

Bachelorarbeit

Systematische Literaturanalyse zu bestehenden Vorgehens- und Referenzmodellen im Bereich Datenanalyse zur Zuverlässigkeitsprognose von Heizsystemen

Zur Erlangung des Grades Bachelor of Science (B. Sc.)

Vorgelegt von:	Dominik Blanke
Matrikelnummer:	212218
Studiengang:	Wirtschaftsingenieurwesen
Ausgabedatum:	23.03.2023
Abgabedatum:	21.06.2023
Erstprüfer:	Dr.-Ing. Dipl.Inform. Anne Antonia Scheidler
Zweitprüfer:	Sahil-Jai Arora

Technische Universität Dortmund

Fakultät Maschinenbau

Fachgebiet IT in Produktion und Logistik

Inhaltsverzeichnis

Abbildungsverzeichnis.....	III
Tabellenverzeichnis.....	IV
Abkürzungsverzeichnis	V
1 Einleitung.....	1
2 Grundlagen zur Zuverlässigkeitsprognose von Heizsystemen mittels Zustandsdaten	3
2.1 Instandhaltung und Instandhaltungsstrategien	3
2.1.1 Ausfallstrategie: reaktive Instandhaltung.....	6
2.1.2 Präventivstrategie: vorbeugende Instandhaltung	7
2.1.3 Inspektionsstrategie: zustandsabhängige Instandhaltung.....	7
2.1.4 Vorausschauende Instandhaltung.....	8
2.2 Zuverlässigkeit	9
2.2.1 Quantitative Zuverlässigkeitskenngrößen	10
2.2.2 Qualitative Zuverlässigkeitsmethoden.....	11
2.3 Referenzmodelle	12
2.4 Wissensentdeckung in Datenbanken.....	14
2.4.1 Vorgehensmodell Knowledge Discovery in Databases (KDD).....	14
2.4.2 Cross Industry Standard Process for Data Mining (CRISP-DM).....	16
2.5 Heizungssysteme.....	18
3 Anforderungsprofil zur Zuverlässigkeitsprognose von Heizungsanlagen.....	21
4 Methodik der systematischen Literaturrecherche.....	25
4.1 Konzepte der systematischen Literaturrecherche.....	25
4.2 Durchführung der Literaturrecherche	27
5 Literaturlauswertung und Darstellung relevanter Ansätze	34
5.1 Machine-Learning Ansatz für vorausschauende Instandhaltung nach Fernandes et al. (2020)	34
5.2 Machine-Learning Ansatz für vorausschauende Instandhaltung nach Bouabdallaoui et al. (2021)	35
5.3 Machine-Learning Framework für nachhaltige Instandhaltung nach Villa et al. (2022).....	35
5.4 Digitaler Zwilling für die vorausschauende Instandhaltung nach Hosamo et al. (2022).....	36
5.5 Machine-Learning Ansatz für die Fehlerdetektion und -diagnose nach Albayati et al. (2023)	37

5.6	Hybrider modellbasierter Ansatz für die Fehlererkennung nach Galvez et al., (2021).....	37
5.7	Prozess zur Fehlerhandhabung nach Mansson et al., (2021)	38
5.8	Fehlerdetektion und -diagnose für die Instandhaltung nach Chen et al. (2022)	39
5.9	Machine-Learning Ansatz für die Zustandsüberwachung nach Prabhu & Chaudhary (2021)	40
5.10	Zuverlässigkeitsmodellierung nach Jayatilleka & Peters (2021).....	40
5.11	Zuverlässigkeitsbewertung nach Ciani et al., (2020)	41
5.12	Datenbasierter Ansatz zur Fehlererkennung und -isolation nach Li et al., (2020).	42
5.13	Machine-Learning Ansatz zur Zustandsprognose nach Yang et al., 2020	43
5.14	Methode zur Fehlerdetektion nach Zhu et al., (2022).....	44
5.15	Prozess zur Fehlerdetektion und -diagnose mit Digitalem Zwilling nach Xie et al., (2023).....	45
5.16	Anwendung des IDAIC-Modells nach Ahern et al., (2023)	45
5.17	Anomalie-Detektion mit Machine-Learning nach Borda et al., (2023)	46
5.18	Transfer-Lern-Methode für Fehlerdetektion und -diagnose nach Martinez-Viol et al. (2022).....	47
5.19	Deep-Learning Ansatz zur Fehlerdetektion und -diagnose nach Haruehansapong et al. (2022)	47
5.20	Gegenüberstellung der relevanten Ansätze mit dem Anforderungsprofil	48
6	Zusammenfassung und Ausblick	50
7	Literaturverzeichnis	51
	Anhang	56

