HOCHSCHULE HEILBRONN

Modellierung der Digitalen Souveränität KI-integrierender Unternehmen in der Nachgelagerten KI-Wertschöpfung: Entwicklung eines Bewertungs-, Gestaltungs- und Reifegradmodells.

Englischer Titel:
Modeling the Digital Sovereignty of AI-Integrating Companies in Downstream AI Value Creation: Development of an Assessment, Design, and Maturity Model.

Masterarbeit
im Masterstudiengang Unternehmensführung,
Fakultät Wirtschaft, Hochschule Heilbronn

vorgelegt von:	Corin Dittmer 3. Fachsemester Matrikelnummer: 210382 E-Mail-Adresse: cdittmer@stud.hs-heilbronn.de
Erstprüfer:	Prof. Dr. Sascha Alpers
Zweitprüfer:	MBA Daniel Nill
Ausgabedatum:	21.03.2026
Abgabedatum:	21.09.2026

Inhaltsverzeichnis

Inhaltsverzeichnis ..... **II**

Abkürzungsverzeichnis ..... **V**

Abbildungsverzeichnis ..... **VI**

1. Einleitung **7**

1.1. Motivation 7

1.2. Zielsetzung 8

1.3. Forschungsfragen 9

1.4. Aufbau der Arbeit 10

2. Theoretische Grundlagen **12**

2.1. Digitale Souveränität 12

2.2. KI Wertschöpfungskette (nachgelagert) 13

2.3. Künstliche Intelligenz und KI-integrierende Unternehmen 14

2.4. Modellierung 15

3. Forschungsdesign **17**

3.1. Allgemeines Vorgehen 17

3.2. Literaturrecherche 18

4. Literaturarbeit **21**

4.1. Stand der Forschung 21

4.1.1. Geopolitische Entwicklung der digitalen Souveränität 22

4.1.2. Infrastrukturelle und technologische digitale Souveränität 23

4.1.3. Nachgelagerte KI-Wertschöpfungskette 23

4.1.4. Technische Ansätze zur Datensouveränität 24

4.2. Vergleichbare Modelle 24

5. Ableitung der Dimensionen digitaler Souveränität **29**

5.1. Methodisches Vorgehen 29

5.2. Dimensionen digitaler Souveränität 30

5.2.1. Technologische und Infrastrukturelle Souveränität 30

5.2.2. Daten- und Informationssouveränität 32

5.2.3. Organisationale Souveränität und KI-Governance 33

5.2.4. Kompetenzen und Akteure 34

5.2.5. Ökosysteme und Machtverhältnisse 36

5.2.6. Regulatorische und Normative Souveränität 37

6. Bewertungsmodell 40
6.1. Ziele und Anforderungen 40
6.2. Modellstruktur 41
6.3. Bewertungsmechanismen 44
6.4. Umsetzung des Bewertungsmodells in der Web-App 45
7. Reifegradmodell 47
7.1. Ziele und Anforderungen 47
7.2. Reifestufen und Messinstrumente 48
7.3. Einordnungsmechanismen und Abhängigkeiten 49
7.4. Analytische Auswertung in der Webanwendung 51
8. Gestaltungsmodell 53
8.1. Ziele und Anforderungen 53
8.2. Berechnung der Priorisierung (Handlungsdruck) 53
8.3. Struktur und Zuordnung der Handlungsempfehlung 55
8.4. Umsetzung und Visualisierung in der Webanwendung 56
9. Anwendungsbeispiel 58
9.1. Vorstellung des Use Cases 58
9.1.1. Unternehmensprofil: MEDfiction GmbH 58
9.1.2. Branchenkontext und Relevanz digitaler Souveränität 59
9.1.3. Profilausprägung: Ist- und Soll-Zustand 60
9.1.4. Bepunktung des Fragebogens 63
9.2. Ergebnisse der Modellierung 67
9.2.1. Bewertungsmodell 67
9.2.2. Reifegradmodell 69
9.2.3. Gestaltungmodell 71
9.3. Ergebnis und Fazit des Anwendungsbeispiels 75
10. Evaluation der Ergebnisse 77
10.1. Vorgehen und Methode 77
10.2. Qualitative Inhaltsanalyse (Experteninterviews) 77
10.3. Ergebnisse der Evaluation 77
11. Ethische, rechtliche und Soziale Überlegungen 78
12. Fazit 80
12.1. Beantwortung der Forschungsfrage 80
12.2. Implikationen 80

12.3. Limitationen 80
12.4. Forschungsausblick 80

Literaturverzeichnis 81

Anhang 89

A. Modellierung mit der Excel-Anwendung 89
B. Übersicht der Handlungsempfehlungen 90

Abkürzungsverzeichnis

KI	Künstliche Intelligenz
KPI	Key Performance Indicator (Leistungskennzahl)
MDR	Medical Device Regulation (EU-Verordnung)
DSGVO	Datenschutz-Grundverordnung
API	Application Programming Interface (Anwendungs-Programmierschnittstelle)
XAI	Explainable AI (Erklärbare KI)
CDR	Corporate Digital Responsibility
MPDG	Medizinprodukterecht-Durchführungsgesetz
CRA	Cyber Resilience Act (EU-Verordnung)
ISMS	Information Security Management System
MLOps	Machine Learning Operations
KMU	Kleine und mittlere Unternehmen
NIS-2	Network and Information Security
SLA	Service Level Agreement
ITIL	Information Technology Infrastructure Library
COBIT	Control Objectives for Information and Related Tech-nologies
SCRM	Supply Chain Risk Management

Abbildungsverzeichnis

Abbildung 1: Beispiel-Struktur des Prompts des Literaturscreening 19
Abbildung 2: Legende für die Konzeptmatrix 20
Abbildung 3: Konzeptmatrix nach Webster und Watson (2002): Stand der Forschung 22
Abbildung 4: Übersicht der Dimensionen digitaler Souveränität 30
Abbildung 5: Übersicht der Fragen des Fragebogens des Bewertungsmodells 42
Abbildung 6: Benutzeroberfläche des Fragebogens in der Webanwendung mit sieben Datenpunkten 43
Abbildung 7: Architektur der Webanwendung 45
Abbildung 8: Bewertungsmodell-Dashboard der Webanwendung (optisch nachträglich angepasst) 46
Abbildung 9: Reifegradmodell-Dashboard der Webanwendung 51
Abbildung 10: Bewertung der Dimensionen für die Reifegradermittlung 52
Abbildung 11: Dashboard des Gestaltungsmodells (Webanwendung) 56
Abbildung 12: Ist- und Soll-Zustands des Anwendungsbeispiels 60
Abbildung 13: Antworten des Fragebogens im Anwendungsbeispiel 64
Abbildung 14: Dimensionsbewertung des Anwendungsbeispiels 65
Abbildung 15: Definition des Soll-Zustands des Anwendungsbeispiels 66
Abbildung 16: Bewertungsmodell - KPIs des Anwendungsbeispiels 67
Abbildung 17: Bewertungsmodell - Dimensionsübersicht des Anwendugnsbeispiels 68
Abbildung 18: Bewertungsmodell - Konfidenz des Anwendungsbeispiels 68
Abbildung 19: Reifegradmodell - Reifegrad des Anwendungsbeispiels 69
Abbildung 20: Reifegradmodell - Schichten & Dimensionen des Anwendungsbeispiels 70
Abbildung 21: Gestaltungsmodell - Priorisierung der Dimensionen im Anwendungsbeispiel 71
Abbildung 22: Gestaltungsmodell - Priorisierung der Kategorien im Anwendungsbeispiel 72
Abbildung 23: Gestaltungsmodell - Handlungsempfehlungen für das Anwendungsbeispiel 73

1. Einleitung

1.1. Motivation

Künstliche Intelligenz (KI) treibt eine neue industrielle, aber auch wissenschaftliche Revolution an, welche sich als Branchen- und Zweckübergreifend herausstellt (vgl. Zuo 2023; Shahade und Deshmukh 2025). KI hat das Potenzial die unternehmerische Wertschöpfung grundlegend zu ändern, indem KI sowohl als effizienzsteigerndes Werkzeug als auch als neuer Produktionsgegenstand dient und menschlich kognitive Aufgaben automatisieren kann (vgl. Zhang und Fu 2021). Dies ermöglicht neue Geschäftsmodelle und Industrien (vgl. Farmakis et al. 2024). Die auf Basis von KI-Technologien entstandene Transformation kann alle wirtschaftliche Aktivitäten von Unternehmen durchdringen (vgl. Mitra et al. 2024). Dabei liegt der Fokus der Wertschöpfung nun mehr auf wissensintensiven, hoch qualifizierten Aktivitäten, wodurch KI zur zentralen Ressource in Unternehmen wird (vgl. Li 2024).

Neben dem Makroökonomischen Wandel entsteht eine strukturelle Veränderung der Integration und Nutzung von KI für unternehmerische Zwecke. Anstatt KI-Systeme selbst zu entwickeln, greifen Unternehmen auf bereits extern entwickelte Modelle oder von dritten bereitgestellte Cloud-Services zurück. Dies sorgt für eine tiefgreifende Umgestaltung der Wertschöpfungskette (vgl. Foster 2025). Aufgrund der enormen Entwicklungskosten von Foundation Models greifen die meisten Unternehmen auf Lösungen von Drittanbietern der vorgelagerten KI-Wertschöpfungskette zurück (vgl. Hynninen und Maunula 2025). Diese Unternehmen werden dadurch zu nachgelagerten Akteuren der KI-Wertschöpfungskette und stimmen bestehende KI-Modelle auf spezifische Aufgaben ab oder integrieren diese in eigene Produkte oder Dienstleistungen (vgl. Foster 2025). Dies führt jedoch zu einer asymmetrischen Souveränität und Machtverteilung innerhalb der globalen KI-Wertschöpfungskette (vgl. Foster 2025; Wamba-Taguimdje et al. 2020).

Die Probleme, die sich durch die Integration externer Technologie hinsichtlich digitaler Souveränität und Kontrolle ergeben, können zu Risiken wie Vendor Lock-Ins, Verlust der Datenhoheit oder Vertrauensverlusten führen (vgl. Hynninen und Maunula 2025; Foster 2025). Unternehmen, die in der nachgelagerten KI-Wertschöpfungskette agieren, können mit erheblichen Abhängigkeiten konfrontiert werden, da diese auf Black-Box Technologie angewiesen sind, die diese wenig oder gar nicht einsehen und nachvollziehen können, was zu intransparenten und schlecht validierbaren Ausgaben oder Ergebnissen führt (vgl. Gurjar et al. 2024; Wamba-Taguimdje et al. 2020). Außerdem nennen Celsi und Zomaya (2025) die Gefahr einer „Algokratie“, bei der die Entscheidungsmacht auf die Daten- und Technologie-Besitzer übergeht. Genauso wie Eigentumsrechte an Daten und die daraus resultierenden Wettbewerbsvorteile konsolidieren sich oft bei großen Technologieanbietern, was negative Folgen für die Innovationsfähigkeit nachgelagerter Akteure haben kann (vgl. Babina et al. 2024).

Trotz der offensichtlichen Relevanz dieser Herausforderungen für die betroffenen Branchen und Forschungsfelder, zeigt die wissenschaftliche Literatur wenig Publikationen zu diesem Thema. Nach Foster (2025) fehlt ein umfassendes Verständnis über die Auswirkungen der Machtstrukturen und Abhängigkeiten auf die nachgelagerte KI-Wertschöpfungskette. Bestehende Literatur bietet oft keine ausreichenden empirischen Belege oder Analysemöglichkeiten für die Governance der Digitalen Souveränität der nachgelagerten KI-Wertschöpfungskette (vgl. Tairov et al. 2024; Gurjar et al. 2024). Auch ist die Komplexität des Zusammenhangs der KI-Systeme mit deren Governance und Werterealisierung unzureichend erforscht (vgl. Enholm et al. 2022) und es mangelt an Erkenntnissen über die strategische Steuerung von technologischem Nutzen von KI und der Wahrung digitaler Souveränität (vgl. Zebec und Stemberger 2024).

Aus dem Zwiespalt zwischen der schnellen technologischen Adoption und dem Fehlen von Steuerungsmechanismen ergibt sich ein Handlungsbedarf. Gegenwärtig mangelt es in Unternehmen, die KI integrieren, an geeigneten Werkzeugen, um ihre Stellung innerhalb der Wertschöpfungskette zu beurteilen oder um ihre digitale Souveränität systematisch zu schützen. Ohne eine Entscheidungsgrundlage droht nachgelagerten Unternehmen der langfristige Verlust ihrer strategischen Unabhängigkeit. Diese Arbeit geht auf das identifizierte Defizit ein, indem ein Modellierungsansatz entwickelt wird, der ein Bewertungs-, Reifegrad- und Gestaltungsmodell für die digitale Souveränität in der nachgelagerten KI-Wertschöpfung beinhaltet.

1.2. Zielsetzung

Diese Forschungsarbeit hat das Ziel, einen Ansatz zur Modellierung und Bewertung der digitalen Souveränität von Unternehmen, die künstliche Intelligenz in ihre Wertschöpfungsprozesse einbinden, zu entwickeln, der sowohl theoretisch fundiert als auch methodisch nachvollziehbar ist. Die Arbeit richtet sich insbesondere an Organisationen, die in der nachgelagerten KI-Wertschöpfungskette tätig sind, indem sie KI-Technologien nutzen, ohne selbst grundlegende KI-Infrastrukturen oder Basismodelle zu erstellen.

Die exponentielle Nutzung externer KI-Technologien, wie Cloud-Diensten, oder vortrainierter Modelle, schafft für viele Unternehmen neue Abhängigkeiten von Technologieanbietern und digitaler bzw. technologischer Infrastruktur. Es gibt jedoch noch keinen konsistenten Ansatz in wissenschaftlicher Literatur, der Unternehmen und Organisationen dabei unterstützt, digitale Souveränität im Hinblick auf diese Abhängigkeiten systematisch zu analysieren, zu bewerten oder strategisch zu verbessern. Dabei sind Diskussionen über digitale Souveränität oft auf politische oder infrastrukturelle Ebene beschränkt und behandeln nur einzelne Aspekte wie Datensouveränität oder technologische Kontrolle.

Demnach ist das Ziel dieser Arbeit, digitale Souveränität auf Ebene von KI-integrierenden Unternehmen konzeptionell zu strukturieren und sie durch wissenschaftliche Modelle operationalisierbar zu machen. Zuerst wird ein theoretisch fundiertes Verständnis von organisationalen Abhängigkeiten in KI-Ökosystemen aufgebaut. Auf Basis dieser Grundlage sollen die wesentlichen Aspekte der digitalen Souveränität bestimmt werden und in einem konzeptionellen Modell vereint werden. Diese Aspekte sollen dann als nächstes in messbare Kriterien umgewandelt werden, die eine systematische Bewertung der digitalen Souveränität ermöglichen. Darüber hinaus wird ein Reifegradmodell erstellt, das verschiedene Entwicklungsstufen der souveränen KI-Integration darstellt und Unternehmen somit eine Orientierung für die Weiterentwicklung ihrer Fähigkeiten bietet.

1.3. Forschungsfragen

Mit der Beantwortung der Haupt-Forschungsfrage **„Wie kann digitale Souveränität von KI-integrierenden Unternehmen in der nachgelagerten KI-Wertschöpfung theoretisch fundiert modelliert und operationalisiert werden?“** sollen konkrete Modelle zur Bewertung und Gestaltung digitaler Souveränität von Organisationen im Kontext der KI-Wertschöpfung entwickelt werden. Um diese Frage allerdings beantworten zu können, sind die folgenden, spezifischen Teilfragen von Nöten:

Durch die erste Teilfrage **„Welche theoretischen Konzepte eignen sich zur Erklärung organisationaler Abhängigkeiten in KI-Ökosystemen?“** sollen die theoretischen Grundlagen der Abhängigkeiten von Unternehmen und Organisationen in der vorgelagerten KI-Wertschöpfungskette geschaffen werden. Da die Nutzung und Integration künstlicher Intelligenz in vielen Fällen über komplexe digitale Ökosysteme erfolgt, versucht die erste untergeordnete Forschungsfrage bestehende theoretische Konzepte zu identifizieren, zu analysieren und Abhängigkeitsstrukturen zu erklären. Hierzu werden relevante Ansätze aus bestehenden Konzepten oder angrenzenden Forschungsfeldern wie beispielsweise IT Governance oder Cloud Computing untersucht. Durch Vergleiche bestehender Perspektiven sollen geeignete Konzepte ermittelt werden, die die spezifischen Abhängigkeitsverhältnisse in der KI-Wertschöpfung beschreiben sollen. Das Ergebnis dieser Forschungsfrage bildet die theoretische Grundlage für weitere Konzepte digitaler Souveränität und sollen zentrale Mechanismen und Einflussfaktoren identifizieren, die für die Modellierung digitaler Souveränität wichtig sind.

Um die Dimensionen digitaler Souveränität für Organisationen zu identifizieren, wird die zweite Teilfrage **„Welche zentralen Dimensionen sind für die Beschreibung digitaler Souveränität auf Ebene KI-integrierender Unternehmen erforderlich?“** gestellt. Während digitale Souveränität meist ein politisches oder normatives Ziel gilt, fehlen bislang zugrundeliegende Dimensionen auf Organisations- und Unternehmensebene. Gerade bei KI-integrierenden Organisationen stellt sich die Frage, welche Faktoren den Umfang der Kontrolle, Entscheidungs- und Gestaltungsfreiheit im Umgang mit externer KI-

Technologie entscheidend sind. Ziel der zweiten untergeordneten Forschungsfrage ist daher die strukturierte Identifikation der Dimensionen, die entscheidend für digitale Souveränität auf Organisationsebene sind. Hierzu werden bestehende Konzepte, Definitionen und verwandte Dimensionen aus akademischer Literatur analysiert und verglichen, um wiederkehrende Merkmale, Einflussfaktoren und Fähigkeiten für verschiedene Ausprägungen digitaler Souveränität zu identifizieren, bevor die gewonnenen Erkenntnisse in konkrete Dimensionen als Grundlage für ein konzeptionelles Modell digitaler Souveränität zusammengeführt werden.

Aufbauend auf der Identifikation relevanter Dimensionen der zweiten untergeordneten Forschungsfrage, soll nun das Konzept in der Praxis anwendbar gemacht werden. Dafür sollen die identifizierten Dimensionen in konkret messbare Kriterien überführt werden. Daher bietet sich die dritte Teilfrage **„Wie kann digitale Souveränität messbar operationalisiert werden?“** an. Ziel der dritten untergeordneten Forschungsfrage ist die Entwicklung eines Mess- und Bewertungsmodells, die eine strukturierte Bewertung der Ausprägung digitaler Souveränität ermöglicht. Hierfür sollen die zuvor identifizierten Dimensionen mit Indikatoren und Bewertungskriterien versehen werden, die unterschiedliche Ausprägungsgrade erfassbar und vergleichbar machen sollen. Ein Bewertungsmodell, das Organisationen und Unternehmen bei der Erfassung und Analyse ihrer aktuellen Position unterstütz, soll dabei das finale Artefakt darstellen.

Die vierte Teilfrage **„Welche Reifegrade souveräner KI-Integration lassen sich auf Grundlage der identifizierten Dimensionen und Bewertungslogiken unterscheiden?“** zielt auf die Erstellung eines Reifegradmodells ab. Da digitale Souveränität kein statischer Zustand ist, können Organisationen und Unternehmen unterschiedliche Entwicklungsstufen hinsichtlich der Kontrolle und Abhängigkeit externer KI-Technologien, erreichen. Ziel der letzten Forschungsfrage ist daher die Identifikation und Definition verschiedener Reifegrade digitaler Souveränität, die aufbauend auf den Bewertungsdimensionen typische Entwicklungsstufen definiert, die wiederum unterschiedliche Ausprägungen der Fähigkeiten, Governance-Strukturen und technischer Kontrolle widerspiegeln. Das resultierende Reifegradmodell soll Organisationen es ermöglichen, den aktuellen Entwicklungsstand einzuordnen und potenzielle Transformationsschritte zu einer verbesserten digitalen Souveränität zu identifizieren. Darüber hinaus soll das Reifegradmodell als analytisches Instrument zur strukturierten Aufnahme und Beschreibung des aktuellen Zustands dienen.

1.4. Aufbau der Arbeit

Dieses Kapitel soll gegen Ende der gesamten Arbeit geschrieben werden, da der Aufbau noch abweichen kann.

Nach der Einleitung folgen zuerst die theoretischen Grundlagen, die ein gemeinsames Bewusstsein für die Begrifflichkeiten und das Themenfeld dieser Arbeit schaffen soll. Jede Definition basiert erst auf reiner Literaturarbeit, welche verschiedenste wissenschaftliche Definitionen zu einer abschließenden, zusammenfassenden Synthese formen. Definiert werden die Begrifflichkeiten der digitalen Souveränität, KI-integrierenden Unternehmen, nachgelagerter KI-Wertschöpfungskette und Modellierung im Kontext dieser Forschungsarbeit. Daraufhin folgt eine detaillierte Ausführung des Forschungsdesigns, welches erst generelle Methoden und dessen Durchführung beschreibt und daraufhin im Detail das Vorgehen und die Struktur der KI-unterstützen Literaturrecherche beleuchtet.

2. Theoretische Grundlagen

Im Folgenden sollen nun die wichtigsten Begriffe und Themen auf Basis wissenschaftlicher Publikationen definiert werden. Nach den literaturbasierten Definitionen folgt jeweils eine Synthese optimiert an das Forschungsthema der Modellierung digitaler Souveränität.

2.1. Digitale Souveränität

Digitale Souveränität beschreibt im Allgemeinen die Fähigkeit von Staaten, Organisationen, Unternehmen und Individuen, die Kontrolle, Autorität und Selbstbestimmung über die digitale Zukunft, Infrastruktur, Vermögenswerte, Technologie und weiteren digitalen Elementen auszuüben (vgl. Calderaro und Blumfelde 2022; Pohle et al. 2025; Tan et al. 2023). Dabei umfasst die digitale Souveränität die Befugnis von Organisationen oder Individuen, unabhängig und autonom über die Existenz, Erstellung und Dynamik digitaler Elemente, wie Daten, Technologie wie Software oder Hardware oder Standards wie Protokolle und zu bestimmen (vgl. Floridi 2020; Roberts 2024; Moerel und Timmers 2021).

Ein fundamentales Element der digitalen Souveränität ist die Datensouveränität, die das Recht über den Zugriff, die Speicherung und die Verwendung entscheidet (vgl. Hummel et al. 2021; Firdausy et al. 2022). Damit beinhaltet die Datensouveränität die Fähigkeit, Datenflüsse zu steuern, die Einhaltung von Gesetzen zu gewährleisten, sowie die Abhängigkeit von externer Hard- und Software zu minimieren, was schlussendlich zu höherer IT-Sicherheit und Resilienz führen soll (vgl. Klare et al. 2025; Aboushady et al. 2024). Hinsichtlich digitaler Unternehmen bedeutet das die Autonomie, digitale Vermögenswerte zu nutzen und (sensible) Daten selbstbestimmt sicher teilen zu können (vgl. Firdausy et al. 2022)

Im Kontext von digitaler Souveränität und künstlicher Intelligenz (KI), findet sich „algorithmischen Souveränität“ (vgl. Abiade 2025) oder KI-Souveränität, die neben der Souveränität der Daten auch die Kontrolle über das Design, die Bereitstellung, die Entscheidungslogik von KI-Systemen beinhaltet (vgl. Calderaro und Blumfelde 2022; Floridi 2020). Auch die Kontrolle über die KI-Modelle, Trainingsdaten aber auch generierten Informationen oder Mehrwerte, sowie der Schutz dieser, vor unbefugter Nutzung durch Dritte (vgl. Firdausy et al. 2022; Tao et al. 2022; Daliparthi et al. 2025).

Für die nachgelagerte KI-Wertschöpfung ist die Reduktion kritischer Abhängigkeiten und Unsicherheiten zwar zentral, allerdings nur schwer umsetzbar, gerade hinsichtlich der Nutzung der führenden großen Sprachmodelle. Hier ist der naheliegendste Ansatz für die Digitale Souveränität, Abhängigkeiten wie Vendor Lack-Ins, von einzelnen dominanten Technologiepartners wie Google oder OpenAI

und deren Black-Box Technologie zu umgehen indem Softwarelösungen explizit modular und anpassbar entwickelt werden (vgl. Janardhanan et al. 2025; Abiade 2025). Es bedarf Strategien zur Diversifizierung von Lieferanten für KI-integrierende Unternehmen der nachgelagerten Wertschöpfungskette, um operative Unabhängigkeit zu wahren und die eigene Handlungs- und Wettbewerbsfähigkeit langfristig zu gewährleisten (vgl. Roberts 2024; Janardhanan et al. 2025; Boeira et al. 2024).

Zusammenfassend wird die Begrifflichkeit „digitale Souveränität“ im Kontext dieser Arbeit als ein relatives, organisationales Konzept in stetigem Wandel verstanden. Für die nachgelagerte KI-Wertschöpfung bedeutet digitale Souveränität die strategische, technologische und informationelle Handlungsfähigkeit, um kritische Abhängigkeitsverhältnisse aktive zu steuern oder zu kontrollieren insbesondere in Anbetracht von Unternehmen und Organisationen. Das bedeutet, dass Organisationen die Integrität der eigenen Wertschöpfungsprozesse trotz Nutzung externer Technologien mithilfe von bewusster Governance, System-architektonischer Modularität oder Lieferantendiversifizierung, wahren.

2.2. KI Wertschöpfungskette (nachgelagert)

Im Allgemeinen kann die KI-Wertschöpfungskette als eine Art Netzwerk von Co-Creation-Strukturen bezeichnet werden, in dem Ressourceninputs und Outputs von KI-Systemen von verschiedenen Akteuren gestaltet werden (vgl. Attard-Frost und Widder 2024). Diese Co-Creation-Strukturen gehen über lineare Lieferketten hinaus, welche sowohl vorgelagerte, nachgelagerte, horizontale Datenflüsse integriert (vgl. Billones et al. 2025). Nachgelagerte Akteure bauen in diesem Ökosystem auf Komponenten, Technologie und Design auf vorgelagerten Ebenen der Wertschöpfungskette auf. Damit können eigene, spezifische Ziele realisiert werden, die allerdings von der vorgelagerten Architektur, Design oder Technologie die Möglichkeiten der nachgelagerten Wertschöpfung beeinflusst, einschränkt und in Abhängigkeit versetzt (vgl. Hopkins et al. 2025).

Der Bereich der nachgelagerten KI-Wertschöpfungskette bezieht sich auf wirtschaftliche Aktivitäten von Unternehmen, die grundlegende Technologie wie Foundationmodels nicht selbst entwickeln, sondern diese nutzen, anpassen, spezialisieren oder weiterentwickeln und in eigene Anwendungen oder Produkte integrieren (vgl. Foster 2025). Nach Heeks und Spiesberger (2024) lässt sich die nachgelagerte Wertschöpfungskette von KI in die Kernelemente Machinelearning Operations (MLOps), Anwendungen und Services unterteilen. Dabei umfasst MLOps Tools und Technologien zur Weiterentwicklung, Datenverarbeitung, Überprüfung, Monitoring und Wartung von KI-Systemen. B2B oder B2C Produkte, die auf Foundation Models aufbauen und branchenspezifische (vertikale) oder generische (horizontale) Anwendung finden, zählen zu dem Kernelement der Anwendung, während Dienstleistungen, die menschliches Wissen und menschlichen Umgang erfordern, zu den Services zählen. Die nachgelagerte Wertschöpfung an sich kann somit als „Letzte Meile“ bezeichnet werden, bei der angepasste

(finetuned) und spezialisierte Foundation Models bereitgestellt als Produkt bereitgestellt werden (vgl. Billones et al. 2025; Pirrone et al. 2025; Pi 2024). Ziel hierbei ist der Transfer von Fähigkeiten und die Integration der KI in Geschäftsprozesse, um eine wirtschaftliche Aufwertung oder Modernisierung zu erreichen (vgl. Foster 2025). Nachgelagerte Akteure nutzen KI direkt oder als Produkt, um final Produktivität zu steigern, Echtzeit-Entscheidungen zu unterstützen und generell Geschäftsergebnisse zu verbessern (vgl. Jacobides et al. 2021).

In der vorgelagerten KI-Wertschöpfungskette gelten Foundation Model Developer wie OpenAI oder Google, Hardware-Provider wie Nvidia oder Oracle und Datenlieferanten wie Datenbroker oder Datenbesitzer (z.B. Internetnutzer) als Akteure (vgl. Foster 2025; Pirrone et al. 2025; Attard-Frost und Widder 2024). In der nachgelagerten KI-Wertschöpfungskette gelten Anwendungsentwickler und Bereitsteller (Deployer), wie Softwareentwickler, die Foundation Models für eigene Produkte finetunen oder anpassen; Service-Provider, die Dienstleistungen wie Beratung oder Produktintegration anbieten; und Endanwender als Akteure (vgl. Pi 2024; Heeks und Spiesberger 2024; Jacobides et al. 2021; Billones et al. 2025).

Zusammenfassend lässt sich der Untersuchungsgegenstand auf die nachgelagerte KI-Wertschöpfung beschränken, die hier als Co-Creation-Ökosystem der letzten Meile definiert wird, in dem Akteure als Anwendungsentwickler, Integratoren, Bereitsteller aber auch Nutzer fungieren. Wesentliches Merkmal der nachgelagerten KI-Wertschöpfung ist die Asymmetrie, die dem System zugrunde liegt, da nachgelagerte Akteure zwar spezifische Wertversprechen realisieren, jedoch Restriktionen durch vorgelagerte Akteure unterliegen wie beispielsweise bestehender Architekturen durch Foundation Models oder genereller technologischer Abhängigkeiten.

2.3. Künstliche Intelligenz und KI-integrierende Unternehmen

Künstliche Intelligenz bezeichnet ein maschinenbasiertes System, welches unter anderem menschliche Intelligenz simuliert und aus Dateneingaben Vorhersagen, Empfehlungen oder Entscheidungen ableiten kann (vgl. Melcher und Vaziri 2025). Auf Basis dessen können auch physische oder virtuelle Umgebungen beeinflusst werden. In weiterer Betrachtung können maschinelle (KI-) Agenten Technologien wie maschinelles Lernen, Deep Learning und generativer KI zur automatisierten Aufgabenlösung verknüpfen indem große Datenmengen verarbeitet werden (vgl. Paschen et al. 2020; Varriale et al. 2025; Ali et al. 2024).

KI-Integrierende Unternehmen bezieht sich in dieser Arbeit auf Unternehmen oder Organisationen der nachgelagerten KI-Wertschöpfungskette, die primär als Betreiber, Weiterentwickler und Anwender tätig sind. Anstatt KI Foundation Models selbst zu entwickeln, erwerben diese Organisationen externe KI-Basistechnologien und integrieren diese tiefgreifend in eigene Geschäftsprozesse, Produkte

oder Dienstleistungen (Melcher und Vaziri 2025; Gujar et al. 2025). Dabei werden meist verschiedene Technologische Bausteine wie vorgefertigte KI- oder ML-Basismodelle kombiniert, ergänzt oder weiterentwickelt, um betriebliche Abläufe zu automatisieren, Informationsaufbereitung zu präzisieren oder Innovations- und Entwicklungsarbeit zu beschleunigen (vgl. Paschen et al. 2020; Ali et al. 2024). Durch diese Integration verschiedener Technologien können spezialisierte, wertschöpfende Innovationen generiert werden, welche nachhaltige Wettbewerbsvorteile durch Optimierung von Entwicklung, Produktion und Prozessen und Steigerung der Produktivität und Marktleistung ermöglichen sollen (vgl. Kretschmer und Orth 2024; Varriale et al. 2025; Kopka und Fornahl 2024).

KI-integrierende Unternehmen wird in dieser Arbeit zusammenfassend als spezifische Organisationsebene definiert, die sich durch Eibettung extern entwickelter KI-Basistechnologie in Prozesse oder Produkte auszeichnet. Der Fokus liegt auf Organisationen, die die Rolle des Integrators oder Anwenders einnehmen und KI nicht als Kernprodukt, sondern als Mittel zur Effizienzsteigerung, Automatisierung und Innovationsgenerierung nutzen. Die Notwendigkeit, externe Innovationskraft zu integrieren, steht hier im Spannungsfeld mit der gleichzeitigen Notwendigkeit, Souveränitätsverluste proaktiv zu managen.

2.4. Modellierung

Hevner et al. (2004) beschreiben den Begriff der Modellierung als einen problemorientierten Ansatz, bei dem Artefakte erstellt und evaluiert werden. Im Allgemeinen stellen Modelle Abstraktionen und Repräsentationen von Problemlösungen dar. Nach Wand und Weber (2002) umfasst die konzeptionelle Modellierung vier Elemente. Die Modellierungsgrammatik beschreibt die Konstrukte und Regeln, die dem Modell zugrunde liegen, während die Methoden zur Anwendung dieser Grammatik dienen. Daneben gibt es das resultierende Modell (-Skript), welches das Produkt des abgebildeten Prozesses ist, sowie der Kontext, auf den das Modell ausgerichtet ist (vgl. Wand und Weber 2002). Ziel der Modellierung ist primär die Erzeugung von Nützlichkeit für praktische Anwendung bzw. Praxis. Daher sollen Modelle in iterativen Entwicklungszyklen aus Konstruktion und Evaluation entwickelt werden (vgl. Hevner et al. 2004).

Für die Ausarbeitung dieser Arbeit sind Bewertungs-, Reifegrad- und Gestaltungsmodelle von besonderer Bedeutung. Ein Bewertungsmodell dient der strukturierten Erfassung und Analyse eines punktuellen Zustands und verfolgt somit einen deskriptiven Ansatz. Ziel ist die messbare Aufzeichnung von Fähigkeiten, Kompetenzen oder Entwicklungsständen in einem definierten Themengebiet, mittels definierter Kriterien oder Dimensionen (vgl. de Bruin et al. 2005). Pöppelbuß und Röglinger (2011)

empfehlen für ein nützliches Bewertungsmodell die Nutzung nachvollziehbarer Kriterien jeder Granularitätsebene sowie eine nachvollziehbare Assessment-Methodik, um konsistente, genaue und wiederholbare Bewertungsergebnisse zu gewährleisten.

Auf den Kriterien bzw. Dimensionen des Bewertungsmodells baut das Reifegradmodell auf, welches eine Abfolge von Reifegraden definiert. Dabei wird ein typischer oder angestrebter Entwicklungspfad repräsentiert, der in diskret abgrenzbare und aufeinander aufbauende Stufen unterteilt ist (vgl. Becker et al. 2009). Dabei wird die Reife häufig mehrdimensional wie etwa durch Inhaltliche, organisatorische und technologische Dimensionen operationalisiert, mit dem Ziel differenzierte Einblicke in gezielte Strategien zu ermöglichen (vgl. Pöppelbuß und Röglinger 2011; de Bruin et al. 2005).

Ein Gestaltungs- oder Vorgehensmodell liefert einen strukturellen oder prozeduralen Rahmen für die Anpassung, Verbesserung oder Entwicklung von Systemen und Prozessen (vgl. Fettke und Loos 2003; Becker et al. 2007). Es beinhaltet konkrete, an einen bestimmten Zielzustand orientierte Handlungsempfehlungen, um von einem Ist-Zustand dorthin zu gelangen (vgl. Pöppelbuß und Röglinger 2011). Durch methodische Konstruktionsprinzipien wie Aggregation, Spezialisierung und Instanziierung lassen sich solche Modelle flexibel anpassen und wiederverwenden, um individuelle oder fallspezifische Lösungen zu generieren (vgl. vom Brocke 2007).

