualitative Auswertungen der Leitlinien, basierend auf einem deduktiv erstellten Kategoriensystem, zu beantworten und diese Ergebnisse anhand von Querauswertungen zu diskutieren und zu reflektieren (vgl. Mayring 1991, S. 213). Im folgenden Kapitel wird daher zunächst der aktuelle Forschungsstand in dem betrachteten Forschungsfeld dargestellt sowie sich hieraus ergebende Folgerungen und Fragestellungen für die vorliegende Arbeit festgehalten. In Kapitel 3 werden das in dieser Arbeit implementierte Forschungsdesign, d. h. die Methode der Scoping Study nach Arksey und O'Malley (2005), die zur qualitativen Untersuchung der Leitlinien der Bundesländer gewählte Methode der strukturierenden qualitativen Inhaltsanalyse nach Mayring (2000) sowie der Prozess der Leitlinien-Auswahl und deren Einordnung in das Kommunikationsmodell nach Mayring (2000) vorgestellt. Anschließend werden in Kapitel 4 die erzielten Untersuchungsergebnisse dargestellt und damit die Forschungsfragen beantwortet. Die erlangten Ergebnisse sowie die Methodik werden dann in Kapitel 5 kritisch beleuchtet und diskutiert. Abschließend folgt ein bilanzierendes Fazit.

2. Forschungsstand

In diesem Kapitel wird in Unterkapitel 2.1 zunächst auf KI und deren Definition bzw. definitorische Schwierigkeiten eingegangen. Anschließend wird in Unterkapitel 2.2 der Forschungsstand zu dem Ethik-Prinzipien-Ansatz für KI und damit einhergehend das Forschungsfeld der Ethik der KI vorgestellt. Unterkapitel 2.3 beleuchtet den Entwicklungs- und Forschungsstand zu ethischer KI in Deutschland. Zuletzt werden

bestehende Forschungslücken aufgezeigt sowie die daraus resultierenden Forschungsfragen der vorliegenden Arbeit vorgestellt.

2.1. Künstliche Intelligenz (KI)

Es gibt eine Vielzahl an verschiedenen Definitionen von KI, da das Feld zum einen sehr breit ist und zum anderen auch für den Begriff der Intelligenz selbst keine einschlägige Definition vorhanden ist (vgl. Buxmann und Schmidt 2019, S. 6). Konsens besteht jedoch darin, dass KI ein Teilgebiet der Informatik ist und, dass zwischen starker und schwacher KI unterschieden werden muss (vgl. Buxmann und Schmidt 2019, S. 6). Starke KI bezeichnet hierbei intelligente Agenten, die Vorgänge des menschlichen Gehirns und menschlichen Verhaltens imitieren und durchführen (vgl. Buxmann und Schmidt 2019, S. 6). Der Ansatz der starken KI ist jedoch lediglich Theorie, da dieser, zumindest bislang, außerhalb der technischen Möglichkeiten liegt (vgl. Hoffmann 2022, S. 6). Unter schwacher KI versteht man solche intelligenten Agenten, die durch Algorithmen bestimmte Problemstellungen bewältigen können und hierbei selbstständig, auf der Grundlage von Big Data, dazulernen (vgl. Buxmann und Schmidt 2019, S. 6f.; vgl. Fetic et al. 2020, S. 9). Auch der vorliegenden Arbeit wird demnach die Definition schwacher KI zugrunde gelegt.

2.2. Ethik-Prinzipien-Ansatz

Das zentrale ethische Bedenken in Bezug auf KI ist deren Fähigkeit eigenständig zu lernen und dadurch immer autonomer handeln zu können (vgl. Ryan und Stahl 2021, S. 63). Sich hieraus ergebende Probleme, die im Feld der Ethik der KI behandelt werden, sind bspw. die Diskriminierung von Kandidat*innen in KI-basierten Recruiting-Prozessen, auf dem Immobilienmarkt oder die Stereotypisierung in Suchmaschinen-Ergebnissen, die durch die beabsichtigte oder unbeabsichtigte Reproduktion von Vorurteilen und Bias der zugrundeliegenden Datensätzen oder der KI-Entwickler*innen entstehen (vgl. Rebstadt et al. 2022, S. 1).

Ebenso wie im Fall der KI, ist jedoch auch für ethische KI bzw. die Ethik der KI bislang keine genaue Definition vorhanden, nicht zuletzt, da es hierfür zuerst eine einschlägige Definition von KI braucht (vgl. Ryan und Stahl 2021, S. 63). Während sich gesellschaftliche, politische und fachkundige Akteur*innen also einig sind, dass die Konzeption, Entwicklung, Implementierung und Auswirkungen von KI ethisch sein müssen, fehlen jedoch genaue Vorstellungen, was dies bedeutet und wie dies umsetzbar wäre (vgl. Jobin et al. 2019, S. 389).

Ein Ansatz, der in diesem Kontext breite Anwendung zur Operationalisierung ethischer KI erfährt, und der auch in der vorliegenden Arbeit implementiert wird, ist der Ethik-Prinzipien-Ansatz.

Der Ethik-Prinzipien-Ansatz lässt sich in die Disziplin der normativen Ethik einordnen, welche Regeln des moralisch verträglichen Handelns formuliert (vgl. Düwell et al. 2011, S. 2f.). Innerhalb der Disziplin gibt es verschiedene Ansätze, die sich in ihrer Art und Weise der Regel-Formulierung und Regel-Anwendung unterscheiden (vgl. Düwell et al. 2011, S. 15). Der für den Ethik-Prinzipien-Ansatz relevante Ansatz ist der der Prinzipienethik, bei dem nur eine einzige Vorschrift, das Moralprinzip, als relevant erachtet wird, wobei dieses in kantianischer Tradition in spezifische Handlungsprinzipien begrenzter Reichweite ausdifferenziert wird, die das Moralprinzip konkretisieren (vgl. Düwell et al. 2011, S. 15f.). Für den Ethik-Prinzipien-Ansatz für KI kann demnach als Moralprinzip formuliert werden, dass die Konzeption, Entwicklung, Implementierung und Folgen bzw. Auswirkungen von KI moralisch verträglich sein müssen. Einzelne Ethik-Prinzipien, welche im Folgenden beispielhaft vorgestellt werden, dienen der Konkretisierung dieses Moralprinzips und stellen damit einen Ansatz dar, den Begriff der ethischen KI zu definieren bzw. zu operationalisieren.

Es lassen sich in der Forschungslandschaft vier zentrale Studien finden, welche ihrerseits Analysen von Ethik-Leitlinien darstellen, mit dem Ziel festzustellen, welche Ethik-Prinzipien ethische KI ausmachen, wie diese interpretiert werden und darüber hinaus, wie sie realisiert werden können. Im Folgenden werden diese vier Beiträge vorgestellt und in Unterkapitel 3.1 in Tabelle 1 deren jeweilige Interpretation der Ethik-Prinzipien dargestellt.

Der erste und am häufigsten zitierte Beitrag, von Jobin et al. (2019), leistet eine (geografische) Analyse von 84 international publizierten Ethik-Leitlinien verschiedener Branchen (bspw. Privatsektor, Regierungsinstitutionen, Wissenschaft und Forschung, intergouvernementale und supranationale Institutionen), die bis April 2019 veröffentlicht wurden. Diese Analyse zielt darauf ab herauszufinden, ob es ein globales Einverständnis in Bezug auf Ethik-Prinzipien für KI und deren Umsetzung gibt. Ergebnis der Inhaltsanalyse der Leitlinien ist die Identifizierung von elf übergreifenden Ethik-Prinzipien: 1) *Transparency*; 2) *Justice, Fairness*; 3) *Non-maleficence*; 4) *Responsibility*; 5) *Privacy*; 6) *Beneficence*; 7) *Freedom, Autonomy*; 8) *Trust*; 9) *Dignity*; 10) *Sustainability* und 11) *Solidarity*, wobei jedoch keines dieser Prinzipien in allen von Jobin et al. analysierten Leitlinien genannt wurde (vgl. Jobin et al. 2019, S. 391). Ein weiteres zentrales Ergebnis der Studie ist die Identifikation einer Konvergenz bei fünf Ethik-

Prinzipien (*Transparency; Justice, Fairness; Non-maleficence; Responsibility* und *Privacy*), wobei jedoch auch diese im globalen Kontext betrachtet, stark unterschiedlich interpretiert, priorisiert, kontextualisiert und umgesetzt werden (vgl. Jobin et al. 2019, S. 389, S. 396).

Der zweite zentrale Beitrag, von Ryan und Stahl (2020), setzt eigener Aussage nach an dieser Kategorisierung der Ethik-Prinzipien von Jobin et al. (2019) an und identifiziert hierzu als Orientierung für Entwickler*innen und Nutzer*innen von KI normative Inhalte für deren Umsetzung in 82 Leitlinien, die auch von Jobin et al. (2019) analysiert wurden, und neun Weiteren (vgl. Ryan und Stahl 2020, S. 61). Ryan und Stahl kamen hierdurch zu dem Ergebnis, dass es zwar eine starke Überschneidung der zentralen angesprochenen Themen und Problematiken in den Leitlinien gibt, jedoch starke Differenzen darin bestehen, auf welche Teilgebiete der KI sich die Leitlinien fokussieren, wie verbindlich sie sind, wie umfangreich sie sind, welche Terminologien genutzt werden und an wen sie sich richten (vgl. Ryan und Stahl 2020, S. 65).

Der nächste zentrale Beitrag, von Fjeld et al. (2020), bietet eine High-Level-Übersicht über die in 36 Ethik-Leitlinien unterschiedlicher Sektoren und Länder genannten Ethik-Prinzipien für KI, um diese sowohl individuell als auch in ihrem Kontext zu verstehen (vgl. Fjeld et al. 2020, S. 4). Hierbei identifizierten sie acht zentrale Ethik-Prinzipien: 1) *Privacy*; 2) *Accountability*; 3) *Safety, Security*; 4) *Transparency, Explainability*; 5) *Fairness, Non-discrimination*; 6) *Human Control of Technology*; 7) *Professional Responsibility* und 8) *Promotion of Human Values* (vgl. Fjeld et al. 2020, S. 4f.). Als weiteres Ergebnis nennen Fjeld et al. den Befund, dass neue Leitlinien tendenziell alle dieser acht Ethik-Prinzipien abdecken, sodass sie auf eine Konvergenz um diese acht Ethik-Prinzipien, im Gegensatz zu den fünf von Jobin et al. (2019) genannten Ethik-Prinzipien schließen (vgl. Fjeld et al. 2020, S. 5). Jedoch identifizieren Fjeld et al. in ihrer Studie eine Lücke zwischen der Artikulierung dieser Ethik-Prinzipien und deren tatsächliche Umsetzung in die Praxis (vgl. Fjeld et al. 2020, S. 66).

Abschließend lässt sich die Studie von Zeng et al. (2019) als zentralen Beitrag nennen, die Verknüpfungen von Prinzipien für ethische KI in internationalen Leitlinien unterschiedlicher Branchen und deren gemeinsame Themen darstellt, um aufzuzeigen, dass es ein umfangreiches Framework braucht, das von der Interaktion und gegenseitigen Vervollständigung verschiedener Ethik-Prinzipien profitiert (vgl. Zeng et al. 2019, S. 1). Zeng et al. haben hierbei zehn Ethik-Prinzipien identifiziert: 1) *Humanity*; 2) *Collaboration*; 3) *Share*; 4) *Fairness*; 5) *Transparency*; 6) *Privacy*; 7) *Security*; 8) *Safety*; 9) *Accountability* und 10) *AGI/ ASI* (Artificial General Intelligence/ Artificial Super

Intelligence) (vgl. Zeng et al. 2019, S. 2). Diese Ethik-Prinzipien werden in verschiedenen Leitlinien größtenteils mit ähnlichem Vokabular beschrieben (vgl. Zeng et al. 2019, S. 3). Jedoch bestehen laut Zeng et al. teilweise Deutungsspielräume in Bezug auf deren Kontext und deren Bedeutung: „For instance, we may ask for ‘transparency’ from the decision-making process of the system out of our fairness concerns. We may also ask for ‘transparency’ from the system to make it more safe, traceable, and controllable” (Zeng et al. 2019, S. 3).

Wie auch diese vier zentralen Studien verdeutlichen, genießt der Ethik-Prinzipien-Ansatz für KI hohe Aufmerksamkeit und eine breite Anwendung. Dennoch weisen einige Forscher*innen, wie Mittelstadt (2019), Morley et al. (2021) oder Rességuier und Rodrigues (2020), auf Problematiken des Ansatzes hin. Im Vergleich zu der lang etablierten Implementierung des Ethik-Prinzipien-Ansatzes in der Medizin, aus der er für KI übernommen wurde, fehlt es dem Feld der KI laut Mittelstadt (2019) bislang an gemeinsamen Zielen und Pflichten, einer gemeinsamen Tradition und gemeinsamen Normen, bewährten Methoden zur Umsetzung der Grundsätze in die Praxis und soliden rechtlichen und beruflichen Rechenschaftsmechanismen (vgl. Mittelstadt 2019, S. 1). Eine Befürchtung ist daher auf der einen Seite, dass die komplexe Ethik-Debatte durch die Formulierung von Ethik-Prinzipien zu stark vereinfacht wird (vgl. Mittelstadt 2019, S. 10). Auf der anderen Seite besteht hierdurch zudem die Problematik, dass Ethik-Prinzipien dann lediglich abstrakte Konzepte darstellen, welche auf unterschiedliche Weise interpretiert werden und aufgrund der fehlenden Klarheit nicht in die Praxis umgesetzt werden können (vgl. Morley et al. 2021, S. 240). Ein Vorwurf, der in Bezug auf diesen Ansatz zur Umsetzung ethischer KI daher häufig laut wird, ist der des ‚Ethics Washings‘, d. h., dass die Aufstellung von Ethik-Prinzipien es erleichtern würde notwendige politische Regulierung von KI zu vermeiden, indem die Ethik-Prinzipien pro forma vorgehalten werden und lediglich eine Selbstverpflichtung durch Organisationen stattfindet (vgl. Rességuier und Rodrigues 2020, S. 2).

2.3.KI-Strategien und ethische Leitlinien in Deutschland

Im Zuge der 1990er-Jahre ist die Diskussion um die politische Regulierung von KI-Systemen in Deutschland innerhalb des politischen Handlungsfeldes der Netz- und Digitalpolitik entstanden (vgl. Goldmann und Bieber 2021, S. 2). Im Jahr 2017 gab es erstmals konzeptionelle Überlegungen zur Entwicklung einer nationalen KI-Strategie im Rahmen der Hightech-Strategie des Bundesministeriums für Bildung und Forschung und des Fachforums Autonome Systeme (vgl. Goldmann und Bieber 2021, S.

2). Nachdem in dem im Juli 2018 veröffentlichten Eckpunkte-Papier der Bundesregierung bereits Schwerpunkte für die Entwicklung einer nationalen KI-Strategie gesetzt wurden, wurde diese im November 2018 veröffentlicht (vgl. Bundesregierung 2018, S. 1ff.; vgl. Bundesregierung 2020, S. 1f.). Die KI-Strategie entstand unter der Federführung des Bundesministeriums für Wirtschaft und Energie, des Bundesministeriums für Arbeit und Soziales und des Bundesministeriums für Bildung und Forschung (vgl. Goldmann und Bieber 2021, S. 3). Sie wurde 2020 fortgeschrieben, um die beschlossenen Maßnahmen in Hinblick auf aktuelle Entwicklungen im Bereich der KI neu zu fokussieren (vgl. Bundesregierung 2020, S. 2). Das Leitbild der KI-Strategie ist „[...] ein europäisches KI-Ökosystem für Innovationen, das die Wettbewerbsfähigkeit der europäischen Forschung und die Wirtschaft ausbaut, die vielfältigen Anwendungsmöglichkeiten von KI in allen Bereichen der Gesellschaft im Interesse der Bürgerinnen und Bürger fördert und auf gemeinsamen europäischen Regeln und Werten basiert“ (Bundesregierung 2020, S. 3). Hierbei sollen der Nutzen für den Menschen, das Gemeinwohl, die Pandemiebekämpfung, Umwelt- und Klimaschutz sowie die europäische und internationale Vernetzung im Mittelpunkt stehen, was das Ziel der Entwicklung und Umsetzung einer ethischen KI verdeutlicht (vgl. Bundesregierung 2020, S. 3, 9f.).

Da diese ethische Gestaltung von KI der Gesellschaft zugutekommen soll, besteht ein Forschungsinteresse an der Einschätzung der Ethik-Prinzipien und ethischer KI allgemein durch die Gesellschaft. Kieslich et al. (2022) untersuchen in einer Studie daher die Wahrnehmung der Wichtigkeit der Ethik-Prinzipien für KI in einem Panel von 1099 deutschen Bürger*innen. Hierbei kamen sie zu dem Ergebnis, dass die Befragten die inkludierten Ethik-Prinzipien (*Explainability, Fairness, Security, Accountability, Accuracy, Privacy, Machine Autonomy*) insgesamt als gleichermaßen wichtig einschätzen (vgl. Kieslich et al. 2022, S. 1). Das zentrale Ergebnis ist daher, dass die Bewertung der Wichtigkeit der Ethik-Prinzipien für KI durch die Befragten nahelegt, dass es einen Konsens zwischen verschiedenen Leitlinien für die Umsetzung ethischer KI braucht, da alle Ethik-Prinzipien als nahezu gleichbedeutend eingeschätzt werden, weshalb unterschiedliche Entwickler*innen und Organisationen Ethik-Prinzipien nicht auf unterschiedliche Weise berücksichtigen und umsetzen sollten (vgl. Kieslich et al. 2022, S. 10). Um dies zu realisieren, betonen verschiedene Autor*innen, dass ein klarer rechtlicher Rahmen benötigt wird (vgl. Kieslich et al. 2022, S. 10f.; vgl. Denkfabrik des Bundesministeriums für Arbeit und Soziales 2022, S. 3). Kieslich et al. halten fest, dass die EU mit dem Vorschlag einer KI-Verordnung durch die Europäische

Kommission (2021) eine Pionierrolle in diesem Gebiet eingenommen hat (vgl. Kieslich et al. 2022, S. 11). In einem zweiten Schritt ihrer Analyse führten Kieslich et al. eine nachträgliche Clusteranalyse durch, die Differenzen zwischen verschiedenen soziodemografisch Gruppen, in deren Einschätzungen der Wichtigkeit der Ethik-Prinzipien offenlegt (vgl. Kieslich et al. 2022, S. 6). Als ein weiteres zentrales Ergebnis der Untersuchung konnte hierdurch die Einschätzung der Befragten von *Accountability* als wichtigstes Ethik-Prinzip festgehalten werden (vgl. Kieslich et al. 2022, S. 10).