Abbildungsverzeichnis

Abbildung 1: Gegenüberstellung der Instandhaltungsziele (vgl. Pawellek, 2013, S. 46)	4
Abbildung 2: Grundmaßnahmen der Instandhaltung (vgl. Strunz, 2012, S. 3)	5
Abbildung 3: Basisstrategien der Instandhaltung (vgl. Strunz, 2012, S. 295)	6
Abbildung 4: Anwendungsbeispiel für ein Referenzmodell (vgl. Arora & Rabe, 2023a, S. 3)	13
Abbildung 5: KDD-Prozess (vgl. Fayyad, et al., 1996, S. 41)	14
Abbildung 6: Ebenen des CRISP-DM Modells (vgl. Chapman, et al., 2000, S. 9)	16
Abbildung 7: Phasen des CRISP-DM-Modells (vgl. Chapman, et al., 2000, S. 13)	17
Abbildung 8: Beispielhafte Heizungsanlage mit zwei Heizkreisen (vgl. Lauckner & Krimmling, 2020, S. 99)	20
Abbildung 9: Literatur-Konzeptmatrix (vgl. Webster & Watson, 2002, S. 17)	25
Abbildung 10: Vorgehen nach vom Brocke et al. (vgl. vom Brocke, et al., 2009, S. 7)	26

Tabellenverzeichnis

Tabelle 1: Basisstrategien im Vergleich (vgl. Strunz, 2012, S. 300)	8
Tabelle 2: Anforderungsprofil für das Zuverlässigkeitsprognose-Modell	21
Tabelle 3: Vorläufige Liste der englischen Keywords für die Literaturrecherche	28
Tabelle 4: Vorläufige Liste der deutschen Keywords für die Literaturrecherche	28
Tabelle 5: Beispielhafte Suchkette	29
Tabelle 6: Finale Liste der Keywords	30
Tabelle 7: Suchfelder auf den Datenbanken	30
Tabelle 8: Literaturübersicht der gefundenen Ansätze	32
Tabelle 9: Gegenüberstellung der Ansätze mit dem Anforderungsprofil	49

Abkürzungsverzeichnis

TWh	Terawattstunden
IoT	Internet of Things
BE	Betrachtungseinheit
LCC	Life Cycle Costs
FMEA	Failure Mode and Effects Analysis
FTA	Fault Tree Analysis
RM	Referenzmodell
KDD	Knowledge Discovery in Databases
CRISP-DM	Cross Industry Standard Process for Data Mining
KI	künstliche Intelligenz
ML	Machine Learning
KLH	Heizungs-, Lüftungs- und Klimatisierung
SCoT	Smart Cloud of Things
LSTM	Long Short-Term Memory
BIM	Building Information Model
FDD	Fault Detection and Diagnosis
ANN	Artificial Neural Network
SVM	Support Vector Machine
K-NN	k-Nearest Neighbor
CMMS	Computerized Maintenance Management System
IM	Infant Mortality
PFMEA	Process Failure Mode and Effect Analysis
DFMEA	Design Failure Mode and Effect Analysis
RBD	Reliability Block Diagram
FDI	Failure Detection and Isolation
CNN	Convolutional Neural Network
TTF	Time To Failure
SVDD	Support Vector Data Description
BoG	Bag of Words
IDAIC	Industrial Data Analysis Improvement Cycle
NN	Neural Network
AFDD	Automatic Fault Detection and Diagnosis
DL	Deep Learning

1 Einleitung

In Deutschland liegt der Gesamtenergieverbrauch jährlich bei rund 2.407 Terawattstunden (TWh) (2021), wobei 670 TWh auf die Haushalte entfallen (AG Energiebilanzen, 2022). Von den 670 TWh entfallen rund 68,2% (ca. 457 TWh) auf die Beheizung von Wohngebäuden (AG Energiebilanzen, 2023). Das entspricht einem Anteil von ca. 19% am Gesamtenergieverbrauch. Eine optimierte Zuverlässigkeitsprognose durch Analyse von Felddnutungsdaten bietet das Potenzial zur Reduzierung des Energieverbrauchs.