Zusammenfassend wird der Begriff der Modellierung in dieser Arbeit als iterativer Problemlösungsprozess verstanden, der drei auf Dimensionen basierende, aufeinander aufbauende Artefakte enthält. Ein Bewertungsmodell soll dem strukturierten Erfassen des Ist-Zustands in Organisationen dienen, während ein Reifegradmodell die erfassten Fähigkeiten in mehrdimensionale und diskrete Reifestufen einteilt. Abgeschlossen wird dieses System durch ein Gestaltungsmodell, welches generelle bzw. normative Handlungsempfehlungen liefert. So soll das theoretische Konstrukt der digitalen Souveränität für praktische Anwendungen messbar, operationalisierbar und strategisch steuerbar gemacht werden.

3. Forschungsdesign

3.1. Allgemeines Vorgehen

Das Forschungsdesign dieser Arbeit folgt dem Design Science Research (DSR) Paradigma nach Hevner et al. (2004). Dabei liegt der Fokus auf der Erstellung und Schaffung praktisch anwendbarer und nutzbarer Artefakte. Nach Hevner et al. ist dabei das fundamentale Ziel die Grenzen menschlicher und organisationaler Fähigkeiten durch die Schaffung neuer Artefakte zu erweitern. In Anbetracht der Neuartigkeit und damit einhergehenden Komplexität des Problems der digitalen Souveränität bei Integration extern entwickelter KI-Modelle bedarf es lösungsorientierte Ansätze. Diese Arbeit versucht das Paradigma nach Hevner et al. (2004) in der iterativen Konstruktion drei aufeinander aufbauender Modelle zu verwirklichen.

Die Umgebung dieser Arbeit bildet das Unternehmensumfeld, aus dem sich konkrete Bedarfe ableiten. Nach Hevner et als. (2004) Problem Relevanz wird hier ein aktuelles und nicht vollständig gelöstes praxisorientiertes Problem adressiert. Die Forschungsaktivitäten sind auf die Bereitstellung technologischer und strategischer Handlungswerkzeuge für Organisationen ausgerichtet, um Asymmetrien in der Wertschöpfung zu messen und strategisch angehbar zu machen.

Um die wissenschaftliche Strenge der Modellierung zu gewährleisten, basiert die Entwicklung der Artefakte bzw. Modelle auf einer umfangreichen Wissensbasis aus wissenschaftlicher Literatur. Dabei dienen bestehende Theorien, Konzepte und Metriken als theoretisches Fundament. Durch die Extraktion und Synthese der in der Literatur enthaltenen Vorarbeiten wird sichergestellt, dass die zu definierenden Dimensionen und Modelle nicht willkürlich gewählt, sondern auf dem aktuellen Stand der wissenschaftlichen Forschung beruhen.

Der Kern der Methode ist ein iterativer Suchprozess, der die Phasen der Erstellung und Evaluation umfasst. Die auf Basis der Wissensgrundlage definierten Dimensionen digitaler Souveränität werden vorerst in prototypische Modelle integriert, welche nach Hevner et als. (2004) Design Bewertung die Anforderungen der Nützlichkeit, Qualität und Wirksamkeit erfüllen müssen. In dieser Arbeit erfolgt die Evaluation mit qualitativer, semistrukturierter Experteninterviews, bei der die Rückmeldungen der Experten iterativen Einfluss auf die Optimierung und Verbesserung der Modelle nimmt, hinsichtlich der Korrektheit, Vollständigkeit, Anwendbarkeit, Verständlichkeit und Realitätsnähe. Damit soll Hevner et als. (2004) formulierter Entwicklungs- und Testzyklus geschlossen werden.

3.2. Literaturrecherche

Die grundlegende Literaturrecherche basiert auf den Literaturdatenbanken IEEE, Google Scholar und Consensus, wird allerdings mit Deep Research von ChatGPT und genereller Internetsuche ergänzt. Die gesuchte Literatur zum Thema digitaler Souveränität in Zusammenhang mit KI wird eingeschränkt durch einen Veröffentlichungszeitraum von 2020 bis 2026, sowie deutscher und englischer Sprache und basiert ausschließlich auf peer-reviewten wissenschaftlichen Veröffentlichungen. Für die Suche werden Freitexte und Schlagwörter für breite Themenfelder zu digitaler Souveränität in Verbindung mit KI, Modellierungsansätzen, Risikoanalysen und Betrachtungsebenen genutzt, da bislang wenig Literatur zu diesem Thema existiert und sich daher eine breite Suche anbietet.

Gefundene Literatur wird manuell nach Titel und Abstract manuell gescreent und gespeichert. Die gesammelte Literatur wird daraufhin von einer eigens entwickelten Bewertungspipeline in Anlehnung zu Dittmer und Alpers (2025) einer ersten Bewertung unterzogen. Im Gegensatz zu Dittmer und Alpers (2025) wird hier allerdings nicht die Qualität der einzelnen Publikationen, sondern die Inhaltlichen Bestandteile durch generative KI analysiert und anschließend gespeichert, da hier die Stärken der KI besser genutzt werden können.

Die Pipeline verwendet eine Gemini 3.1 Pro (Preview) Schnittstelle und nutzt die Funktion des Structured Outputs. Mit dieser Funktion können vorhersagbarere Ergebnisse in einem definierten Datenschema erzielt werden, was das Extrahieren von Daten aus unstrukturierten Texten besser ermöglicht (vgl. Gemini API 2026). Da die Textausgabe strikt an die Struktur gebunden ist, kommt ein vorhersehbares und fehlerfreies Ergebnis zustande. Das Gemini 3.1 Pro (Preview) Modell ist nach einem Benchmark, entwickelt von White et al. (2025), Stand April 2026, eines der am besten performenden KI-Modelle auf dem Markt. Ziel der KI ist die Erstellung einer Übersicht über die Art und relevanten Inhalte der Veröffentlichung. Jede individuelle Veröffentlichung wird in einer unabhängigen API-Anfrage an die KI gesendet und in einer JSON-Datei gespeichert. Durch die Nutzung des Structured Outputs werden mehrere Prompts genutzt, die nacheinander für die KI freigeschaltet werden, so dass das KI-Modell nur relevante Aufforderungen für die individuellen Ausgabe-Aufgaben erhält. Das KI-Modell arbeitet also die Aufgaben nacheinander ab. Dabei wird erst ein übergeordneter Prompt verwendet, der das Thema und generelle Informationen liefert. Die darauffolgenden Aufforderungen für die Einzelnen Unter-Abfragen beziehen dabei den übergeordneten Prompt und das Dokument des wissenschaftlichen Artikels mit ein. Es entsteht folgende Struktur:

Übergeordneter Prompt	Du bist ein sehr kritischer Wissenschaftler, der zum Thema "Modellierung der Digitalen Souveränität KI-integrierender Unternehmen in der nachgelagerten (downstream) KI-Wertschöpfung: Entwicklung eines Bewertungs-, Reifegrad- und Gestaltungsmodells." Literatur untersucht. Bitte halt dich knapp und prägnant und beantworte alle Fragen NUR auf Basis der vorliegenden Literatur. Ziel dieser Analyse ist es die gefundene Literatur zu inhaltlich tief zu untersuchen und dabei alle wichtigen Aspekte für das Forschungsthema herauszuarbeiten. Falls einzelne Aspekte nicht in der vorliegenden Literatur aufzufinden sind, markiere diese mit einem Strich "/".		
Text des Dokuments der Veröffentlichung			
JSON	*Name*	*Typ*	*Description*
Sub-Prompts	Titel	String	Der exakte Haupttitel der Publikation.
	Autor	String	Liste alle Autoren kommagetrennt auf. Wenn unbekannt, schreibe 'Unbekannt'.
	Jahr	Int	Das Erscheinungsjahr als 4-stellige Zahl. Sonst 0.
	Typ	String	Typ der Forschungsarbeit. Wähle hier bitte aus den folgenden aus: conceptual paper, literature review, systematic review, …
	Forschungsmethode	String	Welche Forschungsmethode wird verwendet? Wähle bitte von diesen Beispielen aus: Systematic Literature Review, Fallstudie (Case Study), …
	Genutzte_Theorien_Frameworks	String	Welche Theorien und Frameworks werden hier genutzt? Was könnte davon relevant sein für unser Forschungsthema? Beispiele wären hier: conceptual Framework, maturity model, …
	Forschungskontext	String	In welchem Kontext kommt die Forschung zustande?
	Forschungsziel	String	Welches Ziel verfolgt die Veröffentlichung? Und wie lässt sich dieses Ziel mit dem Ziel unseres Forschungsthema verbinden? Bitte führe genau und präzise aus…
	Schluesselkonzepte_Dimensionen	String	Führt die vorliegende Literatur Schlüsselkonzepte oder Dimensionen an, die für unser Thema relevant sein könnten?...
	Modelle_Messung	Boolean	Werden Modelle erklärt?
	Erklaerung_Modelle_Messung	String	Werden Modelle erklärt? Wenn ja, welche und wie? Wie wird die Messung vorgenommen? Führe übersichtlich, umfangreich und präzise aus.
	…	…	…

Abbildung 1: Beispiel-Struktur des Prompts des Literaturscreening

Zur Optimierung der Extraktionsqualität durch das genutzte KI-Modell wurde „Role Prompting“ angewendet. Durch die Zuweisung der Rolle eines kritischen Wissenschaftlers wird das Sprachmodell angewiesen, das vortrainierte Sprachverständnis explizit Analytisch und Objektive zu halten (vgl. Steinmann und Piazza 2024). Dadurch soll eine tiefgreifende Extraktion der inhaltlich relevanten Elemente der akademischen Veröffentlichungen ermöglicht werden.

Auch wenn durch generative KI erhebliche Effizienzpotenziale für die Literaturrecherche und -Screening möglich sind, besteht das Risiko der Halluzination, bei der faktisch inkorrekte oder nicht im Quelltext enthaltene Details erzeugt werden (vgl. Reddy et al. 2024). Auch können Verzerrung bzw. Bias aus Trainingsdaten zu voreingenommener oder einseitiger Literaturextraktion führen (vgl. Walton

und Watkins 2024). Darüber hinaus besteht beim Einsatz von KI in der Forschung ein potenzieller Verlust der menschlichen Entscheidungsfähigkeit durch „Over-Reliance“, also zu starkem Vertrauen in die KI-Ausgaben, während der eigenen kritischen Analyse vernachlässigt wird (vgl. Ahmad et al. 2023). Um dieser Limitation entgegenzuwirken und wissenschaftliche Validität zu gewährleisten, wird ein strikter Human-in-the-Loop Ansatz verfolgt (vgl. Reddy et al. 2024). Das Strukturierte JSON-Ausgabeformat der KI dient hierbei ausschließlich der effizienten Vorfilterung und Strukturierung der Literatur. Jede in dieser Forschungsarbeit genutzte wissenschaftliche Veröffentlichung wird manuell am originalen Dokument verifiziert. So werden Halluzinationen oder algorithmische Verzerrungen vorgebeugt.

Das Vorgehen der Literaturanalyse orientiert sich an Webster und Watson (2002) und wird auf manuell überprüften Ergebnissen des KI-unterstützen Literaturscreenings angefertigt. Zentral ist hier eine Konzeptmatrix, die neben der Identifikation mit Autorenname und Veröffentlichungsjahr die Betrachtungsebene, Fokus, Beitrag und Risiken der in der Literatur behandelten Themen einordnet. Um eine übersichtliche und informative Darstellung in dieser Arbeit zu gewährleisten, werden in den Kategorien Abkürzungen genutzt. Dabei beschreibt die Betrachtungsebene, welche Stelle der Wertschöpfung die Veröffentlichung betrachtet. Darunter gibt es übergeordnet die politische digitale Souveränität auf nationaler oder regionaler Ebene und untergeordnet die Betrachtung der digitalen Souveränität in der nachgelagerten oder vorgelagerten KI-Wertschöpfung. Der Fokus der Veröffentlichungen unterscheidet zwischen Souveränität der Technologie, Daten oder Infrastruktur, während unter der Kategorie des Beitrags zwischen Bewertungsmodellen, konzeptionellen Frameworks oder Gestaltungsmodellen unterschieden wird. Schlussendlich wird noch aufgezeigt, welche Risiken die einzelnen Veröffentlichungen beschreiben. Hier kommen vor allem Risiken wie Vendor Lock-Ins, Black-Box-Modelle, Machtasymmetrien, Abhängigkeiten und Sicherheitsbedenken vor.

Kategorie	**Beudeutung der Abkürzung**
Betrachtungsebene	**P**olitisch, **N**achgelagert, **V**orgelagert
Fokus	**T**echnologische, **D**aten- & **I**nfrastrukturelle Souveränität
Beitrag	**B**ewertungsmodell (/Reifegradmodell), Konzeptionelles **F**ramework, **G**estaltungsmodell
Risiken	**V**endor Lock-In, **B**lack-Box, **M**achtasymmetrie, **A**bhängigkeiten, **S**ecurity

Abbildung 2: Legende für die Konzeptmatrix

4. Literaturarbeit

Der aktuelle Stand der Forschung wird mithilfe einer umfassenden, mit KI unterstützen Literaturrecherche erhoben. Auf Basis dessen sollen bereits vorhandene Modelle zu verwandten oder ähnlichen Themen beleuchtet und analysiert werden. Die folgenden Unterkapitel erläutern erst die gewählte Vorgehensweise bei der Literaturrecherche und Literaturanalyse, dann den daraus extrahierten Stand der Forschung und schlussendlich eine Betrachtung vergleichbarer Modellierungsansätze.

4.1. Stand der Forschung

Der aktuelle Stand der Forschung wurde mithilfe einer Konzeptmatrix nach Webster und Watson (2002) und einer KI unterstützen Literaturrecherche, angelehnt und erweitert an Dittmer und Alpers (2025), durchgeführt. Insgesamt wurden aus mehr als 140 Veröffentlichungen 26 für diese Arbeit relevanten Veröffentlichungen extrahiert.

Autor & Jahr	Betrachtungsebene	Fokus	Beitrag	Risiken
Bogenstahl & Zinke 2017	N	T, D, I	B	V, B, M, A, S
Jacobides et al. 2021	N, V	T, I	F	V, B, M
Gökalp & Martinez 2021	N	T, D, I	B	S
Kapoor et al. 2021	N, V	T, I	F	V, B, M
Moerel & Timmers 2021	P	T, D, I	F	V, B, M
Mayer & Lu 2022	P	D, I	B	V, A
Gawer 2022	P, N	T, D, I	F	V, B, M, S
Chellapandi et al. 2023	N, V	D, I	F	V, B, M
Fries et al. 2023	P, N	T, D, I	F	V, A, S
Glasze et al. 2023	P	T, D, I	F	V, B, M, S
Pi 2023	P, N, V	T	F	V, B, M
Kopka & Fornahl 2024	N	T	B	V, B
Heeks & Spiesberger 2024	P, N, V	I	F	V
Fratini et al. 2024	P	D, I	F	V, B, M
Farmakis et al. 2024	N	T	F	V, B

Klare & Lechner 2024	N, V	T, D, I	B, F, G	V, B, A
Tairov et al. 2024	N	T, D	F	V, B, M
Stürmer 2024	P	T, D, I	F, G	V, B, A
Ali et al. 2024	N	T	F	V, B, M
Thiel 2024	P	D	F	M, S
Aboushady et al. 2024	V	I	F	V, B, M
Abiade 2025	P, N	T, D, I	G, F	V, B, M, A, S
Celsi & Zomaya 2025	P, N, V	T, D, I	F, G	V, B, M
Hulko et al. 2025	P	T, D, I	F	V, B, M, A, S
Maathuis & Cools 2025	P, N	T, D, I	F, G	B, M, A, S
Wenzelburger & König 2025	P	T, D, I	F	V, M

Abbildung 3: Konzeptmatrix nach Webster und Watson (2002): Stand der Forschung

Der Stand der Forschung im März 2026 zur digitalen Souveränität in der KI-Wertschöpfung generell ist bislang wenig beleuchtet und wird aufgrund der rasch voranschreitenden technologischen, interdisziplinären und organisationalen Entwicklung als sehr dynamisch und fragmentiert angesehen (vgl. Glasze et al. 2023). Aktuell wird das Thema primär im geopolitischen und internationalen Rahmen auf Staaten Ebene diskutiert (vgl. Kapoor et al. 2021).

4.1.1. Geopolitische Entwicklung der digitalen Souveränität

In der wissenschaftlichen Literatur wird digitale Souveränität als ein umkämpftes, dynamisches und mehrdimensionales Konzept beschrieben, welches Daten-, Technologie-, Cybersicherheits- und rechtliche Souveränität beinhaltet (vgl. Hulk et al. 2025). Während frühere Ansätze den Fokus primär auf nationale Ebene legen, wird in neuerer Forschung der Fokus auf Akteure in der Wirtschaft, Zivilgesellschaft und staatlichem Zusammenwirken ausgeweitet (vgl. Fratini et al. 2024; Thiel 2024). Mayer und Lu (2022) sehen eine tiefe Asymmetrie in der globalen Abhängigkeit von digitaler Technologie europäischer Länder, während die USA wesentlich autonomer sind und China stetig aufholt. Dies belegt der Digital Dependance Index (DDI). Diese Abhängigkeiten werden als sehr risikoreich wahrgenommen und bergen das Potenzial eines „digitalen Kolonialismus“, was die Handlungsfähigkeit und Unabhängigkeit europäischer Unternehmen und Staaten beeinträchtigen kann (vgl. Glasze et al. 2023). Um dem entgegenzuwirken, erforscht die wissenschaftliche Gemeinschaft regulierende Instrumente wie dem EU AI Act oder dem Digital Service Act (DAS) um allgemeine Standards für Souveränität in digitalen

Services und Märkten zu etablieren (vgl. Hulko et al., 2025). Eine Diskursanalyse von Wenzelburger und König (2025) zeigt allerdings, dass der Begriff „digitale Souveränität“ eher als politisches Schlagwort genutzt wird, welches die Uneinigkeit in der EU überdecken und mithilfe vager Formulierungen einen Konsens erzeugen soll, obwohl die Interpretation zwischen den Staaten und Parteien individuell ist.

4.1.2. Infrastrukturelle und technologische digitale Souveränität

Auch wenn digitale Plattformen die dominante Organisationsform gilt, liegt ein grundlegendes Ungleichgewicht vor. Digitale Ökosysteme erfordern eine verteilte Wertschöpfung mit einer Vielzahl an unabhängig voneinander Arbeitenden Menschen und Organisationen, resultieren allerdings in einer zentralisierten Wertaneignung, hauptsächlich den Plattformbetreibern und Hyperscalern (vgl. Gawer 2022; Jacobides et al. 2021). Solche risikoreichen Machtkonzentrationen können zu Anreizen für Plattformbetreiber führen, eigenständig und privat Regulierungen durchzusetzen und wettbewerbswidrige Praktiken, „Überwachungskapitalismus“ und Ausnutzung der Plattformnutzer zu betreiben (vgl. Gawer 2022). Um solche Abhängigkeiten und Vendor Lock-Ins zu verhindern, werden der systematische Einsatz von Open Source Software, unabhängiger Cloud-Infrastruktur oder autonomer Rechenkapazität für erhöhte Souveränität empfohlen (vgl. Klare und Lechner 2024; Stürmer 2024).

4.1.3. Nachgelagerte KI-Wertschöpfungskette

Laut Heeks und Spiesberger (2024) ist die Forschung zur Struktur und der Generellen KI-Wertschöpfungskette bislang fragmentiert, nur auf spezifische Anwendungsfälle oder Hardware fokussiert. Daher ist besonders im Bereich der nachgelagerten Wertschöpfung von KI, also Anwendern und Entwicklern, die externe Foundation Models integrieren, ein Forschungsdefizit zu erkennen (vgl. Pi 2023, Heeks und Spiesberger 2024). Gerade bei generativer KI weist die Governance laut Pi (2023) weitreichende Defizite auf. Studien wie die Analyse chinesischer und europäischer Regularien zeigen, dass sich Verantwortungen zwischen Foundation Model Entwicklern bzw. Bereitstellern und nachgelagerten Anwendern oder Integratoren kaum unterscheiden. Insbesondere für kleine und mittlere Unternehmen (KMU) stellt dies eine riskante Einstiegshürde dar, da von ihnen erwartet wird, die Verantwortung für die Risiken ihrer Produkte zu tragen, obwohl sie keinen erweiterten Einblick in die Trainingsdaten oder die technologische Black-Box Architektur haben (vgl. Pi 2023). Andererseits argumentieren Kopka und Fornahl (2024), dass die empirische Forschung belegt, dass die Integration von general purpose KI fundamentale Produktivitätssteigerung insbesondere für KMU ermöglichen kann. Außerdem steigert KI neben Entscheidungspräzision und Automatisierung auch die Innovationsfähigkeit (vgl. Ali et al., 2024).

4.1.4. Technische Ansätze zur Datensouveränität

Chellapandi et al. (2023) schlagen zur Wahrung der Datensouveränität „Federated Learning“ vor. Dieses Konzeptversucht Datenschutz und Sicherheit durch dezentral trainierte KI-Modelle zu ermöglichen, bei dem keine sensiblen Rohdaten übertragen und zentralisiert gespeichert werden. Technologien wie Differential Privacy, homomorphe Verschlüsselung oder Blockchain können ergänzend eingesetzt werden. Die Autoren merken allerdings an, dass Herausforderungen wie Ressourcenbeschränkungen und ungenaue Systemheterogenität aktuell im Fokus der Forschung stehen.

4.2. Vergleichbare Modelle

Aktuell werden einige Modelle und Frameworks in Zusammenhang mit digitaler Souveränität in der Forschung beschrieben. Die hier näher beleuchteten Modelle sollen als Grundlage für die zu entwickelnden Artefakte dieser Forschungsarbeit dienen.

Dimensionen digitaler Souveränität

Um eine greifbarere Darstellung digitaler Souveränität zu ermöglichen, nutzt die Literatur vermehrt Modelle, die auf Schichten oder Dimensionen basieren. Hulko et al. (2025) sehen digitale Souveränität als dynamisches und mehrdimensionales Modell, welches aus den Kernebenen Datensouveränität, technologische Souveränität, Cybersicherheit und rechtlicher Souveränität besteht. Dabei beschreibt Datensouveränität die Kontrolle über das Datenmanagement (also Datensammlung, Datennutzung und Datenspeicherung), technologische Souveränität die Unabhängigkeit der genutzten Infrastruktur aber auch Technologie selbst, Cybersicherheit als Schutz vor Cyberangriffen und digitalen Krisen und rechtliche Souveränität die Fähigkeit, Normen und Regeln zu setzen und durchzusetzen.

In dem Bericht von Stürmer (2024) hingegen, lässt sich indirekt ein technologisches Schichtenmodell erkennen, bei dem allerdings zwischen notwendigen Systemabhängigkeiten, um nahtlosen und effizienten Datenaustausch zu ermöglichen, und kritischer Hersteller-Abhängigkeit wie einem Vendor Lock-In, unterschieden wird. Stürmer (2024) nennt vier aufeinander aufbauende Schichten, bei denen digitale Souveränität verloren gehen kann: Hardware, Software, Daten und KI-Modelle.

Maathuis und Cools (2025) erstellen ein Schichtenmodell, welches von strategischen Ziele bis zur technischen Ausführungsebenen immer granularer wird. Die Autoren formulieren fünf Ebenen: strategische Schicht, Governance und Assurance, Datensouveränität, KI-Souveränität, Cyber Operations und Infrastruktur. Für digitale Souveränität nachgelagerter Unternehmen in der KI-Wertschöpfung kann hier die Trennung von Daten- und KI-Souveränität. Neben dem Schutz der Daten, sollen auch die

integrierten KI-Systeme geschützt werden. Durch beispielsweise einen „Software Bill of Materials“ sollen Risiken in der Lieferkette der Daten und KI-Systeme minimiert werden.

Abiade (2025) hingegen erstellt zur Sicherung oder Wiedererlangung der algorithmischen Souveränität eine mehrschichtige Policy-Architektur, die aus einer institutionellen, technischen und kulturellen Schicht bestehen. Die Besonderheit hier ist die Querverbindung der technologischen Autonomie mit politischen und normativen Rahmenbedingungen. Außerdem Identifiziert Abiade (2025) auch Dimensionen zur Erfassung algorithmischer Souveränität: Operative Kontrolle, Rechtliche Kontrolle, Datensouveränität, Zivile Partizipation, Transparenz und Wiedergutmachung, sowie die Dimension des politischen Nutzens.

Fries et al. (2023) erstellen ein akteurszentriertes Schichtmodell digitaler Souveränität. Dabei werden drei Hauptebenen thematisiert: Staaten bzw. supranationale Institutionen, Organisationen und Individuen. Dabei wird ebenfalls auf eine Wechselwirkung zwischen diesen Schichten eingegangen, bei denen staatliche Regulierungen die Handlungsspielräume von Unternehmen beeinflussen, welche wiederum durch technologische Anpassung die Souveränität des Individuum beeinflussen.

Bewertungs- und Messmodelle

Mayer und Lu (2022) erstellen mit dem Digital Dependence Index (DDI) ein quantitatives Framework zur Messung digitaler Abhängigkeiten auf einer globalen Ebene. Dieses Framework misst das Ausmaß, in dem Akteure in einem Land von ausländisch kontrollierten Daten angewiesen sind. Dabei bewertet der Index Länder auf einer Skala von 0 (vollständig Souverän) bis 1 (vollständig Abhängig). Dabei werden die Faktoren Hardwareimporte, Softwareimporte und Intellektuelles Eigentum bzw. Patentbesitz ausländischer Unternehmen betrachtet. Dieses Framework betrachtet digitale Souveränität als strukturierte globale Abhängigkeitsbeziehung, fokussiert sich allerdings primär auf Volkswirtschaften und geopolitische Messungen.

Fratini et al. (2024) analysieren wissenschaftliche Publikationen zu staatlichen digitalen Souveränitätsstrategien und erstellen dabei vier Idealtypen digitaler Souveränität, die anhand von Kriterien robuster digitaler Governance bewertet werden. Die Veröffentlichung nennt mehrere Governance-Mechanismen, die digitale Souveränität robust machen. Darunter finden sich Testumgebungen für neue Technologien (Regulatory Sandboxes), Skalierbare digitale Infrastruktur, risiko-orientierte Regulierung (Risk based Regulation), Modularisierte digitale Systeme, sowie Redundanzen in der Infrastruktur. Die Autoren benennen schlussendlich als Ziel ein Gleichgewicht zwischen langfristiger Stabilität und kurzfristiger Anpassungsfähigkeit.

Aus der Betrachtung der relevantesten Literatur zu Bewertungs- und Messansätzen wird zwar eine hohe Abstraktionsebene deutlich, jedoch sind diese zu geopolitisch für die Anwendung auf Un-

ternehmensebene. Daher ist der Transfer von staatlichen hin zu unternehmerischen Dimensionen notwendig. Für KI-integrierende Unternehmen fehlt somit in erster Linie eine praxistaugliche Messmöglichkeit, die technologische und strategische Abhängigkeiten für Unternehmen bewertbar und greifbar macht.

Reifegradmodelle

Bogenstahl und Zinke (2017) erstellen ein mehrdimensionales Reifegradmodell. Dieses Modell soll zeigen, dass digitale Souveränität nicht durch einzelne Maßnahmen entstehen kann. Das Modell basiert auf drei aufeinander aufbauenden Dimensionen. Dabei bildet eine leistungsfähige und interoperable IT-Infrastruktur zur Vermeidung von Lock-In-Effekten die Basis. Darauf aufbauend folgt die Kompetenzdimension, die auf Software-, Hardware-, Sicherheits- und Datenkompetenz beruht. Auf diesen beiden Dimensionen baut die dritte Dimension, die Wertschöpfungssouveränität, auf, welches die Möglichkeiten und Kapazitäten zur eigenen digitalen Innovation bedeutet. Unter ökonomischer Betrachtung KI-integrierender Unternehmen sagt dieses Modell aus, dass die Grenzproduktivität von KI-Investitionen sinkt, wenn parallel nicht proportional in Souveränität investiert wird.

Gökalp und Martinez (2021) ermöglichen mit einem „Digital Transformation Capability Maturity Model“ die Messung von organisationalen Fähigkeiten digitale Transformationen durchzusetzen. Dabei werden eine Prozessdimension und eine Fähigkeitsdimension verwendet. Die Prozessdimension ist unterteilt in Governance, IT, Prozesse und Personalmanagement, während die Fähigkeitsdimension diese Prozesse mittels fünf Stufen bewertet. Hinsichtlich nachgelagerter Unternehmen können sich Souveränitätsprozesse wie Lieferantenmanagement, Kontrolle der Abhängigkeiten, Data Governance und Security Management ableiten lassen.

Im Kontext von Daten und KI entwickeln Celsi und Zomaya (2025) eine Implementierungs-Roadmap für ethische KI und nutzen dabei ein Reifegradmodell nach dem DAMA-Framework, um die Reifegrade von Daten- und KI-Anwendungen zu bewerten. Die Implementierungs-Roadmap ist in vier aufeinanderfolgenden Phasen eingeteilt. Die erste Phase ist die Entwurfs- oder Designphase, in der Fairness und Inklusion ermittelt und durchgesetzt werden sollen. Gefolgt von der Entwicklungsphase, in der auf Verantwortlichkeiten und Erklärbarkeit fokussiert wird. Nach der Einsatzphase bei der Transparenz und Sicherheit sichergestellt werden sollen, kommt die letzte Steuerungs- und Überwachungsphase, die Risikobewertung und Mitverantwortung beinhalten.

Das DAMA-Framework ist ein Rahmenwerk für Best Practices im Datenmanagement, welches Potenzial hat auch für digitale Souveränität erweitert zu werden. Das Framework betrachtet die Daten Governance als zentralen Themenbereich, welcher aus zehn Fokusbereichen des Datenmanagement besteht. Darunter befinden sich: Data Architecture, Data Modeling & Design, Data Storage & Opera-

tions, Data Security, Data Integration & Interoperability, Document & Content Management, Reference & Master Data, Data Warehousing & Business Intelligence, Metadata Management und Data Quality (vgl. Sigi 2022).

Gestaltungsmodelle und Frameworks

Klare und Lechner (2024) entwickeln ein sogenanntes „Digital Sovereignty Model from the perspective of IT-Consultancy“ (DSMIC), dass eine Modellierung digitaler Souveränität durch die Strukturierung von Einflussfaktoren in drei Ebenen ermöglicht. Dabei wird unter den Ebenen Unternehmensinterne Themen, Querschnittsthemen und externe Themen unterschieden. Unternehmensinterne Ebenen sind vom Unternehmen direkt steuerbare Faktoren wie beispielsweise IT-Sicherheit durchsetzen, digitale Infrastruktur anpassen oder die Kommunikation verbessern. Querschnittsthemen haben einen Einfluss auf Intern und Extern wie beispielsweise die Berücksichtigung der Nachhaltigkeit, Förderung der Agilität oder Stärkung digitaler Souveränität. Externe Themen sind Faktoren aus dem Umfeld des Unternehmens, die es beeinflussen oder nutzen kann, wie beispielsweise regionale Infrastruktur oder Lieferketten. Diese Faktoren der einzelnen Themenfelder sollen schlussendlich die Relevanz und Priorität der Handlungsfelder gewichten, Anpassung an Bedrohungsszenarien bieten, den Ist-Zustand der digitalen Souveränität des Unternehmens bewerten und schlussendlich als Grundlage für fundierte strategische Entscheidungen dienen. Zusätzlich dazu wird ein Maßnahmenkatalog mit sieben konkreten Maßnahmen beigefügt.

Daneben fokussieren sich andere Veröffentlichungen auch auf Handlungsempfehlungen für die Gestaltung organisationaler Souveränität, die einen direkten Transfer zu KI-integrierenden Unternehmen ermöglichen. Abiade (2025) warnt beispielsweise vor operationellen Risiken, unkontrollierter Systemupdates bei Software-Schnittstellen oder Software-as-a-Service Lösungen. Sollte ein vorgelagerter KI-Entwickler das Recht besitzen, Modelle aus der Ferne zu modifizieren, hat das anwendende Unternehmen seine technische und operative Handlungsfähigkeit verloren und ist damit allen Black-Box Risiken ausgesetzt. Daher fordert Abiade (2025) die Etablierung souveräner IT-Beschaffung, die mindestens Software und Hardware umfasst. Diese souveräne Beschaffung (engl.: „Sovereign Procurement“) muss Beschaffungsrichtlinien, offene Schnittstellen priorisieren, Datenhoheit rechtlich absichern und unabhängigen Audits unterliegen. Fries et al (2023) ergänzen dies mit dem Vorschlag der technologischen Vielseitigkeit wie beispielsweise Multi-Cloud-Strategien oder Verwendung von Open Source Technologien. Zudem muss für eine nachhaltige Gestaltung der digitalen Souveränität insbesondere KI-integrierende Unternehmen umfangreiche

Schulungen für Mitarbeiter ermöglichen, um kritische Evaluation von KI-Systemen zu ermöglichen und Sicherheitsrisiken, bedingt durch fehlende Kompetenz der Mitarbeiter, zu minimieren.

Zur technischen und architektonischen Realisierung von Datensouveränität schlägt Stürmer (2024) die Integration von sogenannten CARE-Prinzipien (Collective Benefit, Authority to Control, Responsibility, Ethics) vor. Ergänzt kann dies durch Aboushady et al. (2024) werden, die vorschlagen, Architekturen nach dem Resilient-Trust-Modell in sicherheitskritischen IoT-, Software und KI-Systemen einzusetzen. Dieses Modell nutzt modulare Sicherheitsbausteine, die auf Resilienz und Vertrauen basieren. Chellapandi et al. (2023) schlagen dezentrales KI-Training, Federated Learning, homomorphe Verschlüsselung, Differential Privacy und Blockchain als Architekturmodell vor.

Zusammenfassend zeigt die momentane Forschung zwar die Existenz fundierter Ansätze für staatliche Akteure (vgl. Mayer und Lu 2022; Fratini et al. 2024) oder zur generellen IT-Transformation (vgl. Bogenstahl und Zinke 2017; Gökalp und Martinez 2021), eine spezifische Modellierung für nachgelagerte KI-integrierende Unternehmen bleibt bislang jedoch aus. Auffallend ist zudem, dass bestehende Frameworks die starken und asymmetrischen Abhängigkeiten vorgelagerter KI-Entwickler oder Ökosystembetreiber nicht berücksichtigen. Daraus ergibt sich ein Forschungsdefizit, das die Notwendigkeit ableiten lässt, politische und unzureichend spezifische Konzepte der digitalen Souveränität für KI-integrierende Unternehmen zu transferieren. Die in diesem Kapitel erlangten Erkenntnisse bieten die Grundlage für die darauffolgende Erstellung der Dimensionen digitaler Souveränität.