Richtet man den Blick von der Bundes- auf die Länderebene, so lässt sich ein zentraler Beitrag von Jobin et al. (2021) festhalten, in dem für dieses Forschungsgebiet der Begriff des KI-Föderalismus eingeführt wird (vgl. Jobin et al. 2021, S. 2).

Aus diesem Beitrag geht hervor, dass alle Bundesländer bis 2017 bereits eine eigene Digital- oder Innovationsstrategie veröffentlicht haben (vgl. Jobin et al. 2021, S. 6). Strategien mit einem Schwerpunkt auf KI und teilweise auch eigenständige KI-Leitlinien erschienen jedoch fast ausschließlich nach der Veröffentlichung der KI-Strategie der Bundesregierung (vgl. Jobin et al. 2021, S. 8). Jobin et al. (2021) konnten daher feststellen, dass Leitlinien auf Länderebene teilweise Themen, Ausdrücke und Rhetoriken ebenso wie Ethik-Prinzipien der KI-Strategie der Bundesregierung übernehmen. Diese werden jedoch an die eigene regionale Identität und die bereits bestehenden Strukturen der Bundesländer angepasst (vgl. Jobin et al. 2021, S. 8, 10). Dies lässt sich unter anderem auf ein Ergebnis der Studie von Goldmann und Bieber (2021) zu KI-Governance in Deutschland zurückführen, in welcher sie festhalten, dass KI-Politik in den Bundesländern klar als Wirtschafts- und Standortpolitik aufgefasst wird (vgl. Goldmann und Bieber 2021, S. 1). Dies führt unter anderem zu starken Variationen und Differenzen der angesprochenen Problematiken und Themen (vgl. Jobin et al. 2021, S. 10). Ein weiterer Grund für die Differenzen in den angesprochenen Aspekten stellen die von Jobin et al. (2021) festgestellten Unterschiede in der Erarbeitungsweise der Dokumente dar. Die von Jobin et al. inkludierten Leitlinien für Hessen, Berlin, Brandenburg und Mecklenburg-Vorpommern wurden bspw. ausgehend von einer Bewertung von Forscher*innen erstellt, während andere Bundesländer ihre Leitlinien z. B. in Kooperation mit privatwirtschaftlichen Partner*innen erarbeitet haben (vgl. Jobin et al. 2021, S. 8). Des Weiteren unterscheidet sich auch das Ausmaß, in dem sich die Bundesländer mit der KI-Governance beschäftigen, was an den starken Variationen des Umfangs und der Anzahl publizierter Leitlinien zu beobachten ist (vgl. Jobin et al. 2021, S. 8). Aufgrund des deutschen Mehrebenen-Regierungssystems und des kooperativen Föderalismus wird diesbezüglich unter den Bundesländern ein

Wettbewerb befeuert, in dem die Bundesländer versuchen durch die Publikation solcher Leitlinien ihre sozioökonomische Stellung zu erhalten oder sogar zu verbessern (vgl. Jobin et al. 2021, S. 6).

Über die beiden Beiträge hinausgehende Untersuchungen von Leitlinien zu spezifischen Ethik-Prinzipien und deren Umsetzung, wie z. B. die im vorangegangenen Unterkapitel vorgestellten Beiträge, blieben auf subnationaler Ebene bislang jedoch aus (vgl. Jobin et al. 2021, S. 1).

2.4.Forschungslücke

Auf der Grundlage der vorangegangenen drei Unterkapitel lassen sich demnach zwei Lücken in der Forschungslandschaft zu ethischer KI identifizieren.

Zum einen hat sich die Forschung bislang darauf konzentriert, den Konsens zwischen verschiedenen Leitlinien zu identifizieren, der benötigt wird, um ethische KI in die Praxis umzusetzen. Hierbei wurde in den zentralen, in Unterkapitel 2.2 vorgestellten Publikationen erwähnt, dass es jedoch offensichtlich Deutungsspielräume bezüglich der Ethik-Prinzipien gibt, die zu Differenzen zwischen den Leitlinien führen. Worin genau die Differenzen der Leitlinien bezüglich der Fragen, was ethische KI ist und wie diese umgesetzt werden kann, bestehen, wurde dennoch bisher nicht explizit untersucht. Da die Beantwortung dieser Fragen jedoch zentral für die Bildung eines Konsenses ist, und damit, um ethische KI tatsächlich zu realisieren, muss diese Lücke geschlossen werden.

Zum anderen wurde herausgearbeitet, dass Analysen, wie sie bspw. von Jobin et al. (2019), Ryan und Stahl (2020), Fjeld et al. (2020) oder Zeng et al. (2019) durchgeführt wurden, auf deutscher Länderebene bislang ausblieben. Da jedoch basierend auf Unterkapitel 2.3 Differenzen zwischen den Leitlinien der Bundesländer zu erwarten sind, ist es notwendig, auch diese Lücke zu schließen.

An diesen beiden Lücken setzt die vorliegende Arbeit an und formuliert zwei Forschungsfragen, die im Rahmen der qualitativen Untersuchung durch Unterfragen ergänzt und hierdurch spezifiziert werden:

1. Welche Differenzen bestehen zwischen den in den Leitlinien der Bundesländer vorgeschlagenen Ethik-Prinzipien für KI?
 - 1.1. Wie werden die Ethik-Prinzipien interpretiert? (vgl. Jobin et al. 2019, S. 396)
 - 1.2. Welche Problematiken thematisieren die Ethik-Prinzipien? (vgl. Jobin et al. 2019, S. 396)

2. Welche Differenzen bestehen zwischen den in den Leitlinien der Bundesländer vorgeschlagenen Maßnahmen zur Umsetzung der Ethik-Prinzipien?

2.1. Welches Ziel hat die Maßnahme? (vgl. Mittelstadt 2019, S. 19)

2.2. An wen richtet sich die Maßnahme? (vgl. Mittelstadt 2019, S. 19)

3. Forschungsdesign

In diesem Kapitel wird die in der vorliegenden Arbeit angewandte Methodik vorgestellt. Es wird eine Scoping Study auf der Basis des methodologischen Frameworks nach Arksey und O'Malley (2005) durchgeführt, was in Unterkapitel 3.1 thematisiert wird. Diese dient als Grundlage für das Kategoriensystem der qualitativen Inhaltsanalyse der Leitlinien der Bundesländer nach Mayring (2000), die in Unterkapitel 3.2 vorgestellt wird. Zuletzt wird in Unterkapitel 3.3 der Prozess der Auswahl der Leitlinien der Bundesländer und deren Einordnung in das Kommunikationsmodell vorgestellt.

3.1. Scoping Study

Um Forschungsbeiträge zu bereits bestehenden Analysen von Ethik-Leitlinien für KI zu identifizieren, die als Grundlage für das Kategoriensystem der qualitativen Inhaltsanalyse der Leitlinien dienen sollen, wird eine Scoping Study auf der Basis des methodologischen Frameworks nach Arksey und O'Malley (2005) durchgeführt (vgl. Arksey und O'Malley 2005, S. 8ff.). Diese definieren die Scoping Study nach Mays et al. (2001) als eine Methode, die darauf abzielt, Forschungsbereiche zu untersuchen, die sehr komplex und neu sind, um deren zentrale Aspekte zu erfassen (vgl. Mays et al. 2001, S. 194).

Die Scoping Study wird in dieser Arbeit als eine umfassende Recherche genutzt, um die relevante Literatur zusammenzutragen, ohne deren Qualität oder Design zu bewerten, da durch deren Ausschluss ggf. zu einem Publikationsbias beigetragen werden kann. Dies stellt einen Vorteil gegenüber der verbreiteteren Methode der Systematic Review dar, der genutzt werden soll, da das betrachtete Forschungsfeld noch relativ neu und dynamisch ist (vgl. Arksey und O'Malley 2005, S. 17).

Wie Arksey und O'Malley beschreiben, kommt es auf die Art der Arbeit bzw. das Ziel der Scoping Study an, in welcher Tiefe diese Einblicke in die identifizierte Literatur gibt (vgl. Arksey und O'Malley 2005, S. 5). In dieser Arbeit dient sie dem Aufbau des Kategoriensystems für die qualitative Untersuchung der identifizierten KI-Leitlinien der Bundesländer anhand der Methode der strukturierenden qualitativen Inhaltsanalyse nach Mayring (2000). In Hinblick auf die beiden Forschungsfragen wird die

Scoping Study daher mit dem Ziel durchgeführt, Forschungsbeiträge zu identifizieren, die sich ihrerseits mit Ethik-Leitlinien für KI auseinandersetzen und hierdurch zu einem Ethik-Prinzipien und zu anderen Maßnahmen zur Umsetzung dieser spezifischen Ethik-Prinzipien identifizieren.

Im Zuge der Scoping Study wurde zunächst eine freie Suche im Inkognito-Modus in den Suchmaschinen Google, Google Scholar und Springer Link zu den Stichworten *AI ethics*, *ethische KI*, *Ethical AI Principles* und *KI Ethik-Prinzipien* durchgeführt. Durch den Inkognito-Modus soll die Verzerrung durch bereits im Browser gespeicherte Daten verhindert werden. In diesem Prozess wurde jeder Link der ersten drei Ergebnis-Seiten berücksichtigt. Die zentralen so identifizierten Forschungsbeiträge, d. h. Forschungsbeiträge, welche mindestens zweimal in anderen Forschungsbeiträgen erwähnt werden, wurden dann ihrerseits auf Connected Papers, einer visualisierenden und explorationsunterstützenden Suchmaschine für wissenschaftliche Fachartikel, eingegeben und hierdurch ähnliche Beiträge identifiziert. Zuletzt wurden manuell Querverweise aufgenommen sowie die Literaturverzeichnisse der bis dahin identifizierten Beiträge untersucht und hierdurch weitere relevante Beiträge identifiziert. Insgesamt wurde in allen Schritten Literatur zum Thema KI-Ethik, die sich auf einen oder mehrere spezifische Aspekte des Themas, wie bspw. Big Data oder Robotik konzentrieren, nicht berücksichtigt (bspw. Royakkers et al. (2018)), sodass das Kategoriensystem auf alle Leitlinien der Bundesländer angewendet werden kann. Ebenso wurde daher Literatur nicht berücksichtigt, die nicht-Ethik-Prinzipien-spezifische Maßnahmen zur Umsetzung ethischer KI vorschlägt (bspw. Rebstadt et al. (2022)).

Somit konnten insgesamt neun Beiträge identifiziert werden, die durch die Untersuchung von Ethik-Leitlinien Ethik-Prinzipien für KI oder Maßnahmen für deren Umsetzung identifiziert haben. In diesen Beiträgen wurden insgesamt 34 Ethik-Prinzipien für KI identifiziert, wobei das Ethik-Prinzip *AGI/ASI* von Zeng et al. (2019) aufgrund zu geringer Informationsbasis bezüglich dessen Interpretation und der hiermit thematisierten Problematik in der vorliegenden Arbeit nicht berücksichtigt wird (vgl. Zeng et al. 2019, S. 2). Da die inkludierten Ethik-Prinzipien inhaltlich starke Überschneidungen aufweisen, wurden sie zu 16 Kategorien zusammengefasst, die Tabelle 1 entnommen werden können. Ebenso werden in Tabelle 1 die Definitionen der in den zentralen Beiträgen genannten Ethik-Prinzipien festgehalten. Eine ausführliche Version der Tabelle, in der alle neun Beiträge und deren Definitionen der Ethik-Prinzipien sowie eine Zusammenfassung von Maßnahmen für deren Umsetzung dargestellt werden,

kann der Übersicht der Ergebnisse der Scoping Study in Anhang 1 entnommen werden.

Tabelle 1: Interpretation der Ethik-Prinzipien in zentralen Forschungsbeiträgen

Ethik-Prinzipien	Jobin et al. (2019)	Ryan und Stahl (2020)	Fjeld et al. (2020)	Zeng et al. (2019)
Collaboration	–	–	–	Zusammenarbeit, Partnerschaft, Kooperation, Dialog
Lawfulness	–	–	–	–
Promotion of Human Values	–	–	Übereinstimmung der Ziele und Zwecke von KI mit den gesellschaftlichen Grundwerten	–
Professional Responsibility	–	–	Professionelles, integriertes Handeln von KI-Entwickler*innen und Anwender*innen	–
Human Control of Technology, Agency & Oversight, Singularity of the Humanum	–	–	Treffen wichtiger Entscheidungen als Aufgabe der Menschen	–
Safety, Security	–	–	Ordnungsgemäßes Funktionieren der KI und Schutz vor Kompromittierung	Sicherheit, Cybersicherheit, Cyberangriff, Hacks
Solidarity	Soziales Netz im Arbeitsmarkt	Förderung von sozialer Sicherheit und Zusammenhalt	–	–
Dignity	Themen der Menschenrechte	Anerkennung der Menschenrechte	–	–
Sustainability	Schutz der Umwelt und Sicherung sozialer Nachhaltigkeit	Schutz der Umwelt und des Klimas	–	–
Trust, Reliability, Robustness	Prozesse vertrauenswürdiger Entwicklung und Implementierung von KI, um das Vertrauen von Kund*innen zu sichern	Vertrauen als Fundament interpersonaler Interaktion und der Gesellschaft	–	–
Freedom, Autonomy	Generelle Förderung von Freiheit und Autonomie	Förderung von Autonomie und Schutz der Freiheit durch KI	–	–

Ethik-Prinzipien	Jobin et al. (2019)	Ryan und Stahl (2020)	Fjeld et al. (2020)	Zeng et al. (2019)
Beneficence, Non-maleficence, Humanity	Vermeidung von vorhersehbarem oder nicht-intendiertem Schaden durch KI; Förderung ‚des Guten‘	Vermeidung von Schaden; Gutes tun	–	Menschlichkeit, Nutzen, Wohlbefinden, menschlicher Wert, Menschenrecht
(Data) Privacy	Datenschutz und Datensicherheit	Schutz der Privatsphäre von Individuen trotz der Arbeit mit Daten	Achtung der Privatsphäre von Individuen durch KI	Privatsphäre, persönliche Informationen, Datenschutz, informiert
Responsibility, Accountability	Handeln mit Integrität und die Klärung der Fragen rechtlicher Haftung für KI	Übernahme moralischer Verantwortung für KI durch KI-Anwender*innen	Adäquate Verteilung der Rechenschaftspflicht für KI	Rechenschaftspflicht, Verantwortung
Justice, Fairness, Equity, Non-discrimination, Share	Vermeidung von Bias und Diskriminierung	Fragen von Schaden und Ungerechtigkeit	Maximierung von Fairness und Inklusivität durch KI	Anteilhabe, Gleichberechtigung, Gerechtigkeit, Ungerechtigkeit
Transparency, Explicability, Explainability, Interpretability	Steigerung der Erklärbarkeit, Interpretierbarkeit und der Offenlegung von Informationen	Transparenz von KI und Transparenz der Organisationen, die KI entwickeln und anwenden	Beaufsichtigung und Interpretierbarkeit der Operationen und Outputs von KI	Transparent, erklärbar, vorhersehbar, nachvollziehbar

Anmerkungen:

Lawfulness wurde in keinem dieser vier zentralen Beiträge genannt, jedoch in den weiteren in der Scoping Study identifizierten Beiträgen. Dieses Ethik-Prinzip wird hier als Einhaltung aller Gesetze und Vorschriften bei der Konzeption, Entwicklung, Implementierung und den möglichen Folgen von KI verstanden. Dies kann Tabelle A1 in Anhang 1 entnommen werden.

Das Ethik-Prinzip *AGI/ASI* von Zeng et al. (2019) wird aufgrund zu geringer Informationsbasis bezüglich dessen Interpretation und der damit thematisierten Problematik nicht berücksichtigt.

Quelle: Eigene Darstellung auf Basis von Jobin et al. 2019, S. 391ff.; Ryan und Stahl 2020, S. 66ff.; Fjeld et al. 2020, S. 4f.; Zeng et al. 2019, S. 2

3.2. Strukturierende qualitative Inhaltsanalyse nach Mayring

Ziel der qualitativen Inhaltsanalyse ist es, das Interpretieren von Texten anhand inhaltsanalytischer Regeln transparent, intersubjektiv nachvollziehbar und reliabel werden zu lassen (vgl. Mayring 2000, S. 602). Mayring unterscheidet drei Grundtechniken der qualitativen Inhaltsanalyse: Zusammenfassung, Explikation und Strukturierung (vgl. Mayring 2010, S. 602). Während die zusammenfassende qualitative Inhaltsanalyse durch die induktive Vorgehensweise zu den Kernaussagen eines Textes gelangt und die explizierende qualitative Inhaltsanalyse unklare Textstellen zu erklären versucht, ist es Ziel der strukturierenden qualitativen Inhaltsanalyse das Textmaterial anhand von deduktiv gebildeten Kategorien zu systematisieren, um hierdurch zentrale Aspekte herauszugreifen (vgl. Mayring 2010, S. 602).