In der heutigen Zeit ist es möglich durch das Internet of Things (IoT) mit moderne Messensoren Daten von auch komplexen Anlagen kontinuierlich zu erfassen und digital auszuwerten (Bauernhansl, et al., 2016). Mittels dieser Informationen kann folglich nicht nur das Heizsystem effizient geregelt und gesteuert werden, sondern bspw. die Zuverlässigkeit der Anlagen optimiert und Zustände prognostiziert werden (Bauernhansl, et al., 2016). Dazu bietet z. B. die vorausschauende Instandhaltung die Möglichkeit durch kontinuierliches Messen und Auswerten von Daten die restliche Lebensdauer von Maschinenkomponenten zu bestimmen (Feldmann, et al., 2017).

Trotz der Verfügbarkeit von Daten, bspw. Nutzungsdaten von Heizsystemen, ist deren Nutzung und Verwertung im Sinne der Wissensentdeckung relativ gering und entspricht nicht den Potenzialen (Arora & Rabe, 2023b). Gründe dafür sind unter anderem hohe Investitionskosten (Karuppiah, et al., 2020) und das Fehlen konkreter Ansatzpunkte für eine erfolgreiche Implementierung (Bauernhansl, et al., 2016). In der Forschung und Industrie gibt es bereits bestehende Referenz- und Vorgehensmodelle zur Entwicklung von Informationsmodellen (Arora & Rabe, 2023a). Diese bestehenden Modelle sind zumeist domänen- und anwendungsspezifisch und unterscheiden sich stark in ihrer Detailebene.

Eine Übersicht über die bestehenden Referenz- und Vorgehensmodelle bietet das Potenzial zur bedarfsgerechten Auswahl und Optimierung dieser Modelle, angepasst an die Anforderungen für den Bereich Zuverlässigkeitsprognose für Heizsysteme. Das Ziel dieser Arbeit ist die Schaffung einer systematischen Übersicht über vorhandene Referenz- und Vorgehensmodelle zur Analyse von Anlagendaten sowie die Bewertung der Anwendbarkeit der Modelle auf Heizsysteme anhand eines Anforderungsprofils.

Um das Ziel dieser Arbeit zu erreichen, wird die Arbeit folgendermaßen gegliedert: Zunächst werden in Kapitel 2 die Grundlagen zur Zuverlässigkeitsprognose von Heizsystemen mittels Zustandsdaten erläutert. Dabei wird auf die verschiedenen Themenbereiche der Arbeit eingegangen. Als erstes wird der Stand der Technik zur Instandhaltung und zu Instandhaltungsstrategien thematisiert. Daraufhin folgt eine Erläuterung des Begriffs Zuverlässigkeit

und es werden verschiedene Kenngrößen der Zuverlässigkeit beschrieben. Im dritten Abschnitt wird der Begriff Referenzmodell erläutert. Anschließend erfolgt eine Vorstellung von Modellen zur Wissensentdeckung in Datenbanken. Abschließend werden in Kapitel 2 die Grundlagen von Heizungsanlagen besprochen. In Kapitel 3 folgt die Erstellung eines Anforderungsprofils an die gesuchten Zuverlässigkeitsprognose-Modelle. In Kapitel 4 wird die Methode zur systematischen Literaturanalyse geschildert. Die gefundenen Ansätze werden in Kapitel 5 kurz dargestellt und dem erstellten Anforderungsprofil aus Kapitel 3 gegenübergestellt. Abschließend erfolgt eine Zusammenfassung der Ergebnisse und ein Ausblick für die zukünftige Forschung.