5. Ableitung der Dimensionen digitaler Souveränität

5.1. Methodisches Vorgehen

Die Ableitung der Dimensionen digitaler Souveränität erfolgt auf Basis der qualitativen Inhaltsanalyse nach Mayring und Fenzel (2019). Dadurch soll eine reproduzierbare und überprüfbare Textanalyse ermöglicht werden, bei der Inhalte der wissenschaftlichen Veröffentlichungen strukturiert schrittweise in Kategorien abgeleitet werden. Die Auswertungseinheit besteht aus identifizierter wissenschaftlicher Literatur, die nach manuellem Screening der KI-unterstützen Literaturübersicht als brauchbar eingestuft wird. Die Kontexteinheit bezieht sich auf den inhaltlichen Hintergrund und Zusammenhang genereller digitaler Souveränität jeder Publikation. Die Kodiereinheit besteht dagegen aus den kleinsten auszuwertenden Bestandteilen wie Schlüsselkonzepten und Aussagen über Abhängigkeiten, Kontrolle oder Governance.

Die eigentliche Extraktion der Dimensionen erfolgt nach Mayring und Fenzel (2019) und fängt mit dem Schritt der Selektion an. Dazu wird, aufbauend auf der KI-unterstützen Literaturrecherche und der daraus resultierenden Literaturübersicht die extrahierten Rohdaten manuell gesichtet. Dokumente, die Titel und Inhalte enthalten, die darauf hindeuten, einen Mehrwert für die Bildung der Dimensionen liefern zu können, werden daraufhin gesammelt. Dabei wird insbesondere auf Aussagen zur Abhängigkeit von Organisationen und generellen Souveränitätsaspekten geachtet. Der zweite Schritt, zur induktiven Kategorienbildung, besteht darin, aus den Publikationen erste prototypische Kategorien abzuleiten, die relevant für die Handlungsfähigkeit KI-integrierender Unternehmen in der nachgelagerten KI-Wertschöpfung sein können. Als nächster Schritt wird eine Bündelung und Generalisierung der zuvor herausgearbeiteten Kategorien vollzogen, bei der übergeordnete Kategorien entstehen. Schlussendlich werden im letzten Schritt werden die Kategorien zu Dimensionen zusammengeführt, welche das finale Kategoriensystem bilden und dabei als konzeptionelles Fundament für das anschließende Bewertungs- und Reifegradmodell dienen.

Zur Sicherstellung der inhaltlichen Validität ist eine methodische Begründung dafür notwendig, weshalb das komplexe Konstrukt der digitalen Souveränität auf exakt sechs Dimensionen eingegrenzt wird. Da der bisherige Diskurs oft geopolitische Aspekte umfasst, konzentriert sich diese Arbeit bei der Kategorienbildung ausschließlich auf die organisationale Mikroebene. Makroökonomische Faktoren, die außerhalb des direkten Handlungsspielraums KI-integrierender Unternehmen liegen, werden ausgeklammert. Die Begrenzung auf genau sechs Dimensionen resultiert aus dem qualitativen Prinzip der theoretischen Sättigung (vgl. Saunders et al., 2018). Im von Mayring und Fenzl (2019) beschriebenen Prozess wurde der Punkt erreicht, an dem neue Literaturquellen keine grundlegend neuen Dimensionen mehr hervorbrachten, sondern sich die gewonnenen Erkenntnisse lediglich als Unterkategorien in das bestehende Raster integrieren lassen.

5.2. Dimensionen digitaler Souveränität

Auf Basis der vorangehenden Literaturrecherche wurde im Rahmen der qualitativen Inhaltsanalyse nach Mayring und Fenzl (2019) die relevantesten Veröffentlichungen selektiert. Im Fokus stehen dabei Publikationen, die sich insbesondere mit Governance, Reifegradmodellen und der nachgelagerten KI- bzw. IT-Wertschöpfung befassen. Das Resultat dieses Syntheseprozesses sind sechs Meta-Dimensionen für digitale Souveränität in der nachgelagerten KI-Wertschöpfungskette mit Fokus auf KI-integrierende Unternehmen. Die folgende Tabelle bietet nun eine Übersicht der Dimensionen mit entsprechenden Kategorien. In den Folgenden Unterkapiteln werden diese dann genauer wissenschaftlich beleuchtet.

Dimension	Kategorien
1. Technologische & Infrastrukturelle Souveränität	1. Infrastrukturelle Kontrolle und Cloud Abhängigkeiten 2. Interoperabilität und Modularität 3. Operative Kontrolle und Resilienz
2. Daten- & Informationssouveränität	1. Eigentum und Kontrolle der Daten und Informationen 2. Datensicherheit und Vertraulichkeit 3. Verfügbarkeit und Transparenz
3. Organisationale Souveränität & KI-Governance	1. Strategische Handlungsfähigkeit 2. Prozessbasierte und strukturelle Governance 3. Risikomanagement
4. Kompetenzen & Akteure	1. Bewertungskompetenz und Entscheidungssouveränität 2. Talentverfügbarkeit und Ressourcen 3. Relationale Praktiken und Domänenwissen
5. Ökosysteme & Machtverhältnisse	1. Strukturelle Macht und Plattform-Abhängigkeit 2. Rollenverteilung im sozio-technischen Ökosystem 3. Kontrolle über Engstellen in der Wertschöpfungskette
6. Regulatorische & Normative Souveränität	1. Regulatorische und rechtliche Compliance 2. Normative Verantwortung 3. Transparenz, Fairness und Verantwortlichkeit

Abbildung 4: Übersicht der Dimensionen digitaler Souveränität

5.2.1. Technologische und Infrastrukturelle Souveränität

Die Dimension der technologischen und infrastrukturellen digitalen Souveränität beschreibt die architektonische, physische und operative Kontrolle einer Partei über die eingesetzten IT-Systeme und KI-Technologien. Da KI-integrierende Unternehmen keine eigene Hardware wie Computer-Chips oder

Basistechnologie wie Foundation Models von Grund auf Entwickeln, bedeutet technologische Souveränität in diesem Aspekt den Bedarf nach Modularität mit architektonisch und operativ gestaltbarer Unabhängigkeit. Nach Analyse identifizierter Literatur lassen sich drei zentrale Kategorien ableiten:

1. Infrastrukturelle Kontrolle und Cloud Abhängigkeiten

Grundlage der KI-Integration ist die Rechen- und Speicherinfrastruktur der Modell-Provider. Sheikh (2022) strukturiert digitale Souveränität in einem vertikalen Schichtenmodell, bei dem besonders die Cloud- und Intelligenzschichten die kritischen Ebenen darstellen. Diese beiden Schichten beinhalten die Entwickler der KI-Modelle und Speicherinfrastruktur selbst und stellen das höchste Risiko zur Abhängigkeit dar. Da nachgelagerte KI-integrierende Unternehmen auf diese wenigen großen Hyperscaler angewiesen sind, besteht ein hohes Risiko von Daten- und Technologie-Lock-Ins (vgl. Armbrust et al. 2010). Die Europäische Kommission (2025) definiert die Technologiesouveränität daher als Fähigkeit, Abhängigkeiten von (meist ausländischen) einzelnen Infrastrukturanbietern zu reduzieren. Hinsichtlich des Bewertungsmodells kann also abgeleitet werden, dass die Souveränität eines Unternehmens davon abhängt, welche Strategie verfolgt wird. Dabei kann übergeordnet zwischen Multi-Cloud Strategien, also verschiedene Anbieter aus verschiedenen Ländern, souveräne und europäische Single-Cloud Strategien, also ein vertrauenswürdiger Anbieter aus naher Umgebung, oder einer gemischten Strategie, also einem Mix aus beiden, entschieden werden, sofern der Einsatz einer selbst verwalteten Infrastruktur nicht infrage kommt.

2. Interoperabilität und Modularität

Um Lock-In-Szenarien vorzubeugen, sollte die IT-Architektur KI-integrierender Unternehmen flexibel gestaltet sein. Fries et al. (2023) identifizieren technologische Standards, Interoperabilität und die daraus resultierende Wahlfreiheit als fundamentale Voraussetzungen für digitale Souveränität. Auf die Praxis bezogen bedeutet das einen modularen Aufbau der Softwarelösung, bei der die Architektur so flexibel und modular gestaltet ist, dass ein einfacher Austausch oder Ersatz einzelner Elemente vorgenommen werden kann, ohne einen Zusammenbruch des Kernsystems zu verursachen. So sollten beispielsweise KI-Schnittstellen von OpenAI schnell durch Schnittstellen von Anthropic oder Google austauschbar sein. Darüber hinaus können auch Open Source Technologien, proprietäre Black-Box Modelle substituieren und zu erhöhter Souveränität führen (vgl. Hynninen und Maunula 2025). Eine technologische Souveränität zeichnet sich demnach durch eine von externen Schnittstellen entkoppelte Architektur ab, die einen schnellen Wechsel zwischen verschiedenen KI-Anbietern ermöglicht.

3. Operative Kontrolle und Resilienz

Die dritte Kategorie adressiert die Aufrechterhaltung des Geschäftsbetriebs bei Nutzung externer KI-Schnittstellen. Hier ist die Operative Kontrolle kritisch (vgl. Abiade 2025). Operative Kontrolle

beschreibt die Abhängigkeit von externen Anbietern hinsichtlich Wartung, Updates und genereller Systemverfügbarkeit, welche schlimmstenfalls die technische Eigenständigkeit des Unternehmens einschränkt. Hopkins et al. (2025) merken an, dass die Kontrolle über KI-Systeme über mehrere Organisationen verteilt sein können, was die Resilienz nachgelagerter Unternehmen gegenüber Ausfällen oder unvorhergesehener Updates stark beeinträchtigt. Aus diesem Grund sollte ein souveränes Unternehmen Mechanismen wie Fallback-Modelle oder lokales Hosting kritischer Systemkomponenten etablieren, um operative Integrität und Aufrechterhaltung des eigenen Produkts oder der eigenen Dienstleistung zu gewährleisten.

5.2.2. Daten- und Informationssouveränität

Daten bilden die Grundlage jeglicher Nutzung von KI. Im Falle nachgelagerter KI-integrierender Unternehmen kommt der Einsatz künstlicher Intelligenz häufig mit fachspezifischem Wissen oder Unternehmensdaten vor, indem Basismodelle trainiert oder feinabgestimmt werden. Daher beschreibt die Daten- und Informationssouveränität die Fähigkeit von Unternehmen, die rechtliche und aktive Kontrolle über diese Daten und Informationsflüsse zu wahren. Auch hier kann sich diese Dimension in drei Kategorien einteilen lassen:

1. Eigentum und Kontrolle der Daten und Informationen

Grundlegend wird die Datensouveränität eines Unternehmens durch die absolute Verfügbarkeit und Entscheidungsgewalt über die eigenen Datenbestände definiert, weshalb Hummel et al. (2021) Kontrolle und Eigentum als zentrale Kategorien der Datensouveränität bezeichnen. Unter Anbetracht KI-integrierender Unternehmen besteht das primäre Risiko also darin, dass proprietäres Fachwissen, durch die Nutzung von externen KI-Schnittstellen ohne Zustimmung oder Wissen in die Trainingsdaten der vorgelagerten Basistechnologieentwickler oder anderer Dritter, gerät. Konkret kann das den eigenen Wettbewerbsvorteil stark schwächen. Beyerer et al. (2018) beschreiben Datensouveränität aus diesem Grund als die selbstbestimmte Kontrolle über die Erhebung, Verarbeitung und Weitergabe organisationsspezifischer Daten. Hinsichtlich des Reifegradmodells bietet es sich demnach an, die Eigentumsverhältnisse der Daten sowohl vertraglich und technisch zu bewerten.

2. Datensicherheit und Vertraulichkeit

Während in 1. die strategische Verfügungsgewalt betrachtet wird, muss hier die technische Sicherheit der Daten gewährleistet sein. Insbesondere Kleine und Mittlere Unternehmen streben laut Hynninen und Maulula (2025) lokale Open Source Lösungen an, um Datensicherheit und Vertraulich-

keit, gerade bei sensiblen Kunden- und Geschäftsdaten, zu wahren. Auch legt die Europäische Kommission (2025) die Sicherheits- und Compliance-Souveränität in ihrem Cloud Sovereignty Framework fest. Dadurch soll der Schutz vor Cyberangriffen und unbefugtem Zugriff garantiert werden. Für nachgelagerte Akteure, die KI integrieren wollen, erfordert dies die Implementierung robuster Sicherheitsarchitekturen, um die Vertraulichkeit der Trainingsdaten, KI-Ein- und Ausgaben zu schützen. Beispiele könnte hier Datenanonymisierung vor API-Nutzung oder Confidential Computing sein.

3. Verfügbarkeit und Transparenz

Nach Augsberg und Gehring (2022) kann Datensouveränität nicht nur als defensives Konstrukt verstanden werden, sondern auch als Funktion zur Verfügbarmachung. Laut den Autoren ist ein Unternehmen nur dann souverän, wenn Datensilos aufgebrochen und die Daten nutzbar gemacht werden können, wie beispielsweise für das Training von KI. Dies erfordert dafür ein hohes Maß an Transparenz über die Datenflüsse und ein Unternehmen muss jederzeit nachvollziehen können, woher welche Daten stammen, wie diese genutzt oder durch KI verändert werden und welche rechtlichen Bedingungen für die Nutzung und Verarbeitung gelten (vgl. Hummel et al. 2021).

5.2.3. Organisationale Souveränität und KI-Governance

Die zuvor beschriebenen Dimensionen sind erst dann umsetzbar, wenn die entsprechenden internen Strukturen und Prozesse darauf abgestimmt sind. Die organisationale Souveränität beschreibt die Fähigkeit eines Unternehmens, KI-Technologien strategisch, strukturiert, sicher und zielgerichtet einzusetzen. In der nachgelagerten KI-Wertschöpfungskette ist eine robuste KI-Governance unverzichtbar, um den Verlust der Handlungsfähigkeit des Unternehmens, bei Integration externer Basistechnologie in Geschäftsprozesse und Produkte, zu verhindern. Auch diese Dimension kann in drei Kategorien untergliedert werden:

1. Strategische Handlungsfähigkeit

Grundlage der organisationalen Souveränität bildet die strategische Ausrichtung durch das Top-Management. Gökalp und Martinez (2021) beschreiben die strategische Governance als einen zentralen Aspekt digitaler Reife. Dabei geht es um die Fähigkeit, eigenständige Strategien zu definieren und technologische Abhängigkeiten bewusst zu managen. Dabei sollte die Wertschöpfungssouveränität, also die Sicherung der Innovations- und Wettbewerbsfähigkeit, trotz Nutzung externer Technologien erhalten bleiben (vgl. Bogenstahl und Zinke 2017). Für KI-integrierende Unternehmen bedeutet das, bewusste die Entscheidungen zu treffen, ob externe Technologie oder Dienstleistung eingekauft oder selbst entwickelt oder bereitgestellt wird. Hier können beispielsweise Effizienzgründe zu einem Zukauf

externer Technologie oder strategische Gründe für ein lokales Open-Source-Modell sprechen. Hinsichtlich eines Bewertungsmodells kann diese Kategorie durch die Existenz einer formalen KI-Strategie, aktives Lieferanten-Management oder Bereitstellung von Budgets für KI-Entwicklung oder KI-Bereitstellung operationalisiert werden.

2. Prozessbasierte und strukturelle Governance

Neben der Strategie (1.) sind die konkreten Prozesse zur Umsetzung und Überwachung dieser nicht zu vernachlässigen. Daher unterteilen Papagiannidis et al. (2023) KI-Governance in prozedurale und strukturelle Praktiken über den gesamten KI-Lebenszyklus hinweg. Dieser KI-Lebenszyklus besteht unteranderem aus der Datenmigration, Evaluation von KI-Systemen bis hin zur vollständigen Automatisierung von Prozessen. Hinsichtlich der nachgelagerten KI-Wertschöpfung befasst sich dies primär mit Machine Learning Operations, welches aufgrund der fehlenden Kontrolle über die externen Basismodelle, eine Implementierung von Prozessen zur Qualitätssicherung und Überwachung erfordert. Darunter zählen beispielsweise Monitoring Prozesse der Modelle in Bezug auf Leistungsabfälle, Erkennung von Halluzinationen oder Management von Schnittstellen-Updates seitens der KI-Provider. Ein hoher Reifegrad kann sich in dieser Dimension durch automatisierte Test- und Validierungssysteme zur Sicherung der resilienten Dienstleistung auszeichnen.

3. Risikomanagement

Die Integration von externen Systemen geht immer mit gewissen Risiken einher. Maathuis und Cools (2025) formulieren den Bedarf eines strikten Risikomanagements und klarer Verantwortlichkeiten, um Bedrohungen frühzeitig zu identifizieren und präventiv zu neutralisieren. Außerdem führen Abhängigkeiten von Kapital und Rechenleistung in der KI-Industrie zu systemischen Risiken und Vulnerabilitäten (vgl. Pirrone et al. 2025). Auf KI-integrierende Unternehmen übertragen bedeutet also organisationale Souveränität die Etablierung von Notfallplänen, die entscheiden, wie beispielsweise Datenvergiftung (Injektion manipulierter Daten in KI-Modelle), Prompt-Manipulation oder plötzliches Wegfallen der KI-Basismodelle, gehandhabt wird. Ein Reifegradmodell kann möglicherweise erfassen, ob KI-spezifische Risikobewertung standardmäßig in herkömmliche IT-Sicherheitsprozesse eingebunden ist.

5.2.4. Kompetenzen und Akteure

Neben technologischer und organisationaler Souveränität ist auch das Verständnis bzw. die Kompetenz zur Bewertung und Steuerung dieser unabdingbar. Die Dimension Kompetenz und Akteure fokussiert sich auf den menschlichen Faktor der digitalen Souveränität. Um keine Wissensasymmetrie

entstehen zu lassen, muss darauf geachtet werden, dass nicht das gesamte Implementierungswissen und die gesamte Fachkompetenz an externen Technologie-, Dienstleistungs- und Softwareprovidern abwandert. Diese Dimension unterteilt sich ebenfalls in drei Kategorien:

1. Bewertungskompetenz und Entscheidungssouveränität

Die Fähigkeit, KI-Systeme und deren Ausgaben kritisch zu hinterfragen, ist für eine reflektierte Unabhängigkeit essenziell. Diese sogenannte „Entscheidungssouveränität" beschreibt die Nachvollziehbarkeit und Beeinflussbarkeit von Entscheidungen autonomer Systeme (vgl. Beyerer et al. 2018). Bogenstahl und Zinke (2017) ergänzen dies mit ihrer „Kompetenzdimension", die die digitale Bewertungs- und Analysefähigkeit der Nutzer und Integratoren in den Mittelpunkt stellt. Anwender in der nachgelagerten KI-Wertschöpfungskette müssen also befähigt sein, KI-generierte Ergebnisse auf Korrektheit und Bias prüfen können. Die von Joecks-Laß (2023) genannten Dimensionen zur digitalen Selbstbestimmung lassen sich hierbei gut auf die Entscheidungssouveränität übertragen. „Kompetenz" bezieht sich auf das Wissen und die Fähigkeit, KI-Ausgaben zu verstehen und zu bewerten. „Informiertheit" beschreibt das Verständnis und den Wissensgrad, Chancen, Risiken und Handlungen von und bei KI-Systemen zu bewerten. Für ein Reifegradmodell kann die Bewertungskompetenz und Entscheidungssouveränität beispielsweise durch den Umfang oder den Automatisierungsgrad von KI-Schulungen oder etablierten Validierungsmechanismen gemessen werden.

2. Talentverfügbarkeit und Ressourcen

In gegenwärtiger, wachsender KI-Industrie ist der Zugang zu hochqualifizierten Fachkräften ein kritischer Engpass. Daher gilt der Mangel an Talenten als eine systemische Verwundbarkeit für Unternehmen (vgl. Pirrone et al. 2025). Für souveränes Agieren benötigen KI-integrierende Unternehmen interne Experten wie Data Scientist oder KI-Ingenieure, die die technische Architektur und Integration externer Technologie an die eigenen Vorhaben anzupassen. Auslagerung dieser Rollen bedeuten einen Verlust technologischer Gestaltungsfähigkeit und damit auch ein Verlust der digitalen Souveränität. Ein Bewertungsmodell sollte daher ein Verhältnis von internen und ausgelagerten KI-Experten und Beratern, sowie Rekrutierungs- und Weiterbildungsmaßnahmen erfassen.

3. Relationale Praktiken und Domänenwissen

Die KI-Integration umfasst mehr als nur das reine IT-Projekt, sondern benötigt eine Verknüpfung zu dem fachlichen Kerngeschäft. Daher betonen Papagiannidis et al. (2023) die Wichtigkeit relationaler bzw. Beziehungsbezogener Praktiken in der KI-Governance. Damit ist insbesondere die Einbindung von Domänenexperten und die Vermittlung von Zielen und Anforderungen über Abteilungsgrenzen hinweg, gemeint. In der nachgelagerten KI-Wertschöpfungskette entsteht der eigentliche Geschäftswert erst

durch die Vereinigung des spezifischen Fachwissens des Unternehmens und den generischen Fähigkeiten der Basistechnologie. Souveränität bedeutet demnach den Verbleib des Domänenwissens innerhalb des Unternehmens aber auch das aktive Einfließen der unternehmensbezogenen Fachkompetenz in die Weiterentwicklung der KI-Modelle. Für ein Reifegradmodell kann diese Kategorie durch die Messung der Abteilungsübergreifenden Zusammenarbeit oder Existenz cross-funktionaler Teams, sowie die Einbindung der Endanwender in den Entwicklungsprozess bewertet werden.

5.2.5. Ökosysteme und Machtverhältnisse

Diese Dimension bezieht sich auf die strategische Positionierung von Unternehmen innerhalb von Plattform-Ökosystemen bzw. plattformabhängiger Wertschöpfung. Genauer beleuchtet wird, wie Unternehmen mit den Machtstrukturen von KI-Anbietern umgehen. Denn Unternehmen sind oft bei Integration von KI nicht autark oder isoliert, sondern werden Teil von komplexen, asymmetrischen Wertschöpfungsnetzwerken, von denen zwar Unternehmen in vielen Hinsichten profitieren können, jedoch starke Abhängigkeiten eingehen müssen. Ein Beispiel hierfür könnte das Ökosystem des Apple Appstores sein, bei dem ein KI-integrierendes Unternehmen zwar einen Vorteil der einfachen Verfügbarkeit für Nutzer hat, jedoch strengen Regularien und Transaktionsbeteiligungen unterliegt. Für Akteure der nachgelagerten KI-Wertschöpfung ist das Verständnis von diesen Ökosystemen essenziell, da die eigene Wettbewerbsfähigkeit hinsichtlich der Gefahr vollständiger Abhängigkeit, auf dem Spiel steht. Auch diese Dimension kann in drei Kategorien aufgeteilt werden.

1. Strukturelle Macht und Plattform-Abhängigkeit

Die globale KI-Wirtschaft, die von wenigen großen Hyperscalern bzw. Technologiekonzernen dominiert wird, wird von Srivastava und Bullock (2024) mittels drei Machtdimensionen analysiert. Diese Machtdimensionen ermöglichen es, die Machtausübung der wenigen Hyperscaler in ihre Bestandteile zu zerlegen. Die größte Gefahr für unternehmerische digitale Souveränität geht laut den Autoren von der strukturellen Macht aus, bei der die Plattformanbieter durch ihre Position und Kontrolle über die materiellen Ressourcen wie Rechenleistung und Basismodell-Technologie direkte Macht auf die Anwender und Integratoren ausüben. Zwar ermöglichen digitale Ökosysteme zwar eine verteilte Wertschöpfung, zentralisiert allerdings die Wertaneignung (Value Capture) jedoch stark bei den Plattform Betreibern (vgl. Gawer 2022). Daher müssen sich KI-integrierende Akteure der Gatekeeper-Funktion dieser Bereitsteller bewusst sein. Für das Bewertungsmodell kann diese Kategorie durch eine Analyse der Konzentration an Lieferanten bzw. Vendor Lock-In-Gefahr, sowie der strategischen Diversifizierung von Technologiepartnern und deren Substituierbarkeit messbar gemacht werden.

2. Rollenverteilung im sozio-technischen Ökosystem

Die Definition der Rolle eines Unternehmens in einem Ökosystem ist voraussetzend für die Bewertung der eigenen digitalen Souveränität. Kapoor et al. (2021) nutzt einen sozio-technischen Ansatz, bei dem die Akteure von Plattform Ökosystemen in die Rollen „Platform Leader“, welche die Anbieter der KI Basis Modelle bzw. Eigentümer der Ökosysteme sind, „Complementors“, welche die Integratoren und Anwendungsentwickler der nachgelagerten Wertschöpfung sind, und Endbenutzer. Bei KI-integrierende Unternehmen, die in der Regel „Complementors“ sind, hängt die Souveränität davon ab, wie die Governance- und Anreizstrukturen von dem „Platform Leader“ ausgearbeitet sind. Eine souveräne Rollenverteilung in dieser Dimension bedeutet also, dass das Unternehmen nicht nur passiv die Produkte und Dienstleistungen des „Platform Leaders“ konsumiert, sondern aktiv an der Gestaltung des Ökosystems durch strategische Partnerschaften und Mitentwickeln an beispielsweise Open-Source-Projekten die Plattform teilnimmt. Ein Beispiel für eine souveräne Rollenverteilung wäre beispielsweise bei Hugging Face, bei der jeder Teilnehmer des Ökosystems Open-Source-Projekte Nutzen und Weiterentwickeln kann (vgl. Kaffee und Jernite 2025).

3. Kontrolle über Engstellen in der Wertschöpfungskette

Gerade auch bei digitaler Souveränität gibt es strategische Engstellen, an denen sich diese entscheidet. Foster (2025) transferiert den „Global Value Chain“-Ansatz auf die KI-Industrie und identifiziert dabei einige Engpässe wie Infrastruktur, Daten und Modelle. Durch die Kontrolle dieser Engpässe, bei denen es nur wenige Anbieter gibt, die überhaupt Kontrolle ausüben können, kann immense Macht und Bindung ausgeübt werden. Gerade KI-integrierende Unternehmen in der nachgelagerten Wertschöpfungskette können Engstellen wie das Entwickeln und Trainieren von Foundationmodels, nicht kontrollieren, was die Souveränität gefährdet. Daher müssen KI-integrierende Unternehmen eigene Engstellen in ihrer Branche oder Nische etablieren, um die fehlende Kontrolle der vorgelagerten Engpässe auszugleichen (vgl. Foster 2025). Hier können beispielsweise dann exklusiver Zugang zu branchenspezifischen Daten oder einzigartige Benutzeroberflächen eigene Engpässe ermöglichen, was andererseits die Souveränität weiterer nachgelagerter Unternehmen oder Kunden reduzieren kann. Für die spätere Bewertung kann das Verhältnis zwischen der Nutzung offener Ökosysteme und dem Aufbau eigener, schwer substituierbarer Systeme, entscheidend sein.

5.2.6. Regulatorische und Normative Souveränität

Digitale Souveränität besteht nicht nur aus technologischen und ökonomische Unabhängigkeitsaspekten, sondern beinhaltet zwangsmäßig rechtliche und ethische Verantwortung, welche teilweise sogar in Wechselwirkung mit ökonomischen Aspekten steht. Da KI-integrierende Unternehmen zwischen KI-Providern und Endkunden stehen, müssen diese sicherstellen, dass die eingesetzte, oftmals

auch importierte Technologie mit gesetzlichen Rahmenbedingungen und gesellschaftlichen Werten übereinstimmt. Diese finale Dimension fokussiert sich auf Compliance, Ethik und den Schutz der Nutzer bzw. Individuen. Wie bei den anderen Dimensionen auch, teilt sich diese in drei Kategorien.

1. Regulatorische und rechtliche Compliance

KI-Systeme generell und die Integration dieser unterliegt zunehmend strengen gesetzlichen Vorgaben wie dem EU AI Act. „Market-Rights-Balance" ist eine Art Dimension digitaler Souveränität, bei der es darum geht, die politischen Ziele der Förderung digitaler Innovationen und gleichzeitig der Schutz der Rechte, zu vereinen (vgl. Fratini et al. 2024). Dafür müssen automatisierte Systeme reguliert und rechtlich beaufsichtigt werden (vgl. Abiade 2025). Für die EU bedeutet das die Einhaltung der Datenschutz-Grundverordnung sowie des EU AI Acts. Diese Verordnungen schränken Unternehmen ein, weshalb ein souveränes Unternehmen sich durch proaktive Compliance Strukturen auszeichnet und nicht reaktive regulatorischen Anforderungen abarbeitet. Um später einen Reifegrad zu ermitteln, kann hierbei gemessen werden, inwieweit regulatorische Compliance-Prozesse in die KI-integration eingebunden sind und wie weit diese beispielsweise beim „AI Act Readiness Assessment" fortgeschritten sind.

2. Normative Verantwortung

Die Souveränität von KI-integrierenden Unternehmen überträgt sich auch auf die Souveränität deren Kunden und Partner, wobei KI-integrierende Unternehmen diese im negativen Sinne auch weiter beeinflussen können. Braun und Hummel (2025), sowie Steinbach (2024) kritisieren ein rein kontrollorientiertes Verständnis von digitaler Souveränität und schlagen daher vor, die Vulnerabilität und die Bedürfnisse von Individuen bzw. von Kunden auch zu berücksichtigen. Für KI-Integrierende Unternehmen bedeutet dies die Übernahme von unternehmerischer digitaler Verantwortung (Corporate Digital Responsibility – CDR). Dabei sollen Unternehmen bei der Integration von KI auf die Wahrung der einer menschen-, werteorientierten, autonomen und souveränen Nutzung achten (vgl. BMV 2026). Für die Modellierung kann hierbei erfasst werden, ob Unternehmen ethische Leitlinien für den Einsatz von KI definiert und praktisch durchsetzt.

3. Transparenz, Fairness und Verantwortlichkeit

Auch für die ethische Nutzung von KI wird technische und prozedurale Transparenz notwendig. Daher sind Fairness, Transparenz und Verantwortlichkeit zentraler Bestandteil der KI-Governance. Celsi und Zomaya (2025) fassen das unter dem Konzept der „Algor Ethik" zusammen (genauer beleuchtet in Kapitel: Ethische, rechtliche und Soziale Überlegungen). Abiade (2025) fordert eine Auditierbarkeit von Algorithmen und die Schaffung von Widerspruchsmöglichkeiten für betroffene Nutzer, sollte KI den ethischen und rechtlichen Vorgaben nicht entsprechen. Gerade bei nachgelagerten Akteuren

der KI-Wertschöpfung liegt hier die Herausforderung, da Blackbox-Modelle weder transparent noch vollkommen verständlich sind. Daher müssen Unternehmen, die KI integrieren, Methoden zur Schaffung einer verständlichen und nachvollziehbaren Entscheidungsfindung der KI implementieren. Wenn dies nicht möglich ist, müssen sie klare Eskalations- und Beschwerdeprozesse für Kunden und Nutzer bereitstellen, falls die KI fehlerhafte oder diskriminierende Ergebnisse liefert. Die Existenz und Wirksamkeit von Maßnahmen, die Transparenz fördern und Widerspruch ermöglichen, kann für ein Bewertungs- oder Reifegradmodell bewertet werden.

6. Bewertungsmodell

Das folgende Kapitel beschäftigt sich neben den Zielen und Anforderungen des Bewertungsmodells auch mit dessen Aufbau und Funktionsweise. Das Modell basiert auf den zuvor definierten und theoretisch hergeleiteten Dimensionen digitaler Souveränität in der nachgelagerten KI-Wertschöpfungskette und überführt diese in ein quantifizierbares Bewertungsmodell. Das Modell fungiert als diagnostisches Instrument zur Feststellung des punktuellen Ist-Zustands eines KI-integrierenden Unternehmens.

Um die Praxistauglichkeit des Modells zu maximieren und einfache Nutzung zu ermöglichen, wurde das Bewertungsmodell primär als interaktive Webanwendung konzipiert. Aus Gründen der Datensicherheit (und „Privacy by Design“) verzichtet die Webanwendung vollständig auf eine serverseitige Datenverarbeitung. Sämtliche Parameter und Eingaben werden über das URL-State-Management im Client (also dem Browser) verarbeitet. Dies ermöglicht es Unternehmen, die Auswertung auf beliebigen Endgeräten durchzuführen und über den URL-Link lokal als Report zu speichern, ohne proprietäre Unternehmensdaten extern preiszugeben.

Anmerkung: Um jedoch Redundanz und Offline-Bewertungen zu ermöglichen, wurde die Berechnungslogik parallel in einer Excel-Mappe abgebildet. Der Aufbau und die Funktionsweise dieses Tools sind im Anhang dokumentiert.

6.1. Ziele und Anforderungen

Das primäre Ziel des Bewertungsmodells digitaler Souveränität KI-integrierender Unternehmen ist die strukturelle Operationalisierbarkeit der zuvor hergeleiteten Dimensionen digitaler Souveränität. Das Modell fungiert als diagnostisches Instrument zur Feststellung eines punktuellen Ist-Zustands eines KI-integrierenden Unternehmens. Im Vordergrund steht die objektive und strukturierte Erfassung vielschichtiger Abhängigkeiten und die Übersetzung dieser in messbare und verständliche Kriterien.

Nach den Anforderungen des Design Science Research Prozesses (vgl. Hevner et al. 2004) gelten für die Erstellung aller Modelle in dieser Arbeit die grundlegenden Gestaltungsanforderungen wie theoretische Konsistenz, Messbarkeit, Praxistauglichkeit und Korrektheit. Zudem sollen für das Bewertungsmodell weitere, für ein Messinstrument spezifische Anforderungen nach Pöppelbuß und Röglinger (2011) gestellt werden. Dabei ist vor allem wichtig, auf Granularität und Operationalisierbarkeit zu achten. Die teilweise schwer greifbaren Dimensionen digitaler Souveränität müssen in messbare Variablen transformiert werden, ohne dabei Informationen zu verlieren. Außerdem muss das Bewertungsmodell objektiv überprüfbar konstruiert werden, so dass unterschiedliche Anwender bei Betrachtun-

gen desselben Unternehmens zu vergleichbaren Bewertungsergebnissen kommen. Hierfür ist eine präzise Formulierung der Indikatoren und Bewertungsmechanismen erforderlich. Darüber hinaus erfordert der hochsensible Kontext von strategischen Unternehmensdaten die zwingende Anforderung, Datensicherheit und Privacy-by-Design zu gewährleisten. Das Bewertungsmodell muss so konzipiert sein, dass eine umfassende Bewertung möglich ist, ohne dass proprietäre Unternehmensdaten an externe Server oder Datenbanken übertragen werden.

6.2. Modellstruktur

Das Bewertungsmodell ist in einer Top-down-Struktur aufgebaut, basierend auf den sechs zuvor definierten Dimensionen mit jeweils drei Kategorien. Jeder Kategorie wird in der Regel eine Prüffrage zugeordnet, welche als konkreter Indikator für die Ausprägung der digitalen Souveränität dient. Diese 20 Fragen bilden das methodische Grundgerüst des Modells.