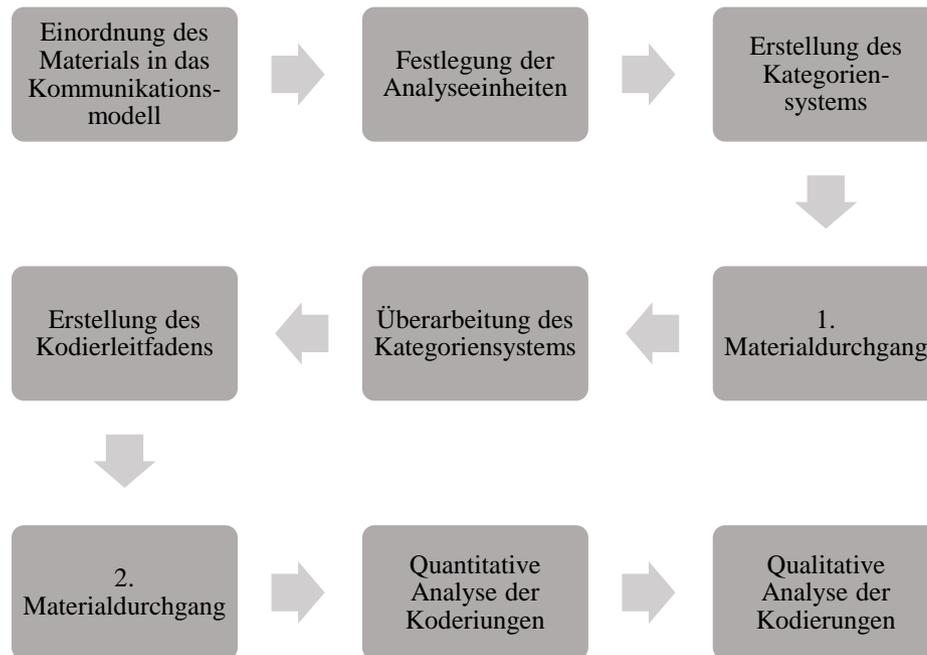
Zur Auswertung der Leitlinien der Bundesländer in der vorliegenden Arbeit wurde die strukturierende qualitative Inhaltsanalyse nach Mayring gewählt, da sie durch ihre

Spezifika der Systematisierung des Materials und der Anfertigung von Querauswertungen am besten geeignet ist, um die Differenzen der Leitlinien zu identifizieren. Wie auch Abbildung 1 zu entnehmen ist, werden die identifizierten Leitlinien im Zuge der strukturierenden qualitativen Inhaltsanalyse zunächst in ein Kommunikationsmodell eingeordnet, d. h. der sozio-kulturelle Hintergrund, die federführenden Akteur*innen, die Textproduktionssituation, Textwirkungen und Zielgruppen der Leitlinien werden identifiziert (vgl. Mayring und Fenzl 2014, S. 546; vgl. Mayring 2010, S. 605). Dies wird im folgenden Unterkapitel dargestellt. Anschließend werden die Analyseeinheiten festgelegt: die Kodiereinheit, d. h. die minimalste Text-Auswertungseinheit, stellen in der vorliegenden Arbeit einzelne Worte dar, die im Zusammenhang mit KI stehen; die Kontexteinheit, d. h. die größte Text-Auswertungseinheit, stellen einzelne Kapitel der Leitlinien zum Thema KI dar, und die Auswertungseinheit, d. h. die Materialportion, der ein Kategoriensystem gegenübergestellt wird, stellen die in dieser Arbeit inkludierten Leitlinien dar (vgl. Mayring und Fenzl 2014, S. 546).

Für die qualitativ-inhaltliche Analyse der Leitlinien erfolgt dann die Erstellung des Kategoriensystems und des Kodierleitfadens. Das Kategoriensystem basiert in der vorliegenden Arbeit auf den durch die Scoping Study identifizierten Ethik-Prinzipien, die in Unterkapitel 3.1 in Tabelle 1 abgebildet sind. Diese wurden zur besseren Nachvollziehbarkeit jeweils zu Kategorien in deutscher Sprache zusammengefasst und zu jeder dieser 16 Ethik-Prinzipien-Kategorien wurde eine Kategorie zu deren Umsetzung erstellt, sodass das Kategoriensystem insgesamt 32 Kategorien umfasst. Dies kann Tabelle 4 in Unterkapitel 4.1 entnommen werden. Das Kategoriensystem wird daraufhin in einem ersten Materialdurchgang induktiv überarbeitet und ergänzt (vgl. Mayring und Fenzl 2014, S. 548). Hierbei wird für jede Kategorie eine Definition (d. h. definierende Merkmale), ein Ankerbeispiel (d. h. typische Merkmale) und Kodierregeln (d. h. Regeln, die die Kategorien voneinander abgrenzen) festgelegt, was dann den Kodierleitfaden darstellt, und der Anhang 3 entnommen werden kann (vgl. Mayring und Fenzl 2014, S. 549ff.). Der nächste Schritt ist der zweite, und finale Materialdurchgang, d. h. die erneute regelgeleitete Zuordnung von Kategorien bzw. Kodierungen zu Textstellen anhand des nun fertiggestellten Kodierleitfadens (vgl. Mayring 2000, S. 5). In der vorliegenden Arbeit wurde hierfür die Software MAXQDA, zur computergestützten qualitativen Inhaltsanalyse der Leitlinien genutzt. Darauf folgt eine einfache quantitative Auswertung der Kodierungen. Abschließend erfolgt die qualitative Analyse der Kodierungen und damit die Beantwortung der

Forschungsfragen. Die Ergebnisse dieser beiden letzten Schritte können Kapitel 4 entnommen werden.

Abbildung 1: Ablaufmodell der strukturierenden qualitativen Inhaltsanalyse nach Mayring (2000)



Quelle: Eigene Abbildung nach Mayring 2000, S. 5

3.3. Kommunikationsmodell nach Mayring

Je Bundesland sollte eine Leitlinie in die Untersuchung aufgenommen werden. Die Auswahl der Leitlinien orientierte sich an der Website der Bundesregierung zur nationalen Strategie für KI, auf welcher Digitalisierungs- bzw. Innovations- und teilweise auch KI-Leitlinien des Bundes und der Länder vorgestellt werden (vgl. Bundesregierung KI o. J., o. S.). Des Weiteren wurden die Websites der Landesregierungen manuell auf die Existenz neuerer Leitlinien überprüft. Falls zwei oder mehr Leitlinien pro Bundesland vorhanden sind und keine spezifische KI-Leitlinie darunter ist, so wurde das Dokument ausgewählt, welches die meisten Treffer bzgl. des Stichwortes *Künstliche Intelligenz* bzw. *KI* aufweist. Falls eine spezifische KI-Leitlinie darunter ist, so wurde diese in jedem Fall ausgewählt. Falls eine Fortschreibung einer älteren Version einer gewählten Leitlinie existiert, so wurde diese neuere Version ausgewählt. Eine Ausnahme hiervon stellt Bayern dar, da die Fortschreibung der Hightech Agenda Bayern als Ergänzung dieser dient, und daher beide Leitlinien inkludiert wurden. Eine weitere Ausnahme stellt das Bundesland Rheinland-Pfalz dar. Da in keiner existierenden Leitlinie des Bundeslandes relevante Informationen in Bezug auf die

Forschungsfragen gefunden werden konnten, wird dieses Bundesland aus den folgenden Untersuchungen und Ausführungen ausgeschlossen (vgl. MWVLW 2021; vgl. Landesregierung Rheinland-Pfalz 2020; vgl. Staatskanzlei Rheinland-Pfalz 2018). Eine Übersicht aller inkludierten Leitlinien kann Tabelle A3 in Anhang 2 entnommen werden.

In Tabelle 2 werden die entsprechend dem beschriebenen Vorgehen ausgewählten Leitlinien der Bundesländer in das Kommunikationsmodell nach Mayring (2000) eingeordnet. Eine ausführliche Version kann Tabelle A4 in Anhang 2 entnommen werden. Ergänzend zu dem Kommunikationsmodell nach Mayring (2000) werden in diesem Unterkapitel, aufgeschlüsselt nach den Bundesländern, deren jeweilige Definition von KI vorgestellt.

Berlin, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt und Thüringen definieren KI als eine Form der Imitation menschlicher Intelligenz (vgl. SenWiEnBe 2021, S. 63; vgl. EM MV 2021, S. 77; vgl. SK Sachsen 2021, S. 4; vgl. MW LSA 2021, S. 68; vgl. TMWWDG 2021, S. 60). Brandenburg definiert KI als ein Verfahren der Datenanalyse (vgl. MWAE 2022, S. 8). Hessen nennt in der Definition beide dieser Komponenten (vgl. Stk Hessen 2021, S. 2). Niedersachsen, Nordrhein-Westfalen und Schleswig-Holstein verstehen KI als ein Teilgebiet der Informatik und als Sammelbegriff für viele verschiedene Technologien (vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 6; vgl. Poretschkin et al. 2021, S. 18f.; vgl. Stk SH 2021, S. 7). Schleswig-Holstein unterscheidet hierbei als einziges Bundesland zwischen starker und schwacher KI (vgl. Stk SH 2021, S.37).

4. Ergebnisse der qualitativen Inhaltsanalyse nach Mayring

In diesem Kapitel werden die Ergebnisse der strukturierenden qualitativen Inhaltsanalyse nach Mayring (2000) vorgestellt. In Unterkapitel 4.1 wird zunächst eine einfache quantitative Auswertung der Kodierungen vorgestellt. In Unterkapitel 4.2 erfolgt dann die Darstellung der Ergebnisse der qualitativen Auswertung der Kodierungen, die sich an den Forschungsfragen orientiert. Wo dies in den Leitlinien der Bundesländer genannt wird, werden demnach die Interpretation der Ethik-Prinzipien und die mit ihnen thematisierten Problematiken genannt. In Bezug auf die zur Umsetzung der Ethik-Prinzipien vorgeschlagenen Maßnahmen werden, wo dies in den Leitlinien genannt wird, die Ziele und Adressat*innen dieser Maßnahmen vorgestellt. Somit werden also die Differenzen zwischen den Bundesländern deskriptiv dargestellt und die Forschungsfragen damit beantwortet.

Tabelle 2: Kommunikationsmodell nach Mayring (2000)

Bundesland	Titel	Jahr	Federführung	Textproduktionssituation	Textwirkungen	Zielgruppe(n)
Baden-Württemberg	Strategiepapier Künstliche Intelligenz - Baden-Württemberg wird Vorreiter für Künstliche Intelligenz	2018	Landesregierung Baden-Württemberg (Landesregierung BW)	Erarbeitung zeitgleich mit der KI-Strategie der Bundesregierung (vgl. Landesregierung BW 2018a, o. S.).	Strebt die Etablierung Baden-Württembergs als globale Leitregion des digitalen Wandels an (vgl. Landesregierung BW 2018b, S. 1f.).	Forschung, Wissenschaft und Wirtschaft (vgl. Landesregierung BW 2018b, S. 1).
Bayern	Hightech Agenda Bayern. Regierungserklärung des Bayerischen Ministerpräsidenten Dr. Markus Söder, MdL, am 10. Oktober 2019 vor dem Bayerischen Landtag	2019	Bayrische Staatskanzlei (BStK)	Situation des Wettbewerbs, sowohl international als auch zwischen den Bundesländern (vgl. BStK 2019, S. 4). Erarbeitung durch Gespräche mit Wissenschaftler*innen und Akteur*innen der Wirtschaft (vgl. BStK 2019, S. 17).	Strebt die technologische Dominanz insbesondere durch die Förderung von KI-Forschung an (vgl. BStK 2019, S. 4).	Forschung, Wissenschaft und Wirtschaft (vgl. BStK 2019, S. 5, S. 15).
	Hightech Agenda Bayern Plus. Bericht aus der Kabinettsitzung	2020	Bayrische Staatskanzlei (BStK)	Fortschreibung und Ergänzung der bestehenden Agenda in der Covid-19-Pandemie (vgl. BStK 2020, S. 2).	Beschleunigung der Umsetzung der Hightech Agenda (vgl. BStK 2020, S. 2).	Forschung, Wissenschaft und Wirtschaft (vgl. BStK 2019, S. 5, S. 15).
Berlin	Grünbuch für die Digitalisierungsstrategie des Landes Berlin	2021	Senatsverwaltung für Wirtschaft, Energie und Betriebe Berlin (SenWiEnBe)	Erarbeitung im Zuge der Covid-19-Pandemie (vgl. SenWiEnBe 2021, S. 4).	Aufstellung von Leitlinien für die digitale Transformation (vgl. SenWiEnBe 2021, S. 8).	Alle Sektoren und Politikfelder (vgl. SenWiEnBe 2021, S. 13).
Brandenburg	Strategie Einsatz von Künstlicher Intelligenz in brandenburgischen Unternehmen	2022	Ministerium für Wirtschaft, Arbeit und Energie des Landes Brandenburg (MWAE)	Die Strategie ist Teil des Prozesses der Erarbeitung einer KI-Strategie des Landes Brandenburg (vgl. MWAE 2022, S. 8).	Wettbewerbsfähigkeit von Unternehmen in Bezug auf den Einsatz von KI (vgl. MWAE 2022, S. 8).	Die strategischen Ziele der EU (vgl. MWAE 2022, S. 8).
Bremen	BREMEN.KI. Strategie Künstliche Intelligenz	2020	Landesregierung Bremen	Entwicklung durch den Wirtschafts- und Wissenschaftsressort (vgl. Landesregierung Bremen 2020, S. 5).	Strebt den Ausbau Bremens KI-Forschung und die stärkere Vernetzung mit der Wirtschaft an (vgl. Landesregierung Bremen 2020, S. 5).	Arbeitsgruppen verschiedener Ressorts, Politik, Gesellschaft, Unternehmen und Institutionen (vgl. Landesregierung Bremen 2020, S. 8, S. 32).

Bundesland	Titel	Jahr	Federführung	Textproduktionssituation	Textwirkungen	Zielgruppe(n)
Hamburg	Digitalstrategie für Hamburg	2020	Senatskanzlei Hamburg	Setzt an bereits etablierten Maßnahmen, Projekten und Entwicklungsbereichen an (vgl. Senatskanzlei Hamburg 2020, S. 6).	Vorantreiben der Digitalisierung (vgl. Senatskanzlei Hamburg 2020, S. 6).	Alle Teile der Stadtgesellschaft (vgl. Senatskanzlei Hamburg 2020, S. 5).
Hessen	Thesepapier des Rats für Digitalethik der Hessischen Landesregierung Vertrauen in KI	2021	Hessische Staatskanzlei (Stk Hessen)	Beruhet auf den Einschätzungen des hessischen Rats für Digitalethik zu der Frage, welcher Beitrag zur vertrauenswürdigen Entwicklung und dem verantwortungsvollen Einsatz von KI aus Sicht der Bürger*innen und Anwender*innen geleistet werden kann (vgl. Stk Hessen 2021, S. 3).	Beratung der hessischen Landesregierung in Bezug auf deren Möglichkeiten zur vertrauenswürdigen Entwicklung und verantwortungsvollem Einsatz von KI (vgl. Stk Hessen 2021, S. 3).	Hessische Landesregierung (vgl. Stk Hessen 2021, S. 2).
Mecklenburg-Vorpommern	Digitale Agenda der Landesregierung Mecklenburg-Vorpommern 2018-2021	2021	Ministerium für Energie, Infrastruktur und Digitalisierung Mecklenburg-Vorpommern (EM MV)	Bearbeitung in verschiedenen Ressorts, dem Digitalisierungsbeirat, in Digitalen Innovationszentren und mit weiteren Akteur*innen (vgl. EM MV 2021, S. 6f.).	Festlegung eines Finanzrahmens für die Digitalisierung (vgl. EM MV 2021, S. 6f.).	Gesellschaft, Wissenschaft, Wirtschaft, Verwaltung (vgl. EM MV 2021, S. 7).
Niedersachsen	KI-Working Paper Niedersachsen	2021	Niedersächsisches Ministerium für Wirtschaft, Arbeit, Verkehr und Digitalisierung (MW Niedersachsen) und Niedersächsisches Ministerium für Wissenschaft und Kultur (MWK Niedersachsen)	KI wurde bereits im Masterplan Digitalisierung aus dem Jahr 2018 als eine Kerntechnologie verankert (vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 3).	Zeigt Handlungsbedarfe und Rahmenbedingungen Bezug auf KI auf (vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 3).	Wirtschaft, Forschung und Gesellschaft (vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 18).
Nordrhein-Westfalen	Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz - KI-Prüfkatalog	2021	Fraunhofer Institut für Intelligente Analyse- und Informationssysteme IAIS (Poretschkin et al.)	Vom Fraunhofer IAIS veröffentlicht, und durch das Ministerium für Wirtschaft, Innovation, Digitalisierung und Energie des Landes Nordrhein-Westfalen gefördert.	Orientierung für Entwickler*innen, um KI vertrauenswürdig zu gestalten und Anleitung für Prüfer*innen zur Untersuchung von KI-Anwendungen strukturiert auf diese Vertrauenswürdigkeit (vgl. Poretschkin et al. 2021, S. 9).	Entwickler*innen und Prüfer*innen von KI (vgl. Poretschkin et al. 2021, S. 9).