2 Grundlagen zur Zuverlässigkeitsprognose von Heizsystemen mittels Zustandsdaten

Im nachfolgenden Kapitel werden die Grundlagen zur Zuverlässigkeitsprognose von Heizsystemen erläutert. Dazu wird zunächst der Stand der Technik in der Instandhaltung beschrieben und Instandhaltungsstrategien vorgestellt. Anschließend wird der Begriff Zuverlässigkeit definiert und es werden verschiedene Zuverlässigkeitskenngrößen präsentiert. Daraufhin werden Definitionen zum Begriff sowie ein Anwendungsbeispiel zu Referenzmodellen gegeben. Im vierten Abschnitt werden Modelle zur Wissensentdeckung in Datenbanken dargestellt. Abschließend werden im letzten Abschnitt des Kapitels Grundlagen zu Heizsystemen erläutert.

2.1 Instandhaltung und Instandhaltungsstrategien

Im folgenden Abschnitt der Arbeit werden Grundlagen und Ziele der Instandhaltung erläutert. Dabei werden die vier Grundmaßnahmen Instandsetzung, Inspektion, Wartung und Verbesserung der Instandhaltung erläutert. Zudem werden grundlegende Strategien der Instandhaltung dargestellt und um die Strategie der vorausschauenden Instandhaltung erweitert.

Die Instandhaltung wird nach der DIN 31051 (2019) wie folgt definiert: Instandhaltung ist die „Kombination aller technischen und administrativen Maßnahmen sowie Maßnahmen des Managements während des Lebenszyklus einer Einheit, die dem Erhalt oder der Wiederherstellung ihres funktionsfähigen Zustandes dient, sodass sie die geforderte Funktion erfüllen kann“ (DIN 31051, 2019, S. 4). Das grundlegende Ziel der Instandhaltung ist es, die Funktion der Betrachtungseinheit (BE) während der Lebensdauer soweit möglich störungsfrei zu gestalten und durch Verzögerung der Abnutzung negative Folgen bei der Nutzung zu verhindern. Somit soll die Instandhaltung zu einer störungsfreien Nutzung einer Anlage beitragen. (Strunz, 2012) Innerhalb der Instandhaltung werden zudem weitere Ziele verfolgt, welche teilweise in Konkurrenz zueinanderstehen (s. Abbildung 1). Zum einen soll die Ausfallzeit, die vorausschauende Instandhaltungszeit sowie die Instandhaltungskosten minimiert werden. Dahingegen ist es das Ziel, die Anlagenverfügbarkeit, die Zuverlässigkeit und die Lebensdauer zu maximieren und die Life-Cycle-Costs (LCC) zu minimieren. (Pawellek, 2013)

Die genannten Instandhaltungskosten lassen sich in direkte und indirekte Kosten unterteilen. Unter indirekten Instandhaltungskosten versteht man Kosten, die keinen direkten Bezug zu der Instandhaltung haben. Darunter fallen die durch Stillstände verursachten Verluste sowie die Verringerung des Wertes der Anlage durch das Auslassen der Instandhaltung. Die indirekten Kosten sind folglich schwierig zu messen, wenn die Anlage keinen Ertrag erzeugt oder der Schaden nicht quantifiziert werden kann. Direkte Instandhaltungskosten sind dage-

gen leicht zu erfassen und stellen alle Kosten dar, die direkt mit dem Instandhaltungsprozess in Verbindung stehen. (Pawellek, 2013) „Der Ausdruck *Life-Cycle-Costs* steht für die *totalen* Kosten, die ein System während seiner Lebensdauer aus Betreibersicht verursacht“ (VDI 2884, 2005, S. 4). Die LCC setzen sich aus den Investitionen, Roh- und Betriebsstoffen sowie den Kosten für die laufenden und langfristigen Instandhaltungsmaßnahmen zusammen (Pawellek, 2013).

Ziele	Minimierung der Ausfallzeit	Minimierung der vorbeugenden IH-Zeit	Minimierung der direkte IH-Kosten	Minimierung der indirekte IH-Kosten	Maximierung der Anlagenverfügbarkeit	Zuverlässigkeitsmaximierung	Lebensdauermaximierung	Life-Cycle-Costs-Minimierung
Minimierung der Ausfallzeit	x	-	-	+	+	+	=	+/-
Minimierung der vorbeugenden IH-Zeit	-	x	+	-	+/-	-	-	+/-
Minimierung der direkte IH-Kosten	-	+	x	-	-	-	=	+/-
Minimierung der indirekte IH-Kosten	+	-	-	x	+	+	=	+/-
Maximierung der Anlagenverfügbarkeit	+	+/-	-	+	x	+	=	+/-
Zuverlässigkeitsmaximierung	+	-	-	+	+	x	=	+/-
Lebensdauermaximierung	=	-	=	=	=	=	x	+/-
Life-Cycle-Costs-Minimierung	+/-	+/-	+/-	+/-	+/-	+/-	+/-	x