Dimension	Kategorie	Frage
1. Technologische & Infrastrukturelle Souveränität	1.1 Infrastrukturelle Kontrolle und Cloud Abhängigkeit	Wie ist die Hosting- und Berechnungsinfrastruktur für KI im Unternehmen strategisch aufgestellt?
	1.2 Interoperabilität & Modularität	Wie hoch ist der Aufwand, die aktuell genutzten KI-Technologien durch eine Alternative (z.B. Open oder Shared Source) zu ersetzen?
	1.3 Operative Kontrolle & Resilienz	Welche technischen Fallback-Mechanismen greifen bei einem Ausfall oder unangekündigten Updates / Downtimes der externen KI-Schnittstellen / Technologie?
2. Daten- & Informationssouveränität	2.1 Eigentum und Kontrolle der Daten und Informationen	Inwiefern ist der kontinuierliche Zugang zu kritischen Trainings- und Optimierungsdaten für die eigene Wertschöpfung langfristig vertraglich oder durch Eigengenerierung gesichert?; Inwiefern ist vertraglich oder technisch ausgeschlossen, dass sensible Unternehmens- oder Kundendaten zum Training externer Basismodelle genutzt werden?
	2.2 Datensicherheit und Vertraulichkeit	Werden technische Schutzmaßnahmen wie Confidential Computing oder Datenanonymisierung angewendet, bevor es zu einer (externen) KI-Verarbeitung via API kommt?; Existiert eine strikte Trennung, welche KI-Systeme für welche Datenklassifizierungen (z.B. öffentlich oder streng vertraulich) genutzt werden dürfen?
	2.3 Verfügbarkeit und Transparenz	Wie transparent und nachverfolgbar / nachvollziehbar ist der Datenfluss von eigenen Datenquellen hin zu KI-Systemen und Datenspeichern?
3. Organisationale Souveränität & KI-Governance	3.1 Strategische Handlungsfähigkeit	Gibt es eine formal definierte Strategie (inkl. definiertem Budget) für die Evaluierung von Make-or-Buy-Entscheidungen bei KI-Systemen?
	3.2 Prozessbasierte und strukturelle Governance	Sind strukturierte Machinelearning-Operations-Prozesse (MLOps) zur Qualitätssicherung und Überwachung der externen KI-Modelle im Einsatz, um beispielsweise gegen Leistungsabfälle oder Halluzinationen vorzugehen?

	3.3 Risikomanagement	Existieren Notfallpläne und Prozesse für KI-spezifische Bedrohungen wie Data Poisoning, Prompt-Injection oder plötzliche Systemdowntimes?
4. Kompetenzen & Akteure	4.1 Bewertungskompetenz & Entscheidungssouveränität	Wie ausgeprägt sind die internen Validierungsmechanismen bzw. Fähigkeiten zur Validierung, KI-generierte Ergebnisse auf Korrektheit und Bias zu prüfen?
	4.2 Talentverfügbarkeit & Ressourcen	Wie gestaltet sich das Verhältnis zwischen interner technologischer Gestaltungsfähigkeit (durch beispielsweise Data Scientists oder Softwarearchitekten) und der Abhängigkeit von externen Beratern oder KI-Dienstleistern?
	4.3 Relationale Praktiken & Domänenwissen	Inwiefern wird das spezifische Domänenwissen des Unternehmens strukturell in die Entwicklung und das Finetuning der KI-Modelle eingebunden?
5. Ökosysteme und Machtverhältnisse	5.1 Strukturelle Macht & Plattform-Abhängigkeit	Wie ausgeprägt ist die Konzentration des Technologie-Portfolios auf einzelne dominante Plattformen (z.B. Hyperscaler, Foundation-Model-Provider oder Hosting Dienste)?
	5.2 Rollenverteilung im sozio-technischen Ökosystem	Welche Rolle immt das Unternehmen in KI-Ökosystemen ein? Eher als passiver Konsument oder beteiligt es sich aktiv an der Gestaltung?
	5.3 Kontrolle über Engstellen in der Wertschöpfungskette	Hat das Unternehmen eigene strategische Engpässe aufgebaut, um selbst Macht über nachgelagerte Unternehmen auszuüben oder die Macht vergelagerter Akteure auszugleichen? (Macht wäre hier bspw. Exklusiver Zugang zu branchenspezifischen oder vertikalen Daten)
6. Regulatorische & Normative Souveränität	6.1 Regulatorische & Rechtliche Compliance	Wie proaktiv ist das Unternehmen auf kommende regulatorische Vorgaben wie den EU AI Act vorbereitet?
	6.2 Normative Verantwortung	Wurden ethische Leitlinien (wie Corporate digital responsibility) für einen werteorientierten Einsatz von KI definiert und praktisch umgesetzt?
	6.3 Transparenz, Fairness & Verantwortlichkeit	Existieren klare Eskalations- und Beschwerdeprozesse, die Kunden nutzen können, falls externe KI-Systeme fehlerhafte oder verzerrte Entscheidungen treffen?

Abbildung 5: Übersicht der Fragen des Fragebogens des Bewertungsmodells

Um eine reproduzierbare und fundierte Beantwortung zu ermöglichen, werden pro Kategorie (Erhebungs-Seite) systematisch neben der jeweiligen Kategorie sieben Datenpunkte abgefragt oder bereitgestellt.

1. **Prüffrage**: Auf Basis der definierten Kategorie wird eine Prüffrage formuliert, welche der Indikator zur Ausprägung digitalen Souveränität für die spezifische Kategorie ist (siehe Abbildung 5).
2. **Quantitativer Ist-Wert**: Die Bewertung erfolgt über einen kontinuierlichen Schieberegler auf einer Likert-Skala.
3. **Beispiel der Skalenlogik**: Jede Prüffrage beinhaltet ein individuelles Beispiel für die zugrundeliegende Skalenlogik, die für die Zustände „0“, „3“ und „5“ Beispiele liefert. Das soll eine reproduzierbare, personenunabhängige und vergleichbare Angabe des aktuellen Zustands digitaler Souveränität unter der spezifischen Kategorie ermöglichen.

4. **Konfidenzbewertung**: Jeder Kategorie ist in dem Fragebogen zudem eine Konfidenz-Abfrage zugeordnet, damit die Organisation oder die ausfüllende Person die Genauigkeit der eigenen Antwort nachvollziehen kann. Dies soll unter anderem Anreize schaffen, welche den Einbezug von Fachexperten der individuellen Fachgebiete (z. B. IT-Architektur, Legal) fordern.
5. **Kategorie-Definition**: Ein Auszug dieser Forschungsarbeit (vgl. Kapitel 5) zur Herleitung der spezifischen Kategorie, um diese nachvollziehbar zu machen, zusätzliche Informationen bereitzustellen und volle Transparenz zu ermöglichen.
6. **Anmerkungen**: Notizen und Anmerkungen können für interne Zwecke hinzugefügt werden. Diese werden ebenfalls in der URL gespeichert.
7. **Literaturquellen**: Zur methodischen Transparenz werden allen abgeleiteten Prüffragen die jeweiligen Quellen aus der Theorie (vgl. Kapitel 5) zugeordnet.

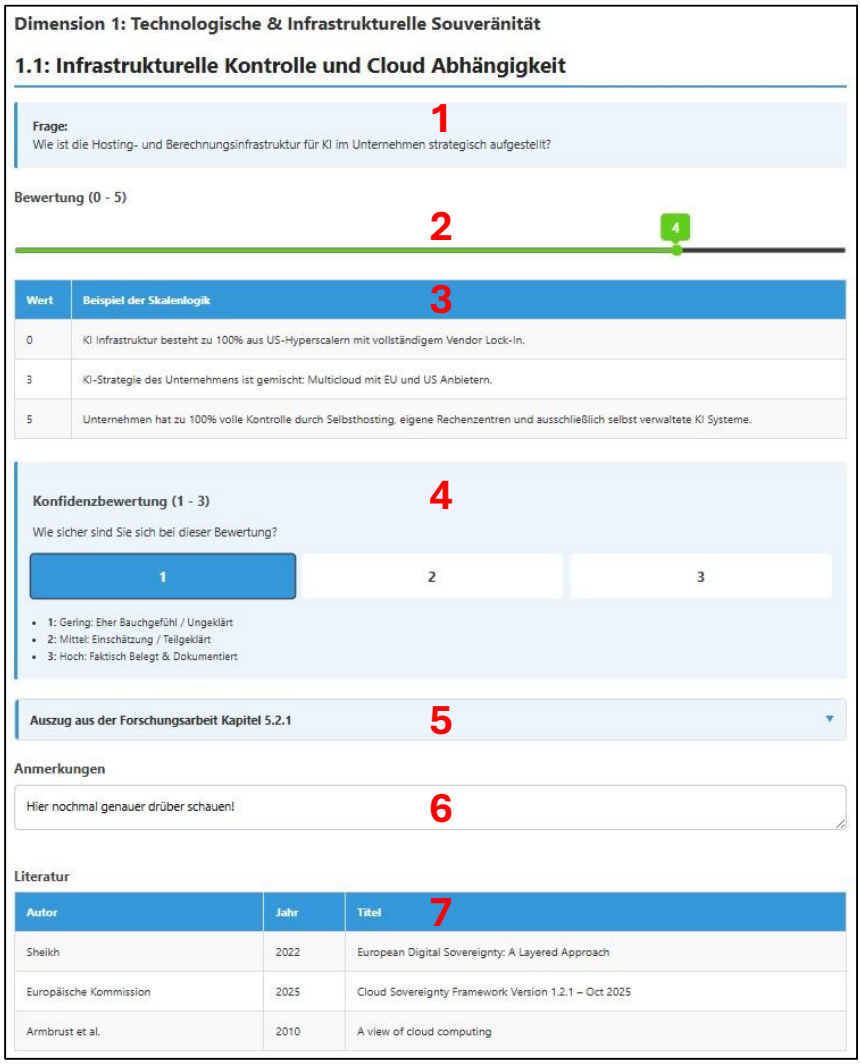

Abbildung 6: Benutzeroberfläche des Fragebogens in der Webanwendung mit sieben Datenpunkten

6.3. Bewertungsmechanismen

Um eine wissenschaftliche und nachvollziehbare Bewertung zu ermöglichen, ist jede der 20 Fragen mit spezifisch zugeschnittenen Ankerbeispielen versehen. Die Skala beschreibt den Zustand der digitalen Souveränität mit 0 als vollständig fehlende Souveränität oder Black-Box-Abhängigkeit und 5 als proaktiv gemanagte, technologische oder strategische Autonomie.

Die Skala nutzt ein geradzahliges Design mit sechs Antwortmöglichkeiten (0 bis 5). Durch den bewussten Verzicht auf eine neutrale Mitte wird ein sogenanntes Forced-Choice-Format erzeugt, welches methodische Verzerrungen durch die „Tendenz zur Mitte“ (Central Tendency Bias) verhindert und den Nutzer zu einer klaren Einschätzung bewegt. Werden in einer Kategorie mehrere Fragen bzw. Indikatoren abgefragt, wird der ungewichtete arithmetische Mittelwert gebildet, um eine Überrepräsentation dieser Kategorie in der finalen Auswertung zu verhindern. Die aggregierten Bewertungen der Kategorien werden für die schlussendliche Auswertung auf Dimensionsebene zusammengeführt und dem Maximalwert gegenübergestellt.

Die Dimensionen sind standardmäßig gleich gewichtet, können allerdings bei branchenspezifischen Besonderheiten auch individuell angepasst werden. Daraus ergibt sich für das Bewertungsmodell ein ungewichteter und gewichteter Souveränitätsscore. Diese Scores lassen sich wie folgt berechnen:

Souveränitätsscore (S) ohne Dimensions-Gewichtung:

$$S_{ungewichtet} = \frac{Summe\ der\ Ist - Werte\ aller\ Dimensionen}{Summe\ der\ Maximalwerte\ aller\ Dimensionen} \times 100$$

Souveränitätsscore (S) mit Dimensions-Gewichtung:

$$S_{gewichtet} = \frac{Summe\ (Ist - Wert\ einer\ Dimension\ \times Gewichtung)}{Summe\ (Maximalwert\ einer\ Dimension\ \times Gewichtung)} \times 100$$

Nach der Gewichtung der sechs Dimensionen können zudem Soll-Werte für jede einzelne Kategorie definiert werden. Die finale Auswertung generiert somit eine multidimensionale Gegenüberstellung des aktuellen Ist-Zustands, mit dem vom Unternehmen definierten Soll-Zustand. Es wird also eine Gap-Analyse durchgeführt. Aufgrund des kontext- oder branchenabhängigen Anspruchs an digitale Souveränität kann demnach eine unternehmensspezifische Definition des Ziel-Zustands vorgenommen werden. Für den Gap- oder Zielerreichungsgrad zur Bestimmung des strategischen Handlungsbedarfs wird der Ist-Zustand dem definierten Soll-Zustand gegenübergestellt. Dieser Zielerreichungsgrad wird ausschließlich gewichtet berechnet:

Zielerreichungsgrad (Z) mit Dimensions-Gewichtung:

$$Z_{gewichtet} = \frac{Summe\ (Ist - Wert\ einer\ Dimension \times Gewichtung)}{Summe\ (Soll - Wert\ einer\ Dimension \times Gewichtung)} \times 100$$

In Zusammenhang mit dem Bewertungs-Score wird die Konfidenz-Skala mit drei Punkten (Bauchgefühl zu faktisch belegt) ausgewertet. Niedrige Konfidenzwerte dienen als analytischer Indikator für organisationale Intransparenz und deuten auf unentdeckte Risikopotenziale, „Blind Spots“ sowie fehlendes Dokumentationswissen innerhalb der KI-Wertschöpfung hin. Die hohe Komplexität der Fragen erfordert zwingend einen interdisziplinären Multi-Befragten-Ansatz. Nur durch diese Konsolidierung unterschiedlicher Domänen lassen sich hohe Konfidenzwerte erzielen und valide Ergebnisse über den Souveränitätsstatus des Unternehmens ableiten. Eine hohe Eingaben-Konfidenz bedeutet in diesem Bewertungsmodell, dass die Bewertung faktisch belegt ist. Die Berechnung geschieht folgendermaßen:

Konfidenz (K) der Eingaben:

$$K = \frac{Summe\ der\ gewählten\ Konfidenzwerte}{Summe\ der\ maximal\ möglichen\ Konfidenzwerte} \times 100$$

6.4. Umsetzung des Bewertungsmodells in der Web-App

Um die Praxistauglichkeit im Rahmen des DSR-Ansatzes zu maximieren, wurde das konzeptionelle Modell in eine vollständig clientseitige Webanwendung (auf Basis von HTML, CSS, JavaScript und JSON) überführt.

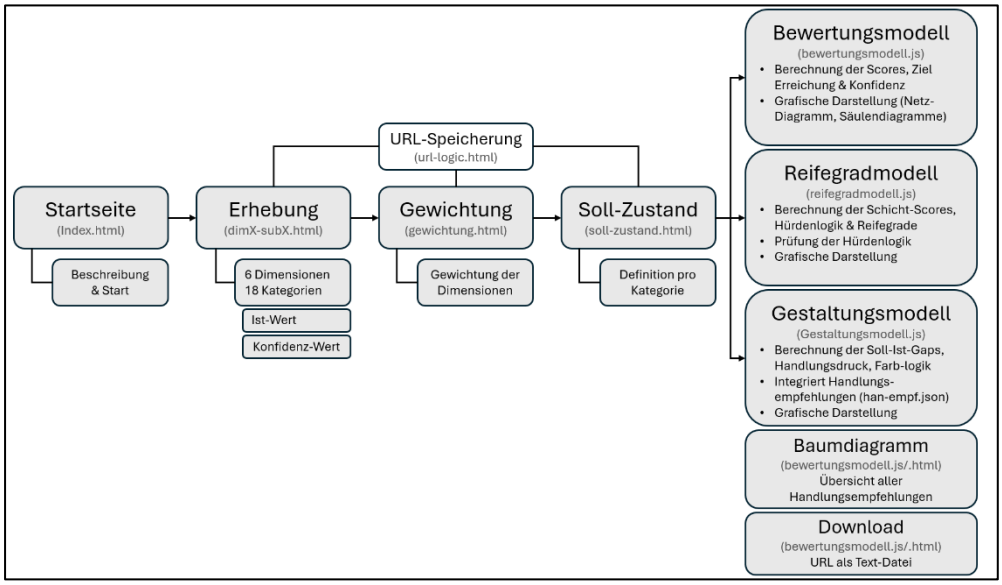

Abbildung 7: Architektur der Webanwendung

Um dem Ansatz von „Privacy-by-Design“ nachzukommen, verfügt die Webanwendung über kein Backend und keine Datenbank. Alle Werte und Eingaben des Nutzers werden Zustandshaltung werden dynamisch über die URL-Parameter gespeichert. Organisationen und Unternehmen müssen für die Nutzung der Webanwendung zwar Internet haben, das Souveränitäts-Assessment wird allerdings lokal (clientseitig) durchgeführt und gespeichert. Dadurch werden keine sensiblen oder proprietäre Daten preisgegeben und die Eingaben können mittels der generierten URL sicher geteilt werden.

Nach Abschluss der 18 Erhebungsschritte sowie der Definition von Gewichtung und Soll-Werten, greift die zentrale Speicherlogik (url-logic) auf die URL-Parameter zu und generiert das Auswertungs-Dashboard des Bewertungsmodells. Dieses visualisiert die Ergebnisse durch dynamische Anzeigen und Diagramme.

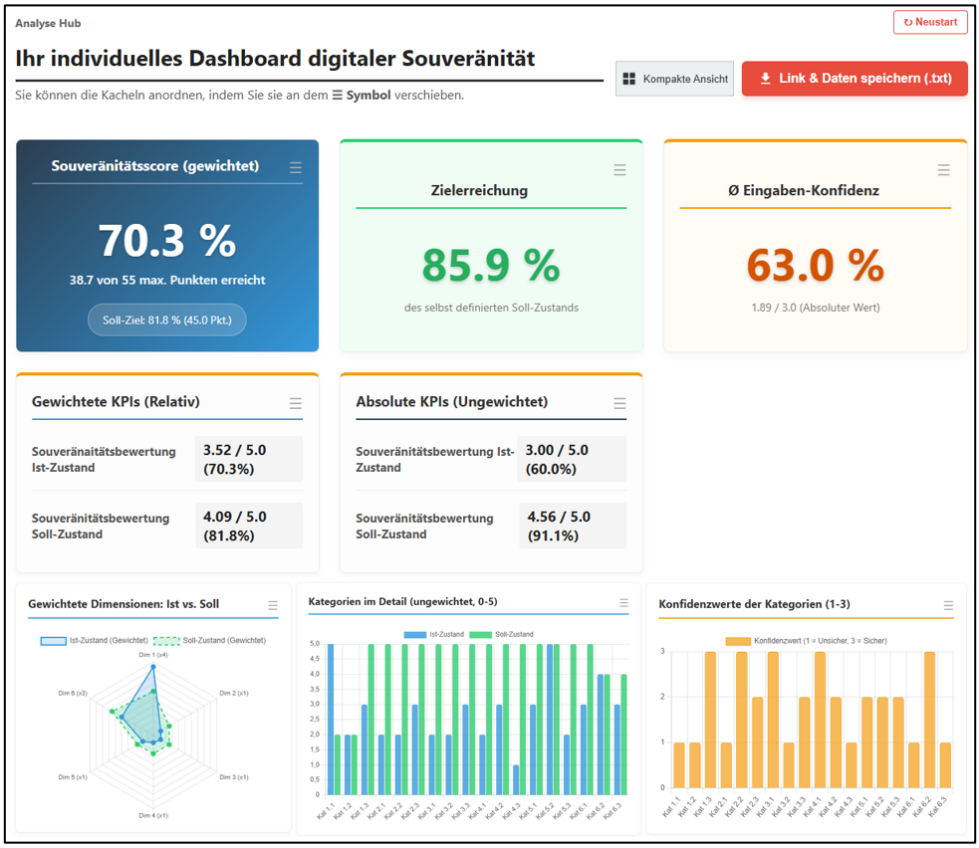

Abbildung 8: Bewertungsmodell-Dashboard der Webanwendung (optisch nachträglich angepasst)

Wie in Abbildung 8 zu sehen, werden die Parameter aus der URL zu einem zusammenfassenden Dashboard organisiert, um eine Bewertung und Analyse des Zustands digitaler Souveränität zu ermöglichen. Zentral dabei sind der Souveränitätsscore und die Zielerreichung, die eine direkte oberflächliche Bewertung der digitalen Souveränität ermöglichen. Die Eingaben-Konfidenz bestätigt die Genauigkeit und faktische Belegung der Aussage des Souveränitäts-Scores, Zielerreichung und weiteren Key Performance Indicators (KPIs). Da in diesem Modellierungsansatz eine Gewichtung der Dimensionen möglich ist, werden die KPIs relativ, also mit Gewichtung, sowie ungewichtet angegeben, um einen Abgleich zu ermöglichen. Daraufhin werden die Eingaben für eine detailliertere Einsicht in Form von Diagrammen angezeigt, um Rückschlüsse auf die zuvor angezeigten Bewertungen zu ermöglichen. Schlussendlich werden noch selbst eingetragene Anmerkungen der Nutzer aus dem Fragebogen angezeigt, sowie eine zusammenfassende Beschreibung jeder Dimension.

7. Reifegradmodell

Während das zuvor beschriebene Bewertungsmodell als deskriptives Instrument zur Datenerhebung und KPI-Berechnung dient, nutzt das Reifegradmodell die erhobenen Daten für eine evaluative Diagnose. Das Modell fungiert dabei als normativer Analyserahmen, der den gemessenen Ist-Zustand eines Unternehmens auf einer evolutionären Skala der digitalen Souveränität verortet. Ziel dieses Kapitels ist es, die konzeptionelle Struktur, die asymmetrischen Einordnungsmechanismen, die Hürdenlogik der Abhängigkeiten sowie die visuelle Repräsentation des Reifegradmodells detailliert zu beschreiben. Diese strukturierte Einordnung dient später als Grundlage für das Gestaltungsmodell und der Priorisierung und Zuordnung der Handlungsempfehlungen.

Anmerkung: Das Reifegradmodell ist ebenfalls Teil der offline Excel-Mappe, jedoch in abgeänderter Form aufgrund der eingeschränkten Möglichkeiten seitens Excel. Der Aufbau und die Funktionsweise dieses Tools sind ebenfalls im Anhang dokumentiert.

7.1. Ziele und Anforderungen

Aufbauend auf der Datenerhebung und dem Bewertungsmodell soll nun das Reifegradmodell die Ergebnisse evaluativ einordnen. Während das Bewertungsmodell die Datenerhebung, also Messung des Ist-Zustands, ermöglicht, soll das Reifegradmodell den gemessenen Zustand auf einer aufeinander aufbauenden, trennscharfen Skala zuteilen. Ziel ist hiermit eine Art Benchmark herzustellen, der es Unternehmen ermöglicht, erkennen zu können, auf welcher Entwicklungsstufe der digitalen Souveränität sie sich aktuell befinden und welche Entwicklungsstufe angestrebt werden soll. Zudem soll eine Vergleichbarkeit zwischen Unternehmen untereinander geschaffen werden.

Neben den übergreifenden Design Science Research Anforderungen nach Hevner et al. (2004), soll das Reifegradmodell auch spezifischen Gestaltungsprinzipien für Reifegrad- / Referenzmodellierung folgen (vgl. Becker et al. 2009). Bei verschiedenen Reifegraden ist besonders die Trennschärfe essenziell. Das Modell muss eindeutig voneinander abgrenzbare Reifestufen definieren, die bei Übergängen der Stufen für Überschneidungsfreiheit sorgen. Außerdem müssen die Reifestufen einer logischen und aufeinander aufbauenden Entwicklungslogik folgen, bei der ersichtlich wird, dass höhere Ausprägungen digitaler Souveränität die Erfüllung grundlegender Fähigkeiten bzw. Reifegrade zwingend voraussetzen.

7.2. Reifestufen und Messinstrumente

Die Struktur des Reifegradmodells orientiert sich an etablierten Normen zur Prozessbewertung wie unteranderem der SPICE-Norm (Software Process Improvement and Capability Determination), welche auch für das „Digital Transformation Capability Maturity Model“ von Gökalp und Martinez (2021) verwendet wird. Wie bei diesem Modell nutzt das in dieser Arbeit entwickelte Reifegradmodell eine fünfstufige Skala.

Die inhaltliche Abgrenzung der Stufen vollzieht sich entlang einer zunehmenden Vernetzung und Autonomie. Nach Bogenstahl und Zinke (2017) entwickelt sich digitale Souveränität von einer anfänglich isolierten Digitalisierung hin zur aktiven Gestaltung kollaborativer Wertschöpfungsnetzwerke. Gleichzeitig erfordern höhere Reifestufen neben der infrastrukturellen und datengetriebenen Technologie eine ausgeprägte normative Governance über den gesamten Lebenszyklus der KI, wie Celsi und Zomaya (2025) in ihrem Modell zur Operationalisierung der „Algor-Ethik“ anmerken. Aus dieser Synthese ergeben sich folgende fünf Reifestufen (in dieser Arbeit auch als Level bezeichnet), die auf Basis des gewichteten Gesamt-Scores (0,0 bis 5,0) zugewiesen werden:

1. **Passiv / Black-Box** (Score 0,00 – 0,99)
    - Das Unternehmen nutzt KI unstrukturiert und operiert in vollständiger Abhängigkeit von externen Providern. Es existieren weder interne Kontrollmechanismen noch ein strategisches Bewusstsein. KI wird als intransparente Black-Box integriert oder konsumiert.
2. **Reaktiv / Sensibilisiert** (Score 1,00 – 1,99)
    - Ein Bewusstsein für Abhängigkeiten und regulatorische Risiken ist vorhanden. Erste Schnittstellen und isolierte Richtlinien existieren, diese werden jedoch ausschließlich reaktiv genutzt, ohne eine ganzheitliche architektonische Verankerung.
3. **Definiert / Strukturiert** (Score 2,00 – 3,49)
    - Strategien, Datenkontrollprozesse und Kompetenzen sind formalisiert und in den operativen Alltag integriert. Es herrscht technologische Interoperabilität und eine definierte Governance (vgl. Gökalp & Martinez 2021).
4. **Proaktiv / Emanzipiert** (Score 3,50 – 4,49)
    - Das Unternehmen agiert hochgradig autonom durch den Einsatz modularer Systeme und proaktiver Multi-Vendor-Strategien. Ethische Vorgaben und Verantwortlichkeit sind „by Design“ im KI-Lebenszyklus verankert.
5. **Souverän / Gestaltend** (Score 4,50 – 5,00)
    - Es herrscht absolute Unabhängigkeit. In Anlehnung an Bogenstahl und Zinke (2017) konsumiert das Unternehmen nicht nur, sondern gestaltet offene Ökosysteme und den Umgang

mit KI aktiv mit (z.B. durch Open-Source-Beiträge) und nutzt eigene Architektur-Engpässe als strategischen Wettbewerbsvorteil.

Die Zuweisung der Score-Intervalle zu den jeweiligen Reifestufen erfolgt bewusst nicht linear, sondern asymmetrisch. Während die unteren Level 1 und 2 eine Standardbandbreite von 1,0 Punkten aufweisen, ist Level 3 (2,00 bis 3,49) mit einer Bandbreite von 1,5 Punkten weiter gefasst. Dies reflektiert die Realität von Verteilungen bei Reifegradmodellen. Stufe 3 stellt ein breiteres organisatorisches Plateau dar, auf dem sich mutmaßlich die Mehrheit der handlungsfähigen KI-integrierenden Unternehmen sammelt.

Der Sprung von diesem definierten Zustand (Level 3) zu einer proaktiven (Level 4) oder gestaltenden Rolle (Level 5) ist mit exponentiellem organisationalem Aufwand verbunden. Dementsprechend ist Level 5 (4,50 bis 5,00) mit einer Bandbreite von lediglich 0,5 Punkten als exklusiver Zustand definiert. Dies korrespondiert mit der theoretischen Vorarbeit von Bogenstahl und Zinke (2017) hinsichtlich der Grenzproduktivität von Souveränitäts- und KI-Investitionen. Das Erreichen der absoluten Autonomie-Spitze erfordert überproportionale strategische und technologische Anstrengungen, weshalb die numerische Hürde für den Eintritt in die höchste Reifestufe stark verengt und trennscharf angesetzt wird.

7.3. Einordnungsmechanismen und Abhängigkeiten

Ein zentraler Baustein des Reifegradmodells ist die Vermeidung einer stark ausgebauten Compliance, ohne allerdings angefangen zu haben die technologische und infrastrukturelle Souveränität durchzusetzen. Eine simple arithmetische Durchschnittsbildung über alle Dimensionen birgt das Risiko, dass Schwächen in der IT-Infrastruktur durch hoch bewertete ethische Richtlinien rechnerisch untergehen. Um dies zu verhindern und die Schaffung einer von Grund auf konsistente digitale Souveränität zu ermöglichen, werden die sechs Dimensionen auf drei aufeinander aufbauende Schichten zugeteilt:

1. **Das Fundament** (Dimension 1 und 2)
    - Technologische und infrastrukturelle Souveränität, sowie Daten- und Informationssouveränität bilden das Fundament, welches unerlässlich ist für ein digital souveränes Unternehmen. Ist das Fundament gegeben, können Prozesse, Kompetenzen, Ökosysteme und Normen souverän ausgestaltet werden.
2. **Die Befähiger** (Dimension 3 und 4)

- Organisationale Souveränität und KI-Governance, sowie die Kompetenzen und Akteure des Unternehmens befähigen das Unternehmen erst souverän zu werden. Ist das technologische Fundament gegeben, können Prozesse, Organisation und Kompetenzen souverän gestaltet werden.

3. **Die Wirkung** (Dimension 5 und 6)
    - Die Teilnahme an Ökosystemen und das Auseinandersetzen mit Machtverhältnissen, sowie die Schaffung und Einhaltung von regulatorischen und normativen Vorgaben ist für Unternehmen meist erst dann ökonomisch und nachhaltig möglich, wenn das Fundament und die Befähiger bereits souverän gestaltet sind.

Der Score der jeweiligen Schichten berechnet sich wie folgt am Beispiel von Schicht 1:

Schichten-Score der ersten Schicht (das Fundament):

$$Score_{Schicht\ 1} = \frac{Ist{-}Wert\ Dimension\ 1 + Ist{-}Wert\ Dimension\ 2}{2}$$

Aufgrund dieser Abhängigkeiten zwischen den Schichten beinhaltet das Reifegradmodell eine deterministische Hürdenlogik, die den finalen Reifegrad beeinflussen kann. Denn wenn Schicht 1 einen aggregierten Score von unter 2,0 aufweist, wird der finale Reife-Score des Unternehmens technisch auf maximal 1,99 (Level 2) gedeckelt, unabhängig davon, wie hoch andere Dimensionen bewertet wurden. Die Begründung liegt darin, dass ohne ein technologisches und datensicheres Fundament keine nachhaltige Skalierung möglich ist. Dabei wird der Reifegrad durch die Fundament-Hürde wie folgt beeinflusst:

Logik des finalen Reifegrads (R_{final}) unter Anbetracht der Fundament-Hürde:

$$R_{final} = \min\left(Gewichteter\ Gesamtscore, 1.99\right)$$

Weist Schicht 2 einen Score von unter 2,0 auf bei intakter Schicht 1, greift die Befähiger-Hürde und es erfolgt eine Begrenzung auf maximal 3,49 (Level 3), da ohne strukturelle Befähiger keine proaktive Emanzipation (Level 4) realisiert werden kann. Diese Logik lässt sich wie folgt darstellen:

Logik des finalen Reifegrads (R_{final}) unter Anbetracht der Befähiger-Hürde:

$$R_{final} = \min\left(Gewichteter\ Gesamtscore, 3.49\right)$$

Zusätzlich implementiert das Modell in der Webanwendung eine analytische Souveränitäts-Warnfunktion, wenn eine der beiden unteren Schichten nicht die Mindestanforderung von einem Schicht-Score von über 2,0 Punkten haben. Dann warnt das System vor einem kritischen Fundament oder einer kritischen Befähiger-Schicht. Dies indiziert, dass gerade an diesen Schichten ausbaubedarf ist, bzw. die Organisation eine sehr hohe Abhängigkeit aufweist.

7.4. Analytische Auswertung in der Webanwendung

Die technische Instanziierung des Reifegradmodells als Webanwendung ermöglicht eine visuelle Auswertung, die direkt in strategische Entscheidungsprozesse überführt werden kann. Das dynamische Dashboard des Reifegradmodells visualisiert den finalen Reifegrad-Score, mögliche Warnungen, sowie Auswertungen zu dem Reifegrad, den Schichten und den Dimensionen.

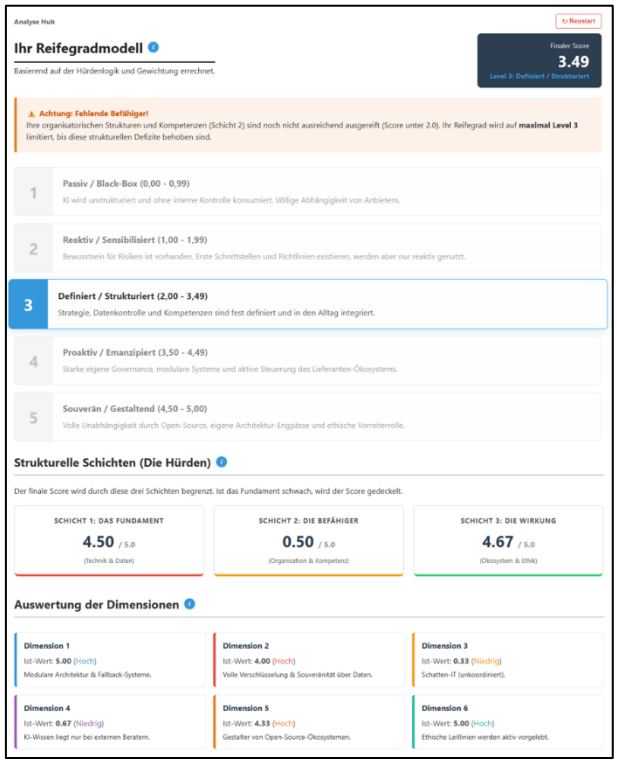

Abbildung 9: Reifegradmodell-Dashboard der Webanwendung

Das Reifegradmodell berechnet den gewichteten Souveränitätsscore auf Basis der Schichten-Scores neu und bezieht die Hürdenlogik mit ein, um auf den finalen Score zu kommen. Dieser Score bestimmt daraufhin den Reifegrad und verursacht, falls Schicht 1 oder Schicht 2 den Schwellwert von 2,0/5,0 unterschreiten, eine Warnung. Für mehr Transparenz und Einsicht werden darunter die Bewertungen der strukturellen Schichten und einzelnen Dimensionen angezeigt, wobei die Dimensionen jeweils mit einer Hoch-, Mittel- oder Niedrig-Bewertung versehen werden. Diesen Bewertungen werden Aussagen bzw. Diagnosen zugeordnet, die von den Beispielen der Skalenlogik abgeleitet sind und lediglich ein Verständnis des Zustands digitaler Souveränität in der jeweiligen Dimension ermöglichen sollen.

Bewertung / Dimension	**Niedrig: Score <2,0**	**Mittel: Score >=2,0, <3,5**	**Hoch: Score >=3,5**
Dimension 1	Harter Vendor-Lock-in	Schnittstellen nutzbar & austauschbar	Modulare Architektur & Fallback-Systeme
Dimension 2	Keine Datenkontrolle	Opt-Outs geregelt	Volle Verschlüsselung & Souveränität über Daten

Dimension 3	Schatten-IT (unkoordiniert)	KI-Strategie existiert	Professionelles Lieferantenmanagement
Dimension 4	KI-Wissen liegt nur bei externen Beratern	Fachabteilungen sind eingebunden	Interne Experten prüfen KI kritisch
Dimension 5	Passiver Konsument	Strategische Auswahl von Partnern	Gestalter von Open-Source-Ökosystemen
Dimension 6	Keine Compliance-Vorgaben	AI Act Readiness geprüft	Ethische Leitlinien werden aktiv vorgelebt

Abbildung 10: Bewertung der Dimensionen für die Reifegradermittlung

8. Gestaltungsmodell

Das folgende Kapitel beschäftigt sich neben den Zielen und Anforderungen des Gestaltungsmodells auch mit dessen Aufbau und Funktionsweise. Dieses Modell rundet den gesamten Bewertungs- und Evaluierungsprozess zur digitalen Souveränität in der nachgelagerten KI-Wertschöpfungskette ab und soll auf Basis der Ergebnisse der vorigen Modelle Gestaltungsmöglichkeiten und Handlungsempfehlungen aufzeigen. Während das Bewertungsmodell den Zustand deskriptiv erfasst und das Reifegradmodell diesen evaluativ einordnet, wandelt das Gestaltungsmodell die identifizierten Lücken in priorisierte Handlungsempfehlungen um.