Bundesland	Titel	Jahr	Federführung	Textproduktionssituation	Textwirkungen	Zielgruppe(n)
Saarland	Strategie für Innovation und Technologie Saarland (Innovationsstrategie) - StrategiePlus (2021-2023)	2021	Staatskanzlei des Saarlandes (Stk Saarland)	Basiert auf der Analyse von Statistiken und Förderdaten, Gutachten, Strategien und der Konsultation von Stakeholder*innen und Expert*innen (vgl. Stk Saarland 2021, S. 5).	Dient der Steigerung der bis dahin festgelegten Innovationsaktivitäten und als Grundlage für Neuauflagen der Strategie im Jahr 2023 (vgl. Stk Saarland 2021, S. 27f.).	Akteur*innen der Bereiche Informatics, Life Science & Material Science und Smart Production & Automotive (vgl. Stk Saarland 2021, S. 2).
Sachsen	Künstliche Intelligenz. Strategie für den Freistaat Sachsen	2021	Sächsische Staatskanzlei (SK Sachsen)	Eingebettet sowohl in die Strategien des Bundes und der Europäischen Union als auch in die bestehenden Digital- und Innovationsstrategien Sachsens (vgl. SK Sachsen 2021, S. 3).	Ausbau der KI-Infrastruktur im Land Sachsen (vgl. SK Sachsen 2021, S. 10).	KI-Akteur*innen, Forschung, Wirtschaft, Bürger*innen, Bildung und die Verwaltung (vgl. SK Sachsen 2021, S. 10ff., S. 20, S. 30, S. 34, S. 38).
Sachsen-Anhalt	Künstliche Intelligenz - Digitale Agenda für das Land Sachsen-Anhalt	2021	Ministerium für Wirtschaft, Wissenschaft und Digitalisierung des Landes Sachsen-Anhalt (MW LSA)	KI als immanenter Bestandteil der Digitalen Agenda eingebettet (vgl. MW LSA 2021, S. 3).	Digitalisierung und KI in der Gesellschaft fördern (vgl. MW LSA 2021, Präambel).	Bürger*innen, Wirtschaft, Wissenschaft, Bildung, Verwaltung, Medien, Kultur (vgl. MW LSA 2021, S. 1).
Schleswig-Holstein	Künstliche Intelligenz – Strategische Ziele und Handlungsfelder für Schleswig-Holstein	2021	Staatskanzlei Schleswig-Holstein (Stk SH)	Fortschreibung der Strategie aus dem Jahr 2019, die die erste KI-Strategie eines deutschen Bundeslandes war (vgl. Stk SH 2021, S. 3).	Erneuerung der ersten Version (vgl. Stk SH 2021, S. 3).	Bürger*innen Schleswig-Holsteins und alle „[...] KI-Interessierte [...]“ (Stk SH 2021, S. 3).
Thüringen	Thüringer Strategie für die Digitale Gesellschaft - Update 2021	2021	Thüringer Ministerium für Wirtschaft, Wissenschaft und Digitale Gesellschaft (TMWWDG)	Erarbeitung in Zusammenarbeit mit externen Interessenvertreter*innen aus Wirtschaft, Wissenschaft, Politik, Intermediär*innen sowie Bürger*innen (vgl. TMWWDG 2021, S. 4).	Gestaltung des Handelns der Landesregierung in digitalen Angelegenheiten (vgl. TMWWDG 2021, S. 4).	Bürger*innen (vgl. TMWWDG 2021, S. 4).

Anmerkung: Eine ausführliche Version des Kommunikationsmodells kann Tabelle A4 in Anhang 2 entnommen werden.

Quelle: Eigene Darstellung nach Mayring 2000, S. 3 auf Basis der Leitlinien (siehe Leitlinienverzeichnis)

Im Folgenden wird hierbei von Codes bzw. Kodierungen im Kontext der empirischen Arbeit am Material gesprochen, wobei diese identisch mit den Kategorien des Kategoriensystems sind, die Tabelle 4 entnommen werden können.

4.1. Quantitative Auswertung

Tabelle 3 zeigt, dass in der Leitlinie Nordrhein-Westfalens mit 68 Codes die meisten Codes vergeben wurden. Für die Leitlinie des Saarlandes wurden mit einem Code die wenigsten Codes aller inkludierten Bundesländer vergeben.

Tabelle 3: Anzahl verbogener Codes pro Leitlinie(n) eines Bundeslandes

Liste der Bundesländer	Anzahl verbogener Codes in den Leitlinien der Bundesländer
Baden-Württemberg	21
Bayern	3
Berlin	18
Brandenburg	18
Bremen	14
Hamburg	11
Hessen	19
Mecklenburg-Vorpommern	3
Niedersachsen	18
Nordrhein-Westfalen	68
Saarland	1
Sachsen	28
Sachsen-Anhalt	14
Schleswig-Holstein	31
Thüringen	6
Gesamt	273

Quelle: Eigene Berechnung auf Basis der Leitlinien (siehe Leitlinienverzeichnis)

Insgesamt wurden 273 Codes vergeben, was ebenfalls aus Tabelle 3 hervorgeht. Tabelle 4 kann entnommen werden, wie viele Codes pro Kategorie vergeben wurden und in wie vielen Leitlinien unterschiedlicher Bundesländer die jeweilige Kategorie genannt wird. In mindestens der Hälfte der Leitlinien der Bundesländer werden demnach die Kategorien *Fairness und Nicht-Diskriminierung*, *Datenschutz und Privatsphäre*, *Transparenz und Nachvollziehbarkeit*, *Nachhaltigkeit*, *Sicherheit*, *Zusammenarbeit* sowie *Menschliche Kontrolle und Menschliche Singularität* thematisiert. Hingegen wurden keine Codes für das Ethik-Prinzip *Zusammenhalt* vergeben.

Insgesamt sind 68% aller vergebenen Codes Kodierungen von Ethik-Prinzipien und damit nur 32% der vergebenen Codes Kodierungen von Kategorien zur Umsetzung

der Ethik-Prinzipien. *Vertrauen und Verlässlichkeit* ist hierbei das einzige Ethik-Prinzip, bei dem die Anzahl vergebener Codes für die Kategorie zur Umsetzung des Ethik-Prinzips höher ist als die Anzahl vergebener Codes für die Kategorie des Ethik-Prinzips selbst. Die Kategorie *Umsetzung des Vertrauens und der Verlässlichkeit* weist insgesamt die zweithöchste Anzahl an Codes vor, wobei diese Kategorie jedoch nur für vier Bundesländer kodiert wurde. Mit 19 vergebenen Codes ist daneben noch *Umsetzung der Fairness und Nicht-Diskriminierung* als vergleichsweise häufig kodierte Kategorie zur Umsetzung eines Ethik-Prinzips hervorzuheben. Für die Kategorien *Umsetzung der Rechtmäßigkeit*, *Umsetzung der Anerkennung gesellschaftlicher Grundwerte*, *Umsetzung des Zusammenhalts* und *Umsetzung der Achtung der Menschenrechte* wurden keine Codes vergeben.

Tabelle 4: Anzahl vergebener Codes pro Kategorie

Kategorien	Konstituierende Ethik-Prinzipien	Vergebene Codes	Vorkommen in unterschiedlichen Leitlinien
Zusammenarbeit	Collaboration	15	8
Umsetzung der Zusammenarbeit		6	1
Rechtmäßigkeit	Lawfulness	7	5
Umsetzung der Rechtmäßigkeit		0	0
Anerkennung gesellschaftlicher Grundwerte	Promotion of Human Values	4	4
Umsetzung der Anerkennung gesellschaftlicher Grundwerte		0	0
Berufliche Verantwortung	Professional Responsibility	4	3
Umsetzung der beruflichen Verantwortung		3	2
Menschliche Kontrolle und Menschliche Singularität	Human Control of Technology, Agency & Oversight, Singularity of the Humanum	10	8
Umsetzung der menschlichen Kontrolle und menschlichen Singularität		7	6
Sicherheit	Safety, Security	17	10
Umsetzung der Sicherheit		7	3
Zusammenhalt	Solidarity	0	0
Umsetzung des Zusammenhalts		0	0
Achtung der Menschenrechte	Dignity	5	4
Umsetzung der Achtung der Menschenrechte		0	0
Nachhaltigkeit	Sustainability	17	10
Umsetzung der Nachhaltigkeit		1	1
Vertrauen und Verlässlichkeit	Trust, Reliability, Robustness	15	6

Kategorien	Konstituierende Ethik-Prinzipien	Vergebene Codes	Vorkommen in unterschiedlichen Leitlinien
Umsetzung des Vertrauens und der Verlässlichkeit		21	4
Freiheit und Autonomie	Freedom, Autonomy	11	4
Umsetzung der Freiheit und Autonomie		5	3
Gemeinwohl und Schadensvermeidung	Beneficence, Non-maleficence, Humanity	6	4
Umsetzung des Gemeinwohls und der Schadensvermeidung		2	2
Datenschutz und Privatsphäre	(Data) Privacy	20	9
Umsetzung des Datenschutzes und der Privatsphäre		6	4
Verantwortung und Rechenschaftspflicht	Responsibility, Accountability	7	6
Umsetzung der Verantwortung und Rechenschaftspflicht		1	1
Fairness und Nicht-Diskriminierung	Justice, Fairness, Equity, Non-discrimination, Share	29	10
Umsetzung der Fairness und Nicht-Diskriminierung		19	9
Transparenz und Nachvollziehbarkeit	Transparency, Explicability, Explainability, Interpretability	19	9
Umsetzung der Transparenz und Nachvollziehbarkeit		9	3

Quelle: Eigene Berechnung auf Basis von Fjeld et al. 2020, S. 4f.; Floridi und Cowls 2019, S. 7ff.; Europäische Kommission 2018, S. 10ff.; Jobin et al. 2019, S. 391ff.; Ryan und Stahl 2020, S. 66ff.; Fetic et al. 2020, S. 26ff.; Axente 2021, o. S.; Kazim und Koshiyama 2021, S. 8ff.; Zeng et al. 2019, S. 2; sowie auf Basis der Leitlinien (siehe Leitlinienverzeichnis)

4.2. Qualitative Auswertung

Zusammenarbeit

Das Ethik-Prinzip *Zusammenarbeit* kann in den Leitlinien von Baden-Württemberg, Bayern, Berlin, Brandenburg, Hessen, Sachsen und Sachsen-Anhalt wiedergefunden werden. Hierbei wird es interpretiert als die Relevanz von Zusammenarbeit innerhalb eines Bundeslandes, zwischen den Bundesländern, der Bundesländer mit dem Bund sowie auf europäischer und internationaler Ebene ebenso wie die Relevanz der Zusammenarbeit verschiedener Branchen, in den Bereichen Finanzierung und Forschung. Mit diesem Ethik-Prinzip thematisieren verschiedene Leitlinien unterschiedliche Problematiken. Zum einen wird in der Leitlinie von Berlin beschrieben, dass den Problematiken fehlenden Schutzes von Verbraucher*innen und fehlenden Vertrauens in KI durch Zusammenarbeit entgegengewirkt werden soll (vgl. SenWiEnBe 2021, S.

40). Zum anderen thematisieren Brandenburg und Sachsen hingegen die Problematik der fehlenden Sicherung der Wettbewerbsfähigkeit von, und Unterstützung für Unternehmen durch fehlende Zusammenarbeit (vgl. MWAE 2022, S. 37, S. 41; vgl. SK Sachsen 2021, S. 31).

Umsetzung der Zusammenarbeit

Konkrete Maßnahmen für die Umsetzung des Ethik-Prinzips *Zusammenarbeit* werden nur in der Leitlinie Baden-Württembergs vorgeschlagen. Baden-Württemberg nennt hier Projekte zur Forschung zu KI, die in Zusammenarbeit mit verschiedenen Forschungseinrichtungen, Unternehmen sowie gemeinsam mit europäischen, insbesondere französischen, Partner*innen umgesetzt werden sollen, wie bspw. Cyber Valley oder die ELLIS-Initiative (vgl. Landesregierung BW 2018b, S.5). Hierbei wird auch das Anstreben einer Kooperation mit der Bundesregierung, welche das Ziel eines deutsch-französischen Netzwerkes für KI verfolgt, betont (vgl. Landesregierung BW 2018b, S. 5).

Rechtmäßigkeit

Für das Ethik-Prinzip *Rechtmäßigkeit* wurden Codes in den Leitlinien der Bundesländer Berlin, Hamburg, Niedersachsen, Sachsen und Thüringen vergeben. Die zentrale Problematik, die in den Leitlinien mit dem Ethik-Prinzip thematisiert wird, ist die Frage der Vereinbarkeit von KI-Anwendungen mit bestehenden rechtlichen Standards, die auch für KI gelten müssen, und damit einhergehend die Gefahr deren Aushöhlung. In den Leitlinien von Hamburg und Sachsen wird *Rechtmäßigkeit* daher als Möglichkeit zum Schutz der Verbraucher*innen und zur Stärkung deren Vertrauen in KI interpretiert (vgl. Senatskanzlei Hamburg 2020, S. 57; vgl. SK Sachsen 2021, S. 31). Hingegen wird *Rechtmäßigkeit* in der Leitlinie Thüringens in Hinblick auf die hierdurch entstehenden Hemmnisse für die wirtschaftliche Entwicklung von KI thematisiert (vgl. TMWWDG 2021, S. 11).

Umsetzung der Rechtmäßigkeit

Für die Kategorie *Umsetzung der Rechtmäßigkeit* wurden keine Codes vergeben, d. h. in keiner der Leitlinien werden Maßnahmen für die Umsetzung des Ethik-Prinzips *Rechtmäßigkeit* vorgeschlagen.

Anerkennung gesellschaftlicher Grundwerte

Das Ethik-Prinzip *Anerkennung gesellschaftlicher Grundwerte* wird in den Leitlinien von Baden-Württemberg, Hamburg, Nordrhein-Westfalen und Schleswig-Holstein thematisiert. Interpretiert wird es in allen Leitlinien als die Wichtigkeit der Vereinbarkeit der Konzeption, Entwicklung, Implementierung und der möglichen Folgen von KI mit gesellschaftlichen und humanistischen Wertvorstellungen. Zwei verschiedene Problematiken werden mit diesem Ethik-Prinzip thematisiert. Baden-Württemberg sieht das mögliche Nicht-Standhalten Deutschlands im weltweiten Wettbewerb um KI als problematisch an (vgl. Landesregierung BW 2018b, S. 4). Hamburg thematisiert mit dem Ethik-Prinzip hingegen die Problematik, dass der Mensch aus dem Mittelpunkt der Veränderungen durch KI gerückt werden könnte (vgl. Senatskanzlei Hamburg 2020, S. 57).

Umsetzung der Anerkennung gesellschaftlicher Grundwerte

Für die Kategorie *Umsetzung der Anerkennung gesellschaftlicher Grundwerte* wurden keine Codes vergeben.

Berufliche Verantwortung

Das Ethik-Prinzip *Berufliche Verantwortung* lässt sich in den Leitlinien der Bundesländer Brandenburg, Sachsen-Anhalt und Schleswig-Holstein wiederfinden, wobei in den Leitlinien Brandenburgs und Schleswig-Holsteins hierunter die Konsultation von KI-Stakeholder*innen verstanden wird (vgl. MWAE 2022, S. 54; vgl. Stk SH 2021, S. 11). Sachsen-Anhalt fasst unter diesem Ethik-Prinzip sowohl die Folgenabschätzung als auch die Konsultation von KI-Stakeholder*innen und den Schutz von Verbraucher*innen (vgl. MW LSA 2021, S. 15, S. 58). Spezifische Problematiken werden nicht genannt.

Umsetzung der beruflichen Verantwortung

Sachsen-Anhalt und Schleswig-Holstein geben Maßnahmen zur Umsetzung der beruflichen Verantwortung an. Sachsen-Anhalt nennt hier den Marktwächter Digitale Welt, der der Folgenabschätzung und Lösung möglicher Missstände dient (vgl. MW LSA 2021, S. 58). Schleswig-Holstein setzt auf die Konsultation eines von dem Bundesland etablierten KI-Expert*innenrates (vgl. Stk SH 2021, S. 21).

Menschliche Kontrolle und Menschliche Singularität

Das Ethik-Prinzip *Menschliche Kontrolle und Menschliche Singularität* wird in den Leitlinien der Bundesländer Baden-Württemberg, Berlin, Brandenburg, Bremen, Hessen, Niedersachsen, Nordrhein-Westfalen und Schleswig-Holstein thematisiert. Hierbei werden zwei unterschiedliche Interpretationen genannt. Zum einen wird das Ethik-Prinzip als die Wichtigkeit des Menschen als Mittelpunkt des Veränderungsprozesses um KI verstanden und betrifft damit die Problematik der Verdrängung des Menschen aus diesem Zentrum, was auf die Leitlinien Baden-Württembergs, Berlins und Brandenburgs zutrifft (vgl. Landesregierung BW 2018b, S. 5; vgl. SenWiEnBe 2021, S. 58; vgl. MWAE 2022, S. 57). Zum anderen verstehen Bremen, Hessen und Niedersachsen das Ethik-Prinzip als die Wichtigkeit menschlicher Kontrolle von KI, d. h. sie thematisieren die Problematik unerwünschter Folgen von KI aufgrund von fehlender menschlicher Kontrolle bzw. Aufsicht (vgl. Landesregierung Bremen 2020, S. 5; vgl. Stk Hessen 2021, S. 2; vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 11). Nordrhein-Westfalen und Schleswig-Holstein nennen in ihren Leitlinien beide Interpretationen und Problematiken (vgl. Poretschkin et al. 2021, S. 47; vgl. Stk SH 2021, S. 11).

Umsetzung der menschlichen Kontrolle und menschlichen Singularität

Zur Umsetzung dieses Ethik-Prinzips werden einige unterschiedliche Maßnahmen in den Leitlinien Baden-Württembergs, Berlins, Brandenburgs, Hamburgs, Nordrhein-Westfalens und Sachsens genannt. Baden-Württemberg setzt auf eine transdisziplinäre Kooperation der Geistes- und Sozialwissenschaften mit den Technikwissenschaften sowie den Bürger*innen, während Berlin und Brandenburg auf die Qualifizierung des (zukünftigen) KI-Personals durch Aus- und Weiterbildungen setzen, sodass die Bürger*innen dieser Bundesländer in die Gestaltung des Veränderungsprozesses einbezogen werden können (vgl. Landesregierung BW 2018b, S. 5; vgl. SenWiEnBe 2021, S. 58; vgl. MWAE 2022, S. 57). Hamburg, Nordrhein-Westfalen und Sachsen schlagen hingegen Maßnahmen zur Kontrolle und Überwachung vor, wie bspw. das Überprüfen des Ergebnisses von KI-Anwendungen durch Fachpersonal, das Eingreifen in, oder sogar Abschalten von KI-Anwendungen aufgrund (potenzieller) Personen- und Sachschäden und die Überprüfung des (Normal-)Betriebs durch qualifiziertes Personal, anhand von übersichtlichen Dokumentationen (vgl. Senatskanzlei Hamburg 2020, S. 37; vgl. Poretschkin et al. 2021, S. 53ff.; vgl. SK Sachsen 2021, S. 31).