- konkurrierend
 = indifferent
 + komplementär
 +/- Teilweise konkurrierend und komplementär

Abbildung 1: Gegenüberstellung der Instandhaltungsziele (vgl. Pawellek, 2013, S. 46)

Zu den Aufgaben der Instandhaltung zählen unter anderem die vier Grundmaßnahmen Wartung, Inspektion, Instandsetzung und Verbesserung (s. Abbildung 2). Darüber hinaus ist es Aufgabe der Instandhaltung unter anderem auch potenzielle Störungen besser sichtbar zu machen und das Ausfallverhalten zu analysieren sowie das generelle Verhindern von Störungen. (DIN 31051, 2019)

Die DIN 31051 (2019) definiert die Grundmaßnahmen und listet darüber hinaus Aufgaben und Maßnahmen, die bei den entsprechenden Grundmaßnahmen anfallen, auf.

Instandsetzung

Instandsetzung ist die „physische Maßnahme, die ausgeführt wird, um Funktionen eines Fehlerhaften Objektes wiederherzustellen“ (DIN 31051, 2019, S. 6).

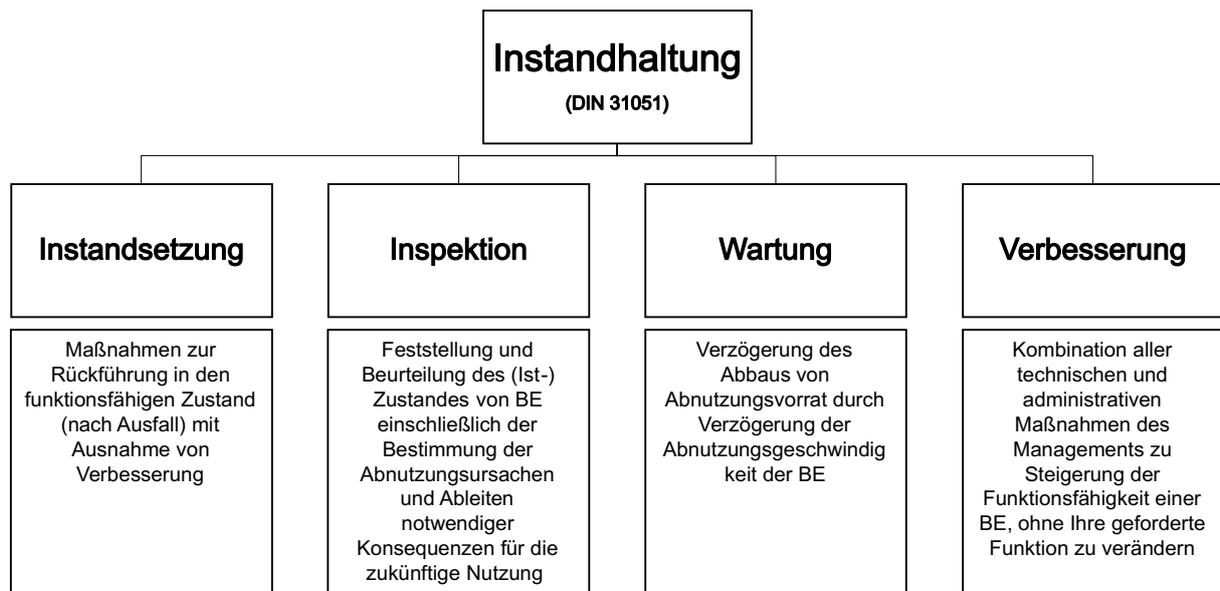


Abbildung 2: Grundmaßnahmen der Instandhaltung (vgl. Strunz, 2012, S. 3)

Inspektion

Unter Inspektion versteht man die „Prüfung auf Konformität der maßgeblichen Merkmale eines Objekts, durch Messung, Beobachtung oder Funktionsprüfung“ (DIN 31051, 2019, S. 5).