8.1. Ziele und Anforderungen

Das Gestaltungsmodell dieser Forschungsarbeit soll als Bindeglied zwischen den Zustandsanalysen der vorigen Modelle und den strategischen Transformationszielen KI-integrierender Unternehmen dienen. Ziel dabei soll sein, nachgelagerte Akteure mit einem Handlungsrahmen zu unterstützen, bei dem Konkrete Maßnahmen und Strukturen ableitbar sind, um Unternehmen auf höhere Reifestufen der digitalen Souveränität zu bringen.

Auch hier soll neben den Design Science Research Anforderungen das Gestaltungsmodell spezifischen Anforderungen gerecht werden. Zum einen muss das Modell die abstrakten Ergebnisse der vorherigen Modelle in konkrete, implementierbare und realistische Handlungsempfehlungen übersetzen. Dabei muss das Modell dann nicht nur aufzeigen, welche Maßnahmen wichtig sind, sondern auch eine strategische Reihenfolge ermöglichen, die beispielsweise nach Dringlichkeit sortiert sind. Um zudem auf individuelle Unternehmensressourcen eingehen zu können, sind die Handlungsempfehlungen generalisiert formuliert. Sie versteifen sich nicht auf spezifische Branchen oder singuläre Softwareprodukte, sondern bieten modulare Lösungswege, die an den jeweiligen Kontext adaptiert werden können, aber Eigeninitiative des anwendenden Unternehmens erfordern. Das Modell muss also flexibel sein, um angestrebte Ziele und Reifestufen mit modularen und individualisierbaren Lösungswegen zu ermöglichen.

8.2. Berechnung der Priorisierung (Handlungsdruck)

Um eine Überforderung durch eine Vielzahl an gleichgewichteten Maßnahmen bzw. Handlungsempfehlungen zu vermeiden, berechnet das Gestaltungsmodell für jede Dimension und Kategorie einen mathematischen Handlungsdruck. Dieser Wert determiniert die Dringlichkeit einer Maßnahme

und setzt sich aus den Faktoren Leistungslücke (Gap), der anwenderspezifischen Gewichtung der Dimension, sowie einem strukturellen Schichten-Multiplikator, zusammen. Dabei berechnet sich die Leistungslücke (Gap) aus der Differenz zwischen dem eingestellten Soll-Wert und dem gemessenen Ist-Wert. Ist der Ist-Wert größer oder gleich dem Soll-Wert, beträgt die Lücke 0.

Leistungslücke (Gap) der Dimension (Dim):

$$Gap_{Dim} = \max(0, Soll_{\emptyset\, Dim} - Ist_{\emptyset\ Dim})$$

Leistungslücke (Gap) der Kategorie (Kat):

$$Gap_{Kat} = \max(0, Soll_{Kat} - Ist_{Kat})$$

Daraus wird der Handlungsdruck für die Dimensionen und Kategorien gleich berechnet.

Handlungsdruck (H):

$$H = Gap \times Gewichtung_{Dim} \times Multiplikator_{Schicht}$$

Um die Hürdenlogik und die Wichtigkeit stabiler grundlegender Schichten auf das Gestaltungsmodell zu transferieren, wird ein Schichtmultiplikator eingeführt. Eine ungewichtete Priorisierung würde beispielsweise dazu führen, dass Defizite in strategischen Randbereichen identisch priorisiert werden wie kritische Sicherheitslücken in der Basis-Infrastruktur. Maßnahmen in höheren Schichten wie z.B. ethische Richtlinien bleiben wirkungslos, wenn das Fundament z.B. bei der Datensicherheit kompromittiert ist. Um diese Abhängigkeiten abzubilden, implementiert das Modell einen Multiplikator, der Fundament-Lücken mathematisch forciert. Nach ausgiebigen Tests für eine sinnvolle Priorisierung und Darstellung der Handlungsempfehlungen der Dimensionen und Kategorien ergeben sich folgende Schichtmultiplikatoren:

- **Schicht 1 (Fundament): M = 1,5**
  Höchste Priorität, da technologische Defizite systemkritisch sind
- **Schicht 2 (Befähiger): M = 1,2**
  Hohe Priorität zur Sicherstellung der organisationalen Handlungsfähigkeit
- **Schicht 3 (Wirkung): M = 1,0**
  Normale Priorität, da normative Gestaltung erst auf intakten Systemen greift

8.3. Struktur und Zuordnung der Handlungsempfehlung

Um an mehreren organisationalen Handlungsfeldern anzugreifen und eine ganzheitliche Transformation zu ermöglichen, folgt das Gestaltungsmodell einem sozio-technischen Ansatz. Jede Handlungsempfehlung ist systematisch in drei Sub-Ebenen untergliedert:

1. **Strategische Ebene:** Fokus auf Verträge, Positionierung, Budgetierung und Richtlinien
2. **Organisatorische Ebene:** Fokus auf Prozesse, Rollen, Schulungen und Validierungsmethoden
3. **Technische Ebene:** Fokus auf Softwarearchitektur, Hosting, Verschlüsselung und Schnittstellen

Die Zuordnung der spezifischen Handlungsempfehlungen (aus einer integrierten Wissensdatenbank: han-empf.json) erfolgt dynamisch auf Basis des erreichten Ist-Werts in der jeweiligen Kategorie. Je nach Leistungsklasse (Low, Mid, High) werden passgenaue Maßnahmenpakete ausgeliefert, die aus den Skalenlogiken abgeleitet und praxistauglich generalisiert wurden. Die Abgrenzungslogik orientiert sich an den Beispielen der Skalenlogik und einer gleichmäßigen Einteilung der Bewertungsmöglichkeiten des Fragebogens:

- **Niedrig (Ist-Wert < 2):** Die Maßnahmen zielen auf die Beseitigung existenzbedrohender Abhängigkeiten und die Etablierung von Basis-Fähigkeiten ab.
- **Mittel (2 < Ist-Wert < 4):** Die Maßnahmen fokussieren sich auf Standardisierung, Prozess-Governance und den Ausbau technologischer Interoperabilität.
- **Hoch (4 < Ist-Wert < 5):** Maßnahmen dienen der Optimierung, der proaktiven Marktgestaltung und der absoluten Autonomie.
- **Kontrollmechanismus (Ist-Wert = 5):** Bewertet sich eine Organisation in einer Kategorie mit maximaler Punktzahl, werden alle Maßnahmenpakete (Niedrig, Mittel und Hoch) zur Überprüfung und zum Abgleich angezeigt. Zudem erscheint eine Meldung, die zur kritischen Überprüfung und Abgleich aller umgesetzten Maßnahmen auffordert.

Hinweis: Alle Handlungsempfehlungen sind im Anhang einzusehen.

8.4. Umsetzung und Visualisierung in der Webanwendung

Die technische Instanziierung des Gestaltungsmodells in der Webanwendung übersetzt die komplexen Berechnungen in ein intuitives, handlungsorientiertes Dashboard.

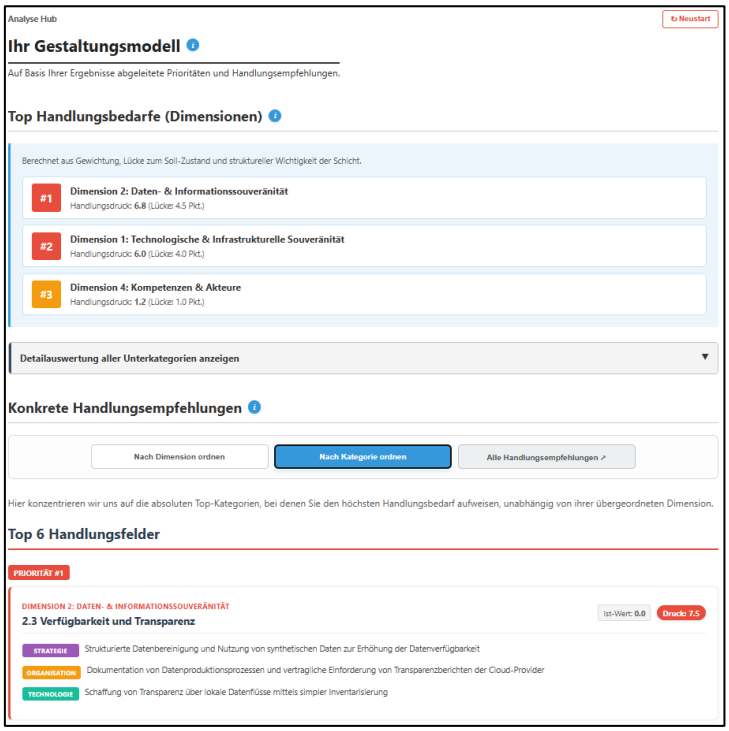

Abbildung 11: Dashboard des Gestaltungsmodells (Webanwendung)

Um den berechneten Handlungsdruck und die daraus resultierende Priorisierung greifbarer zu machen, ordnet die Webanwendung den berechneten Handlungsdruck-Werten eine relative Ampellogik zu. Die Dringlichkeit wird stets in Relation zum höchsten im Unternehmen identifizierten Handlungsdruck gesetzt. Dadurch wird sichergestellt, dass auch bei dimensionsübergreifender gleichbleibender Performance die größten relativen Schwachstellen hervorgehoben werden. Die Ampellogik ist wie folgt eingestellt und beeinflusst nur die Darstellungsebene:

- **Rot (Akuter Handlungsbedarf):** Die oberen 40 % der Handlungsdrücke (relativ zum angegebenen Maximalwert).
- **Gelb (Moderater Handlungsbedarf):** Die mittleren 45% der Handlungsdrücke.
- **Grün (Gering / Kein Handlungsbedarf):** Die unteren 15% der Handlungsdrücke.

Unter der Darstellung der Handlungsbedarfe der Dimensionen und der ausklappbaren Ansicht der Handlungsbedarfe der Kategorien (Detailansicht), bietet die Benutzeroberfläche zudem dynamische Sortiermechanismen an. Anwender können wählen, ob die Listung der Handlungsempfehlungen nach Handlungsdruck der Dimensionen gruppiert angezeigt wird (wobei pro Dimension die zugehörigen Kategorien und deren dreistufige Handlungsempfehlungen gelistet werden) oder strikt absteigend

nach Handlungsdruck der einzelnen Kategorien sortiert wird. Letzteres generiert eine sofort anwendbare, priorisierte Roadmap für das Management. Darüber hinaus können alle Handlungsempfehlungen in einem Top-Down Baumdiagramm in einem separaten Tap geöffnet werden. Schlussendlich werden die der eingegebenen Bewertung zugeordneten Handlungsempfehlungen sortiert angezeigt, wobei nur die höchstpriorisierten sechs Kategorien (oder zwei Dimensionen mit je drei Kategorien) direkt angezeigt werden. Die restlichen Kategorien oder Dimensionen sind aus Gründen der Übersichtlichkeit in ausklappbaren Leisten hinterlegt.

9. Anwendungsbeispiel

Um die qualitative Evaluierung durch Experten zu unterstützen und die Funktionsweise der erstellten Artefakte dieser Arbeit zu demonstrieren, wird in diesem Kapitel ein Anwendungsbeispiel ausgearbeitet und durch die Modelle hinsichtlich der digitalen Souveränität bewertet. Im Mittelpunkt dieses Anwendungsbeispiels steht ein fiktives KI-integrierendes Unternehmen, da die momentane wissenschaftliche, aber auch öffentliche Datenlage nicht genug Informationen über Souveränitätszustände und -Bemühungen preisgibt, um eine vollständige und wahrheitsgetreue Modellierung digitaler Souveränität von KI-integrierenden Unternehmen in der nachgelagerten KI-Wertschöpfungskette zu ermöglichen.

9.1. Vorstellung des Use Cases

Das Anwendungsbeispiel bzw. Untersuchungsobjekt, um die Funktion und Validität der entwickelten Modelle zu beweisen ist das fiktive Unternehmen MEDfiction GmbH. Dieses Unternehmen versucht einen klassischen Vertreter deutscher KMU darzustellen und ist im stark regulierten Medizintechnik-Sektor tätig, bei dem digitale Souveränität eine substanzielle Rolle spielt. In der KI-wertschöpfungskette zählt die MEDfiction GmbH primär als Integrator und Anwender, also als nachgelagerter Akteur. Dabei wird proprietäres Domänenwissen mit Foundation-Models verknüpft, um branchenspezifische Mehrwerte zu generieren.

9.1.1. Unternehmensprofil: MEDfiction GmbH

- **Sitz:** Heilbronn, Deutschland
- **Rechtsrahmen:** unterliegt DE- & EU-Recht – DSGVO, MPDG, MDR, CRA, NIS-2, EU AI Act *
- **Unternehmensgröße:** 55 Mitarbeiter, Fokus auf med. Hard- & Softwareentwicklung (und Vertrieb)
- **Geschäftsmodell:** Entwicklung und Vertrieb von Software as a Medical Device (SaMD).
- **Kernprodukt:** Radiologisches Assistenzmodul mit Software, das MRT-Scans mittels künstlicher Intelligenz analysiert, um Diagnose-Prozesse zu beschleunigen und frühzeitige Erkennung von Auffälligkeiten in der MRT-Bildgebung ermöglichen.
- **Rolle in der KI-Wertschöpfungskette:** Nachgelagerter Integrator - MEDfiction trainiert und optimiert externe Foundation-Modelle eines US-Hyperscalers über lizenzierte Schnittstellen mit proprietären Daten und integriert sie in eigene Produkte.

Basiert grob auf den folgenden Unternehmen:

- *mediaire (Berlin): Entwickelt KI-gestützte Software zur automatisierten Erkennung von Auffälligkeiten in der MRT-Bildgebung, insbesondere für die Neurologie (z. B. Multiple Sklerose).*
- *FUSE-AI (Hamburg): Bietet mit der Plattform Sherlog eine intelligente Software an, die Radiologen dabei unterstützt, Läsionen und Tumore auf medizinischen Bildern schneller zu identifizieren.*
- *Aignostics (Berlin): Dreht sich um die computergestützte Pathologie. Ihre KI analysiert komplexe Gewebeschnitte für die Präzisionsmedizin und die Krebsforschung.*

9.1.2. Branchenkontext und Relevanz digitaler Souveränität

Der deutsche Gesundheitssektor strebt aktuell eine tiefgreifende, datengetriebene digitale Transformation an, welche durch gesetzliche Initiativen des Bundesministeriums für Gesundheit (BMG) unterstützt wird. Dadurch steigt zugleich auch der Druck auf Technologieanbieter, hochverfügbare, sichere und interoperable Lösungen zu liefern (vgl. BMG 2026).

Dem hochsensiblen medizinischen Umfeld liegt jedoch eine signifikante Diskrepanz zwischen verschiedenen Ebenen der digitalen Souveränität zugrunde. Es sollte zwingend zwischen Daten-, Betriebs- und technischer Souveränität unterschieden werden (vgl. Wenger 2025). Die Datensouveränität, also die rechtssichere Verarbeitung und Kontrolle von Daten, ist im Gesundheitssektor durch strikte Compliance-Vorgaben oftmals noch gewährleistet. Die höchste Stufe, die technische Souveränität, stellt für KMU nach Wenger (2025) jedoch eine massive Hürde dar, da die Vermeidung von Vendor-Lock-Ins bei Hardware- und Software-Stacks angesichts der Marktmacht globaler Hyperscaler kaum realisierbar erscheint. Diese Problematik wird auch von führenden Akteuren der Gesundheitsbranche wie der Fresenius SE adressiert. Das Unternehmen strebt zunehmend strategische Allianzen an, um souveräne, interoperable und KI-gestützte Gesundheitsplattformen zu etablieren. Diese sollen offene Industriestandards nutzen und höchste Anforderungen an die Datenhoheit und Souveränität erfüllen (vgl. Fresenius 2026). Als Teil der deutschen KMU in der Gesundheitsbranche steht die MEDfiction GmbH vor sehr ähnlichen Herausforderungen, verfügt allerdings über weitaus weniger Ressourcen. Daher wären eine gezielte, multidimensionale Messung, Bewertung und Gestaltung der eigenen digitalen Souveränität ein existenzsichernder Schritt.

9.1.3. Profilausprägung: Ist- und Soll-Zustand

Die strategischen, organisationalen und technischen Einzelheiten des Unternehmens werden für die Strukturierte Evaluierung durch die Modelle direkt entlang der sechs Dimensionen aufgeschlüsselt. Je Dimension wird zwischen dem Ist-Zustand des Unternehmens und der strategischen Ambition, also Soll-Zustand, unterschieden. Die folgende Tabelle zeigt zusammenfassend die darauffolgende Ausführliche Beschreibung der einzelnen Dimensionen der MEDfiction GmbH.

Dimension	Ist-Zustand (Status quo)	Soll-Zustand (Strategisches Ziel)
1.	• Sole-Source-Ansatz • KI-APIs sind hartcodiert (hoher Aufwand bei Wechsel) • kein lokaler Fallback bei Ausfällen	• Hybride Architektur (Container/EU-Cloud) • modularer Aufbau • Open-Weight-Modelle als operatives Fallback-System
2.	• Strikte vertragliche Opt-Outs für Trainingsdaten • hoher Transparenzgrad • Pseudonymisierung vor API-Nutzung	• Beibehaltung der Datenhoheit • Implementierung von Confidential Computing für absolute technische Verschlüsselung
3.	• Ad-hoc Beschaffung ohne formale KI-Strategie/Budget • manuelle Tests statt MLOps • keine spezifischen KI-Risikopläne	• Definiertes KI-Budget • automatisierte MLOps (Überwachung von Model Drift) • Verankerung von KI-Bedrohungen im ISMS
4.	• 80 % Abhängigkeit von externen IT-Beratern • fehlende Validierungsmechanismen für Bias • starke cross-funktionale Teams	• Insourcing-Strategie (Ziel: 70 % interne Experten) • Aufbau von "AI Literacy" und systematischen Bias-Audits
5.	• Passiver Konsument im Hyperscaler-Ökosystem (hohes Konzentrationsrisiko) • proprietäre "Data Moat" vorhanden	• Transformation zum Mitgestalter (Open-Source-Engagement) • aktive strategische Nutzung der eigenen Daten-Engstelle
6.	• Hohe Basis-Compliance (MDR/DSGVO) • reaktive EU AI Act Vorbereitung • Black-Box-KI ohne Eskalationsprozesse	• Proaktive AI Act Compliance • Etablierung ethischer CDR-Leitlinien • Nutzung von Explainable AI (XAI) für Transparenz

Abbildung 12: Ist- und Soll-Zustands des Anwendungsbeispiels

Dimension 1: Technologie und Infrastruktur

Aktuell ist die MEDfiction GmbH stark von einem einzelnen US-Hyperscaler abhängig. Das system ist einerseits historisch gewachsen, weshalb proprietäre KI-Schnittstellen hartcodiert in das Kernsystem integriert sind, andererseits hat das Unternehmen bislang nur die Ressourcen gehabt, ein Foundation-Modell selbst zu trainieren und feinabzustimmen. Auch wenn die KI-Schnittstelle leicht durch die eines anderen KI-Anbieters ersetzt werden kann, würde das Trainieren und Feinabstimmen eines

anderen KI-Modells viel Zeit und Ressourcen verschlingen. Diese Sole-Source-Strategie führt folglich zu einem erheblichen Vendor-Lock-In, bei dem es auch keine operativen Fallback-Möglichkeiten hinsichtlich der KI-Integration gibt. Die restliche Software- und Hardwarearchitektur des Produkts ist selbst entwickelt und hat keine direkte Abhängigkeit zu anderen Organisationen. Die generelle KI-Nutzung der Mitarbeiter für den Arbeitsalltag hingegen geschieht über klassische Unternehmenslizenzen, auch bei einem großen KI-Hyperscaler. Diese allgemeinen KI-Modelle müssen nicht individuell trainiert oder abgestimmt werden und die Lizenzen können jederzeit gekündigt oder gewechselt werden.

Zukünftig strebt das Unternehmen zwar kein vollständiges Selbsthosting der KI oder keine vollständige und kostenintensive Multi-Cloud-Strategie an, sondern eine hybride, modulare Architektur. Durch beispielsweise Containerisierung (wie Docker) und leicht austauschbare KI-Schnittstellen soll das System modularisiert werden. Ziel dabei soll es sein, kritische oder sensible Bestandteile und Prozesse auf souveräne europäische Clouds bzw. Rechenzentren auszulagern und ein lokal gehostetes, transparentes aber selbst trainiertes Open-Weight-Modell (ein frei zugängliches KI-Modell) als ausfallsichere Notlösung bereitzustellen.

Dimension 2: Daten und Informationen

MEDfiction bezieht die Trainingsdaten für das externe KI-Modell bereits pseudonymisiert oder anonymisiert durch Kooperationen mit Krankenhäusern und Kunden, sowie Forschungsnetzwerken. Hier besteht bislang eine hohe Abhängigkeit von den Datenlieferanten, da es aktuell noch keine Verpflichtung für Krankenhäuser oder ähnliches gibt, Daten für Zwecke der Gesundheitsforschung zu teilen. Dies soll sich allerdings 2029 mit dem Inkrafttreten der Sekundärnutzung des European Health Data Space (EHDS) ändern, bei dem beispielsweise Krankenhäuser dazu verpflichtet werden, bestimmte elektronische Gesundheitsdaten anonymisiert über eine staatlich regulierte Zugangsstelle mit medizinisch forschenden Unternehmen zu teilen (vgl. Europäische Kommission 2026). Hinsichtlich der Daten-Weitergabe an KI-Hersteller ist das Unternehmen aufgrund der Branchenregulationen sehr souverän. Der Ist-Zustand zeichnet sich durch strikte B2B-Enterprise-Verträge aus, die mittels Opt-Outs rechtlich und technisch ausschließen, dass Patientendaten zum Training der externen Foundation Models genutzt werden. Zukünftig soll dieses hohe Niveau der Datensouveränität beibehalten und durch den Einsatz von Hardware-basiertem Confidential Computing (verschlüsselte Verarbeitung) technisch zur absoluten Vertraulichkeit und Souveränität ausgebaut werden.

Dimension 3: Organisation und Governance

Gegenwärtig reagiert die MEDfiction GmbH im Bereich der KI-Governance sehr reaktiv. Die Entwicklung musste schnell gestartet und umgesetzt werden, weshalb die Lizensierung der KI-Modelle ohne strategische Evaluation der Anbieter, Make-or-Buy-Entscheidungen oder ohne spezifische Budgetplanung für technologische unabhängigkeitsfördernde Maßnahmen erfolgte. Auch struktu-

rierte Machine Learning Operations (MLOps) zur Überwachung der Modellqualität sind nicht vollständig implementiert. Angestrebt wird die Etablierung eines zentralen IT-Portfolio-Managements bei dem bspw. eine KI-spezifische Risikomatrix (inkl. Bedrohungen wie Model Drift oder Prompt-Injection) in das Information Security Management System (ISMS) integriert und automatisierte MLOps-Pipelines zur Qualitätskontrolle implementiert werden.

Dimension 4: Kompetenz und Akteure

Das medizinische Domänenwissen ist im Unternehmen hervorragend ausgeprägt und fließt durch cross-funktionale Teams (Radiologen und IT) in die Entwicklung ein. Es besteht jedoch eine kritische technische Abhängigkeit, bei der der aktuelle Zustand zeigt, dass die KI-Architektur zu 80 % von externen Beratungsfirmen aufgebaut wurde und dass es intern an der methodischen Kompetenz zur systematischen Überprüfung von KI-Bias fehlt. Strategisch soll ein Insourcing der Kernkompetenzen erfolgen. Dazu hat das Unternehmen kürzlich mit umfangreichen Schulungs- und Weiterentwicklungsmaßnahmen angefangen. Das Unternehmen strebt ein Verhältnis von 70 % interner zu 30 % externer KI-Expertise an. Zu diesem Zweck sollen eigene Data Scientists und AI-Engineers rekrutiert oder weitergebildet und interne „AI Literacy"-Programme weiter ausgerollt werden.

Dimension 5: Ökosystem und Machtverhältnisse

Das Unternehmen agiert momentan als rein passiver Konsument im Ökosystem des Hyperscalers, was ein hohes asymmetrisches Risiko birgt. Positiv ist jedoch, dass die MEDfiction GmbH durch ihre umfangreiche Datenbank mit über einer Million fachärztlich annotierten MRT-Scans eine eigene strategische Engstelle (Data Moat) in der Wertschöpfungskette geschaffen hat. Der Soll-Zustand versucht, diese Daten-Engstelle aktiv als Wettbewerbsvorteil und Verhandlungsmittel zu nutzen. Zudem soll die Rolle im Ökosystem durch die proaktive Beteiligung an medizinischen Open-Source-Gemeinschaften vom Technologie-Nehmer zum Mitgestalter transformiert werden.

Dimension 6: Regulationen und Normen

Als Medizintechnik-Unternehmen verfügt das MEDfiction über eine exzellente Compliance hinsichtlich der Medical Device Regulation (MDR) und DSGVO, da das Unternehmen so gut wie nur anonymisierte oder pseudonymisierte Daten verarbeitet. Die Durchsetzung des bisherigen EU AI Act und die Vorbereitung auf kommende Schritte werden ebenfalls vorbereitet, sofern die externe Technolgie das zulässt. Denn technologisch gesehen stehen die Experten des Unternehmens und Anwender bei KI-Fehlentscheidungen vor einer technologischen Black-Box, da es an Erklärbarkeit seitens des KI-Providers mangelt. Es wurden zwar Feedback-Mechanismen für Kunden eingerichtet, die Fehlentscheidungen der KI an MEDfiction-Entwickler übermitteln, jedoch bleiben diese Rückmeldungen intern und weder Kunden noch KI-Provider werden davon in Kenntnis gesetzt. Zukünftig soll eine Corporate Digital Responsibility Erklärung formuliert werden, die den werteorientierten Einsatz von KI in ihren Produkten regelt. Ergänzend strebt das Unternehmen die Implementierung rudimentärer Explainable AI (XAI)-

Dashboards an, um die ärztliche Nachvollziehbarkeit und Verantwortlichkeit auf Systemebene zu gewährleisten.

9.1.4. Bepunktung des Fragebogens

Auf Basis der Unternehmens-Persona und weiteren Angaben wird nun der Fragebogen ausgefüllt, die Dimensionen gewichtet und ein Soll-Zustand definiert. Die Eingaben, die auf Basis dieser Information getätigt wurden, werden in den drei folgenden Tabellen angezeigt:

Kategorie	Punkte	Konfidenz	Begründung
1.1 Infrastrukturelle Kontrolle und Cloud Abhängigkeit	1	3	100 % Sole-Source-Abhängigkeit bei einem US-Hyperscaler, es gibt keine eigene KI-Infrastruktur, allerdings ist die restliche Software- und Hardware-Architektur selbst entwickelt (daher 1 Punkt).
1.2 Interoperabilität & Modularität	1	3	Proprietäre APIs sind hartcodiert im Kernsystem verankert, ein Wechsel würde Monate der Anpassung und des Trainings erfordern. Allerdings ist der technische Wechsel schnell (daher 1 Punkt).
1.3 Operative Kontrolle & Resilienz	0	3	Es existiert kein operativer Fallback, ein API-Ausfall führt zum sofortigen Funktionsverlust des Diagnose-Moduls.
2.1 Eigentum und Kontrolle der Daten und Informationen	4	3	Der Zugang ist durch einige Kooperationen mit Krankenhäusern aktuell gesichert und es wurde ein Datenbestand von über einer Millionen MRT-Scans aufgebaut. Dennoch besteht eine geringe Abhängigkeit (daher 4 Punkte).
	4	3	Strikte B2B-Enterprise-Verträge und vertragliche Opt-Outs sichern den Schutz der Patienten- und Unternehmensdaten vor externer Verwertung vertraglich absolut ab (4 Punkte).
2.2 Datensicherheit und Vertraulichkeit	3	3	Daten-Pseudonymisierung in Kliniken ist in diesem Fall Standard, Hardware-gestützte Verschlüsselung im RAM des Providers (Confidential Computing) wird bislang vom Hyperscaler nicht angeboten (deshalb 3 Punkte).
	5	3	Die Zuweisung von Datenklassen zu KI-Systemen ist klar definiert und wird technisch durch das Access-Management strikt getrennt (5 Punkte).
2.3 Verfügbarkeit und Transparenz	4	3	Der Datenfluss ist aufgrund der medizinischen Audit-Pflichten (MDR) lückenlos und für Prüfinstanzen nachvollziehbar dokumentiert, allerdings ist das KI-Modell eine Blackbox (daher 4 Punkte).
3.1 Strategische Handlungsfähigkeit	1	1	Externe Modelle wurden historisch ad-hoc lizenziert. Eine formalisierte Strategie oder Budgets für technologische Autonomie fehlt allerdings noch größtenteils (daher 1 Punkt).
3.2 Prozessbasierte und strukturelle Governance	1	2	Es erfolgen meistens unstrukturierte, manuelle Tests, eine automatisierte Echtzeit-Überwachung der Modellqualität fehlt komplett (1 Punkt).

3.3 Risikomanagement	1	2	Ein zertifiziertes ISMS existiert, berücksichtigt KI-Schwachstellen und entsprechende Eskalationspläne bisher nicht (1 Punkt).
4.1 Bewertungskompetenz & Entscheidungssouveränität	2	3	Mediziner prüfen auf offensichtliche fachliche Fehler und die externen, endanwendenden Radiologen sind für die finale Kontrolle der Ausgaben zuständig. Methodische und statistische Bias-Audits der Algorithmen fehlen intern (daher 2 Punkte).
4.2 Talentverfügbarkeit & Ressourcen	2	2	Etwa 80 % der KI-Architektur wurden extern durch Consultants aufgebaut und es herrscht eine hohe Abhängigkeit von fremdem Know-how hinsichtlich der KI. Die medizinischen Kenntnisse sind größtenteils inhouse (daher 2 Punkte).
4.3 Relationale Praktiken & Domänenwissen	4	3	Radiologen sind strukturell eng in das Finetuning eingebunden und die cross-funktionale Zusammenarbeit ist hoch ausgeprägt (4 Punkte).
5.1 Strukturelle Macht & Plattform-Abhängigkeit	0	3	Es besteht eine extreme Konzentration auf einen Anbieter (Single-Point-of-Failure) und damit absolute Plattform-Abhängigkeit (0 Punkte).
5.2 Rollenverteilung im sozio-technischen Ökosystem	0	3	Das Unternehmen agiert ausschließlich als passiver Konsument, es gibt bislang noch keine Mitgestaltung in offenen Netzwerken (0 Punkte).
5.3 Kontrolle über Engstellen in der Wertschöpfungskette	3	3	Eine exklusive MRT-Datenbank (Data Moat) existiert, wird aber strategisch noch nicht aktiv als Verhandlungshebel gegenüber Providern monetarisiert (3 Punkte).
6.1 Regulatorische & Rechtliche Compliance	3	3	Die MDR- und DSGVO-Basis-Compliance ist exzellent und der EU AI Act wird so gut es geht durchgesetzt und vorbereitet. Allerdings fehlt es an Erklärbarkeit seitens des KI-Providers (daher 3 Punkte).
6.2 Normative Verantwortung	3	3	Allgemeine medizinethische Grundregeln gelten und der unternehmenseingene Code of Conduct wird eingehalten, allerdings wurde eine ausformulierte und spezifische CDR-Richtlinie für KI wurde bislang nicht vollständig ausgearbeitet (3 Punkte).
6.3 Transparenz, Fairness & Verantwortlichkeit	2	3	Die KI fungiert gegenüber Ärzten (Kunden) als Black-Box, diese können allerdings an Feedback an MEDfiction bei Fehlerhaften KI-Ausgaben übermitteln. Dies dient allerdings nur zur Verbesserung des KI-Trainings (daher 2 Punkte).

Abbildung 13: Antworten des Fragebogens im Anwendungsbeispiel

Dimension	Gewichtung	Begründung
1. Technologische & Infrastrukturelle Souveränität	8	Ausfälle der Infrastruktur stoppen Diagnoseprozesse. Aufgrund begrenzter KMU-Ressourcen ist absolute technische Autonomie jedoch unrealistisch.
2. Daten- & Informationssouveränität	10	Der Schutz hochsensibler Patientendaten ist das absolute Fundament. Ein Verlust der Datenhoheit führt zum sofortigen Entzug der Marktzulassung.

3. Organisationale Souveränität & Governance	7	MLOps und formale Strategien sind wichtig für die Skalierung und Qualitätssicherung, gefährden das Tagesgeschäft kurzfristig jedoch weniger als Datenlecks.
4. Kompetenzen & Akteure	8	Internes KI-Know-how ist essenziell, um die Nachvollziehbarkeit klinischer KI-Bewertungen sicherzustellen und Abhängigkeiten von externen Beratern zu reduzieren.
5. Ökosysteme & Machtverhältnisse	5	Eine aktive Ökosystem-Gestaltung ist langfristig erstrebenswert, jedoch für ein KMU mit 55 Mitarbeitern zunächst sekundär gegenüber Compliance und Datensicherheit.
6. Regulatorische & Normative Souveränität	10	Fehlende Compliance (MDR, EU AI Act) impliziert ein existenzielles rechtliches Risiko. Ohne normative Sicherheit ist die KI-Integration in Krankenhäusern ausgeschlossen.

Abbildung 14: Dimensionsbewertung des Anwendungsbeispiels

Kategorie	Soll-Punkte	Begründung des strategischen Zielbilds (Soll)
1.1 Infrastrukturelle Kontrolle und Cloud Abhängigkeit	4	Umstieg auf eine hybride Architektur bei der kritische Module auf EU-Servern laufen, der Hyperscaler federt nur noch Spitzenlasten ab.
1.2 Interoperabilität & Modularität	4	Einführung von Container-Technologien (z. B. Docker), um KI-Schnittstellen bei Anbieterwechsel flexibel austauschen zu können.
1.3 Operative Kontrolle & Resilienz	3	Implementierung eines lokal gehosteten Open-Weight-Modells als garantiertes Backup-System.
2.1 Eigentum und Kontrolle der Daten und Informationen	5	Der vollumfängliche rechtliche und technische Schutz der Trainingsdaten vor externer Verwertung bleibt höchste Priorität.
2.2 Datensicherheit und Vertraulichkeit	5	Ergänzung der Pseudonymisierung durch Confidential Computing zur Verschlüsselung von "Data in Use" beim KI-Provider.
2.3 Verfügbarkeit und Transparenz	4	Beibehaltung der lückenlosen MDR-konformen Dokumentation der Datenflüsse.
3.1 Strategische Handlungsfähigkeit	4	Aufbau eines IT-Portfolio-Managements mit dediziertem Budget zur Evaluierung europäischer KI-Alternativen.
3.2 Prozessbasierte und strukturelle Governance	4	Etablierung automatisierter MLOps-Pipelines zur proaktiven, systematischen Detektion von Leistungsabfällen oder Fehleranomalien (durch bspw. Model Drift).
3.3 Risikomanagement	4	Vollständige Inklusion von spezifischen KI-Sicherheitsrisiken (Data Poisoning) in das zertifizierte ISMS.
4.1 Bewertungskompetenz & Entscheidungssouveränität	4	Aufbau interner methodischer "AI Literacy", um Bias und Diskriminierung in Algorithmen-Ausgaben statistisch zu auditieren.
4.2 Talentverfügbarkeit & Ressourcen	4	Strategisches Insourcing mit dem Zie eine interne Abdeckung von 70 % der Kernkompetenzen durch eigene KI-Ingenieure abzudecken.
4.3 Relationale Praktiken & Domänenwissen	5	Institutionalisierung der cross-funktionalen Teams. Das Domänenwissen der Ärzte soll die technische Entwicklung führen.
5.1 Strukturelle Macht & Plattform-Abhängigkeit	3	Diversifizierung der Anbieter. Evaluierung von Open-Source-Plattformen (z. B. Hugging Face) zur Risikominimierung.