Sicherheit

Das Ethik-Prinzip *Sicherheit* wird in den Leitlinien der Bundesländer Baden-Württemberg, Bayern, Berlin, Niedersachsen, Nordrhein-Westfalen, Saarland, Sachsen, Sachsen-Anhalt und Schleswig-Holstein genannt. Baden-Württemberg, Bayern, Berlin, das Saarland, Sachsen und Sachsen-Anhalt nennen hierbei die Problematik von Cyberangriffen auf oder durch KI (vgl. Landesregierung BW 2018b, S. 4; vgl. BStK 2019, S. 7; vgl. SenWiEnBe 2021, S. 39; vgl. Stk Saarland 2021, S. 15; vgl. SK Sachsen 2021, S. 39; vgl. MW LSA 2021, S. 58). Bremen, Niedersachsen und Schleswig-Holstein hingegen nennen, dass fehlende Sicherheit in Bezug auf KI Problematiken mit sich bringen kann, präzisieren diese jedoch nicht weiter (vgl. Landesregierung Bremen 2020, S. 14; vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 11; vgl. Stk SH 2021, S. 8). Die Leitlinie Nordrhein-Westfalens geht als Einzige ausführlicher auf die Thematik ein und beschreibt verschiedene Bereiche von KI-Sicherheit und Problematiken für diese Sicherheit (vgl. Poretschkin et al. 2021, S. 26, S. 116). Insgesamt wird das Ethik-Prinzip jedoch in allen Leitlinien als Sicherheit in Bezug auf die Konzeption, Entwicklung, Implementierung und möglichen Folgen von KI interpretiert.

Umsetzung der Sicherheit

In den Leitlinien von Baden-Württemberg, Niedersachsen und Nordrhein-Westfalen werden konkrete Maßnahmen für die Umsetzung des Ethik-Prinzips *Sicherheit* angegeben. Baden-Württemberg schlägt vor, Cybersicherheit durch verschiedene Forschungsprojekte, auch durch die Unterstützung der Bundesregierung, zu garantieren, um mittelständische Unternehmen und die Kritische Infrastruktur zu unterstützen und zu sichern (vgl. Landesregierung BW 2018b, S. 4). Niedersachsen schlägt die Erweiterung der Rolle der Landesbeauftragten für den Datenschutz unter anderem auf den Bereich der Cybersicherheit vor (vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 16). Nordrhein-Westfalen schlägt einen umfangreichen Maßnahmenkatalog zur Umsetzung des Ethik-Prinzips für drei verschiedene Risiko-Gebiete in Bezug auf Sicherheit vor: Es werden Vorgehen beschrieben, die den Umgang mit Daten allgemein (bspw. die Dokumentation über den Umgang mit den Trainingsdaten), die Einbettung in das Technologie-Ökosystem (bspw. die Dokumentation darüber, wo Daten gespeichert werden sollen) und den Umgang mit Daten bei dem Betrieb von KI (bspw. die Dokumentation über die Verwaltung von Zugangs- und Zugriffsberechtigungen) beschreiben, um Fehler bzw. Ausfälle von KI-Anwendungen zu vermeiden oder möglichst sicher mit ihnen umgehen zu können (vgl. Poretschkin et al. 2021, S. 126ff.).

Zusammenhalt

Für die Kategorie *Zusammenhalt* wurden keine Codes vergeben, d. h. das Ethik-Prinzip wird in keiner der inkludierten Leitlinien thematisiert.

Umsetzung des Zusammenhalts

Für die Kategorie *Umsetzung des Zusammenhalts* wurden keine Codes vergeben.

Achtung der Menschenrechte

Das Ethik-Prinzip *Achtung der Menschenrechte* wird in allen Leitlinien, in denen hierfür Codes vergeben wurden, d. h. den Leitlinien der Bundesländer Baden-Württemberg, Niedersachsen, Schleswig-Holstein und Thüringen, konsistent als die Herausforderung der Vereinbarkeit der Entwicklungen im Zusammenhang mit KI und der Berücksichtigung der Menschen- und Grundrechte interpretiert (vgl. Landesregierung BW 2018b, S. 4; vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 11; vgl. Stk SH 2021, S. 8; vgl. TMWWDG 2021, S. 62). Auch die hiermit thematisierte Problematik, die Missachtung der Menschen- und Grundrechte im Zusammenhang mit KI, ist für diese Leitlinien konsistent.

Umsetzung der Achtung der Menschenrechte

Für die Kategorie *Umsetzung der Achtung der Menschenrechte* wurden keine Codes vergeben.

Nachhaltigkeit

Das Ethik-Prinzip *Nachhaltigkeit* wird in den Leitlinien der Bundesländer Brandenburg, Bremen, Hamburg, Hessen, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Sachsen, Sachsen-Anhalt, Schleswig-Holstein und Thüringen thematisiert. Größtenteils werden in Bezug auf das Ethik-Prinzip lediglich die Stichwörter des Klima- und Umweltschutzes oder der Nachhaltigkeit genannt (vgl. MWAE 2022, S. 41; vgl. Landesregierung Bremen 2020, S. 14; vgl. Senatskanzlei Hamburg 2020, S. 54; vgl. EM MV 2021, S. 68; vgl. Poretschkin et al. 2021, S. 8; vgl. MW LSA 2021, S. 57). In den weiteren vier Leitlinien, die auf dieses Ethik-Prinzip eingehen, werden darüber hinaus auch konkrete Bereiche genannt, in denen KI zur Reduktion des Ressourcenverbrauchs beitragen kann, wie bspw. bei der Planung ressourcenschonender Produktionsprozesse (vgl. Stk Hessen 2021, S. 2; vgl. SK Sachsen 2021, S. 4; vgl. Stk SH 2021, S. 3; vgl. TMWWDG 2021, S. 62).

Umsetzung der Nachhaltigkeit

Schleswig-Holstein nennt als einziges Bundesland eine Maßnahme zur Umsetzung der Nachhaltigkeit von bzw. durch KI. Nämlich die Akzentuierung des Themas Nachhaltigkeit in der Klimaforschung, in Bezug auf die Energiewende und in der Landwirtschaft (vgl. Stk SH 2021, S. 16).

Vertrauen und Verlässlichkeit

Das Ethik-Prinzip *Vertrauen und Verlässlichkeit* wird in den Leitlinien Brandenburgs, Hessens, Niedersachsens, Nordrhein-Westfalens, Sachsens und Schleswig-Holsteins thematisiert. Insgesamt zielen alle Leitlinien hierbei auf die Problematik fehlenden Vertrauens gegenüber KI ab. In den Leitlinien von Niedersachsen und Sachsen wird das Ethik-Prinzip daher als Importanz der Stärkung dieses Vertrauens interpretiert (vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 11; vgl. SK Sachsen 2021, S. 9). Laut den Leitlinien von Brandenburg und Hessen ist das Ethik-Prinzip zentral, da KI-Wertschöpfung überhaupt erst möglich sei, wenn Menschen Vertrauen in KI haben (vgl. MWAE 2022, S. 37; vgl. Stk Hessen 2021, S. 2). Auch laut Nordrhein-Westfalen und Schleswig-Holstein ist die Vertrauenswürdigkeit von KI zentral, wobei Schleswig-Holstein hierbei insbesondere die Relevanz der technischen Robustheit und Nordrhein-Westfalen zusätzlich die Zuverlässigkeit von KI betont (vgl. Stk SH 2021, S. 11; vgl. Poretschkin et al. 2021, S. 11).

Umsetzung des Vertrauens und der Verlässlichkeit

Laut Brandenburg muss zur Umsetzung des Vertrauens und der Verlässlichkeit eine verantwortungsvolle Technikgestaltung und die Zertifizierung von Qualitätsmerkmalen und Ethikstandards durchgeführt werden, mit der bereits Forscher*innen und Entwickler*innen beauftragt worden sind (vgl. MWAE 2022, S. 11). Hessen schlägt die Stärkung der (interdisziplinären) Forschung und Technikreflexion sowie die Möglichkeit der freien Wahl von KI-Anbieter*innen als sinnvolle Maßnahmen zur Stärkung des Vertrauens von Bürger*innen in KI und der Position Hessens im internationalen Wettbewerb vor (vgl. Stk Hessen 2021, S. 3f.). Die Leitlinie Nordrhein-Westfalens, als ‚Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz‘, nennt verschiedene Maßnahmen in drei Bereichen: 1) Daten (bspw. die Dokumentation des Umgangs mit Unsicherheiten in Datensätzen), 2) KI-Komponente (bspw. die Wahl einer geeigneten Methode zur Unsicherheitsbewertung), 3) Einbettung (bspw. die Überprüfung der Auswirkungen der Unsicherheiten), die insbesondere auf die Robustheit und

Verlässlichkeit von KI-Anwendungen abzielen (vgl. Poretschkin et al. 2021, S. 1). In der Leitlinie Sachsens wird besonders deutlich, dass die Umsetzung des Vertrauens und der Verlässlichkeit die Erfüllung weiterer Ethik-Prinzipien voraussetzt: „Dazu muss KI entwickelt und angewendet werden, die auf den in der EU geltenden Standards wie Datenschutz, Datensicherheit, Gleichberechtigung, Diskriminierungsfreiheit, Diversität und Teilhabe beruht“ (SK Sachsen 2021, S. 9).

Freiheit und Autonomie

Die Bundesländer Baden-Württemberg, Brandenburg, Hessen und Nordrhein-Westfalen thematisieren das Ethik-Prinzip *Freiheit und Autonomie* in ihren Leitlinien. In den Leitlinien der Bundesländer Baden-Württemberg, Brandenburg und Hessen wird mit dem Ethik-Prinzip die Problematik thematisiert, dass technologische Abhängigkeiten, insbesondere von Ländern mit anderen Wertesystemen, vermieden werden müssen, sodass das Ethik-Prinzip als Freiheit und Autonomie der KI-Systeme interpretiert wird (vgl. Landesregierung BW 2018b, S. 4; vgl. MWAE 2022, S. 4; vgl. Stk Hessen 2021, S. 4). In der Leitlinie Hessens wird mit dem Ethik-Prinzip darüber hinaus die Problematik des fehlenden Schutzes der Freiheit und die eingeschränkte digitale Selbstbestimmung Einzelner und von Unternehmen thematisiert und es wird damit als die Notwendigkeit der Freiheit Einzelner interpretiert (vgl. MWAE 2022, S. 41). Nordrhein-Westfalen geht auf beide Aspekte ein und interpretiert es als die Autonomie sowohl des Menschen als auch der KI-Anwendung (vgl. Poretschkin et al. 2021, S. 24).

Umsetzung der Freiheit und Autonomie

In der Leitlinie Hessens werden verpflichtende Zulassungs- und Prüfverfahren, (freiwillige) Zertifizierungs- oder Normierungsansätze, Kennzeichnungs- und Offenlegungspflichten sowie der Dialog mit Bürger*innen, Wirtschaftseinrichtungen und Unternehmen als Maßnahmen zur Umsetzung des Ethik-Prinzips vorgeschlagen (vgl. Stk Hessen 2021, S. 3). Niedersachsen setzt auf das europäische Projekt GAIA-X, um *Freiheit und Autonomie* als eine der Dimensionen vertrauenswürdiger KI, wie sie von der EU angestrebt wird, umzusetzen (vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 16f.). Nordrhein-Westfalen schlägt für die Umsetzung dieses Ethik-Prinzips dieselben Maßnahmen wie bei der Kategorie *Umsetzung der menschlichen Kontrolle und menschlicher Singularität* vor, da die beiden Ethik-Prinzipien in der Leitlinie als eine Zieldimension vertrauenswürdiger KI verstanden werden (vgl. Poretschkin et al. 2021, S. 53ff.).

Gemeinwohl und Schadensvermeidung

Das Ethik-Prinzip *Gemeinwohl und Schadensvermeidung* wird in allen Leitlinien, in denen hierzu Codes vergeben wurden, d. h. Baden-Württemberg, Berlin, Mecklenburg-Vorpommern und Schleswig-Holstein, als die Wichtigkeit einer Entwicklung einer KI, die dem Gemeinwohl dient, interpretiert und zielt auf die Problematik negativer Folgen für die Gesellschaft durch KI ab (vgl. Landesregierung BW 2018b, S. 4; vgl. SenWiEnBe 2021, S. 58; vgl. EM MV 2021, S. 6; vgl. Stk SH 2021, S. 21).

Umsetzung des Gemeinwohls und der Schadensvermeidung

Brandenburg und Schleswig-Holstein schlagen Maßnahmen zur Umsetzung des Ethik-Prinzips vor. Brandenburg setzt hierbei auf die Berücksichtigung von Kriterien für gute Arbeit im KI-Design im Zuge des Change-Managements, um „[...] Arbeitsbedingungen zu verbessern, psychische und physische Arbeitsbelastungen zu reduzieren oder auch redundante Tätigkeitsinhalte abzulösen“ (MWAE 2022, S. 57). Schleswig-Holstein sieht hingegen die Beratung bezüglich des gemeinwohlverträglichen Einsatzes von KI durch Expert*innen vor (vgl. Stk SH 2021, S. 21).

Datenschutz und Privatsphäre

Thematisiert wird das Ethik-Prinzip *Datenschutz und Privatsphäre* in den Leitlinien von Berlin, Brandenburg, Hessen, Niedersachsen, Nordrhein-Westfalen, Sachsen, Sachsen-Anhalt, Schleswig-Holstein und Thüringen. In den Leitlinien der Bundesländer Berlin, Brandenburg, Hessen und Sachsen werden mit dem Ethik-Prinzip Problematiken in den Bereichen Informationssicherheit und Datenschutz thematisiert (vgl. SenWiEnBe 2021, S. 39; vgl. Stk Hessen 2021, S. 3; vgl. SK Sachsen 2021, S. 31). In der Leitlinie Niedersachsens werden hingegen Problematiken der Privatsphäre oder Privacy-by-Design genannt. Nordrhein-Westfalen, Sachsen-Anhalt und Schleswig-Holstein gehen auf beide genannten Problematiken ein (vgl. Poretschkin et al. 2021, S. 14, S. 142f.; vgl. MW LSA 2021, S. 55; vgl. Stk SH 2021, S. 11). Insgesamt wird das Ethik-Prinzip jedoch in allen diesen Leitlinien als der Schutz sensibler Daten bei der Entwicklung und dem Betrieb von KI interpretiert. Hingegen werden Datenschutz und Privatsphäre in der Leitlinie Thüringens in Hinblick auf die hierdurch entstehenden Hemmnisse für die wirtschaftliche Entwicklung von KI thematisiert (vgl. TMWWDG 2021, S. 11).

Umsetzung des Datenschutzes und der Privatsphäre

Zur Umsetzung des Datenschutzes und der Privatsphäre werden in der Leitlinie Nordrhein-Westfalens Maßnahmen für vier Risiko-Gebiete vorgeschlagen: 1) Daten (bspw. deren Anonymisierung oder Pseudonymisierung), 2) KI-Komponente (bspw. die Sparsamkeit mit Daten bei der Modellbildung), 3) Einbettung (bspw. Maßnahmen zur Verhinderung der Ausspähung von Daten), 4) Maßnahmen für den Betrieb von KI (bspw. Vorgaben zur Speicherung und Löschung von Daten) (vgl. Poretschkin et al. 2021, S. 148ff.). Sachsen-Anhalt nennt als Maßnahmen die Implementierung einheitlicher Mindestsicherheitsstandards für die Informationssicherheit zwischen beteiligten Stellen der Landesverwaltung Sachsen-Anhalts, die enge Zusammenarbeit dieser Stellen sowie die Einrichtung eines ressortübergreifenden Informationssicherheitsmanagementsystems (vgl. MW LSA 2021, S. 56). Schleswig-Holstein und Thüringen bringen in ihren Leitlinien an, dass sie innerhalb des europäischen Binnenmarktes der Daten agieren wollen, um die europäischen Standards für Datenschutz umzusetzen (vgl. Stk SH 2021, S. 22; vgl. TMWWDG 2021, S. 11).

Verantwortung und Rechenschaftspflicht

Dieses Ethik-Prinzip wird in den Leitlinien Brandenburgs, Hessens, Niedersachsens, Nordrhein-Westfalens und Schleswig-Holsteins als Haftung für KI interpretiert. Hessen und Niedersachsen nennen hierbei die Problematik fehlender „[...] Haftungsregeln und Versicherungsmechanismen für KI-basierte Dienste und Produkte [...]“ (Stk Hessen 2021, S. 3; vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 11). Nordrhein-Westfalen sieht Herausforderungen in Bezug auf die Verantwortung für die Qualität von KI ebenso wie die Rechenschaftspflicht, die auch Schleswig-Holstein betont (vgl. Poretschkin et al. 2021, S. 11; vgl. Stk SH 2021, S. 21). Im Gegensatz zu diesen Interpretationen des Ethik-Prinzips als das verantwortungsvolle Handeln der die KI konzipierenden, entwickelnden und implementierenden Personen, interpretieren Sachsen und Thüringen *Verantwortung und Rechenschaftspflicht* in Bezug auf die KI-Komponente und sprechen damit von verantwortungsbewusster KI (vgl. SK Sachsen 2021, S. 14; vgl. TMWWDG 2021, S. 43).

Umsetzung der Verantwortung und Rechenschaftspflicht

In der Leitlinie Niedersachsens wird eine Maßnahme zur Umsetzung der Verantwortung und Rechenschaftspflicht genannt. Diese wurde bereits zu der Kategorie *Umsetzung der Sicherheit* vorgestellt und zielt auf die stärkere Verankerung der

Datenverantwortung in der niedersächsischen Landesverwaltung bzw. bei der Landesbeauftragten für den Datenschutz ab (vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 16).

Fairness und Nicht-Diskriminierung

Das Ethik-Prinzip *Fairness und Nicht-Diskriminierung* wird in den Leitlinien von Bayern, Berlin, Brandenburg, Bremen, Hamburg, Niedersachsen, Nordrhein-Westfalen, Sachsen, Sachsen-Anhalt und Schleswig-Holstein genannt. Mit dem Ethik-Prinzip thematisieren diese Leitlinien die Problematik, dass Vorurteile und Diskriminierung durch KI reproduziert oder sogar verstärkt werden, insbesondere durch Ausschlussmechanismen und Hierarchien, die im Bereich der Digitalisierung vorherrschen. Berlin, Brandenburg, Bremen, Hamburg, Niedersachsen, Nordrhein-Westfalen, Sachsen, Sachsen-Anhalt und Schleswig-Holstein interpretieren das Ethik-Prinzip als die Gleichbehandlung, Antidiskriminierung und Fairness gegenüber Individuen im Veränderungsprozess um KI, mit besonderem Schwerpunkt auf den Digital Gap zwischen Frauen und Männer (vgl. SenWiEnBe 2021, S. 39; vgl. MWAE 2022, S. 1; vgl. Landesregierung Bremen 2020, S. 5; vgl. Senatskanzlei Hamburg 2020, S. 41; vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 16; vgl. Poretschkin et al. 2021, S. 23; SK Sachsen 2021, S. 9; vgl. MW LSA 2021, S. 57; vgl. Stk SH 2021, S. 11; vgl. Landesregierung Bremen 2020, S. 5). Bayern hingegen interpretiert das Ethik-Prinzip als die Gerechtigkeit bzw. das gleiche Profitieren von KI zwischen Stadt und Land (vgl. BStK 2020, S. 5).