Wartung

Bei der Wartung werden „Maßnahmen zur Verzögerung des Abbaus des vorhandenen Abnutzungsvorrats“ (DIN 31051, 2019, S. 5) durchgeführt.

Verbesserung

Unter Verbesserung versteht man die „Kombination aller technischen und administrativen Maßnahmen sowie Maßnahmen des Managements zur Steigerung der immanenten Zuverlässigkeit und/oder Instandhaltbarkeit und/oder Sicherheit eines Objektes, ohne seine ursprüngliche Funktion zu ändern“ (DIN 31051, 2019, S. 6).

Als Instandhaltungsstrategie versteht man die „Vorgehensweise des Managements zur Erreichung der Instandhaltungsziele“ (DIN EN 13306, 2017, S. 9). Instandhaltungsstrategien sollen in erster Linie dazu beitragen, ein festgelegtes Ziel zu erreichen (Strunz, 2012). Zudem sollen sie helfen, das Hauptaugenmerk der zu treffenden Maßnahmen auf relevante Elemente zu werfen (Pawellek, 2013). Die Instandhaltungsmaßnahmen, wie z. B. die Intervallart und die benötigten Ressourcen, sollen innerhalb der Instandhaltungsstrategie kostenoptimiert gestaltet werden (Strunz, 2012; Ryll & Freund, 2010). In der Literatur werden für die Instandhaltung zumeist drei Basisstrategien beschrieben: die Ausfallstrategie, die Präventivstrategie und die Inspektionsstrategie (Pawellek, 2013; Ryll & Freund, 2010; Strunz, 2012).

Strunz (2012) ergänzt die Strategien um die Modernisierungsstrategie (s. Abbildung 3). Ryll & Freund (2010) ergänzen bereits die vorausschauende Instandhaltung als Weiterentwicklung der Inspektionsstrategie.

Im Folgenden werden die Ausfallstrategie, die Präventivstrategie, die Inspektionsstrategie sowie die vorausschauende Instandhaltungsstrategie näher erläutert.



Abbildung 3: Basisstrategien der Instandhaltung (vgl. Strunz, 2012, S. 295)

Welche Strategie angewendet wird, hängt von verschiedenen Faktoren ab (Pawellek, 2013). Unter anderem sind das Ausfallrisiko und das Ausfallverhalten sowie die damit einhergehenden ökonomischen Schäden entscheidend (Strunz, 2012). Zu den weiteren Faktoren zählen die Kosten, Produktqualität, Ausfallverhalten, Qualifikation der Mitarbeiter, gesetzliche Vorschriften und geplante Verfügbarkeit (Hodapp, 2018). Außerdem spielt die Menge an vorliegenden Informationen und Daten zum Verhalten der Anlage eine wichtige Rolle (Strunz, 2012). Aufgrund der Komplexität der Anlagen und den Umweltbedingungen ist die Auswahl einer passenden Strategie schwierig. Zudem sind vorhandene Instandhaltungsstrategien häufig über die Zeit entwickelt worden und beruhen auf subjektiven Erfahrungen. Sie entsprechen daher oft nicht den aktuellen Anforderungen und sind schwierig nachvollziehbar aufgrund von schlechter oder unvollständiger Dokumentation. (Pawellek, 2013)

2.1.1 Ausfallstrategie: reaktive Instandhaltung

Die Ausfallstrategie zeichnet sich dadurch aus, dass der Ausfall der betrachteten Anlage abgewartet wird. Somit wird bei dieser Strategie die Anlagenverfügbarkeit maximal ausgenutzt und kein Aufwand in vorbeugende Maßnahmen wie Wartung und Inspektion investiert. (Strunz, 2012; Ryll & Freund, 2010) Der Prozess der Instandhaltung, also auch die Planung,

wird erst nach dem Eintreten des Ausfalls gestartet (Pawellek, 2013). Sie ist mit geringen Kosten verbunden und benötigt zudem wenig Planung. Nachteil dieser Strategie ist, dass es aufgrund von langen Wartezeiten auf die Instandhaltungsmaßnahmen oder einer lang andauernden Fehlersuche zu einem langen Stillstand der Anlage kommen kann. Infolgedessen kann es wegen des Zeitdruckes zu einer fehlerhaften Instandhaltung kommen. Da der Ausfall der Maschine plötzlich auftritt, bedarf es zudem einer längeren Vorbereitungszeit der Maßnahmen. Durch den Ausfall können außerdem Folgeschäden auftreten, die zu erheblichen Mehrkosten führen. (Strunz, 2012) Die Strategie eignet sich aus genannten Gründen vor allem für Anlagen oder Bauteile, welche keine wichtige Rolle oder Bedeutung haben (Pawellek, 2013; Strunz, 2012).