5.2 Rollenverteilung im sozio-technischen Ökosystem	4	Strategische Partizipation in medizinischen Open-Source-Initiativen zur Etablierung am Markt als anerkannter Complementor.
5.3 Kontrolle über Engstellen in der Wertschöpfungskette	4	Aktive Nutzung der exklusiven MRT-Datenbank als strategisches Verhandlungsmittel gegenüber Providern und dem Wettbewerb.
6.1 Regulatorische & Rechtliche Compliance	5	Verankerung des EU AI Acts in der frühen Entwicklungsphase ("Compliance by Design") zur Vermeidung späterer Zulassungsprobleme.
6.2 Normative Verantwortung	4	Ausarbeitung und aktives Vorleben einer CDR für den Umgang mit künstlicher Intelligenz.
6.3 Transparenz, Fairness & Verantwortlichkeit	4	Implementierung von Explainable AI (XAI) wie z. B. Saliency Maps, um Ärzten die Nachvollziehbarkeit der KI-Entscheidung zu ermöglichen.

Abbildung 15: Definition des Soll-Zustands des Anwendungsbeispiels

9.2. Ergebnisse der Modellierung

Nach der systematischen Eingabe der Parameter der MEDfiction GmbH in die Webanwendung werden im Folgenden die Ergebnisse der Modellierung präsentiert. Die Auswertung findet nach der Logik der Webanwendung und bisherigen Vorgehensweise statt. Von der deskriptiven Messung des Bewertungsmodells, über die evaluative Einordnung des Reifegradmodells bis hin zu normativen Handlungsempfehlungen des Gestaltungsmodells.

9.2.1. Bewertungsmodell

Das Bewertungsmodell diagnostiziert den aktuellen Zustand der digitalen Souveränität der MEDfiction GmbH und liefert eine punktuelle, quantitative Bestandsaufnahme. Durch die Verarbeitung der Angaben zu den 20 Prüffragen generiert das Dashboard zentrale KPIs , die das Verhältnis von Abhängigkeit zu Autonomie messbar machen.

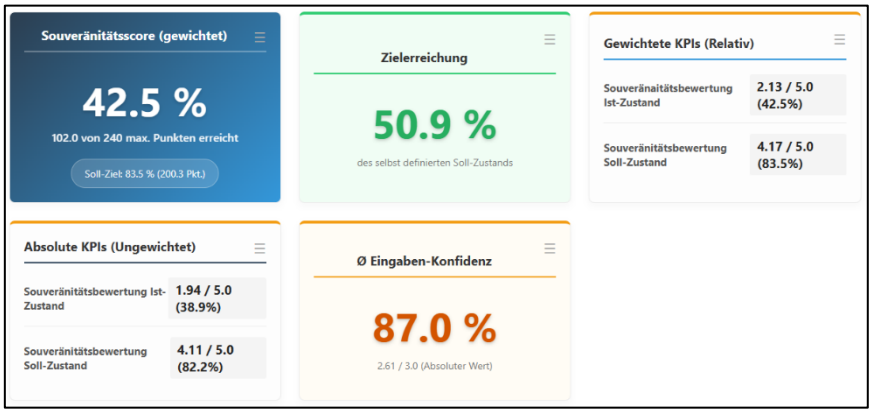

Abbildung 16: Bewertungsmodell - KPIs des Anwendungsbeispiels

Mit den vorherigen Eingaben hat die MEDfiction GmbH einen Souveränitätsscore von 42,5% (102 von maximal 240 erreichbaren Punkten) erreicht. Damit hat das Unternehmen knapp die Hälfte zu dem von ihm definierten Soll-Zustand erreicht, was die Gap-Analyse mit einem Zielerreichungswert von 50,9% bestätigt und auf einen erheblichen strategischen Handlungsbedarf hindeutet. Die MEDfiction GmbH hat einen unternehmensspezifischen Soll-Zustand von 83,5 % (200,3 Punkte bzw. 4,17/5,0) definiert. Das KMU hat mit der Souveränitätsbewertung von 2,13/5,0 folglich erst knapp die Hälfte des Weges zur angestrebten technologischen und organisationalen Autonomie absolviert. Der absolute (ungewichtete) Ist-Wert fällt mit 38,9 % (1,94/5,0) noch geringer aus. Diese Differenz indiziert, dass das Unternehmen in den Bereichen, die es als strategisch überlebenswichtig priorisiert hat (insbesondere Dimension 2 mit einer Gewichtung von 10), etwas besser abschneidet als in Randbereichen. Den-

noch ist die Organisation auf einem generell niedrigen Souveränitätsniveau, das von starken technologischen, aber auch organisationalen Abhängigkeiten geprägt ist. Zudem ist das Unternehmen insgesamt zu 87% sicher, vollständig, faktisch belegte Aussagen getroffen zu haben.

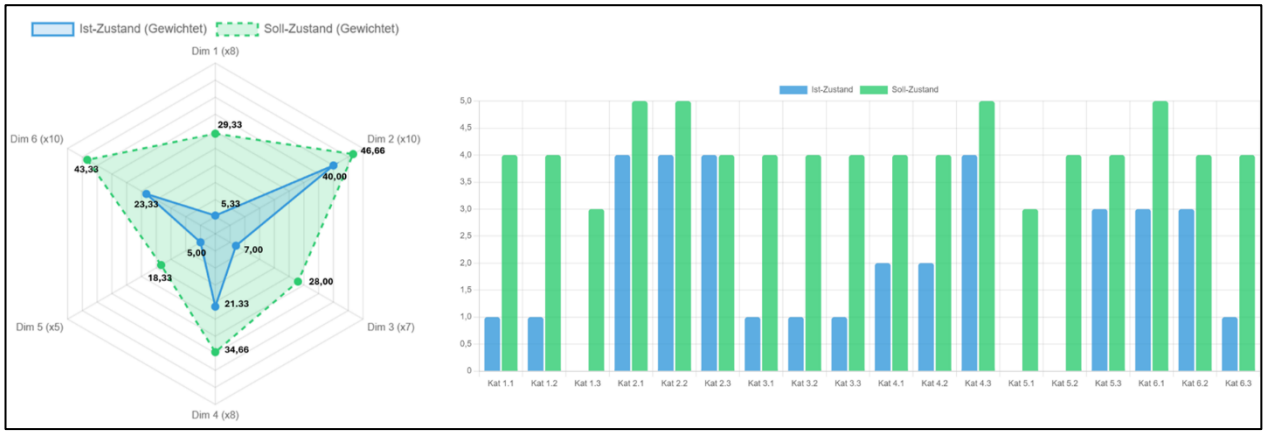

Abbildung 17: Bewertungsmodell - Dimensionsübersicht des Anwendugnsbeispiels

Die gewichteten Dimensionen (visualisiert im Netz-Diagramm) und die Auswertung auf Kategorie-Ebene (Säulendiagramm) verdeutlichen das stark fragmentierte Profil des Unternehmens. Hier wird deutlich, dass die fundamentale Stärke des Unternehmens mit 40,00 von 46,00 angestrebten Punkten in der Daten- und Informationssouveränität liegt. Durch regulatorische Vorgaben wie der Medical Device Regulation oder DSGVO, ist das Unternehmen unabhängig von KI-bezogenen Regulierungen gezwungen, Datenflüsse transparent zu halten und vertragliche Opt-Outs strikt durchzusetzen.

Allerdings wird auch deutlich sichtbar, dass das Unternehmen kritische Defizite aufweist. In den Dimensionen 1: Technologie und Infrastruktur, 3: Organisation und KI-Governance, sowie 5: Ökosysteme und Machverhältnisse zeigt das Netzdiagramm starke Diskrepanzen im Vergleich zum Soll-Zustand. In Dimension 1 werden lediglich 5,33 von angestrebten 29,33 Punkten erreicht, was ein Beweis für den harten Vendor-Lock-in durch hartcodierte APIs und fehlende Fallback-Mechanismen ist. Auch Dimension 3 mit nur 7,00 von 28,00 Soll-Punkten deutet auf Schatten IT und unkoordinierte Entwicklung ohne strategische Planung hin. Dimension 5 hingegen zeigt in dem Säulendiagramm zwar, dass eine Kontrolle über Engstellen bzw. schwer substituierbaren proprietären Daten vorliegt, allerdings keine Partizipation oder strukturelle Macht in Ökosystemen ausgeübt wird. Mit 5,00 von 18,33 Punkten hat das Unternehmen also eher die Position eines passiven Technologie-Konsumenten.

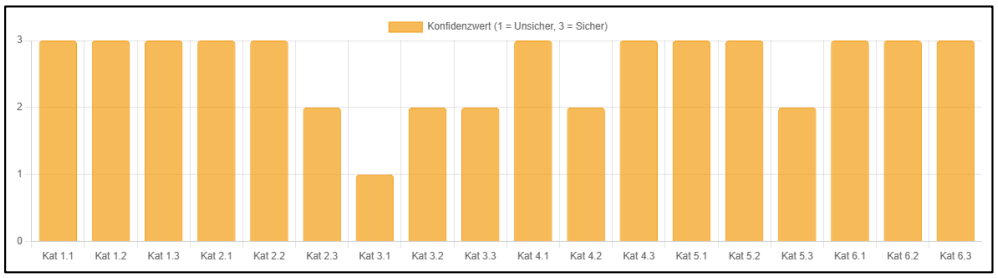

Abbildung 18: Bewertungsmodell - Konfidenz des Anwendungsbeispiels

Ein zentrales Element des entwickelten Modells ist die Erfassung der Eingaben-Konfidenz, um organisationale Blind Spots zu identifizieren und eine Validitätsprüfung durch eine Eingaben-Konfidenz zu ermöglichen. Die durchschnittliche Eingaben-Konfidenz liegt bei sehr hohen 87,0 % (absoluter Wert: 2,61 von 3,0). Dies belegt, dass die berechneten Scores vertrauenswürdig und valide sind, sowie auf faktisch belegter Dokumentation basieren. Die visuelle Auswertung der Konfidenzwerte auf Kategorieebene (Konfidenz-Säulendiagramm) bestätigt zudem die theoretischen Annahmen der Persona, dass die meisten Fragen durchgehend mit höchster Sicherheit (Wert 3) beantwortet werden konnten. Allerdings sinkt die Konfidenz im Bereich der Datenverfügbarkeit und Governance (Kategorien 2.3 bis 3.3) auf den Wert 2 („teilweise geklärt“) und bei der strategischen Handlungsfähigkeit (Kategorie 3.1) sogar auf den Wert 1 („Bauchgefühl“). Dies zeigt, dass das Unternehmen im Bereich der KI-Strategie, MLOps und Risiko-Governance über keine formalisierten Dokumente oder Metriken verfügt. Dies führt zu einer hohen organisationalen Ungewissheit.

9.2.2. Reifegradmodell

Aufbauend auf der quantitativen Datenerhebung überführt das Reifegradmodell den berechneten Ist-Zustand in eine fünfstufige Skala. Dieses Artefakt dient als Analyserahmen, der den aktuellen Status der digitalen Souveränität mithilfe einer integrierten Hürdenlogik auch strukturelle Abhängigkeiten zwischen den Dimensionen aufdeckt.

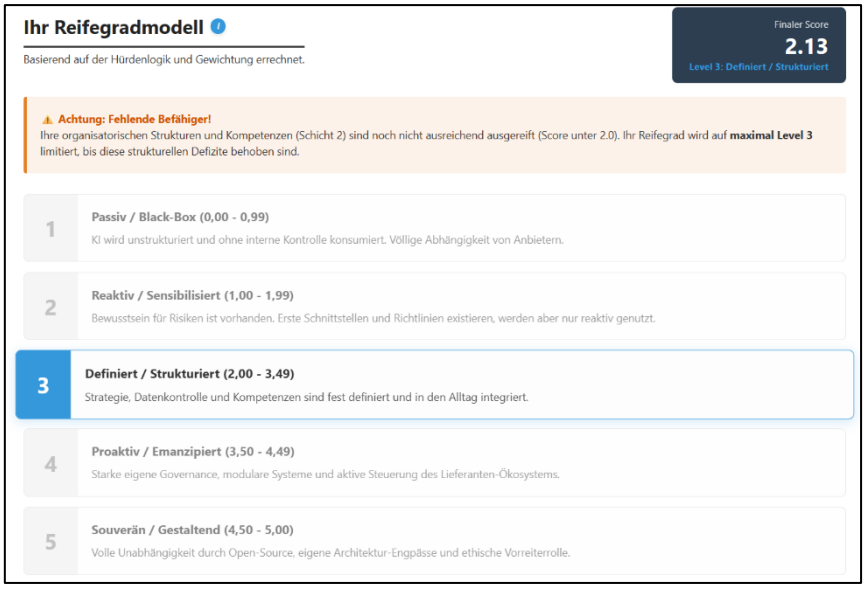

Abbildung 19: Reifegradmodell - Reifegrad des Anwendungsbeispiels

Mit einem final gewichteten Score von 2,13 ordnet das Modell die MEDfiction GmbH in Level 3 (Definiert / Strukturiert) ein, was mit dem Souveränitätsscore von 42,5% übereinstimmt. Durch die Skalenlogik des Modells bedeutet dies, dass Strategien, Datenkontrolle und Kompetenzen in Teilbereichen fest definiert sind. Unter Einbezug der Erkenntnisse des Bewertungsmodells fällt auf, dass die

Einstufung primär von der enormen Stärke in den Dimensionen 2 (Daten und Informationen) und 4 (Kompetenzen und Akteure) getragen wird, während andere Bereiche kaum ausgeprägt sind. Um einen methodischen Aggregationseffekt, also das Überblenden technologischer Defizite durch hervorragende datenschutzrechtliche Rahmenbedingungen, zu verhindern, greift hier die analytische Schichten- und Hürdenlogik des Modells.

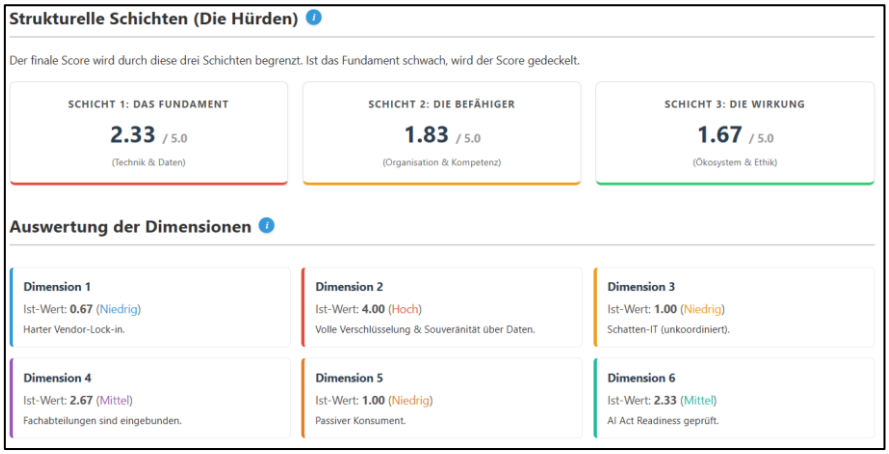

Abbildung 20: Reifegradmodell - Schichten & Dimensionen des Anwendungsbeispiels

Die Auswertung der drei architektonischen Schichten offenbart das strukturelle Ungleichgewicht der Organisation. Das technologische Fundament überschreitet den kritischen Schwellenwert (zu einer Warnung und Begrenzung des Reifegrads) von 2,0 knapp. Dies resultiert allerdings aus der Tatsache, dass der schwerwiegende Vendor-Lock-in in Dimension 1 (0,67 Punkte) durch die herausragenden Werte der Daten- und Informationssouveränität (Dimension 2: 4,00 Punkte) kompensiert wird.

In der mittleren Schicht deckt das Modell dann ein systemkritisches Defizit auf, bei dem die Kombination aus unkoordinierter (Schatten-) IT in der KI-Governance (Dimension 3: 1,00 Punkte) und der teilweisen Einbindung von Fachabteilungen (Dimension 4: 2,67 Punkte) zu einem Schichtenwert von 1,83 führt. Da die zweite Schicht nun den Wert von 2,00 unterschreitet, wird eine Warnung angezeigt und die Softwareapplikation demonstriert die implementierte, deterministische Hürdenlogik des Reifegradmodells. Obwohl der Gesamtscore (oder Finale Score) von 2,13 die Limitierungsgrenze von 3,49 (Obergrenze Level 3) in diesem Fall nicht überschreitet, setzt der Algorithmus hier eine strategische Grenze. Das Modell signalisiert dem Unternehmen eindeutig, dass isolierte Investitionen in die oberen Schichten (z.B. eine noch bessere EU AI Act Compliance) die Organisation nicht auf Level 4 oder 5 anheben werden, solange die internen Befähiger-Schichten (Governance oder Akteure) keine Mindeststabilität aufweisen.

Zwar weist die dritte Wirkungs-Schicht die niedrigste Bewertung auf, ist nach der Hürdenlogik aber auch die Schicht mit der niedrigsten Priorität, da hier die Auswirkungen von souveränitätssteigernden Maßnahmen durch die fehlende Befähiger-Basis ab einem Gewissen Punkt limitiert ist.

Zur besseren Übersicht und Verständnis der Schichtenbewertung aber auch zur tiefergehenden Ursachenanalyse liefert das Reifegradmodell abschließend automatisierte, generalisiert deskriptive Diagnosen für jede der sechs Dimensionen. Hier wird beispielsweise Dimension 1 mit einem Wert von 0,67 als niedrig klassifiziert während Dimension 2 mit einem Wert von 4,00 mit hoch bewertet wird. Dadurch soll der Kompensationseffekt bei der Mittelwertbildung der Fundament-Schicht aufgedeckt werden.

Zusammenfassend validiert die Anwendung des Reifegradmodells auf die Beispielanwendung die Trennschärfe und analytische Nutzbarkeit des Modells. Es zeigt präzise auf, dass die MEDfiction GmbH zwar über ein valides Datenfundament verfügt, der Weg zu einer echten, emanzipierten digitalen Souveränität (Level 4 oder 5) jedoch eine fundamentale Restrukturierung der organisationalen Governance erfordert.

9.2.3. Gestaltungmodell

Das finale Artefakt des iterativen Evaluierungsprozesses digitaler Souveränität ist das Gestaltungsmodell. Während die vorangegangenen Modelle den Ist-Zustand deskriptiv erfassen und evaluativ einschränken, transformiert das Gestaltungsmodell die identifizierten Leistungslücken in normative, priorisierte Handlungsempfehlungen. Das Ziel besteht darin, der MEDfiction GmbH eine strukturierte Transformations-Roadmap zur Erreichung des definierten Soll-Zustands zur Verfügung zu stellen.

Um eine Überforderung durch eine unüberschaubare Menge an Maßnahmen zu verhindern, berechnet der Algorithmus der Webanwendung den Handlungsdruck. Dieser Wert determiniert die Dringlichkeit und setzt sich aus der Leistungslücke, der unternehmensspezifischen Dimensionsgewichtung sowie dem Schichtmultiplikator zusammen. Über eine relative Ampellogik werden die kritischsten Schwachstellen (die oberen 40 % der Handlungsdrücke) als rote Felder markiert.

Abbildung 21: Gestaltungsmodell - Priorisierung der Dimensionen im Anwendungsbeispiel

Die Analyse des Handlungsdrucks offenbart eine eindeutige strategische Reihenfolge für das Management der MEDfiction GmbH und ermöglicht eine Priorisierung auf Dimensions- und Kategorie-Ebene. Auf Dimensionsebene identifiziert das System folgende Top 3:

1. **Dimension 1 (Technologie & Infrastruktur):** Mit einem Handlungsdruck von 36,00 und einer Lücke zum Soll-Zustand von 3 Punkten besteht ein akuter Bedarf, der rot markiert wird. Dimension 1 hebt sich besonders durch den Schichtmultiplikator hervor, da die fundamentale Schicht besonders hoch priorisiert wird (M = 1,5).
2. **Dimension 3 (Organisation & KI-Governance):** Aufgrund der niedrigeren Priorisierung der Befähiger-Schicht (M = 1,2) wird der Handlungsdruck mit einer Lücke von 3,0 zum Soll-Zustand auf 25,2 berechnet. Dennoch ist der Handlungsbedarf enorm und wird ebenfalls Rot angezeigt.
3. **Dimension 6 (Regulatorische & Normative Souveränität):** Mit einem Handlungsdruck von 20,0 und einer Lücke zum Soll-Zustand von 2,0 ist der Bedarf an Maßnahmen im Vergleich zu Dimension 1 und Dimension 3 moderat und wird daher gelb gekennzeichnet. Auch wenn Dimension 5 eine größere Lücke von 2,66 aufweist, wird durch die niedrigere Priorisierung (5 vs. 10) der Handlungsdruck bei Dimension 5 geringer ausfallen.

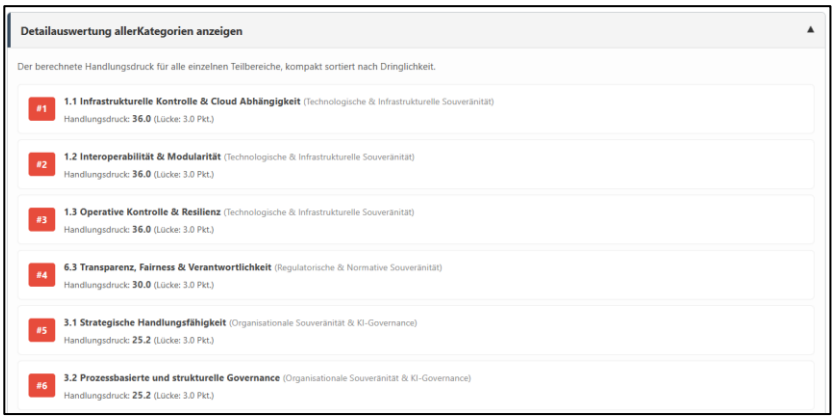

Abbildung 22: Gestaltungsmodell - Priorisierung der Kategorien im Anwendungsbeispiel

Unter Anbetracht der tiefergreifenden Kategorie-Ebene fällt auf, dass Obwohl Kategorie 3.1 einen höheren Schichtmultiplikator aufweist, Kategorie 6.3 (Transparenz & Fairness) mit einem Handlungsdruck von 30 vor den Governance-Kategorien (Druck: 25,2) eingestuft wird. Da die MEDfiction GmbH die normative Compliance (Dimension 6) mit dem Maximalwert 10 gewichtete, zieht der Algorithmus diese Soll-Ist-Diskrepanz priorisiert heran. Ansonsten ist die Priorisierung der ersten Dimension aufgrund des Schichtmultiplikators nachvollziehbar bei einer Lücke von jeweils drei Punkten.

Prio	Kategorie	Art	Handlungsempfehlung
1	1.1 Infrastrukturelle Kontrolle und Cloud Abhängigkeit	Strategie	Identifikation von Sole-Source-Abhängigkeiten; Wahl des am wenigsten einschränkenden IT-Beschaffungsmittels
		Organisation	Implementierung eines Supply Chain Risk Managements (SCRM) für alle bezogenen IT-Komponenten
		Technologie	Etablierung von "Security by Design" und Prüfung bestehender Cloud-Infrastrukturen auf ausländische Abhängigkeiten

2	1.2 Interoperabilität & Modularität	Strategie	Priorisierung von offenen Standards und Interoperabilität bei jeder neuen Software-Beschaffung
		Organisation	Verbot von hartcodierten (proprietären) API-Integrationen, um künftige Migrationskosten zu senken
		Technologie	Strikte Nutzung von offenen, nicht-proprietären APIs als System-Schnittstellen
3	1.3 Operative Kontrolle & Resilienz	Strategie	Identifizierung kritischer KI-Systeme und Festlegung von Toleranzgrenzen für Ausfälle (BSI-Richtlinien)
		Organisation	Definition analoger (menschlicher) Aufsichts- und Notfallprozesse bei KI-Versagen
		Technologie	Etablierung regelmäßiger System-Backups und Sicherung der Datenverfügbarkeit
4	6.3 Transparenz, Fairness & Verantwortlichkeit	Strategie	Anerkennung, dass mangelnde Transparenz das Vertrauen der Nutzer (und damit die Geschäftsgrundlage) zerstört
		Organisation	Einforderung regelmäßiger Transparenzberichte (z.B. zu staatlichen Datenzugriffen) von den Cloud-Providern
		Technologie	Etablierung rudimentärer Logging-Systeme zur grundlegenden Dokumentation von KI-Entscheidungen
5	3.1 Strategische Handlungsfähigkeit	Strategie	Beendigung von "Schatten-IT" (Ad-hoc-Projekten) durch ein klares Management-Commitment und die Formulierung einer expliziten Digital-Souveränitäts-Strategie
		Organisation	Zuweisung dezidierter Budgets und Ressourcen für das Management von Abhängigkeiten
		Technologie	Etablierung eines zentralen IT-Portfoliomanagements zur Inventarisierung aller im Unternehmen genutzten KI-Tools und APIs
6	3.2 Prozessbasierte und strukturelle Governance	Strategie	Etablierung von "Enterprise Governance of IT" als strategischem Partner, nicht als reiner Dienstleister
		Organisation	Definition klarer Rollen, Verantwortlichkeiten und IT-Beschaffungsrichtlinien unter Berücksichtigung von Souveränitätskriterien
		Technologie	Einführung von ITIL- (Information Technology Infrastructure Library) oder COBIT-basierten (Control Objectives for Information and Related Technologies) Basis-Prozessen zur strukturierten Erfassung von KI-Service-Level-Agreements (SLAs)

Abbildung 23: Gestaltungsmodell - Handlungsempfehlungen für das Anwendungsbeispiel

Für jede Kategorie liefert das Gestaltungsmodell dynamisch passgenaue auf Strategie, Organisation und Technologie ausgerichtete Handlungsempfehlungen. Da die Angaben der MEDfiction GmbH in den kritischen Dimensionen überwiegend Werte von 0 oder 1 ausweisen, gibt das Gestaltungsmodell die grundlegenden Low-Level-Handlungsempfehlungen aus. Diese zielen auf die Behebung existenzieller struktureller Defizite ab. Im Folgenden wird für die sechs Kategorien mit dem höchsten Handlungsbedarf analytisch dargelegt, wie die dreiteiligen Handlungsempfehlungen in den Bereichen Strategie, Organisation und Technologie den unzureichenden Ist-Zustand beheben und den Weg zum definierten Soll-Zustand vorbereiten.

Kategorie 1.1: Infrastrukturelle Kontrolle und Cloud-Abhängigkeit (Handlungsdruck 36,0)

Das Unternehmen agiert derzeit mit einem vollständigen Sole-Source-Ansatz bei einem US-Hyperscaler. Da eine sofortige Migration der gesamten Infrastruktur (Soll-Zustand mit hybrider Architektur) für ein KMU operativ nahezu unmöglich ist, setzt das Modell auf die Schaffung von Transparenz. Strategisch fordert es die Identifikation von Single-Source-Abhängigkeiten. Organisatorisch wird dies durch die Einführung eines SCRM ergänzt. Technologisch werden „Security by Design" und die Prüfung auf ausländische Abhängigkeiten vorgeschlagen. Bevor die MEDfiction GmbH in souveräne EU-Clouds investiert, muss sie das genaue Ausmaß der Abhängigkeit (z.B. durch ein SCRM) kartografieren. Dies bildet das zwingende Fundament für zukünftige Architektur-Entscheidungen.

Kategorie 1.2: Interoperabilität und Modularität (Handlungsdruck 36,0)

Aktuell sind proprietäre KI-Schnittstellen fest im Sourcecode des medizinischen Kernsystems veranktert. Daher thematisiert das Gestaltungsmodell die Kernursache des Vendor-Lock-Ins und schlägt vor, auf strategischer Ebene offene Standards in der Neubeschaffung jeglicher Software (-komponenten) zu priorisieren. Auf organisatorischer Ebene wird für die Zukunft ein striktes Verbot von hartcodierten API-Integrationen empfohlen, gepaart mit der technologischen Vorgabe, offene, nicht proprietäre APIs zu nutzen. Das Verbot von Hardcoding stoppt den Aufbau weiterer technischer Abhängigkeiten und ist die Voraussetzung, um den Soll-Zustand mit Modularisierungsmaßnahmen (wie bspw. Docker) überhaupt erst technisch zu ermöglichen.

Kategorie 1.3: Operative Kontrolle und Resilienz (Handlungsdruck 36,0)

Der Ist-Zustand des Unternehmens ist die fehlende Existenz von operativen Fallback-Systemen mit weitreichenden Folgen wie sofortigem Diagnose-Ausfall durch API-Downtimes. Da bislang noch keine Fallback-Systeme existieren, empfiehlt das Gestaltungsmodell vorerst Maßnahmen zur Schadensbegrenzung. Strategisch wird die Festlegung von Teleranzgrenzen für Ausfällen nahegelegt, die durch die Definition organisatorischer analoger (bzw. menschlicher) Notfallprozesse unterstützt wird. Technologisch werden zudem regelmäßige System-Backups zur Sicherung der Datenverfügbarkeit empfohlen. Für ein Medizintechnikunternehmen ist diese Empfehlung essenziell für die Patientensicherheit. Falls die Technik versagt und kein redundantes KI-System etabliert ist, muss der „Human-in-the-Loop" (also der Arzt) den Betrieb mithilfe vordefinierter analoger Prozesse übernehmen. Daher muss dieses Modell ein hohes Verständnis für kritische Systemausfälle zeigen.

Kategorie 6.3: Transparenz, Fairness und Verantwortlichkeit (Handlungsdruck 30,0)

Zurzeit agiert der KI-Provider, als Black-Box und es gibt weder Möglichkeiten zur Erklärbarkeit der KI-Ausgaben (z.B. XAI), noch richtige Beschwerdeprozesse bei algorithmischer Verzerrung. Deshalb sieht das Gestaltungsmodell eine strategische Sensibilisierung vor, die verdeutlicht, dass mangelnde Transparenz die Geschäftsgrundlage zerstört. Organisatorisch soll das Unternehmen Transparenzberichte vom Provider einfordern. Technologisch wird die Etablierung grundlegender Logging-Systeme

zur Dokumentation von KI-Entscheidungen empfohlen. Bevor nämlich komplexe technische Maßnahmen (wie beispielsweise visuelle Dashboards mit Saliency Maps) implementiert werden können, muss die Nachvollziehbarkeit grundlegend protokolliert werden. Ein Logging von Entscheidungen ist die technologische Grundvoraussetzung für jede zukünftige Explainable AI. Des Weiteren kann die Einforderung von Provider-Berichten das Unternehmen rechtlich gegenüber den Krankenhäusern schützen.

Kategorie 3.1: Strategische Handlungsfähigkeit (Handlungsdruck 25,2)

Bislang hat die MEDfiction GmbH ihre aktuellen produktiven Systeme ohne formale Strategie, Budget oder Bewertung alternativer KI-Systeme aufgebaut. Das System empfiehlt einen harten kulturellen Bruch in Form der Beendigung der Schatten-IT durch ein klares strategisches Commitment des Managements. Dies wird organisatorisch durch die Zuweisung dezidierter Budgets und technologisch durch den Aufbau eines zentralen IT-Portfoliomanagements zur Inventarisierung begleitet. Ein Unternehmen kann nämlich keine Make-or-Buy-Entscheidungen treffen, wenn es keinen Überblick über die aktuell genutzten Tools hat. Das IT-Portfoliomanagement schafft diese Transparenz, während die Budgetzuweisung die organisatorische Handlungsfähigkeit erst herstellt.

Kategorie 3.2: Prozessbasierte und strukturelle Governance (Handlungsdruck 25,2)

Aufgrund der aktuellen Ausgangslage des Unternehmens gibt es keine etablierten MLOps-Praktiken und die Modellqualität wird nur manuell oder unstrukturiert überwacht. Um eine automatisierte Qualitätssicherung zu ermöglichen, empfiehlt das Gestaltungsmodell die Etablierung grundlegender Governance-Strukturen. Strategisch soll sich die IT zur „Enterprise Governance of IT“ entwickeln, um ein Business-IT-Alignment sicherzustellen. Organisatorisch sollen klare Rollen und IT-Beschaffungsrichtlinien unter unternehmensspezifischen Souveränitätskriterien definiert werden, während technologisch ITIL- oder COBIT-basierte Basisprozesse zur Erfassung von SLAs etabliert werden sollen. Hier erkennt das Modell, dass Automatisierung (durch beispielsweise MLOps) nicht ohne standardisierte Service-Prozesse funktioniert. Bevor Echtzeit-Überwachungspipelines gebaut werden, müssen die Service-Level-Agreements mit dem Provider überhaupt erst strukturiert erfasst und Verantwortlichkeiten im Entwickler-Team zugewiesen werden.

9.3. Ergebnis und Fazit des Anwendungsbeispiels

Die durchgängige Anwendung des Bewertungs-, Reifegrad- und Gestaltungsmodells auf die fiktive MEDfiction GmbH demonstriert die Funktionsweise und Praxistauglichkeit der in dieser Arbeit entwickelten Artefakte. Das Anwendungsbeispiel belegt, dass sich das oftmals abstrakte Konstrukt der digitalen Souveränität erfolgreich auf die organisationale Mikroebene nachgelagerter KI-Akteure transferieren und operationalisieren lässt.

Der Evaluierungsprozess hat die strukturelle Inkonsistenz offengelegt, mit der mittelständische Integratoren in hochregulierten Branchen durchaus konfrontiert sein können. Mithilfe des Bewertungsmodells konnte gezeigt werden, dass eine hervorragende Daten- und Informationssouveränität (die durch MDR und DSGVO erzwungen wird) nicht zwangsläufig mit technologischer Unabhängigkeit einhergeht. Der errechnete Zielerreichungsgrad von 50,9 % verdeutlicht den erheblichen Handlungsbedarf zwischen dem aktuellen Zustand und der strategischen Zielsetzung.

Das Reifegradmodell lieferte daraufhin eine evaluative Einordnung und verhinderte durch seine integrierte, deterministische Hürdenlogik einen Aggregationseffekt, bei dem ein arithmetischer Mittelwert erhebliche Schwachstellen kaschiert hätte. Obwohl das Datenfundament exzellent bewertet wurde, deckte das Modell die unzureichende Reife der organisationalen „Befähiger" (z. B. Schatten-IT oder fehlende MLOps) auf und limitierte den finalen Reifegrad korrekterweise auf Level 3 („Definiert/Strukturiert"). Diese harte Begrenzung signalisiert den Verantwortlichen im Unternehmen, dass eine Skalierung ohne solides strategisches Fundament nicht möglich ist.