Umsetzung der Fairness und Nicht-Diskriminierung

Konkrete Maßnahmen zur Umsetzung der Fairness und Nicht-Diskriminierung werden in den Leitlinien der Bundesländer Baden-Württemberg, Berlin, Bremen, Hamburg, Niedersachsen, Nordrhein-Westfalen, Sachsen-Anhalt und Schleswig-Holstein vorgeschlagen. In den Leitlinien der Bundesländer Baden-Württemberg, Berlin, Bremen und Hamburg wird die Förderung von Frauen in den Bereichen Aus- und Weiterbildung sowie bei der Stellenvergabe im Berufskontext als Maßnahme genannt (vgl. Landesregierung BW 2018b, S. 3; vgl. SenWiEnBe 2021, S. 58; vgl. Landesregierung Bremen 2020, S. 5; vgl. Senatskanzlei Hamburg 2020, S. 41). Berlin strebt zudem die Sensibilisierung der Öffentlichkeit und der Verwaltung sowie die Berücksichtigung der Risiken bei automatisierten Entscheidungsprozessen in der Verwaltung und dem nach außen gerichteten Handeln des Landes Berlin an, unter anderem durch den

rechtlichen Rahmen durch das Landesantidiskriminierungsgesetz (vgl. SenWiEnBe 2021, S. 39ff.). Bremen, Hamburg, Niedersachsen und Sachsen nennen als Maßnahme die Einsetzung heterogener KI-Gestaltungsteams (vgl. Landesregierung Bremen 2020, S. 5; vgl. Senatskanzlei Hamburg 2020, S. 41; vgl. Senatskanzlei Hamburg 2020, S. 41; vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 16; vgl. SK Sachsen 2021, S. 32). Hamburg nennt darüber hinaus die ausgewogene Besetzung der digitalisierungsrelevanten Gremien, das Monitoring und die Evaluation automatisierter Entscheidungssysteme, in denen eine Diskriminierung nach Geschlecht erwartet werden kann, sowie den Austausch zwischen Senats- und Fachbehörden (vgl. Senatskanzlei Hamburg 2020, S. 41). Niedersachsen nennt als Maßnahme die Unterstützung der europäischen Projekte GAIA-X und Europäische Datenräume (vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 16). Auch für dieses Ethik-Prinzip werden in der Leitlinie Nordrhein-Westfalens Maßnahmen für vier Risiko-Gebiete vorgeschlagen: 1) Daten (bspw. die Überprüfung der Daten auf Bias-Freiheit), 2) KI-Komponente (bspw. die faire Modellbildung), 3) Einbettung (bspw. die faire Weiterverarbeitung der Daten) und 4) Maßnahmen für den Betrieb (bspw. Kontrolle der Ausgaben im Betrieb) (vgl. Poretschkin et al. 2021, S. 42ff.). Sachsen-Anhalt nennt den Kompetenzaufbau und die Sensibilisierung der Bevölkerung, die Verbreitung von Best Practice für Daten, Infrastruktur, Software und Hardware sowie die barrierefreie Gestaltung der Websites öffentlicher Stellen und deren Überprüfung als geplante Maßnahmen (vgl. MW LSA 2021, S. 56). Zuletzt lassen sich für Schleswig-Holstein die Sicherstellung der Diskriminierungsfreiheit bei der Auswahl von Testdaten, die Änderungen im Rechtsrahmen der Zulassung von automatisierten Entscheidungen und die Folgeabschätzungen von KI-Anwendungen als vorgeschlagene Maßnahmen festhalten (vgl. Stk SH 2021, S. 15).

Transparenz und Nachvollziehbarkeit

Dieses Ethik-Prinzip wird in den Leitlinien von Brandenburg, Bremen, Hamburg, Hessen, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Sachsen und Schleswig-Holstein thematisiert. Insgesamt lässt sich eine für alle diese Bundesländer konsistente Interpretation, als die Transparenz bei der Konzeption, Entwicklung, Implementierung und den möglichen Folgen von KI, festhalten. Hierbei thematisiert Hessen mit dem Ethik-Prinzip die Problematik bestehender Vorbehalte gegenüber KI, die durch deren transparente Gestaltung ausgeräumt werden sollen (vgl. Stk Hessen 2021, S. 3). In anderen Leitlinien wird die Problematik fehlender Nachvollziehbarkeit bzw.

Verständlichkeit und Erklärbarkeit von KI genannt (vgl. MWAE 2022, S. 11; vgl. Senatskanzlei Hamburg 2020, S. 57; vgl. Poretschkin et al. 2021, S. 24; vgl. SK Sachsen 2021, S. 31; vgl. Stk SH 2021, S. 11).

Umsetzung der Transparenz und Nachvollziehbarkeit

Berlin, Nordrhein-Westfalen und Sachsen nennen Maßnahmen zur Umsetzung der Transparenz und Nachvollziehbarkeit. Berlin nennt die Möglichkeit, dass sich Bürger*innen über ihre Datenschutz- und Informationsfreiheitsrechte bei der Beauftragten für Datenschutz informieren können (vgl. SenWiEnBe 2021, S. 39). In der Leitlinie Nordrhein-Westfalens werden auch hier wieder mehrere Maßnahmen für die verschiedenen Risikobereiche genannt: 1) Daten (bspw. Trainings- und Testdaten für Nutzer*innen und Betroffene zugänglich und interpretierbar machen), 2) KI-Komponente (bspw. die Interpretierbarkeit des ML-Modells gewährleisten), 3) Einbettung (bspw. die Begründungen von Entscheidungen kommunizieren) und 4) Maßnahmen für den Betrieb (bspw. einen Prozess zur Beantwortung von Benutzeranfragen einführen) (vgl. Poretschkin et al. 2021, S. 68ff.). Zuletzt können für Sachsen Open-Source-Entwicklungen der eingesetzten Algorithmen und Daten sowie die Schaffung von Wissensangeboten zur Umsetzung der Transparenz und Nachvollziehbarkeit genannt werden, die zum Ziel haben, dass Verbraucher*innen gegenüber KI-Anwendenden, Unternehmen und Verwaltungen ihre Rechte und Optionen kennen (vgl. SK Sachsen 2021, S. 31ff.).

5. Diskussion

In der vorliegenden Arbeit wurde herausgearbeitet, dass die inkludierten Leitlinien der Bundesländer von unterschiedlichen Akteur*innen, in unterschiedlichen Situationen erarbeitet wurden und auf unterschiedlichen Definitionen von KI basieren. Hierdurch entstehen Deutungsspielräume in Bezug auf das Konzept ethischer KI, die sich in den Differenzen zwischen den Bundesländern äußern. Die beiden Forschungsfragen, die nach den spezifischen Differenzen zwischen den Leitlinien der Bundesländer zum einen in Bezug auf die Ethik-Prinzipien und zum anderen in Bezug auf die vorgeschlagenen Maßnahmen für deren Umsetzung fragen, wurden anhand der Deskription der Ergebnisse der qualitativen Inhaltsanalyse in Unterkapitel 4.2 beantwortet.

In diesem Kapitel werden die Ergebnisse der qualitativen Inhaltsanalyse sowie die Methodik diskutiert. Für deren Diskussion wird in diesem Kapitel auf die quantitativen Auswertungen aus Unterkapitel 4.1, den Forschungsstand aus Kapitel 2 und das Kommunikationsmodell aus Unterkapitel 3.3 zurückgegriffen.

5.1.Diskussion der Ergebnisse

Es lässt sich zunächst generell festhalten, dass in mindestens der Hälfte der Leitlinien aller Bundesländer die Ethik-Prinzipien *Sicherheit, Nachhaltigkeit, Fairness und Nicht-Diskriminierung, Datenschutz und Privatsphäre, Transparenz und Nachvollziehbarkeit, Zusammenarbeit* sowie *Menschliche Kontrolle und Menschliche Singularität* (Aufzählung in Reihenfolge der Häufigkeit) genannt werden. Somit kann von einer Konvergenz um diese sieben Ethik-Prinzipien in den analysierten Leitlinien gesprochen werden. Die folgende Diskussion wird sich daher vorwiegend auf diese sieben Ethik-Prinzipien beziehen.

Differierende Interpretationen

Für die Ethik-Prinzipien *Datenschutz und Privatsphäre, Fairness und Nicht-Diskriminierung* sowie *Menschliche Kontrolle und Menschliche Singularität* wurden klare Differenzen bei den Interpretationen der verschiedenen Bundesländer identifiziert.

So wird das Ethik-Prinzip *Datenschutz und Privatsphäre* zwar von der Mehrheit der Bundesländer, die es thematisieren, als die Thematik des Schutzes sensibler Daten bei der Konzeption, Entwicklung und Implementierung von KI interpretiert. Die Interpretation des Ethik-Prinzips in der Leitlinie Thüringens weicht hiervon jedoch ab, da *Datenschutz und Privatsphäre* hier als Hemmnis für die wirtschaftliche Entwicklung verstanden wird. Betrachtet man das Kommunikationsmodell, so fällt auf, dass wirtschaftliche Akteur*innen an der Erarbeitung der Leitlinie Thüringens mitgewirkt haben, was einen möglichen Erklärungsansatz für diese abweichende Interpretation darstellt.

Auch für das Ethik-Prinzip *Fairness und Nicht-Diskriminierung* wurden zwei unterschiedliche Interpretationen festgehalten. Von der Mehrheit der Bundesländer, die das Ethik-Prinzip thematisieren, wird es als Gleichbehandlung, Antidiskriminierung und Fairness gegenüber Individuen im Prozess der Konzeption, Entwicklung, Implementierung und auch bei den möglichen Folgen von KI interpretiert, während Bayern es hingegen als die Gerechtigkeit bzw. das gleiche Profitieren von KI zwischen Stadt und Land thematisiert. Eine mögliche Erklärung für diese abweichende Interpretation Bayerns könnte die starke Wettbewerbsorientierung des Bundeslandes, insbesondere gegenüber Baden-Württemberg, als „[...] stärkster Mitbewerber [...]“, sein (BStK 2019, S. 6). Da der Einwohneranteil im ländlichen Raum in Bayern mit 56% erheblich höher ist als in Baden-Württemberg mit 34%, liegt es nahe, dass sich der Fokus Bayerns auf den ländlichen Raum aus der Wettbewerbsorientierung des Bundeslandes ergibt, die auch bereits in Unterkapitel 2.3 erwähnt wurde (vgl. Staatsministerium Baden-

Württemberg o. J., o. S.; vgl. Bayerisches Staatsministerium der Finanzen und für Heimat 2020, S. 12).

Zuletzt lassen sich auch für das Ethik-Prinzip *Menschliche Kontrolle und Menschliche Singularität* zwei unterschiedliche Interpretationen identifizieren. Zum einen wird darunter die Importanz des Menschen als Mittelpunkt des Veränderungsprozesses durch KI verstanden, zum anderen die Notwendigkeit menschlicher Kontrolle bzw. Aufsicht von KI. Nordrhein-Westfalen und Schleswig-Holstein nennen hingegen beide Interpretationen, sodass hier die überdurchschnittliche Auseinandersetzung Nordrhein-Westfalens, aber auch Schleswig-Holsteins mit dem Thema der KI-Governance bzw. ethischer KI betont werden soll. An dieser Stelle soll daher auch nochmal darauf hingewiesen werden, dass die Leitlinie Nordrhein-Westfalens, im Auftrag des Bundeslandes, durch die Fraunhofer-Gesellschaft erarbeitet wurde und daher anzunehmen ist, dass sie auf der Grundlage eines umfangreicheren Forschungsstandes im Bereich KI als die anderen inkludierten Leitlinien erarbeitet wurde, und damit vermutlich auch speziell in Bezug auf den Ethik-Prinzipien-Ansatz erstellt wurde. Somit kann auch die überdurchschnittlich hohe Anzahl vergebener Codes für Nordrhein-Westfalen, wie sie Tabelle 3 entnommen werden kann, begründet werden. Insgesamt kann festgehalten werden, dass sich die Differenzen der Interpretationen der Ethik-Prinzipien größtenteils auf die Differenzen der die Leitlinien erarbeitenden Akteur*innen zurückführen lassen.

Entsprechend der differierenden Interpretationen unterscheiden sich bei den Ethik-Prinzipien *Datenschutz und Privatsphäre, Fairness und Nicht-Diskriminierung* sowie *Menschliche Kontrolle und Menschliche Singularität* auch die zu deren Umsetzung vorgeschlagenen Maßnahmen der Bundesländer. Eine ausführliche Diskussion der zur Umsetzung der Ethik-Prinzipien vorgeschlagenen Maßnahmen folgt am Ende dieses Unterkapitels.

Thematisierung unterschiedlicher Problematiken

Die Interpretationen der Ethik-Prinzipien *Sicherheit* sowie *Transparenz und Nachvollziehbarkeit* sind für alle Bundesländer, die die Ethik-Prinzipien thematisieren konsistent, wobei jedoch Differenzen in Bezug auf die mit dem jeweiligen Ethik-Prinzip thematisierten Problematiken bestehen.

Für *Sicherheit* lässt sich allgemein festhalten, dass das Ethik-Prinzip auf der Grundlage der quantitativen Auswertung zwar sehr relevant zu sein scheint, jedoch in den Leitlinien nur oberflächlich durch die Nennung einzelner Stichworte (bspw. Sicherheit,

Cybersicherheit etc.) behandelt wird, wobei diese alle auf die Sicherheit der Konzeption, Entwicklung, Implementierung und möglichen Folgen von KI abzielen (vgl. Tab. 4). Differenzen zwischen den Bundesländern werden in Bezug auf die mit dem Ethik-Prinzip thematisierten Problematiken sichtbar. Während einige Bundesländer die Gefahr von Cyberangriffen für KI-Systeme nennen, nennen andere Bundesländer lediglich die Problematik fehlender Sicherheit allgemein. Nordrhein-Westfalen nennt als einziges Bundesland mehrere Problemdimensionen der Sicherheit in Bezug auf KI. Auch hier ist wieder auf die Rolle der Fraunhofer-Gesellschaft für den Inhalt der Leitlinie hinzuweisen.

Auch für das Ethik-Prinzip *Transparenz und Nachvollziehbarkeit* ist eine konsistente Interpretation, als Transparenz bei der Konzeption, Entwicklung, Implementierung und den möglichen Folgen von KI, festzuhalten, wobei auch hier dennoch verschiedene Problematiken in unterschiedlichen Leitlinien thematisiert werden. Die Mehrheit der Bundesländer nennt die Problematik fehlender Nachvollziehbarkeit bzw. Verständlichkeit und Erklärbarkeit von KI. Hessen hingegen nennt die Problematik bestehender Vorbehalte gegenüber KI, die durch deren transparente Gestaltung ausgeräumt werden soll. Diese Abweichung Hessens von der Mehrheit der Bundesländer lässt sich vermutlich auf die Erarbeitungsweise Hessens Leitlinie zurückführen, da sie auf den Einschätzungen des hessischen Rats für Digitalethik beruht und aus Sicht der Bürger*innen erstellt wurde, da diese die größten Vorbehalte gegenüber KI haben (vgl. Stk Hessen 2021, S. 3).

Insgesamt können die Ausführungen in Unterkapitel 2.3 einen Erklärungsansatz dafür liefern, dass sich die mit den Ethik-Prinzipien thematisierten Problematiken, trotz gleicher Interpretation, unterscheiden. Da die Leitlinien der Bundesländer Themen, Ausdrücke und Rhetoriken der KI-Strategie der Bundesregierung übernehmen und diese dann an ihre jeweilige regionale Identität anpassen, werden die Ethik-Prinzipien in den Leitlinien ähnlich interpretiert, wobei sich aufgrund der unterschiedlichen Wirtschafts- und Standortpolitik Differenzen insbesondere in Bezug auf die thematisierten Problematiken und zur Umsetzung vorgeschlagenen Maßnahmen ergeben. Dementsprechend legt bspw. Baden-Württemberg den Fokus der mit den Ethik-Prinzipien thematisierten Problematiken und der vorgeschlagenen Maßnahmen auf Forschungsanliegen und Forschungsvorhaben in Anschluss an das Forschungskonsortium Cyber Valley als zentrale Anlaufstelle für den Bereich KI in Baden-Württemberg. Auch hier tragen daher die differierenden, die Leitlinien erarbeitenden Akteur*innen, maßgeblich zu den Differenzen bei.

Konsens

Für die Ethik-Prinzipien *Nachhaltigkeit* und *Zusammenarbeit* sind jeweils über alle Bundesländer, die die Ethik-Prinzipien thematisieren, hinweg konsistente Interpretationen und dieselben Problematiken genannt worden.

Im Fall des Ethik-Prinzips *Zusammenarbeit* könnte dies auf die, im Vergleich mit anderen Ethik-Prinzipien, relativ geringe theoretische Abstraktheit zurückgeführt werden, da Zusammenarbeit bzw. Kooperation ein praxisnahes, bekanntes (Ethik-)Prinzip ist.

Im Fall des Ethik-Prinzips *Nachhaltigkeit* hingegen wurde genau eine solche theoretische Abstraktheit daran festgestellt, dass die Leitlinien nur selten über die Nennung der Stichworte Nachhaltigkeit, Klima- bzw. Umweltschutz und Ressourcenschonung hinausgehen. Das Ausbleiben von Differenzen lässt sich in diesem Fall daher darauf zurückführen, dass diese Abstraktheit bzw. Knappheit die Feststellung von Differenzen kaum zulässt.