Die abschließende Transformation der Messdaten in eine priorisierte Roadmap validiert die Leistungsfähigkeit des Gestaltungsmodells. Die detaillierte Plausibilitätsprüfung belegt, dass der Algorithmus durch die dynamische Gewichtung und den Schichtmultiplikator präzise Basismaßnahmen (Low Level) aussteuert. Diese sind zwingend erforderlich, um die historischen Strukturdefizite (wie API-Hardcoding, Schatten-IT oder fehlendes Logging) der MEDfiction GmbH zu bereinigen. Die ausgegebenen Maßnahmenpakete zielen nicht nur auf technische Anpassungen ab, sondern betrachten das Problem aus einer interdisziplinären Perspektive, die Strategie, Organisation und Technologie umfasst. Sie bilden das logische und notwendige Fundament, auf dem die anspruchsvolleren Transformationsziele wie Containerisierung, Confidential Computing oder MLOps in nachgelagerten Phasen iterativ aufgebaut werden können.

Aus methodischer Sicht nach Hevner et al. (2004) erbringt dieses Anwendungsbeispiel den Nachweis der Nützlichkeit der konstruierten Artefakte. Praxisprobleme wie der fehlenden Mess- und Steuerbarkeit asymmetrischer Abhängigkeiten in der KI-Wertschöpfungskette wurden für die MEDfiction GmbH gelöst. Das Tool identifizierte nicht nur riskante Vendor-Lock-ins, sondern deckte mithilfe der Konfidenzbewertung auch organisationale Intransparenzen (also Blind Spots) innerhalb der Organisation auf.

Zusammenfassend transformiert der entwickelte Modellierungsprozess die digitale Souveränität von einem passiven Compliance-Risiko in einen steuerbaren strategischen Wettbewerbsvorteil. Die MEDfiction GmbH verfügt nun über ein datengestütztes Diagnose- und Steuerungsinstrument, um ihre Rolle im Ökosystem vom passiven Technologie-Konsumenten zum emanzipierten, aktiven Mitgestalter europäischer Gesundheits-KI zu entwickeln.

10. Evaluation der Ergebnisse

10.1. Vorgehen und Methode

10.2. Qualitative Inhaltsanalyse (Experteninterviews)

10.3. Ergebnisse der Evaluation

11. Ethische, rechtliche und Soziale Überlegungen

Digitale Souveränität kann als mehrdimensionales Konzept verstanden werden, das ethische, rechtliche und soziale Konsequenzen mit sich tragen kann (vgl. Braun und Hummel 2025). Laut Tairov et al. (2024) kann die unreflektierte Übernahme und Integration von KI-Systemen externer Anbieter erhebliche Risiken wie Diskriminierung, Datenschutzverletzungen bis hin zu reputablen und monetären Schäden inhibieren. Ein robustes Gestaltungs- und Reifegradmodell für digitale Souveränität muss daher ausnahmslos fair, transparent, verantwortlich sowie inklusiv sein.

Eines der zentralen ethischen Themen der KI-Nutzung ist die Übernahme potenziell voreingenommener KI-Modelle, die Verzerrung durch Vorurteile, Sprache oder suboptimaler Trainingsdaten ausgeben (vgl. Shahade und Deshmukh 2025). Dies kann zu diskriminierenden Ergebnissen in kritischen Bereichen wie der Personalbeschaffung oder Kreditvergabe führen (vgl. Tairov et al. 2024). Um digitale Souveränität zu ermöglichen, müssen Organisationen nicht nur auf technische Performance, sondern auch ethische KI-Entwicklung und -Integration sicherstellen. Hier ist es wichtig, standardmäßig Fairness, Rechenschaft und Nachverfolgbarkeit zu garantieren, was die Sicherstellung bestmöglicher Datenqualität, Datendiversität und kontinuierlicher technischer Überprüfung und Einsicht erfordert (vgl. Shahade und Deshmukh 2025). Ein weiteres Risiko ist das Entstehen einer „Algokratie“, die durch eine Zentralisierung der Entscheidungsmacht auf die Wenigen, die die Daten und Technologie kontrollieren, entstehen kann (vgl. Celsi und Zomaya 2025). Celsi und Zomaya (2025) fordern daher das Konzept der „Algor-Ethik“, die ein Rahmenwerk ist, das ethische Kontrollmechanismen über den gesamten KI-Lebenszyklus etabliert. Dabei wird die menschliche Würde, Verantwortung und Aufsicht durch den Menschen betont. Digitale Souveränität muss sich daher normativ an der Freiheit und Abhängigkeit des Menschen bzw. der Organisation orientieren und inklusive Designs unterstützen (vgl. Braun und Hummel 2025).

Die Entwicklung und das Training von KI-Modellen erfordert exorbitante Datenmengen, was Datenschutzverletzungen zu einem omnipräsenten Risiko macht. Daher müssen robuste Daten-Governance-Frameworks implementiert werden, die Prinzipien wie „Privacy by Default“ oder Datenminimierung und Nutzerkontrolle gewährleisten (vgl. Shahade und Deshmukh 2025). Der rechtliche Rahmen für digitale Souveränität bezieht sich hauptsächlich nur auf den Schutz von natürlichen Personen bei demokratischen und Grundrechten, wie durch die EU-DSGVO (vgl. Fratini et al. 2024). Global herrscht Stand 2026 eine fragmentierte Regulierungslandschaft und -Konsens, weshalb sich Unternehmen an risikobasierten Ansätzen wie dem EU AI Act orientieren müssen, der risikoreiche KI-Sytemen mit strengen Vorgaben zu Transparenz und menschlicher Aufsicht belegt (vgl. Celsi und Zomaya 2025; Fratini et al. 2024). Um Datenschutz und eine souveräne nachgelagerte KI-Wertschöpfung technologisch zu si-

chern, schlagen Chellapandi et al. (2023) den Ansatz des „Federated Learning“ vor, der ein kollaboratives Modelltraining ohne serverseitige Zentralisierung sensibler Rohdaten ermöglicht. Ergänzend können Open-Source Technologien oder Lösungen einen Vendor Lock-In reduzieren, birgt allerdings die Herausforderung eines hohen personellen und monetären Aufwands, sowie hoher Eigenverantwortung insbesondere in der IT-Sicherheit (Klare und Lechner 2024).

Die Nutzung und Integration von künstlicher Intelligenz wirft Fragen hinsichtlich Machtasymmetrien und geopolitischen Abhängigkeiten auf. Da nur wenige große Unternehmen diese Technologie beherrschen, droht eine Zentralisierung von Reichtum und Marktmacht, was sozioökonomische Ungleichheit begünstigt und verschärfen kann (vgl. Shahade und Deshmukh 2025). Darüber hinaus stellt die Abhängigkeit von ausländischen digitalen Technologie-Importen ein erhebliches operatives, sicherheitstechnisches und geopolitisches Risiko dar (vgl. Klare und Lechner 2024; Fratini et al. 2024). Daher sollte eine marktorientierte Modellierung digitaler Souveränität Strukturen beinhalten, die Monopolbildung verhindern und dabei breite Partizipation und Daten-Solidarität fördern (vgl. Braun und Hummel 2025; Fratini et al. 2024). Im Kontext von Unternehmen verändert KI die Arbeitsumgebung tiefgreifend. Durch Automatisierung monotoner Aufgaben, sowie Verbesserung der Arbeits- und generellen Sicherheit kann KI-Tätigkeiten mit höherer Wertschöpfung fördern (vgl. Tairov et al. 2024). Allerdings überwiegt das Risiko der Arbeitsplatzverdrängung, was Schulungs- und Weiterbildungsprogramme auf Kosten von Unternehmen erforderlich macht, um diskriminierungsfreien Zugang zu dieser Technologie zu gewährleisten (vlg. Shahade und Deshmukh 2025).

Digitale Souveränität ist eng verknüpft mit ethischen, rechtlichen und sozialen Rahmenbedingungen. Um Vertrauen, Rechtskonformität und gesellschaftliche Akzeptanz zu erlangen, müssen Akteure in der nachgelagerten KI-Wertschöpfung internationale Standards adaptieren, sofern diese regional und global etabliert wurden, sowie dezentrale und Privatsphäre schützende Technologien ausbauen und einsetzen.

12. Fazit

12.1. Beantwortung der Forschungsfrage

12.2. Implikationen

12.3. Limitationen

12.4. Forschungsausblick

Literaturverzeichnis

Abiade, S.F., 2025. Algorithmic Sovereignty and the New Security Dependencies: How Foreign AI Surveillance Technologies Reshape Domestic Autonomy in the Global South. World Journal of Advanced Research and Reviews, 27(2), S. 162–180. doi: 10.30574/wjarr.2025.27.2.2845.

Aboushady, H. et al., 2024. Trusted SMEs for Sustainable Growth of Europeans Economical Backbone to Strengthen the Digital Sovereignty: The KDT Resilient Trust Project. In: 2024 27th Euromicro Conference on Digital System Design (DSD). Paris, Frankreich, S. 620–627. doi: 10.1109/DSD64264.2024.00088.

Ahmad, S.F., Han, H., Alam, M.M. et al., 2023. Correction: Impact of artificial intelligence on human loss in decision making, laziness and safety in education. Humanit Soc Sci Commun 10, 361 (2023). doi: 10.1057/s41599-023-01842-4

Ali, M., Khan, T., Khattak, M. und Sener, I., 2024. Synergizing AI and business: Maximizing innovation, creativity, decision precision, and operational efficiency in high-tech enterprises. Journal of Open Innovation: Technology, Market, and Complexity. doi: 10.1016/j.joitmc.2024.100352.

Armbrust, M., Fox, A., Griffith, R., et al., 2010. A view of cloud computing. Commun. ACM 53, 4 (April 2010), 50–58. doi: 10.1145/1721654.1721672.

Attard-Frost, B. und Widder, D., 2025. The ethics of AI value chains. Big Data & Society, 12. doi: 10.1177/20539517251340603.

Augsberg, S. und Gehring, P., 2022. Datensouveränität: Positionen zur Debatte. Campus Verlag GmbH, Frankfurt am Main, Deutschland, S. 125-129. Verfügbar unter: https://library.oapen.org/bitstream/handle/20.500.12657/86493/9783593451947.pdf?sequence=1&isAllowed=y (letzter Zugriff: 12.04.2026)

Babina, T. et al., 2024. Artificial intelligence, firm growth, and product innovation. Journal of Financial Economics. doi: 10.1016/j.jfineco.2023.103745.

Becker, J., Delfmann, P. und Rieke, T., 2007. Referenzmodellierung — Perspektiven für die effiziente Gestaltung von Softwaresystemen. In: Becker, J., Delfmann, P. und Rieke, T. (Hrsg.), Effiziente Softwareentwicklung mit Referenzmodellen. Heidelberg: Physica-Verlag HD. doi: 10.1007/978-3-7908-1994-6_1.

Becker, J., Knackstedt, R. und Pöppelbuß, J., 2009. Entwicklung von Reifegradmodellen für das IT-Management – Vorgehensmodell und praktische Anwendung. Wirtschaftsinformatik. doi: 10.1007/11576-009-0167-9.

Beyerer, J., Müller-Quade, J. und Reussner, R., 2018. Karlsruher Thesen zur Digitalen Souveränität Europas. Datenschutz Datensich 42, 277–280. doi: 10.1007/s11623-018-0940-2.

Billones, R.K.C. et al., 2025. AI Ecosystem and Value Chain: A Multi-Layered Framework for Analyzing Supply, Value Creation, and Delivery Mechanisms. Technologies, 13(9), 421. doi: 10.3390/technologies13090421.

BMV, Bundesministerium der Justiz und Verbraucherschutz, 2026. Gemeinsam Verantwortung übernehmen – Die CDR-Initiative: Corporate Digital Responsibility. Verfügbar unter: https://www.bmjv.de/DE/themen/verbraucherschutz/digitaler_verbraucherschutz/cdr/cdr_node.html (letzter Aufruf: 14.04.2026)

Bogenstahl, C. und Zinke, G., 2017. Digitale Souveränität – ein mehrdimensionales Handlungskonzept für die deutsche Wirtschaft, S. 65-80. doi: 10.1007/978-3-662-55788-4_2

Boeira, D.F.F., Scheid, E.J., Franco, M.F., Zembruzki, L. und Granville, L.Z., 2024. Traffic Centralization and Digital Sovereignty: An Analysis Under the Lens of DNS Servers. In: NOMS 2024-2024 IEEE Network Operations and Management Symposium. Seoul, Republik Korea, S. 1–9. doi: 10.1109/NOMS59830.2024.10575700.

Braun, M. und Hummel, P., 2025. Is digital sovereignty normatively desirable? Information, Communication & Society, 28(10), S. 1721–1734. doi: 10.1080/1369118X.2024.2332624.

Bues, P., 2025. Unlocking the Future of Data Security: Confidential Computing as a Strategic Imperative. White Paper verfügbar unter: https://confidentialcomputing.io/wp-content/uploads/sites/10/2025/11/US53866125.pdf (letzter Aufruf: 29.05.2026)

Calderaro, A. und Blumfelde, S., 2022. Artificial intelligence and EU security: the false promise of digital sovereignty. European Security, 31(3), S. 415–434. doi: 10.1080/09662839.2022.2101885.

Celsi, L. und Zomaya, A., 2025. Perspectives on Managing AI Ethics in the Digital Age. Information, 16(4), 318. doi: 10.3390/info16040318.

Chellapandi, V.P. et al., 2024. Federated Learning for Connected and Automated Vehicles: A Survey of Existing Approaches and Challenges. IEEE Transactions on Intelligent Vehicles, 9(1), S. 119–137. doi: 10.1109/TIV.2023.3332675.

Daliparthi, V.S.S.A., Tutschku, K., Kebande, V.R. und Momen, N., 2025. Digital Sovereignty for Collaborative AI Engineering: A Survey. IEEE Access, 13, S. 216438–216465. doi: 10.1109/ACCESS.2025.3647085.

de Bruin, T., Freeze, R., Kulkarni, U. und Rosemann, M., 2005. Understanding the Main Phases of Developing a Maturity Assessment Model. In: ACIS 2005 Proceedings. Verfügbar unter: https://aisel.aisnet.org/acis2005/109.

Enholm, I. et al., 2021. Artificial Intelligence and Business Value: a Literature Review. Information Systems Frontiers, 24, S. 1709–1734. doi: 10.1007/s10796-021-10186-w.

Europäische Komission, Public Health, 2026. Verordnung über den Europäischen Gesundheitsdatenraum (EHDS). Luxembourg. Verfügbar unter: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_de (letzter Zugriff: 08.06.2026)

Europäische Kommission, Directorate-General for digital Services, 2025. Cloud Sovereignty Framework Version 1.2.1 – Oct 2025. Luxembourg. Verfügbar unter: https://commission.europa.eu/document/download/09579818-64a6-4dd5-9577-446ab6219113_en (letzter Zugriff: 11.04.2026)

Farmakis, T. et al., 2024. Innovations in Manufacturing Business Models Enabled by Digital and AI-driven Transformation: A Conceptual Framework. In: 2024 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC). Funchal, Portugal, S. 1–8. doi: 10.1109/ICE/ITMC61926.2024.10794358.

Fettke, P. und Loos, P., 2003. Classification of reference models: a methodology and its application. Information Systems and e-Business Management, 1, S. 35–53. doi: 10.1007/BF02683509.

Firdausy, D.R., De Alencar Silva, P., Van Sinderen, M. und Iacob, M.-E., 2022. Towards a Reference Enterprise Architecture to enforce Digital Sovereignty in International Data Spaces. In: 2022 IEEE 24th Conference on Business Informatics (CBI). Amsterdam, Niederlande, S. 117–125. doi: 10.1109/CBI54897.2022.00020.

Floridi, L., 2020. The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU. Philosophy & Technology, 33, S. 369–378. doi: 10.1007/s13347-020-00423-6.

Foster, C., 2025. Openness in AI and downstream governance: A global value chain approach. arXiv preprint. doi: 10.48550/arXiv.2509.10220.

Fratini, S. et al., 2024. Digital Sovereignty: A Descriptive Analysis and a Critical Evaluation of Existing Models. Digital Society (DISO), 3, 59. doi: 10.1007/s44206-024-00146-7.

Fresenius, 2026. KI im Gesundheitswesen: SAP und Fresenius beschleunigen die digitale Gesundheitsversorgung. Verfügbar unter: https://www.fresenius.com/de/node/7288 (letzter Zugriff: 07.06.2026)

Fries, I., Greiner, M., Hofmeier, M., et al., 2023. Towards a Layer Model for Digital Sovereignty: A Holistic Approach. In: Hämmerli, B., Helmbrecht, U., Hommel, W., Kunczik, L., Pickl, S. (eds) Critical Information Infrastructures Security. CRITIS 2022. Lecture Notes in Computer Science, vol 13723. Springer, Cham. doi: 10.1007/978-3-031-35190-7_9.

Gawer, A., 2022. Digital platforms and ecosystems: remarks on the dominant organizational forms of the digital age. Innovation, 24(1), S. 110–124. doi: 10.1080/14479338.2021.1965888.

Gemini API, 2026. Dokumentation: Strukturierte Ausgaben. Verfügbar unter: https://ai.google.dev/gemini-api/docs/structured-output?hl=de&example=recipe (letzter Zugriff: 10.04.2026)

Glasze, G. et al., 2023. Contested Spatialities of Digital Sovereignty. Geopolitics, 28(2), S. 919–958. doi: 10.1080/14650045.2022.2050070.

Gökalp, E. und Martinez, V., 2022. 'Digital transformation maturity assessment: development of the digital transformation capability maturity model', International Journal of Production Research, 60(20), pp. 6282–6302. doi: 10.1080/00207543.2021.1991020.

Grete, P., Kestermann, C., 2020. Cloud Computing im Spannungsfeld von Souveränität, Cybersicherheit und Ökonomie. Datenschutz Datensich 46, 632–636. doi: 10.1007/s11623-022-1672-x

Gujar, P., Panyam, S. und Paliwal, G., 2025. AI Integrated Product Development: Building Sustainable Competitive Advantage. IEEE Engineering Management Review, 53(1), S. 73–81. doi: 10.1109/EMR.2024.3475408.

Gurjar, K. et al., 2024. An Analytical Review on the Impact of Artificial Intelligence on the Business Industry: Applications, Trends, and Challenges. IEEE Engineering Management Review, 52, S. 84–102. doi: 10.1109/emr.2024.3355973.

Heeks, R. und Spiesberger, P., 2024. Constructing an AI Value Chain and Ecosystem Model. Manchester Centre for Digital Development Working Paper, 109. doi: 10.13140/RG.2.2.11981.86246.

Hevner, A.R., March, S.T., Park, J. und Ram, S., 2004. Design Science in Information Systems Research. Management Information Systems Quarterly, 28, S. 75–105.

Hopkins, A. et al., 2025. AI Supply Chains: An Emerging Ecosystem of AI Actors, Products, and Services. arXiv preprint. doi: 10.48550/arXiv.2504.20185.

Hulko, G. et al., 2025. The politics of digital sovereignty and the European Union's legislation: navigating crises. Frontiers in Political Science, 7, 1548562. doi: 10.3389/fpos.2025.1548562.

Hummel, P., Braun, M., Tretter, M. und Dabrock, P., 2021. Data sovereignty: A review. Big Data & Society, 8(1). doi: 10.1177/2053951720979777.

Hynninen, T. und Maunula, S., 2025. Towards Enabling Privacy-First AI Solutions for SMEs Using Open-Source LLMs. In: 2025 MIPRO 48th ICT and Electronics Convention. Opatija, Kroatien, S. 1915–1919. doi: 10.1109/MIPRO65660.2025.11131816.

Jacobides, M.G. et al., 2021. The Evolutionary Dynamics of the Artificial Intelligence Ecosystem. Strategy Science, 6(4), S. 412–435. doi: 10.1287/stsc.2021.0148.

Janardhanan, S., Samonaki, M., Heegaard, P.E., Kellerer, W. und Mas-Machuca, C., 2025. Network Sovereignty: A Novel Metric and Its Application on Network Design. IEEE Transactions on Reliability, 74(2), S. 2927–2941. doi: 10.1109/TR.2024.3434560.

Joecks-Laß, D., 2023. Begriffliche Grundlagen der digitalen Selbstbestimmung. In: Digitale Selbstbestimmung. Springer Gabler, Wiesbaden. doi: 10.1007/978-3-658-43132-7_2.

Kaffee, L., Jernite, Y., 2025. Open Source AI: A Cornerstone of Digital Sovereignty. Hugging Face Community Article. Verfügbar unter: https://huggingface.co/blog/frimelle/sovereignty-and-open-source (letzer Zugriff: 13.04.2025).

Kapoor, K. et al., 2021. A socio-technical view of platform ecosystems: Systematic review and research agenda. Journal of Business Research, 128, S. 94–108. doi: 10.1016/j.jbusres.2021.01.060.

Klare, M. und Lechner, U., 2024. A Reference Model to Strengthen Digital Sovereignty in Companies. In: ICSOB Companion. Verfügbar unter: https://api.semanticscholar.org/CorpusID:276673683.

Klare, M., Lechner, U. und Fritzsche, A., 2025. Digital sovereignty through sovereign clouds? In: International Conference on AI and the Digital Economy (CADE 2025). Venedig, Italien, S. 14–19. doi: 10.1049/icp.2025.2947.

Kopka, A. und Fornahl, D., 2023. Artificial intelligence and firm growth — catch-up processes of SMEs through integrating AI into their knowledge bases. Small Business Economics, 62, S. 63–85. doi: 10.1007/s11187-023-00754-6.

Kretschmer, M. und Orth, R., 2025. Integrating GenAI into Corporate Knowledge Management: A Socio-Technical Approach to Overcoming Knowledge Management Challenges. In: European Conference on Knowledge Management. doi: 10.34190/eckm.26.1.3768.

Li, L., 2024. Research on the Impact of Artificial Intelligence and Internationalization on Sustainable Development of Enterprises. In: 2024 7th International Conference on Artificial Intelligence and Big Data (ICAIBD). Chengdu, China, S. 538–542. doi: 10.1109/ICAIBD62003.2024.10604496.

Maathuis, C., und Cools, K., 2025. Digital Sovereignty Control Framework for Military AI-based Cyber Security. doi: 10.48550/arXiv.2509.13072.

Mayer, M. und Lu, Y.-C., 2023. Digital Autonomy? Measuring the Global Digital Dependence Structure. SSRN. doi: 10.2139/ssrn.4404826.

Mayring, P. und Fenzl, T., 2019. Qualitative Inhaltsanalyse. In: Baur, N., Blasius, J. (eds) Handbuch Methoden der empirischen Sozialforschung. Springer VS, Wiesbaden. doi: 10.1007/978-3-658-21308-4_42

Melcher, P. und Vaziri D., 2025. AI Competencies for Companies: Definition, Taxonomy and Regulatory Classification. World Journal of Artificial Intelligence and Robotics Research. doi: 10.63620/mkwjairr.2025.1012.

Mitra, A. et al., 2024. Adoption of Artificial Intelligence Enabling Value Creation in Business. In: 2024 4th International Conference on Robotics, Automation and Artificial Intelligence (RAAI). Singapur, S. 213–220. doi: 10.1109/RAAI64504.2024.10949520.

Moerel, E.M.L. und Timmers, P., 2021. Reflections on Digital Sovereignty. EU Cyber Direct, Research in Focus series. Verfügbar unter: https://ssrn.com/abstract=3772777.

Papagiannidis, E., Enholm, I.M., Dremel, C. et al., 2023. Toward AI Governance: Identifying Best Practices and Potential Barriers and Outcomes. Inf Syst Front 25, 123–141. doi: 10.1007/s10796-022-10251-y.

Paschen, U., Pitt, C. und Kietzmann, J., 2020. Artificial intelligence: Building blocks and an innovation typology. Business Horizons, 63, S. 147–155. doi: 10.1016/j.bushor.2019.10.004.

Pi, Y., 2024. Missing Value Chain in Generative AI Governance China as an example. arXiv preprint. doi: 10.48550/arXiv.2401.02799.

Pirrone, C. et al., 2025. Exploring Vulnerability in AI Industry. arXiv preprint. doi: 10.48550/arXiv.2510.23421.

Pohle, J., Nanni, R. und Santaniello, M., 2025. Unthinking Digital Sovereignty: A Critical Reflection on Origins, Objectives, and Practices. Policy & Internet, 16(4), S. 666–671. doi: 10.1002/poi3.437.

Pöppelbuß, J. und Röglinger, M., 2011. What Makes a Useful Maturity Model? A Framework of General Design Principles for Maturity Models and its Demonstration in Business Process Management. In: ECIS 2011 Proceedings. Verfügbar unter: https://aisel.aisnet.org/ecis2011/28.

Reddy, G., Pavan Kumar, Y. & Prakash, K.P., 2024. Hallucinations in Large Language Models (LLMs). In: 2024 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania, 2024. pp.1-6. doi: 10.1109/eStream61684.2024.10542617.

Roberts, H., 2024. Digital sovereignty and artificial intelligence: a normative approach. Ethics and Information Technology, 26(70). doi: 10.1007/s10676-024-09810-5.

Saunders, B., Sim, J., Kingstone, T. et al., 2018. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 52, 1893–1907. doi: 10.1007/s11135-017-0574-8

Shahade, A.K. und Deshmukh, P.V., 2025. A Critical and Scientific Overview of Artificial Intelligence in the Industrial Revolution. In: 2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT). Bengaluru, Indien, S. 783–788. doi: 10.1109/IDCIOT64235.2025.10914988.

Sheikh, H., 2022. European Digital Sovereignty: A Layered Approach. DISO 1, 25. doi: 10.1007/s44206-022-00025-z.

Sigi, A., 2024. Designing Data Governance With DAMA DMBOK Framework. Jurnal Teknobisnis, 8(2), S. 79–89. doi: 10.12962/j24609463.v8i2.1408.

Srivastava, S., und Bullock, J., 2024. AI, Global Governance, and Digital Sovereignty. ArXiv. doi: 10.48550/arXiv.2410.17481

Steinbach, A., 2024. Digitale Souveränität. List Forum 50, S. 51–74. doi: 10.1007/s41025-023-00252-3

Steinmann, N., Piazza, A., 2024. KI-basierte Textkreation im Content Marketing: Design und Evaluation eines effektiven Prompts. HMD 61, 402–417.doi: 10.1365/s40702-024-01058-3.

Stürmer, M., 2024. Technologische Perspektive der digitalen Souveränität: Blick auf die Schweiz, internationale Trends sowie Empfehlungen für die „Strategie Digitale Souveränität der Schweiz“. Bericht zuhanden des Eidgenössischen Departements für auswärtige Angelegenheiten (EDA). Berner Fachhochschule, Wirtschaft, Public Sector Transformation. doi: 10.24451/arbor.22222.

Sorrentino, E., Spagnuolo, A.F., Portaro, A., Taverniti, M. und Cardillo, E., 2025. Towards Semantic Coherence: Understanding Cybersecurity and Digital Sovereignty in the Automotive Landscape. In: 2025 IEEE 31st International Symposium on On-Line Testing and Robust System Design (IOLTS). Ischia, Italien, S. 1–5. doi: 10.1109/IOLTS65288.2025.11116851.

Tairov, I. et al., 2024. Review of AI-Driven Solutions in Business Value and Operational Efficiency. Economics Ecology Socium. doi: 10.61954/2616-7107/2024.8.3-5.

Tan, K.L., Chi, C.-H. und Lam, K.-Y., 2023. Survey on Digital Sovereignty and Identity: From Digitization to Digitalization. ACM Computing Surveys, 56(3), Artikel 61. doi: 10.1145/3616400.

Tao, Y., Yang, S. und Ge, H., 2022. Comparative Study on Data Sovereignty Guarantee Technology. In: 2022 IEEE 13th International Symposium on Parallel Architectures, Algorithms and Programming (PAAP). Peking, China, S. 1–6. doi: 10.1109/PAAP56126.2022.10010593.

Thiel, T., 2024. Forschungsdiskurse zu digitaler Souveränität. Politische Vierteljahresschrift, 65, S. 835–839. doi: 10.1007/s11615-024-00566-7.

Varriale, V., Cammarano, A., Michelino, F. und Caputo, M., 2023. Critical analysis of the impact of artificial intelligence integration with cutting-edge technologies for production systems. Journal of Intelligent Manufacturing, 36, S. 61–93. doi: 10.1007/s10845-023-02244-8.

vom Brocke, J., 2007. Design Principles for Reference Modelling. Reusing Information Models by Means of Aggregation, Specialisation, Instantiation, and Analogy. In: Fettke, P. und Loos, P. (Hrsg.), Reference Modelling for Business Systems Analysis. Hershey, PA, USA: Idea Group Publishing, S. 47–75. doi: 10.4018/9781599040547.ch003.

Walton, R.O., Watkins, D.V., 2024. The use of generative AI in research: a production management case study from the aviation industry. J Market Anal. doi: 10.1057/s41270-024-00317-y

Wamba-Taguimdje, S. et al., 2020. Influence of artificial intelligence (AI) on firm performance: the business value of AI-based transformation projects. Business Process Management Journal, 26, S. 1893–1924. doi: 10.1108/bpmj-10-2019-0411.

Wand, Y. und Weber, R., 2002. Research commentary: Information systems and conceptual modeling - a research agenda. Information Systems Research, 13, S. 363–376. doi: 10.1287/isre.13.4.363.69.

Webster, J. und Warson, R.,2002. Analyzing the past to prepare for the future: Writing a literature review. In: MIS Quarterly Vol. 26 No. 2, S. 13-23. doi: 10.2307/4132319

Wenger, F., 2025. Digitale Souveränität und Autonomie – eine Spurensuche und Auslegeordnung. Verfügbar unter: https://www.bedag.ch/de/aktuelles/meldungen/Digitale-Souveraenitaet-und-Autonomie-eine-Spurensuche-und-Auslegeordnung.php?highlight=souver%C3%A4nit%C3%A4t (letzter Zugriff: 07.06.2026)

White, C., Dooley, S., Roberts, M., Pal, A., Feuer, B., Jain, S., Shwartz-Ziv, R., Jain, N., Saifullah, K., Naidu, S., Hegde, C., LeCun, Y., Goldstein, T., Neiswanger, W. & Goldblum, M., 2025. 'LiveBench: A Challenging, Contamination-Free LLM Benchmark'. doi: 10.48550/arXiv.2406.19314 [Ergebnisse sind auf der LiveBench-Website einzusehen: https://livebench.ai/# (letzter Zugriff: 25.04 2026)].

Wenzelburger, G. und König, P.D., 2025. Sending Signals or Building Bridges? Digital Sovereignty in EU Communicative and Co-Ordinative Discourse. JCMS: Journal of Common Market Studies, 63(2), S. 526–547. doi: 10.1111/jcms.13638.

Zebec, A. und Stemberger, M., 2024. Creating AI business value through BPM capabilities. Business Process Management Journal, 30, S. 1–26. doi: 10.1108/bpmj-07-2023-0566.

Zhang, Y. und Fu, Y., 2021. Modeling the Economic and Economic Impact of Artificial Intelligence Data Feature Analysis. In: 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA). Coimbatore, Indien, S. 1412–1415. doi: 10.1109/ICECA52323.2021.9676067.

Zuo, X., 2023. Research on the Application of Artificial Intelligence Technology in Economic Management in the Information Age. In: 2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA). Chongqing, China, S. 1737–1741. doi: 10.1109/ICIBA56860.2023.10165033.

Anhang

A. Modellierung mit der Excel-Anwendung

Grundlegend ist die Struktur der Excel-Anwendung zur Modellierung digitaler Souveränität für KI-integrierende KMU ähnlich zu der Struktur der Webanwendung.

- Navigiert wird dabei über die unteren Tabellenblätter in Excel.
- Gestartet wird mit der Einleitung, die den Ablauf, die Funktion und eine Übersicht über die gesamte Modellierung gibt.
- Der Fragebogen besteht aus 6 Tabellenblättern, die den Namen der jeweiligen Dimension tragen. Diese beinhalten jeweils die 3 Kategorien mit ihren Prüffragen und Konfizenz-Abfragen.
- Daraufhin können die Dimensionen gewichtet und Soll-Zustände der Kategorien definiert werden.
- BITTE NUR DIE ROTEN ZELLEN AUSFÜLLEN!
- Das Bewertungsmodell ist zweigeteilt: Ist-Zustand (mit Konfidenzanalyse) und Soll-Ist-Vergleich. Die Eingaben aus dem Fragebogen werden automatisch übernommen, hier also bitte nichts ändern.
- Das Reifegradmodell zeigt den Reifegrad dynamisch in Farbe an, jedoch ist die Hürdenlogik aus der Webanwendung nicht implementiert (hier bitte auch nichts ändern).
- Das Gestaltungsmodell zeigt die Handlungsempfehlungen ebenfalls dynamisch an, ordnet allerdings die Dimensionen nicht nach Handlungsdruck (hier bitte ebenfalls nichts ändern).
- Alle Handlungsempfehlungen können in dem gleichnamigen Tabellenblatt eingesehen werden.

Die Modellierung der Excel-Anwendung ist im Anhang unter der Excel-Datei:

Souveränitätsmodellierung.xlsx zu finden.

Anmerkung: Je nach Dokumententransfer, Excelversion oder Glück können die Referenzen und Formeln der Excel Version durcheinanderkommen. Daher ist eine Kontrolle der Ausgaben des Gestaltungsmodells besonders relevant.

B. Übersicht der Handlungsempfehlungen

Die Handlungsempfehlungen sind in der JSON-Datei han-empf.json der Webanwendung hinterlegt und werden dynamisch ausgewählt in dem Gestaltungsmodell oder dem Baumdiagramm angezeigt.

Dimension	Kategorie	Reifegrad	Ebene	Handlungsempfehlung	Quelle
1: Technologische & Infrastrukturelle Souveränität	1.1 Infrastrukturelle Kontrolle & Cloud-Abhängigkeiten	Low	Strategisch	Identifikation von Sole-Source-Abhängigkeiten; Wahl des am wenigsten einschränkenden IT-Beschaffungsmittels	Davis & Cobb 2010
			Organisatorisch	Implementierung eines Supply Chain Risk Managements (SCRM) für alle bezogenen IT-Komponenten	Maathuis & Cools 2025
			Technisch	Etablierung von "Security by Design" und Prüfung bestehender Cloud-Infrastrukturen auf ausländische oder kritische Abhängigkeiten	European Commission 2025
		Mid	Strategisch	Aufbau von diversifizierten Multi-Cloud- oder Hybrid-Strategien zur Risikostreuung	Grete & Kestermann 2022; Celtekligil 2020
			Organisatorisch	Nutzung standardisierter Zertifizierungen (z.B. C5) zur objektiven Risikobewertung von Cloud-Providern	Grete & Kestermann 2022
			Technisch	Parallelbetrieb von Legacy-Systemen und neuen Cloud-Lösungen (Ambidexterity) zur Sicherung der Betriebskontinuität	Plekhanov et al. 2023
		High	Strategisch	Nutzung offener Cloud-Ökosysteme (z.B. Sovereign Cloud Stack) für maximale Portabilität	Stürmer 2024; Pohle & Thiel 2020
			Organisatorisch	Investitionen in den Aufbau lokaler oder national geprüfter Infrastrukturkapazitäten	Roberts et al. 2021
			Technisch	Lokales Hosting von Open-Source LLMs (z.B. Llama) auf eigener Hardware oder Nutzung von souveränen Bereitstellern	Hynninen & Maunula 2025; Maathuis & Cools 2025
	1.2 Interoperabilität und Modularität	Low	Strategisch	Priorisierung von offenen Standards und Interoperabilität bei jeder neuen Software-Beschaffung	Paschen et al. 2020
			Organisatorisch	Verbot von hartcodierten (proprietären) API-Integrationen, um künftige Migrationskosten zu senken	Klare et al. 2025
			Technisch	Strikte Nutzung von offenen, nicht-proprietären APIs als System-Schnittstellen	European Commission 2025; Song 2019
		Mid	Strategisch	Tiefe Einbettung von KI in Kernprozesse anstatt nur oberflächlicher Chatbot-Integrationen	Gujar et al. 2025
			Organisatorisch	Förderung von Open-Source-Technologien in Entwicklungs- und Review-Prozessen	Klare et al. 2025
			Technisch	Modularisierung der Architektur durch Container-Technologien (z.B. Docker) für portable Deployments	Papagiannidis et al. 2023; Hynninen & Maunula 2025
		High	Strategisch	Aktive Gestaltung von Interdependenzen im Wertschöpfungsnetzwerk statt rein linearer Lieferketten	Attard-Frost & Widder 2024
			Organisatorisch	Mitarbeit an Open-Source-Rahmenwerken und Etablierung europäischer Schnittstellen-Regeln	Weiss 2022
			Technisch	Einsatz dynamischer Modell-Partitionierung (Splitting von Modellen zur Lastverteilung auf eigene und externe Knoten)	Wu et al. 2024
	1.3 Operative Kontrolle und Resilienz	Low	Strategisch	Identifizierung kritischer KI-Systeme und Festlegung von Toleranzgrenzen für Ausfälle (BSI-Richtlinien)	Klare et al. 2025
			Organisatorisch	Definition analoger (menschlicher) Aufsichts- und Notfallprozesse bei KI-Versagen	Hanisch et al. 2023
			Technisch	Etablierung regelmäßiger System-Backups und Sicherung der Datenverfügbarkeit	Klare et al. 2025
		Mid	Strategisch	Aufbau interner "Control Towers" für die orchestrierte Steuerung von KI-Operationen	Wamba-Taguimdje et al. 2020
			Organisatorisch	Implementierung standardisierter Risiko-Governance gemäß ISO/IEC 42001	Celsi & Zomaya 2025
			Technisch	Aufbau von lokalen Fallback-Systemen zur Überbrückung bei Cloud-Ausfällen	Papagiannidis et al. 2023; Floridi 2020
		High	Strategisch	Übergang zu Zero-Trust-Prinzipien für alle genutzten KI-Modelle und Schnittstellen	Maathuis & Cools 2025
			Organisatorisch	Hybride Governance: Nahtloser Wechsel zwischen automatisierten, augmentierten und analogen Steuerungsmodi	Hanisch et al. 2023

			Technisch	Implementierung standardisierter, maschinenlesbarer Usage-Control-Policies (z.B. IDS Lucon) für systemübergreifende Resilienz	Tao et al. 2022
2: Daten- & Informationssouveränität	2.1 Eigentum und Kontrolle der Daten und Informationen	Low	Strategisch	Identifikation von kritischen Daten-Abhängigkeiten hinsichtlich Abhängigkeit essenzieller KI-Funktionen von externen Daten-Monopolisten; Anerkennung unternehmenseigener Datenbestände als strategischen Engpass und kritische Wissensressource	Celtekligil 2020
			Organisatorisch	Nachverhandlung bestehender Data-Sourcing-Verträge zur Sicherung langfristiger Zugriffsrechte (Vermeidung plötzlicher Vertragskündigungen); Schaffung klarer vertraglicher Rahmenbedingungen, um implizite Datenmonetarisierung durch Cloud-Provider zu verhindern	Foster 2025
			Technisch	Etablierung standardisierter Ingestion-Schnittstellen, um externe Datenlieferanten im Notfall schnell austauschen zu können; Definition strikter Zugriffsrechte und Limitierung der Datenübergabe an externe APIs	Hanisch et al. 2023
		Mid	Strategisch	Diversifizierung der Datenbeschaffung durch den strategischen Anschluss an branchenspezifische Datenräume (z. B. Mobility Data Space); Anwendung des DAMA-Frameworks zur Strukturierung und Kartierung aller Datenflüsse	Celsi & Zomaya 2025
			Organisatorisch	Etablierung von Data-Sharing-Agreements mit strategischen Partnern, um Abhängigkeiten von reinen Datenbrokern zu reduzieren; Etablierung von Datenmanagement-Rollen und Berechtigungsstrukturen (z.B. Einsatz von Datentreuhändern)	Papagiannidis et al. 2023; Weiss 2022
			Technisch	Implementierung von Systemen zur Generierung synthetischer Daten, um Lücken in den externen Trainingsdaten eigenständig schließen zu können; Implementierung von Mechanismen zur irreversiblen Datenlöschung sowie Sicherstellung der exklusiven kryptografischen Schlüsselkontrolle	European Commission 2025
		High	Strategisch	Aufbau eines "Data Moats" (Strategischer Wettbewerbsvorteil durch proprietäre Daten) mit Nutzung exklusiver, nur dem Unternehmen vorliegender Daten als massiven Wettbewerbsvorteil gegenüber Konkurrenten; Daten nicht nur schützen, sondern als Teil kollektiver Strukturen für solidarische Datenteilung begreifen (CARE-Prinzipien: Collective Benefit, Authority to Control, Responsibility, and Ethics)	Braun & Hummel 2025; Stürmer 2024
			Organisatorisch	Orchestrierung von "Data Trusts" (Datentreuhändern), bei denen das eigene Unternehmen die Regeln des Daten-Zuflusses definiert; Förderung von Data Spaces und gemeinschaftsbasierten Eigentumsmodellen zur kollektiven Selbstbestimmung	Weiss 2022; Costa Barbosa et al. 2024
			Technisch	Aufbau vollautomatisierter, proprietärer Daten-Feedback-Loops, bei denen Nutzer-Interaktionen direkt und exklusiv in die Verbesserung der eigenen Modelle fließen; Einsatz von "Compute-to-Data"-Ansätzen und Self-Sovereign Identity (SSI) zur vollautonomen Identitätskontrolle	Weiss 2022
	2.2 Datensicherheit und Vertraulichkeit	Low	Strategisch	Verankerung eines "Privacy-First"-Ansatzes in allen KI-Projekten	Hynninen & Maunula 2025
			Organisatorisch	Vorab-Klassifizierung aller Datenbestände vor der Einbindung in externe KI-Umgebungen	Tao et al. 2022; Maathuis & Cools 2025
			Technisch	Systematischer Einsatz datenschutzfreundlicher Anonymisierungs- und Pseudonymisierungstechnologien	Gujar et al. 2025
		Mid	Strategisch	Fokusverschiebung auf den Schutz von "Data in Use" (Daten während der Verarbeitung)	Bues 2025
			Organisatorisch	Etablierung strenger Kontrollmechanismen für die DSGVO/GDPR-Compliance, inkl. Prüfung auf CLOUD-Act-Risiken	Kaloudis et al. 2023; Grete & Kestermann 2022
			Technisch	Nutzung hardware-basierter Trusted Execution Environments (TEEs) / Confidential Computing zur Isolierung kritischer Workloads	Bues 2025; Tao et al. 2022
		High	Strategisch	Striktes Management der Rechtsprechung (Jurisdiction) über die Datenhaltung	Pohle & Thiel 2020; Maathuis & Cools 2025
			Organisatorisch	Aufbau vertrauenswürdiger Daten-Ökosysteme auf Basis von Blockchain-Coupling zur manipulationssicheren Nachweisführung	Klare et al. 2023
			Technisch	Dezentrales Modelltraining mittels Federated Learning (FL), Differential Privacy und Secure Multiparty Computation (SMPC)	McMahan et al. 2017; Klare & Lechner 2024
	2.3 Verfügbarkeit und Transparenz	Low	Strategisch	Strukturierte Datenbereinigung und Nutzung von synthetischen Daten zur Erhöhung der Datenverfügbarkeit	Kretschmer & Orth 2024
			Organisatorisch	Dokumentation von Datenproduktionsprozessen und vertragliche Einforderung von Transparenzberichten der Cloud-Provider	Attard-Frost & Widder 2024; Grete & Kestermann 2022
			Technisch	Schaffung von Transparenz über lokale Datenflüsse mittels simpler Inventarisierung	Grete & Kestermann 2022
		Mid	Strategisch	Implementierung einer "Data-driven Governance" zur Einbindung von Daten in Kern-Entscheidungsprozesse	Shirwa et al. 2025; Rahaman & Mannan 2025
			Organisatorisch	Etablierung lückenloser Audit-Trails und Transparenzregeln für die abteilungsübergreifende Datennutzung	Celsi & Zomaya 2025

			Technisch	Nutzung von Leistungs-Dashboards, die IT- und Business-KPIs integrieren, zur Überwachung der Datenströme	Rahaman & Mannan 2025
		High	Strategisch	Positionierung als Knotenpunkt im Informationsnetzwerk durch Entwicklung eigener Kanäle für Datenaustausch (Open Data / OGD)	Stürmer 2024; Plekhanov et al. 2023
			Organisatorisch	Bereitstellung eigener proprietärer Datensätze zur Förderung lokaler oder partnerschaftlicher Innovationen	Roberts et al. 2021
			Technisch	Einsatz von Blockchain-basierten Systemen zur Gewährleistung unveränderbarer Daten-Provenienz und Auditierung	Hanisch et al. 2023
3: Organisationale Souveränität & KI-Governance	3.1 Strategische Handlungsfähigkeit	Low	Strategisch	Beendigung von "Schatten-IT" (Ad-hoc-Projekten) durch ein klares Management-Commitment und die Formulierung einer expliziten Digital-Souveränitäts-Strategie	Klare & Lechner 2024; Gujar et al. 2025
			Organisatorisch	Zuweisung dezidierter Budgets und Ressourcen für das Management von Abhängigkeiten	Celtekligil 2020
			Technisch	Etablierung eines zentralen IT-Portfoliomanagements zur Inventarisierung aller im Unternehmen genutzten KI-Tools und APIs	Rahaman & Mannan 2025
		Mid	Strategisch	Verknüpfung der KI-Strategie mit übergeordneten Geschäfts- und Nachhaltigkeitszielen	Kretschmer & Orth 2024; Shirwa et al. 2025
			Organisatorisch	Etablierung einer "Controlled Agility": Balance zwischen agiler Dezentralisierung und zentraler, leitplankengestützter Kontrolle	Rahaman & Mannan 2025; Plekhanov et al. 2023
			Technisch	Nutzung von Leistungs-Dashboards zur transparenten Überwachung von IT- und Business-KPIs im Rahmen der KI-Integration	Rahaman & Mannan 2025
		High	Strategisch	Übergang zu "Cybernetischer Steuerung": Den technologischen Wandel aktiv lenken, statt nur darauf zu reagieren	Floridi 2020
			Organisatorisch	Gründung eines Open Source Programme Office (OSPO) zur zentralen strategischen Koordination quelloffener Initiativen	Stürmer 2024
			Technisch	Integration von föderierten, dezentralen Frameworks, um lineare Lieferketten-Abhängigkeiten aufzubrechen	Shirwa et al. 2025
	3.2 Prozessbasierte und strukturelle Governance	Low	Strategisch	Etablierung von "Enterprise Governance of IT" als strategischem Partner, nicht als reiner Dienstleister	Rahaman & Mannan 2025
			Organisatorisch	Definition klarer Rollen, Verantwortlichkeiten und IT-Beschaffungsrichtlinien unter Berücksichtigung von Souveränitätskriterien	Klare et al. 2023; Stürmer 2024
			Technisch	Einführung von ITIL- (Information Technology Infrastructure Library) oder COBIT-basierten (Control Objectives for Information and Related Technologies) Basis-Prozessen zur strukturierten Erfassung von KI-Service-Level-Agreements (SLAs)	Papagiannidis et al. 2023
		Mid	Strategisch	Anpassung von KI-Projekten an die jeweilige Entwicklungsphase des Unternehmens (z.B. Nutzung eines "AI-Business Fit" Reifegradmodells)	Costa Barbosa et al. 2024
			Organisatorisch	Implementierung cross-funktionaler Governance-Komitees aus Business- und IT-Vertretern	Rahaman & Mannan 2025
			Technisch	Etablierung automatisierter Test- und Validierungsprozesse (MLOps) sowie Agent-based Workflows zur strukturierten Qualitätskontrolle von KI-Outputs	Papagiannidis et al. 2023; Hynninen & Maunula 2025
		High	Strategisch	Etablierung partizipativer Governance-Strukturen, die eine Mitgestaltung durch Zivilgesellschaft und externe Akteure zulassen	Costa Barbosa et al. 2024
			Organisatorisch	Kombination aus bürokratischer Stabilität und netzwerkartiger Flexibilität ("strategische Agilität") zur Bewältigung von Turbulenzen	Fratini et al. 2024
			Technisch	Hybride Governance-Systemarchitektur: Ein nahtloser Mix aus automatisierten Kontrollsystemen und analogen (menschlichen) Eskalationsinstanzen	Hanisch et al. 2023
	3.3 Risikomanagement	Low	Strategisch	Explizite Definition des eigenen Risikoappetits hinsichtlich der KI-Nutzung	Grete & Kestermann 2022
			Organisatorisch	Bewertung der Auswirkungen von KI auf bestehende Kernkompetenzen im Rahmen des kontinuierlichen strategischen Risikomanagements	Paschen et al. 2020
			Technisch	Durchsetzung grundlegender IT-Sicherheits-Baselines und standardisierter Zertifizierungen (C5 Cloud Computing Compliance Controls Catalogue, ISO 27001) für eingesetzte KI-Dienste	Grete & Kestermann 2022
		Mid	Strategisch	Implementierung eines umfassenden AI Management Systems (z.B. gemäß ISO/IEC 42001) über den gesamten KI-Lebenszyklus	Celsi & Zomaya 2025
			Organisatorisch	Definition expliziter Eskalationsregeln (Human-in-the-loop / Human-on-the-loop) und Notfallpläne bei Providerausfällen	Maathuis & Cools 2025; Papagiannidis et al. 2023
			Technisch	Anpassung der Risikokontrollen an "Non-IID" und unbalancierte Datensätze, die typisch für verteilte KI-Architekturen sind	McMahan et al. 2017
		High	Strategisch	Abkehr von rein unilateralen Kontrollansprüchen; Anerkennung systemischer und sozialer Bedrohungen als Teil der Risikomatrix	Braun & Hummel 2025; Attard-Frost & Widder 2024
			Organisatorisch	Institutionalisierung von "Algorithmic Impact Assessments" zur kontinuierlichen Überwachung gesellschaftlicher KI-Risiken	Kaloudis et al. 2023
			Technisch	Verknüpfung von KI und Blockchain-Technologie zur manipulationssicheren, unveränderbaren Protokollierung von Audit-Trails	Tao et al. 2022

4: Kompetenz & Akteure	4.1 Bewertungskompetenz und Entscheidungssouveränität	Low	Strategisch	Investitionen in grundlegende "Digital Literacy" bei Führungskräften und "Strategic Literacy" beim IT-Personal	Rahaman & Mannan 2025
			Organisatorisch	Durchführung unternehmensweiter Basis-Schulungen (Prompt Engineering, Medienkompetenz) zur Reduktion von "AI-Phobia"	Papagiannidis et al. 2023; Kretschmer & Orth 2024
			Technisch	Bereitstellung standardisierter Leitfäden und Tools, um externe KI-Ergebnisse (z.B. Halluzinationen) auf grundlegende Plausibilität zu prüfen	Tao et al. 2022
		Mid	Strategisch	Internen Kompetenzaufbau priorisieren, um Abhängigkeiten von externen IT-Beratern beim Technologie-Einkauf zu reduzieren	Celtekligil 2020
			Organisatorisch	Fachübergreifende Zusammenarbeit (Legal, IT, Business) institutionalisieren, um technische Anforderungen gegen strategische Risiken souverän abzuwägen	Grete & Kestermann 2022
			Technisch	Befähigung der IT-Sicherheitsteams, maschinenlesbare Usage-Control-Policies selbstständig zu definieren und zu auditieren	Tao et al. 2022
		High	Strategisch	Aufbau der Kapazität zur "Technology Appropriation" – Technologien nicht nur nutzen, sondern ihre Kern-Funktionsweise vollständig beherrschen	Costa Barbosa et al. 2024
			Organisatorisch	Institutionalisierung von transdisziplinären Dialogen zwischen Managern, Ethikern und IT-Architekten	Celsi & Zomaya 2025
			Technisch	Etablierung interner Forschungs- & Entwicklungsprojekte zur Validierung komplexer Algorithmen auf Basis von "Ethics-by-Design"	Klare et al. 2023; Celsi & Zomaya 2025
	4.2 Talentverfügbarkeit und Ressourcen	Low	Strategisch	Den menschlichen Faktor als kritischste Erfolgsressource für KI anerkennen und eine gezielte Talent-Rekrutierungsstrategie aufsetzen	Wamba-Taguimdje et al. 2020
			Organisatorisch	Systematische Adressierung des Fachkräftemangels durch Hands-on-Workshops und Weiterbildungsprogramme	Bues 2025
			Technisch	Implementierung niederschwelliger HR-Management-Tools zur Identifikation und Verfolgung digitaler Kompetenzlücken im Unternehmen	Hynninen & Maunula 2025
		Mid	Strategisch	Formung eines dedizierten Talent-Mixes (Verknüpfung von technologischem KI-Verständnis mit spezifischem Domänenwissen) anstelle von reinen IT-Monokulturen (Data Scientists + Ingenieure + Fachkräfte)	Gujar et al. 2025
			Organisatorisch	Wissenstransfer aktiv managen, um temporäre Überlastungen der Kern-Belegschaft durch neue KI-Anforderungen zu verhindern	Klare et al. 2025
			Technisch	Aufbau interner technischer "Sandbox"-Umgebungen oder "AI-Laboratories" für praxisnahes, risikofreies Experimentieren	Hynninen & Maunula 2025
		High	Strategisch	Aufbau von 'sozialem intellektuellem Kapital' durch strategische Kooperationen mit Universitäten und Forschungseinrichtungen	Kaloudis et al. 2023
			Organisatorisch	Aufbau und Sicherung eines souveränen Talentpools auf nationaler/europäischer Ebene zur Reduzierung globaler Personalanbhängigkeiten	European Commission 2025
			Technisch	Investitionen in den internen Kompetenzaufbau zum autarken Hosting und Fine-Tuning von "Open-Weight" KI-Modellen	Foster 2025
	4.3 Relationale Praktiken und Domänenwissen	Low	Strategisch	KI-Initiativen aus der reinen IT-Sicht befreien; KI muss reale Business-Probleme lösen	Gujar et al. 2025
			Organisatorisch	Zwingende Einbindung von Domänenexperten (Fachabteilungen) bereits in die frühe Anforderungsdefinition von KI-Tools	Papagiannidis et al. 2023
			Technisch	Etablierung agiler Kommunikations-Tools zur Reibungsreduktion zwischen IT-Entwicklern und Fachanwendern	Tao et al. 2022
		Mid	Strategisch	Paradigmenwechsel zu "AI-Human Synergy": Menschen als "Business Executor" beibehalten, die KI nur als Co-Pilot/Augmentation nutzen	Kretschmer & Orth 2024
			Organisatorisch	Etablierung ständiger interdisziplinärer Teams, die technisches Verständnis kontinuierlich mit domänenspezifischem Fachwissen verschmelzen	Gujar et al. 2025; Plekhanov et al. 2023
			Technisch	Manuelle Datenklassifizierung (z.B. bei komplexen Industriedaten) gezielt durch ML-gestützte Auditierungssysteme unterstützen	Tao et al. 2022
		High	Strategisch	Aktive Steuerung der Governance von Start-ups und externen Dienstleistern durch gezielte Partnerschafts-Netzwerke	Hanisch et al. 2023
			Organisatorisch	Aufbau eines Umfelds für "Knowledge Spillovers": Gezielte Inklusion diverser Stakeholder (Zivilgesellschaft, externe Forscher) zur Vermeidung von Betriebsblindheit	Song 2019; Attard-Frost & Widder 2024
			Technisch	Bereitstellung standardisierter Metadaten und "Self-Descriptions", um die Austauschbarkeit und externe Kollaboration in Datenräumen zu fördern	Weiss 2022
5: Ökosysteme & Machtverhältnisse	5.1 Strukturelle Macht und Plattform-Abhängigkeit	Low	Strategisch	Identifikation von "Quasi-Souveränen" (Big Tech oder Hyperscaler) im eigenen Portfolio und kritische Analyse bestehender Lock-in-Risiken	Roberts 2024; Floridi 2020
			Organisatorisch	Transparenz über die gesamte Lieferkette herstellen, inklusive vertraglicher Auditrechte bei Unterauftragnehmern	European Commission 2025
			Technisch	Konsequente Modularisierung der Architekturen, um die Abhängigkeit von zentralen Plattform-Engstellen ("Choke Points") der Hyperscaler zu verringern	Foster 2025; Kretschmer & Orth 2024
		Mid	Strategisch	Aufbau strategischer Allianzen und Diversifizierung der Technologiepartner zur Stärkung der eigenen Verhandlungsposition	Celtekligil 2020; Fratini et al. 2024
			Organisatorisch	Auswahl von Partnern, die ähnliche Souveränitäts- und europäische Werte teilen, um Abhängigkeiten aus Nicht-EU-Staaten zu reduzieren	Klare et al. 2023; Roberts et al. 2021
			Technisch	Einsatz föderierter Architekturen (z.B. Federated Learning), um die Notwendigkeit zentraler Cloud-Datenspeicher bei Hyperscalern zu reduzieren	McMahan et al. 2017

		High	Strategisch	Gezielte staatliche/industrielle Förderung und Nutzung von "Digital Public Goods" (offenen digitalen Gütern)	Stürmer 2024
			Organisatorisch	Anschluss an branchenspezifische, dezentrale Ökosysteme (z.B. Catena-X) zur gemeinsamen Kontrolle der Wertschöpfungskette	Weiss 2022
			Technisch	Etablierung direkter Peer-to-Peer bzw. V2V-Kommunikationsarchitekturen, um zentrale Plattform-Gatekeeper vollständig zu umgehen	Wu et al. 2024
	5.2 Rollenverteilung im sozio-technischen Ökosystem	Low	Strategisch	Die eigene Rolle kritisch reflektieren: Vermeidung der passiven Konsumentenrolle ("Consumer-User-Dependence")	Floridi 2020; Roberts 2024
			Organisatorisch	Vermeidung der vollständigen Offenlegung von Domänenwissen an externe Plattformanbieter zum Schutz des eigenen Geschäftsmodells	Papagiannidis et al. 2023
			Technisch	Bei der Systemauswahl auf offene Standards achten, um überhaupt eine Marktchance für Drittanbieter (Agents) im Ökosystem offenzuhalten	Song 2019
		Mid	Strategisch	"Open Washing" (scheinbare Offenheit von Providern) kritisch prüfen und den Trade-off zwischen Kontrolle und Fähigkeiten aktiv steuern	Foster 2025
			Organisatorisch	Etablierung einer kollaborativen Governance, die Interessenskonflikte zwischen Regierungen, Firmen und Zivilgesellschaft im Ökosystem ausgleicht	Shirwa et al. 2025
			Technisch	Sicherstellung der technischen Interoperabilität zwischen verschiedenen KI-Regimen, um den Wechsel zwischen Allianzen zu ermöglichen	Celsi & Zomaya 2025
		High	Strategisch	Als "Gegengewicht" zur Plattformökonomie agieren, indem aktiv internationale Governance-Standards mitgestaltet werden	Celsi & Zomaya 2025; Costa Barbosa et al. 2024
			Organisatorisch	Fokus auf kollektive, partizipatorische Strukturen (Gemeinwohl) statt privater Monopolisierung von Ressourcen	Braun & Hummel 2025
			Technisch	Aktive Beteiligung an Open-Source-Initiativen und Bereitstellung eigener Open-Source-Module für die Community	Weiss 2022; Hynninen & Maunula 2025
	5.3 Kontrolle über Engstellen in der Wertschöpfungskette	Low	Strategisch	Identifikation von Nischen, in denen das Unternehmen eigene Stärken ausspielen kann, statt horizontal mit Gatekeepern zu konkurrieren	Celtekligil 2020
			Organisatorisch	Analyse der Lieferkette zur Identifikation kritischer, fremdgesteuerter Engstellen (SCRM)	Maathuis & Cools 2025
			Technisch	Stopp von oberflächlichen Integrationen (z.B. reine "GPT-Wrapper") hin zu echten wertschöpfenden Anwendungen	Gujar et al. 2025
		Mid	Strategisch	Fokussierung auf den Aufbau von schwer ersetzbaren Systemen ("Competitive Moats") zur Differenzierung im Wettbewerb	Plekhanov et al. 2023; Gujar et al. 2025
			Organisatorisch	Etablierung von Nischenstrategien, um einseitige Abhängigkeiten von "Digital Giants" aktiv zu stören	Kaloudis et al. 2023
			Technisch	Tiefe Einbettung von KI in Kernprozesse durch die Kombination eigener proprietärer Daten mit externen Modellen (z.B. RAG-Systeme)	Gujar et al. 2025
		High	Strategisch	Umkehr der Machtverhältnisse: Das Unternehmen nutzt eigene spezifische Assets ('Engstellen') zur Machtausübung im Netzwerk	Floridi 2020; Davis & Cobb 2010
			Organisatorisch	Strategische Neupositionierung im Ökosystem als unverzichtbarer Dienstleister für andere Knoten (B2B-Enabler)	Wu et al. 2024
			Technisch	Nutzung von Confidential Computing als technologische Barriere, um eigene IP/Modelle selbst auf fremden Gatekeeper-Plattformen zu schützen	Bues 2025
6: Regulatorische & Normative Souveränität	6.1 Regulatorische und rechtliche Compliance	Low	Strategisch	Compliance nicht als Barriere, sondern als zwingenden strategischen Erfolgsfaktor zur Sicherung der Souveränität anerkennen	Klare et al. 2023
			Organisatorisch	Systematische Prüfung der rechtlichen Exposition gegenüber Drittstaaten-Gesetzen (z.B. US CLOUD Act) und Erfüllung grundlegender Normen (DSGVO)	European Commission 2025; Grete & Kestermann 2022
			Technisch	Umsetzung der "Data Minimization" (Datensparsamkeit) sowie lokale Datenverarbeitung zur Reduktion regulatorischer Angriffsflächen	McMahan et al. 2017; Hynninen & Maunula 2025
		Mid	Strategisch	Proaktive Integration künftiger Vorgaben (z.B. EU AI Act, DORA - Digital Operational Resilience Act) in die frühe Planungsphase von KI-Projekten	Celsi & Zomaya 2025; Bues 2025
			Organisatorisch	Einführung institutionalisierter "Algorithm Impact Assessments" zur Bewertung rechtlicher und gesellschaftlicher KI-Risiken	Kaloudis et al. 2023
			Technisch	Architektonische Trennung von sensiblen Trainingsdaten und Modell-Updates (z.B. Federated Learning) zur Wahrung der DSGVO-Compliance by Design	Wu et al. 2024
		High	Strategisch	Aktive Einbindung in institutionelle Netzwerke, um regulatorische Rahmenbedingungen (Lobbying/Standardsetzung) im eigenen Sinne mitzugestalten	Celtekligil 2020; Stürmer 2024
			Organisatorisch	Etablierung von "Safety Harbours" und regulatorischen Sandboxes zur rechtssicheren Förderung von Deep-Tech-Innovationen	Fratini et al. 2024
			Technisch	Automatisierte Compliance-Nachweise durch "Verifiable Credentials" oder kryptografische Attestierung zur Transparenz gegenüber Regulatoren	Weiss 2022; Bues 2025
	6.2 Normative Verantwortung	Low	Strategisch	Definition erster ethischer Unternehmens-Leitlinien für die KI-Entwicklung und -Nutzung	Klare et al. 2023
			Organisatorisch	Das Top-Management muss normative Verantwortung übernehmen; Compliance darf nicht nur an IT-Entwickler delegiert werden	Klare et al. 2023
			Technisch	Proaktive Prüfung der Trainingsdaten auf Bias und Diskriminierung vor der Produktivschaltung	Paschen et al. 2020
		Mid	Strategisch	Ausrichtung digitaler Initiativen an übergeordneten ökologischen und sozialen Nachhaltigkeitszielen	Shirwa et al. 2025; Plekhanov et al. 2023

			Organisatorisch	Einführung von "Ethics-by-Design"-Checklisten für Entwickler und Etablierung ethischer Sourcing-Praktiken in der Lieferkette	Celsi & Zomaya 2025; Attard-Frost & Widder 2024
			Technisch	Einsatz hardware-basierter Absicherungen, um die Integrität und ethische Vertrauenswürdigkeit der Workloads ("Data in Use") zu garantieren	Bues 2025
		High	Strategisch	Etablierung einer umfassenden Corporate Digital Responsibility (CDR), die auf kollektive Teilhabe und das Gemeinwohl ausgerichtet ist	Costa Barbosa et al. 2024; Braun & Hummel 2025
			Organisatorisch	Berücksichtigung der Stimmen von Betroffenen (Minderheiten, Endnutzer) bei der KI-Entwicklung zur aktiven Verhinderung von Machtasymmetrien	Braun & Hummel 2025
			Technisch	Umsetzung von "Privacy by Design" durch IT-Systeme, die Verwendungszwecke und zeitliche Befristungen von Daten technisch (z.B. im Code) erzwingen	Tao et al. 2022
	6.3 Transparenz, Fairness und Verantwortlichkeit	Low	Strategisch	Anerkennung, dass mangelnde Transparenz das Vertrauen der Nutzer (und damit die Geschäftsgrundlage) zerstört	Shirwa et al. 2025; Song 2019
			Organisatorisch	Einforderung regelmäßiger Transparenzberichte (z.B. zu staatlichen Datenzugriffen) von den Cloud-Providern	Grete & Kestermann 2022
			Technisch	Etablierung rudimentärer Logging-Systeme zur grundlegenden Dokumentation von KI-Entscheidungen	Papagiannidis et al. 2023
		Mid	Strategisch	Verankerung von "Accountability-by-Design": Jede KI-Entscheidung muss einer klaren menschlichen / juristischen Instanz zugeordnet werden	Hanisch et al. 2023
			Organisatorisch	Schließung von "Legitimitätslücken" durch unabhängige, externe technische Audits der genutzten Algorithmen	Roberts 2024
			Technisch	Integration von Explainable AI (XAI) und Dashboards, um KI-Generierung für Kunden und Sachbearbeiter nachvollziehbar zu machen	Papagiannidis et al. 2023
		High	Strategisch	Vermeidung einer "Algokratie" durch die feste Verankerung eines prinzipiellen "Human Override" (menschliche Überstimmungs-Macht)	Hanisch et al. 2023
			Organisatorisch	Einrichtung institutionalisierter Eskalations- und Beschwerdeprozesse bei algorithmischer Diskriminierung für betroffene Nutzer	Braun & Hummel 2025
			Technisch	Nutzung von Blockchain oder Distributed-Ledger-Technologien (DLT), um unveränderbare, manipulationssichere Audit-Trails von Transaktionen und Modellanpassungen zu garantieren	Klare et al. 2023; Tao et al. 2022

Literaturverzeichnis erweitert für diese Tabelle / Handlungsempfehlungen:

Bues, P., 2025. Unlocking the Future of Data Security: Confidential Computing as a Strategic Imperative. White Paper verfügbar unter: https://confidentialcomputing.io/wp-content/uploads/sites/10/2025/11/US53866125.pdf (letzter Aufruf: 29.05.2026)

Celtekligil, K., 2020. Resource Dependence Theory. In: Dincer, H., Yüksel, S. (eds) Strategic Outlook for Innovative Work Behaviours. Contributions to Management Science. Springer, Cham. doi: 10.1007/978-3-030-50131-0_7

Costa Barbosa, A., Herlo, B. und Joost, G., 2024. Digital Sovereignty between perils of hegemonic agendas and possibilities of alternative approaches. doi: 10.18617/liinc.v20i2.7312

Davis, G. und Cobb, A., 2010. Chapter 2 Resource dependence theory: Past and future, Stanford's Organization Theory Renaissance, 1970–2000, doi: 10.1108/S0733-558X(2010)0000028006

Grete, P. und Kestermann, C., 2020. Cloud Computing im Spannungsfeld von Souveränität, Cybersicherheit und Ökonomie. Datenschutz Datensich 46, 632–636. doi: 10.1007/s11623-022-1672-x

Hanisch, M., Goldsby, C., Fabian, N., Oehmichen, J., 2023. Digital governance: A conceptual framework and research agenda, Journal of Business Research, Volume 162, 113777, ISSN 0148-2963. doi: 10.1016/j.jbusres.2023.113777.

Kaloudis, M., & Chmielarz, G. und Kucęba, R., 2023. Managing digital sovereignty: The importance of control and ownership in the digital age using the examples of Poland, the Czech Republic and Germany. 148-155. doi: 10.17512/CUT/9788371939563/23.

Klare, M., Verlande, L., Greiner, M., Lechner, U., 2023. How Blockchain and Artificial Intelligence influence Digital Sovereignty. In: Papadaki, M., Rupino da Cunha, P., Themistocleous, M., Christodoulou, K. (eds) Information Systems. EMCIS 2022. Lecture Notes in Business Information Processing, vol 464. Springer, Cham. doi: 10.1007/978-3-031-30694-5_1

McMahan, H., Moore, E., Ramage, D., Arcas, B., 2017. Communication-Efficient Learning of Deep Networks from Decentralized Data. Doi: arxiv.org/abs/1602.05629

Plekhanov, D., Franke, H., Netland, T., 2023. Digital transformation: A review and research agenda, European Management Journal, Volume 41, Issue 6, Pages 821-844, ISSN 0263-2373. Doi: 10.1016/j.emj.2022.09.007.

Pohle, J. und Thiel, T., 2020. Digital sovereignty. Internet Policy Review, 9(4). Doi: 10.14763/2020.4.1532

Rahaman, A., Mannan, A., 2025. IT Governance and Business Strategy Alignment in the Era of Digital Transformation, Vol 3 Issue 4, J. state gov. mass media 2025. doi: 10.64907/xkmf.v3i4.jsgmm.2

Roberts, H., 2021. Safeguarding European values with digital sovereignty: an analysis of statements and policies. Internet Policy Review, 10(3). Doi: 10.14763/2021.3.1575

Shirwa, A. et al., 2025. A cooperative governance framework for sustainable digital transformation in construction: The role of digital enablement and digital strategy, Results in Engineering, Volume 25, 104139, ISSN 2590-1230, doi: 10.1016/j.rineng.2025.104139.

Song, A., 2019. The Digital Entrepreneurial Ecosystem - a critique and reconfiguration. Small Bus Econ 53, 569–590. Doi: 10.1007/s11187-019-00232-y

Weiss, A., 2022. Gaia-X – Grundlagen für den Aufbau föderierter, digitaler Ökosysteme nach europäischen Regeln. Datenschutz Datensich 46, 227–232. doi: 10.1007/s11623-022-1593-8

Wu, C., Fan, H., Wang, K., Zhang, P., 2024. Enhancing Federated Learning in Heterogeneous Internet of Vehicles: A Collaborative Training Approach. Electronics, 13, 3999. Doi: 10.3390/electronics13203999