Es schlägt jeweils nur ein Bundesland Maßnahmen zur Umsetzung dieser Ethik-Prinzipien vor. Somit kann keine Aussage darüber getroffen werden, ob Konsens in Bezug auf Ethik-Prinzipien auch Konsens in Bezug auf zur Umsetzung vorgeschlagene Maßnahmen garantiert.

Umsetzung der Ethik-Prinzipien

Bei den vorgeschlagenen Maßnahmen zur Umsetzung der Ethik-Prinzipien werden die Differenzen zwischen den Bundesländern besonders deutlich, was wie zuvor beschrieben größtenteils auf die Differenzen der Wirtschafts- und Standortpolitik der Bundesländer sowie die die Leitlinien erarbeitenden Akteur*innen zurückgeführt werden kann. Die Differenzen werden zum einen anhand der differierenden Ziele und Adressat*innen der Maßnahmen deutlich, zum anderen aber auch insbesondere in Bezug darauf, ob überhaupt Maßnahmen, Ziele oder Adressat*innen in den Leitlinien genannt werden.

Generell konnte festgestellt werden, dass selten spezifische Ziele von Maßnahmen genannt werden, sondern, dass die vorgeschlagenen Maßnahmen generell dem Ziel dienen, das jeweilige Ethik-Prinzip, der jeweiligen Interpretation entsprechend, umzusetzen. Damit können die differierenden Maßnahmen in Folge der differierenden Interpretationen von Ethik-Prinzipien erklärt werden.

Es fällt des Weiteren auf, dass insgesamt nur wenige Maßnahmen im Vergleich zu den vergebenen Codes für die Ethik-Prinzipien genannt werden. Damit scheint das

Bedenken zumindest verhärtet, dass Ethik-Prinzipien teilweise abstrakte Konzepte darstellen, welche auf unterschiedliche Weise interpretiert und verwendet werden und denen eine Klarheit fehlt, die es zur Umsetzung in die Praxis braucht (vgl. Morley et al. 2021, S. 240). Dies ist bspw. auch daran zu erkennen, dass die Problematik, dass der Mensch aus dem Mittelpunkt des Veränderungsprozesses durch KI gerückt werden könnte, sowohl in Bezug auf das Ethik-Prinzip *Anerkennung gesellschaftlicher Grundwerte*, als auch in Bezug auf das Ethik-Prinzip *Menschliche Kontrolle und Menschliche Singularität* thematisiert wird. Da die Problematik zwei verschiedenen Ethik-Prinzipien zugeschrieben wird, scheint unklar zu sein, wie in der Praxis damit umgegangen werden sollte. Dies schlägt sich bei *Anerkennung gesellschaftlicher Grundwerte* durch das Ausbleiben von Maßnahmen zur Umsetzung nieder. Zu der Kategorie *Umsetzung der menschlichen Kontrolle und menschlichen Singularität* werden zwar Maßnahmen zur Einbindung der Bürger*innen in die Gestaltung des Veränderungsprozesses genannt, wobei aufgrund fehlender Ausführungen und theoretischer Unklarheit des Konzepts menschenzentrierter KI offenbleibt, ob diese tatsächlich ausreichend sind, um zu verhindern, dass die Menschen aus dem Zentrum des Veränderungsprozesses durch KI verdrängt werden (vgl. Kieslich et al. 2022, S. 3).

Andererseits zeigt sich jedoch am Beispiel der Ethik-Prinzipien *Nachhaltigkeit* und *Zusammenarbeit*, dass selbst wenn Konsens in Bezug auf die Interpretation und die mit dem Ethik-Prinzip thematisierten Problematik besteht, kaum Maßnahmen zur Umsetzung vorgeschlagen werden, sodass es fraglich bleibt, ob der Ethik-Prinzipien-Ansatz ohne Einbettung in ein tatsächliches Governance-Ökosystem überhaupt Wirkung entfalten kann (vgl. Fjeld et al. 2020, S. 5). Der folgende Befund verstärkt dieses Bedenken.

Sowohl Jobin et al. (2019) als auch Fjeld et al. (2020) identifizieren in ihren Beiträgen, in Übereinstimmung mit der vorliegenden Arbeit, ebenfalls eine Konvergenz um die Ethik-Prinzipien *Transparenz und Nachvollziehbarkeit*, *Fairness und Nicht-Diskriminierung* sowie *Datenschutz und Privatsphäre*. Fjeld et al. (2020) identifizieren zudem eine Konvergenz um das Ethik-Prinzip *Menschliche Kontrolle und Menschliche Singularität*. Darüber hinaus wird in beiden Beiträgen jedoch, im Gegensatz zu der vorliegenden Arbeit, auch eine Konvergenz um das Ethik-Prinzip *Verantwortung und Rechenschaftspflicht*, und bei Fjeld et al. (2020) des Weiteren um *Berufliche Verantwortung* sowie *Anerkennung gesellschaftlicher Grundwerte* festgehalten. Auffällig ist, dass das Ethik-Prinzip *Verantwortung und Rechenschaftspflicht* in den in der vorliegenden Arbeit analysierten Leitlinien jedoch, insbesondere auch mit Blick auf die

Ergebnisse der Befragung deutscher Bürger*innen von Kieslich et al. (2022), in der das Ethik-Prinzip im Durchschnitt als das wichtigste Ethik-Prinzip bewertet wurde, vergleichsweise selten kodiert wurde. Gleiches ist auch für das Ethik-Prinzip *Berufliche Verantwortung* festzuhalten (vgl. Tab. 1). Auch in Bezug auf die Maßnahmen zur Umsetzung aller in der qualitativen Inhaltsanalyse inkludierten Ethik-Prinzipien fällt auf, dass, wenn Maßnahmen genannt werden, sich diese insgesamt selten an eine bestimmte Gruppe oder Person richten, und noch seltener ausdrücklich Verantwortung für die Konzeption, Entwicklung, Implementierung oder die möglichen Folgen von KI an diese übertragen wird. Die Leitlinie Nordrhein-Westfalens stellt insofern eine Ausnahme von diesem Umstand dar, als dass sie sich ausdrücklich als Prüfkatalog an Prüfer*innen und Entwickler*innen von KI richtet und die Verantwortung für die vorgeschlagenen Maßnahmen daher auf diese spezifischen Personengruppen übertragen wird. Dies ist in den restlichen Leitlinien nicht der Fall, da sie sehr diffuse Zielgruppen wie die Wissenschaft, Wirtschaft oder Gesellschaft nennen, wie auch Unterkapitel 3.3 entnommen werden kann, die zudem nur selten über die zur Durchführung der vorgeschlagenen Maßnahmen benötigten Kompetenzen verfügen. Wie Unterkapitel 4.2 entnommen werden kann, sind für diese Leitlinien lediglich für vier Ethik-Prinzipien Maßnahmen festzuhalten, die explizit Verantwortung an eine bestimmte Person oder Gruppe übertragen: In Bezug auf die *Umsetzung der menschlichen Kontrolle und menschlichen Singularität* in der Leitlinie Hamburgs, zur Umsetzung von *Sicherheit* in der Leitlinie Niedersachsens, zur *Umsetzung des Vertrauens und der Verlässlichkeit* in der Leitlinie Brandenburgs und zur Umsetzung von *Verantwortung und Rechenschaftspflicht* in der Leitlinie Niedersachsens (vgl. Senatskanzlei Hamburg 2020, S. 37; vgl. MW Niedersachsen und MWK Niedersachsen 2021, S. 16; vgl. MWAE 2022, S. 11). Dies deckt sich demnach mit den Bedenken, die auch in Unterkapitel 2.2 zusammengefasst sind, nämlich dass der Ethik-Prinzipien-Ansatz Ethics Washing, d. h. den Missbrauch der Aufstellung von Ethik-Prinzipien zur Selbstverpflichtung als Ersatz für tatsächliche Regulierung, begünstigen könnte (vgl. Rességuier und Rodrigues 2020, S. 2). Verhärtet wird dieses Bedenken zudem in Hinblick auf den Befund, dass keine Leitlinie Maßnahmen zur Umsetzung von *Rechtmäßigkeit*, d. h. zur Einhaltung der geltenden Gesetze und Vorschriften bei der Konzeption, Entwicklung, Implementierung und den möglichen Folgen von KI, vorschlägt.

5.2.Diskussion der Methodik

In Bezug auf die Auswahl der Leitlinien der Bundesländer soll hier reflektiert werden, dass es sich bei den Leitlinien der Bundesländer um graue Literatur handelt und deren Erhebung daher deutlich weniger replizierbar ist, als bspw. die Erhebung peer-review-ter Beiträge aus wissenschaftlichen Datenbanken. Durch die Beschreibung des genauen Verfahrens deren Erhebung und des Verfahrens deren Auswahl in Unterkapitel 3.3, wurde jedoch versucht, dem entgegenzuwirken.

In Bezug auf die Methode der strukturierenden qualitativen Inhaltsanalyse nach Mayring ist zu bedenken, dass die klassischen Gütekriterien der Sozialwissenschaften (Reliabilität, Validität und Objektivität) bei qualitativen Inhaltsanalysen generell nur schwer zu erfüllen sind (vgl. Mayring 2015, S. 124f.). Reliabilität wird in qualitativen Inhaltsanalysen häufig durch Interkoderreliabilität, d. h. die Kodierung durch mehrere Personen und den Vergleich derer Ergebnisse, geprüft und Validität durch deren Zirkularität, d. h. das Heranziehen von Material außerhalb der eigenen Untersuchung als Gütemaßstab (vgl. Mayring 2015, S. 124f.). Diese Verfahren sind aufgrund des Umfangs der vorliegenden Arbeit nicht umsetzbar. Die typischen Grenzen der Inhaltsanalyse, d. h. die Schwierigkeit der Überprüfbarkeit und der Umsetzung der Reliabilität und Validität gelten damit auch für die vorliegende Arbeit. Durch die Benennung dieser Grenzen, durch das regelgeleitete Vorgehen und durch die Beschreibung der Ergebnisse der qualitativen Inhaltsanalyse soll die intersubjektive Nachvollziehbarkeit dennoch bestmöglich umgesetzt, und den Grenzen der Methode entgegengewirkt werden (vgl. Ruin 2019, S. 11).

Des Weiteren ergeben sich Schwierigkeiten aus dem Untersuchungsmaterial, d. h. aus den Leitlinien der Bundesländer, da die Vergleichbarkeit durch den unterschiedlichen Umfang sowie die Erarbeitung durch unterschiedliche Institutionen und Akteur*innen herabgesetzt wird. Dies lässt sich nicht umgehen und muss daher, wie im vorangegangenen Unterkapitel, mitreflektiert werden. Nicht zuletzt, da auch dieser Umstand für sich ein Ergebnis dieser Arbeit darstellt.

Zuletzt soll darauf eingegangen werden, dass das Kategoriensystem, das der qualitativen Analyse der Leitlinien zugrunde liegt, immer auch die mit dessen Hilfe erzielten Ergebnisse beeinflusst. In dieser Arbeit ist dies insbesondere zu beachten, da einige Ethik-Prinzipien, wie in Unterkapitel 3.2 beschrieben wird, zu einer Kategorie zusammengefasst wurden. Hierdurch können sich zusätzliche Deutungsspielräume innerhalb der Kategorien ergeben, was am Beispiel der Kategorie *Vertrauen und Verlässlichkeit* besonders gut nachvollziehbar ist. Durch die Zusammenfassung von *Trust* und

Robustness in dieser Kategorie fallen sowohl Themen des menschlichen Vertrauens in KI als auch der technischen Robustheit von KI in diese Kategorie und können zu der Feststellung von Differenzen zwischen Bundesländern führen. Ebenso zeigt sich am Beispiel von *Vertrauen und Verlässlichkeit* besonders gut, dass Ethik-Prinzipien sich häufig gegenseitig bedingen oder miteinander verflochten sind, wodurch sich Überschneidungen der Kategorien ergeben können. Zur Umsetzung von *Vertrauen und Verlässlichkeit* bspw. müssen laut der Leitlinie Sachsens zunächst die Ethik-Prinzipien *Datenschutz und Privatsphäre* sowie *Fairness und Nicht-Diskriminierung* realisiert werden (vgl. SK Sachsen 2021, S. 9). Wie jedoch Unterkapitel 4.2 am Beispiel des Ethik-Prinzips *Rechtmäßigkeit* zu entnehmen ist, bestehen wie vermutet auch bei nicht-zusammengefassten Ethik-Prinzipien-Kategorien Deutungsspielräume, woraus sich Differenzen zwischen den Leitlinien der Bundesländer ergeben. Während also mitreflektiert werden muss, dass sich festgestellte Differenzen eventuell zum Teil auf das Kategoriensystem zurückführen lassen könnten, muss dennoch festgehalten werden, dass es auf die Natur der Ethik-Prinzipien als abstrakte Konstrukte zurückzuführen ist, dass sich Deutungsspielräume überhaupt ergeben.

6. Fazit

Abschließend soll die Beantwortung der Forschungsfragen zusammenfassend dargelegt werden. Außerdem sollen die zentralen Ergebnisse dieser Arbeit dargestellt, und davon ausgehend Schlussfolgerungen und Fragen für zukünftige Forschung in diesem Gebiet aufgestellt werden.

Wie in Unterkapitel 5.2 dargelegt, muss hierbei die wissenschaftliche Güte der Ergebnisse aufgrund der fehlenden Reliabilitäts- und Validitätsprüfungen kritisch reflektiert werden. Wird dies reflektiert, so können sie dennoch als Ausgangspunkt für die Bildung eines Konsenses der Bundesländer bezüglich der Ethik-Prinzipien und Maßnahmen für deren Umsetzung bzw. für weitere Forschung mit diesem Ziel, dienen.

Die erste Forschungsfrage zielt auf die Differenzen zwischen den in den Leitlinien der Bundesländer vorgeschlagenen Ethik-Prinzipien für KI ab. Hierbei wurde zwischen den jeweiligen Interpretationen der Ethik-Prinzipien und den mit ihnen thematisierten Problematiken unterschieden. Betrachtet man bezüglich dieser ersten Forschungsfrage die sieben zentralen, zwischen den Bundesländern konvergierenden Ethik-Prinzipien *Sicherheit, Nachhaltigkeit, Fairness und Nicht-Diskriminierung, Datenschutz und Privatsphäre, Transparenz und Nachvollziehbarkeit, Zusammenarbeit* und *Menschliche Kontrolle und Menschliche Singularität*, so lässt sich festhalten, dass Differenzen bei

den Interpretationen von drei Ethik-Prinzipien, *Datenschutz und Privatsphäre, Fairness und Nicht-Diskriminierung* sowie *Menschliche Kontrolle und Menschliche Singularität*, bestehen. Für die Ethik-Prinzipien *Sicherheit* und *Transparenz und Nachvollziehbarkeit* sind die jeweiligen Interpretationen für alle Bundesländer konsistent, die mit ihnen thematisierten Problematiken unterscheiden sich jedoch. Bei den Ethik-Prinzipien *Nachhaltigkeit* und *Zusammenarbeit* wurde hingegen Konsens sowohl in Bezug auf die Interpretation als auch die thematisierten Problematiken identifiziert.

Betrachtet man diese Befunde in Hinblick auf die zweite Forschungsfrage, die nach den Differenzen zwischen den in den Leitlinien der Bundesländer vorgeschlagenen Maßnahmen zur Umsetzung der Ethik-Prinzipien fragt, so lässt sich festhalten, dass sowohl die unterschiedliche Interpretation von Ethik-Prinzipien als auch das Thematisieren unterschiedlicher Problematiken mit einem Ethik-Prinzip zu Differenzen der von den Bundesländern vorgeschlagenen Maßnahmen führen. Da für *Nachhaltigkeit* und *Zusammenarbeit* jeweils nur ein Bundesland Maßnahmen vorschlägt, lässt sich keine Aussage darüber treffen, ob Konsens in Bezug auf Ethik-Prinzipien zu Konsens in Bezug auf die zur Umsetzung vorgeschlagenen Maßnahmen garantiert.

Generell ist, im Vergleich zu der Anzahl vergebener Codes bezüglich der Ethik-Prinzipien, eine sehr geringe Anzahl vorgeschlagener Maßnahmen für deren Umsetzung festzuhalten.

Die zweite Forschungsfrage zielt darüber hinaus darauf ab, zu identifizieren, welche Differenzen zwischen den Bundesländern bestehen, in Bezug darauf, welches Ziel die jeweiligen Maßnahmen verfolgen und an wen sich die vorgeschlagenen Maßnahmen der Bundesländer richten. Hierzu lässt sich zusammenfassend festhalten, dass die vorgeschlagenen Maßnahmen der Bundesländer generell dem Ziel dienen, das jeweilige Ethik-Prinzip, der jeweiligen Interpretation entsprechend, umzusetzen. Da teilweise jedoch nur einzelne Stichworte in Bezug auf ein Ethik-Prinzip genannt werden, kann diese Frage nicht für alle Ethik-Prinzipien bzw. nicht für alle Bundesländer beantwortet werden. Des Weiteren kann festgehalten werden, dass generell selten Adressat*innen von Maßnahmen explizit genannt werden. Mit Blick auf das Kommunikationsmodell, dem die Zielgruppen der Leitlinien entnommen werden können, kann aufgezeigt werden, dass die Kompetenzen der Zielgruppen der Leitlinien selten mit den zur Durchführung der vorgeschlagenen Maßnahmen benötigten Kompetenzen übereinstimmen. Zudem werden hier nur sehr vage Gruppen, wie bspw. allgemein Wissenschaftler*innen genannt. Damit können auch die Zielgruppen der Leitlinien nicht als Adressat*innen der Maßnahmen gewertet werden. Ausgenommen hiervon ist lediglich

die Leitlinie Nordrhein-Westfalens, die sich an Entwickler*innen und Prüfer*innen von KI richtet. Dieses Ergebnis sollte in Hinblick auf den Befund betrachtet werden, dass das Ethik-Prinzip *Verantwortung und Rechenschaftspflicht* in der vorliegenden Arbeit im Vergleich mit den anderen Ethik-Prinzipien, aber auch im internationalen Vergleich mit anderen Forschungsbeiträgen, selten kodiert wurde. Das Bedenken von Ethics Washing durch den Ethik-Prinzipien-Ansatz verhärtet sich demnach, da die genannten Maßnahmen zur Umsetzung der Ethik-Prinzipien kaum an bestimmte Personen adressiert sind, damit also auch kaum Verantwortung für KI an Personen übergeben wird, und zudem keine Maßnahmen zur Umsetzung der *Rechtmäßigkeit* von KI vorgeschlagen werden.

Da nicht die KI selbst in die Verantwortung für deren ethische Konzeption, Entwicklung, Implementierung und Folgen gestellt werden kann, bleibt die Umsetzung ethischer KI fraglich, solange keine spezifischen Personen oder Personengruppen in der (formellen) Rechenschaftspflicht hierfür stehen. Ein Ansatz zur Überwindung dieses Problems stellt der Vorschlag der KI-Verordnung der Europäischen Kommission (2021) dar, die durch die Schaffung eines rechtlichen Rahmens zu einheitlichen Lösungen dieser Differenzen bzw. Lücken führen kann. Solange keine solche einheitliche Lösung tatsächlich beschlossen wird, d. h. die Differenzen der Leitlinien nicht überwunden werden, haben diese maßgebliche negative Konsequenzen für die Wirkung des Ethik-Prinzipien-Ansatzes.

Mit Blick auf diese Ergebnisse und den aktuellen Forschungsstand lässt sich daher festhalten, dass sich zukünftige Forschung, insbesondere auch auf deutscher Länder-ebene, verstärkt mit der spezifischen Umsetzung der Ethik-Prinzipien auseinandersetzen sollte. Fragen, die aus der vorliegenden Arbeit hervorgehen bzw. hier nicht beantwortet werden konnten, sind bspw. ‚Kann Konsens in Bezug auf die Interpretationen und die mit den Ethik-Prinzipien thematisierten Problematiken zu Konsens in Bezug auf die Maßnahmen für deren Umsetzung führen?‘, oder auch ‚Braucht es einen solchen Konsens in Bezug auf die Maßnahmen tatsächlich für die erfolgreiche Realisierung ethischer KI?‘.

Literaturverzeichnis

- Arksey, Hilary und Lisa O'Malley. 2005. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology* 8(1): S. 1-28. <https://eprints.whiterose.ac.uk/1618/1/Scopingstudies.pdf>. Zugegriffen: 02.07.2022.
- Axente, Maria Luciana. 2021. Ethical AI: 10 principles the world (mostly) agrees on — and what to do about them. <https://www.pwc.com/us/en/tech-effect/ai-analytics/how-to-make-ai-ethical.html>. Zugegriffen: 02.07.2022.
- Bayerisches Staatsministerium der Finanzen und für Heimat. 2020. HEIMATBERICHT 2020. Entwicklungen in Stadt und Land. https://www.stmfh.bayern.de/heimat/Heimatbericht_2020.pdf. Zugegriffen: 02.07.2022.
- Bundesregierung KI. O. J. KI-Strategien von Bund und Ländern. <https://www.ki-strategie-deutschland.de/home.html>. Zugegriffen: 02.07.2022.
- Bundesregierung. 2018. Eckpunkte der Bundesregierung für eine Strategie Künstliche Intelligenz. https://www.bmwi.de/Redaktion/DE/Downloads/E/eckpunktepapier-ki.pdf?__blob=publicationFile&v=10. Zugegriffen: 02.07.2022.
- Bundesregierung. 2020. Strategie Künstliche Intelligenz der Bundesregierung. Fortschreibung 2020. https://www.ki-strategie-deutschland.de/files/downloads/201201_Fortschreibung_KI-Strategie.pdf. Zugegriffen: 02.07.2022.
- Buxmann, Peter und Holger Schmidt. 2019. Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens. In *Künstliche Intelligenz. Mit Algorithmen zum wirtschaftlichen Erfolg*, hrsg. Peter Buxmann und Holger Schmidt, S. 3-19. Berlin: Springer Gabler.
- Denkfabrik des Bundesministeriums für Arbeit und Soziales. 2022. Netzwerk KI in der Arbeits- und Sozialverwaltung. Eine Wertegrundlage für den KI-Einsatz in der Arbeits- und Sozialverwaltung. <https://www.denkfabrik-bmas.de/projekte/ki-in-der-verwaltung/welche-werte-braucht-eine-moderne-verwaltung-bei-der-nutzung-von-ki>. Zugegriffen: 06.07.2022.
- Düwell, Marcus, Christoph Hübenthal, und Micha H. Werner. 2011. Einleitung. In *Handbuch Ethik*, hrsg. Marcus Düwell, Christoph Hübenthal und Micha H. Werner, S. 1-23. Stuttgart: J. B. Metzler.
- Europäische Kommission. 2018. Ethik-Leitlinien für eine vertrauenswürdige KI. <https://op.europa.eu/de/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-de/format-PDF>. Zugegriffen: 02.07.2022.
- Europäische Kommission. 2021. Verordnung des Europäischen Parlaments und des Rates. Zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz (Gesetz über Künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206>. Zugegriffen: 06.07.2022.
- Fetic, Lajla, Torsten Fleischer, Paul Grünke, Thilo Hagendorf, Sebastian Hallensleben, Marc Hauer, Michael Herrmann, Rafaela Hillerbrand, Carla Hustedt, Christoph Hubig, Andreas Kaminski, Tobias Krafft, Wulf Loh, Philipp Otto und Michael Puntschuh. 2020. From Principles to Practice. An interdisciplinary framework to operationalise AI ethics. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/WKIO_2020_final.pdf. Zugegriffen: 08.07.2022.
- Floridi, Luciano und Josh Cowls. 2019. A Unied Framework of Five Principles for AI in Society. *Harvard Data Science Review* 2019(1.1): S. 1-17.
- Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Christopher Nagy und Madhulika Srikumar. 2020. Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI. The

- Berkman Klein Center for Internet & Society Research Publication No. 2020-1: S. 1-71.
- Goldmann, Anne und Christoph Bieber. 2021. KI-Governance - Künstliche Intelligenz als Gegenstand politischer Steuerung. <https://www.cais.nrw/ki-gov/>. Zugegriffen: 02.07.2022.
- Hoffmann, Christian Hugo. 2022. A philosophical view on singularity and strong AI. *AI & SOCIETY* (2022): S. 1-18.
- Jobin, Anna, Licinia Güttel, Laura Liebig und Christian Katzenbach. 2021. AI Federalism: Shaping AI Policy within States in Germany. <http://export.arxiv.org/abs/2111.04454>. Zugegriffen: 02.07.2022.
- Jobin, Anna, Marcello Lenca und Effy Vayena. 2019. The global landscape of AI ethics Leitlinien. *Nature Machine Intelligence* 1(9): S. 389-399.
- Jobin, Anna. 2020. Ethische Künstliche Intelligenz – von Prinzipien zu Prozessen. In *Digitaler Wandel und Ethik*, hrsg. Markus Hengstschläger und dem Rat für Forschung und Technologieentwicklung, S. 143-159. Elsbethen: Ecowing.
- Kazim, Emre und Adriano Soares Koshiyama. 2021. A high-level overview of AI ethics. *Patterns* 2(9): S. 1-12.
- Kieslich, Kimon, Birte Keller und Christopher Starke. 2022. Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data & Society* 9(1): S. 1-15.
- Landesregierung BW. 2018a. Vorreiter für Künstliche Intelligenz. <https://www.baden-wuerttemberg.de/de/service/presse/pressemitteilung/pid/vorreiter-fuer-kuenstliche-intelligenz/>. Zugegriffen: 02.07.2022.
- Mayring, Philipp und Thomas Fenzl. 2014. Qualitative Inhaltsanalyse. In *Handbuch Methoden der empirischen Sozialforschung*, hrsg. Nina Baur und Jörg Blasius, S. 543-556. Wiesbaden: Springer Fachmedien Wiesbaden.
- Mayring, Philipp. 1991. Qualitative Inhaltsanalyse. In *Handbuch qualitative Forschung: Grundlagen, Konzepte, Methoden und Anwendungen*, hrsg. Uwe Flick, Ernst von Kardoff, Heiner Keupp, Lutz von Rosenstiel und Stephan Wolff, S. 209-213. München: Beltz.
- Mayring, Philipp. 2000. Qualitative Inhaltsanalyse. *Forum: Qualitative Social Research* 1(2): S. 1-10.
- Mayring, Philipp. 2010. Qualitative Inhaltsanalyse. In *Handbuch qualitative Forschung*, hrsg. Günter Mey und Katja Mruck, S. 601-613. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mayring, Philipp. 2012. Qualitative Inhaltsanalyse – ein Beispiel für Mixed Methods. In *Mixed Methods in der empirischen Bildungsforschung*, hrsg. Michaela Gläser-Zikuda, Tina Seidel, Carsten Rohlf, Alexander Gröschner, Sascha Ziegelbauer, S. 27-36. Münster: Waxmann.
- Mayring, Philipp. 2015. *Qualitative Inhaltsanalyse. Grundlagen und Techniken*, 12. Auflage. Weinheim und Basel: Beltz Verlag.
- Mays, Nicholas, Emilie Roberts und Jennie Popay. 2001. Synthesising research evidence. In *Methods for studying the delivery and organisation of health services*, hrsg. Pauline Allen, Pauline Allen, Nick Black, Aileen Clarke, Naomi Fulop, Stuart Anderson, S. 188-220. London: Routledge.
- Mittelstadt, Brent. 2019. Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence* 1(11): S. 1-19.
- Morley, Jessica, Anat Elhalal, Francesca Garcia, Libby Kinsey, Jakob Mökander und Luciano Floridi. 2021. Ethics as a Service: A Pragmatic Operationalisation of AI Ethics. *Minds and Machines* 2021(31): S. 239-256.

- Orwat, Carsten. 2020. Risks of Discrimination through the Use of Algorithms. https://www.antidiskriminierungsstelle.de/EN/homepage/_documents/download_diskr_risiken_verwendung_von_algorithmen.pdf?__blob=publication-File&v=1. Zugegriffen: 02.07.2022.
- Rebstadt, Jonas, Henrik Kortum, Laura Sophie Gravenmeier, Birgid Eberhardt, Oliver Thomas. 2022. Non-Discrimination-by-Design: Handlungsempfehlungen für die Entwicklung von vertrauenswürdigen KI-Services. *HMD Praxis der Wirtschaftsinformatik* 59, S. 495–511.
- Rességuier, Anaïs und Rowena Rodrigues. 2020. AI ethics should not remain toothless! A call to bring back the teeth of ethics. *Big Data & Society* 7(2): S. 1-5.
- Royakkers, Lambér, Jelte Timmer, Linda Kool und Rinie van Est. 2018. Societal and ethical issues of digitization. *Ethics and Information Technology* 2018(20): S. 127-142.
- Ruin, Sebastian. 2019. Kategorien als Ausdruck einer ausgewiesenen Beobachter_innenperspektive? Ein Vorschlag für eine qualitativere qualitative Inhaltsanalyse. *Forum: Qualitative Social Research* 20(3): S. 1-30.
- Ryan, Mark und Bernd Carsten Stahl. 2021. Artificial intelligence ethics Leitlinien for developers and users: clarifying their content and normative implications. *Journal of Information, Communication and Ethics in Society* 19(1): S. 61-68.
- Staatsministerium Baden-Württemberg. O. J. Starker ländlicher Raum. <https://www.baden-wuerttemberg.de/de/bw-gestalten/nachhaltiges-baden-wuerttemberg/laendlicher-raum/>. Zugegriffen: 02.07.2022.
- Zeng, Yi, Cunqing Huangfu und Enmeng Lu. 2019. Linking Artificial Intelligence Principles. https://www.researchgate.net/publication/329608120_Linking_Artificial_Intelligence_Principles. Zugegriffen: 02.07.2022.

Leitlinienverzeichnis

- BStK. 2019. Hightech Agenda Bayern. Regierungserklärung des Bayerischen Ministerpräsidenten Dr. Markus Söder, MdL, am 10. Oktober 2019 vor dem Bayerischen Landtag. https://www.bayern.de/wp-content/uploads/2019/10/hightech_agenda_bayern.pdf. Zugegriffen: 02.07.2022.
- BStK. 2020 Hightech Agenda Bayern Plus. Bericht aus der Kabinettsitzung. <https://www.bayern.de/wp-content/uploads/2020/10/200914-ministerrat.pdf>. Zugegriffen: 02.07.2022.
- EM MV. 2021. Digitale Agenda der Landesregierung Mecklenburg-Vorpommern 2018-2021. <https://www.digitalesmv.de/digitale-agenda-mv>. Zugegriffen: 02.07.2022.
- Landesregierung Brandenburg. 2021. Digitalprogramm des Landes Brandenburg 2025. Digital. Vernetzt. Gemeinsam. https://digitalesbb.de/wp-content/uploads/2021/12/211214_dp25_Konsultationsfassung.pdf. Zugegriffen: 02.07.2022.
- Landesregierung Bremen. 2020. BREMEN.KI. Strategie Künstliche Intelligenz. https://www.senatspressestelle.bremen.de/sixcms/media.php/13/20210415_WFB_KI-Brosch%C3%BCre_32_Seiten.pdf. Zugegriffen: 02.07.2022.
- Landesregierung BW. 2018b. Strategiepapier Künstliche Intelligenz. Baden-Württemberg wird Vorreiter für Künstliche Intelligenz. https://www.baden-wuerttemberg.de/fileadmin/redaktion/m-stm/intern/dateien/publikationen/Anlage_zu_PM_114_Strategiepapier_KI.pdf. Zugegriffen: 02.07.2022.
- Landesregierung Rheinland-Pfalz. 2020. KI-Agenda verabschiedet - Rheinland-Pfalz Vorreiter der Schlüsseltechnologie Künstliche Intelligenz.

- <https://www.rlp.de/de/aktuelles/einzelansicht/news/News/detail/ki-agenda-verabschiedet-rheinland-pfalz-vorreiter-der-schluesselftechnologie-kuenstliche-intelligenz/>. Zugegriffen: 02.07.2022.
- MW LSA. 2021. Digitale Agenda für das Land Sachsen-Anhalt. <https://digital.sachsen-anhalt.de/>. Zugegriffen: 02.07.2022.
- MW Niedersachsen und MWK Niedersachsen. 2021. KI-Working Paper Niedersachsen. https://www.mw.niedersachsen.de/download/168696/KI-Working_Paper_Niedersachsen_nicht_barrierefrei_.pdf. Zugriffen: 02.07.2022.
- MWAE. 2022. Strategie Einsatz von Künstlicher Intelligenz in brandenburgischen Unternehmen. <https://mwae.brandenburg.de/media/bb1.a.3814.de/KI-Strategie-Wirtschaft-Brandenburg.pdf>. Zugegriffen: 02.07.2022.
- MWVLW. 2021. Fortschreibung der Regionalen Innovationsstrategie Rheinland-Pfalz. [https://efre.rlp.de/foerderperiode-2021-2027/fortschreibung-regionale-innovationsstrategie-ris#:~:text=Aktuell%20laufen%20die%20Vorbereitungen%20f%C3%BCr,%20DPfalz%20\(RIS\)%20fortgeschrieben](https://efre.rlp.de/foerderperiode-2021-2027/fortschreibung-regionale-innovationsstrategie-ris#:~:text=Aktuell%20laufen%20die%20Vorbereitungen%20f%C3%BCr,%20DPfalz%20(RIS)%20fortgeschrieben.). Zugegriffen: 02.07.2022.
- Poretschkin, Maximilian, Anna Schmitz, Maram Akila, Linara Adilova, Daniel Becker, Armin B. Cremers, Dirk Hecker, Sebastian Houben, Michael Mock, Julia Rosenzweig, Joachim Sicking, Elena Schulu, Angelika Voss und Stefan Wrobe. 2021. Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz. KI-Prüfkatalog. Sankt Augustin: Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS.
- Senatskanzlei Hamburg. 2020. Digitalstrategie für Hamburg. <http://static.hamburg.de/fhh/epaper/digitalstrategie/>. Zugegriffen: 02.07.2022.
- SenWiEnBe. 2021. Grünbuch für die Digitalisierungsstrategie des Landes Berlin. <https://digitalstrategie.berlin.de/haupt/de/inhalte-zum-download/>. Zugegriffen: 14.04.2022.
- SK Sachsen. 2021. Künstliche Intelligenz. Strategie für den Freistaat Sachsen. https://www.smart.esachsen.de/download/KI_Strategiebroschuere_Aufgabe_2_Doppelseiten_neu.pdf. Zugegriffen: 02.07.2022.
- Staatskanzlei Rheinland-Pfalz. 2018. Strategie für das digitale Leben. Rheinland-Pfalz Digital – Wir vernetzen Land und Leute. https://www.digital.rlp.de/fileadmin/Redaktion/Pdf/Strategie_fuer_das_Digitale_Leben_RLP.pdf. Zugegriffen: 02.07.2022.
- Stk Hessen. 2021. Thesenpapier des Rats für Digitalethik der Hessischen Landesregierung Vertrauen in KI. https://digitales.hessen.de/sites/digitales.hessen.de/files/2021-06/20210205_thesenpapier_ethikrat_vertrauen_in_ki.pdf. Zugegriffen: 02.07.2022.
- Stk Saarland. 2021. Strategie für Innovation und Technologie Saarland (Innovationsstrategie). StrategiePlus (2021-2023). https://www.saarland.de/Shared-Docs/Downloads/DE/stk/wissenschaft-forschung-technologie/dld_strategie-plus.pdf?__blob=publicationFile&v=2. Zugegriffen: 02.07.2022.
- Stk SH. 2021. Künstliche Intelligenz – Strategische Ziele und Handlungsfelder für Schleswig-Holstein. Version 2.0. https://transparenz.schleswig-holstein.de/dataset/b5e2601c-eb2b-459c-9e52-b80894be2a81/resource/d4ea6473-6dd0-4340-af9f-3f9d9f42c6c9/download/ki_strategie_download.pdf. Zugegriffen: 02.07.2022.
- TMWWDG. 2021. Thüringer Strategie für die Digitale Gesellschaft. Update 2021. https://wirtschaft.thueringen.de/fileadmin/th6/digitales/3_Anlage_2_DigitalstrategieUpdate2021.pdf. Zugegriffen: 02.07.2022.