2.1.2 Präventivstrategie: vorbeugende Instandhaltung

Die vorbeugende Instandhaltung zeichnet sich dadurch aus, dass die Instandhaltung mittels Planung in bestimmten, festgelegten Intervallen durchgeführt wird. Der aktuelle Zustand der Anlage wird nicht betrachtet. Die Festlegung der Intervalle kann, je nach Anlagenart, nach einer Anzahl von Stunden im Betrieb, einer bestimmten Produktionsmenge oder nach einem Zeitintervall geschehen. (Pawellek, 2013; Ryll & Freund, 2010) Ein Vorteil dieser Strategie ist die hohe Planbarkeit und dadurch verringerte Stillstandzeiten während der Instandhaltung. Zudem sinkt das Risiko eines plötzlich auftretenden Ausfalls und den möglichen, damit einhergehenden Folgeschäden. Allerdings hat die Präventivstrategie auch Nachteile. Da die Instandhaltung der Elemente nach einem festgelegten Intervall durchgeführt wird und der aktuelle Zustand nicht betrachtet wird, kann es zu nicht notwendigen Kosten kommen, da noch funktionsfähige Elemente durch neue ersetzt werden. Zudem ist die Strategie aufgrund der Planung kostenintensiv und aufwendig, da Wissen zum Ausfallverhalten, der Nutzungsintensität und der Nutzungsdauer vorliegen muss, um die Instandhaltungsintervalle festzulegen. (Strunz, 2012) Die Methode wird angewendet, wenn Vorgaben aus Gesetzen eine regelmäßige Instandhaltung verlangen oder ein Ausfall einer Einheit schwerwiegende Folgen hat (Pawellek, 2013).

2.1.3 Inspektionsstrategie: zustandsabhängige Instandhaltung

Bei der zustandsabhängigen Instandhaltung wird eine Betrachtungseinheit erst erneuert, wenn sie eine bestimmte Grenze des Abnutzungsvorrats erreicht hat. Um festzustellen, wann diese Grenze erreicht ist, wird regelmäßig der Ist-Zustand der Einheit ermittelt. Auf der Basis der ermittelten Ist-Zustände wird der Abnutzungsverlauf prognostiziert. Zur Ermittlung des Ist-Zustandes kann eine beliebige Methode eingesetzt werden. Diese kann frühzeitig geplant werden. (Pawellek, 2013; Ryll & Freund, 2010) Ein Hauptvorteil der Strategie ist,

dass die Lebenszeit der Betrachtungseinheit weitestgehend ausgenutzt wird, bis sie kurz vor dem Ausfall steht. Damit einhergehend sind die Kosten für Ersatzteile geringer. Durch die regelmäßige Erfassung des Ist-Zustandes werden zudem Fehler frühzeitig erkannt und das Risiko eines Ausfalls wird gesenkt. Die Instandhaltungsmethode ist allerdings mit einem hohen Kosten- und Personalaufwand verbunden, da die Erfassung des Ist-Zustandes aufwendig ist und qualifiziertes Personal bedarf. (Strunz, 2012) Die Strategie findet Anwendung, wenn die Abnutzung bzw. der Abnutzungsvorrat gemessen werden kann und der Aufwand aus wirtschaftlichen Aspekten vertretbar ist (Ryll & Freund, 2010).

Tabelle 1 zeigt einen Vergleich der drei beschriebenen Instandhaltungsbasisstrategien anhand von verschiedenen Bewertungskriterien.

Tabelle 1: Basisstrategien im Vergleich (vgl. Strunz, 2012, S. 300)

Bewertungskriterium	Ausfallmethode	Präventivmethode	Inspektionsmethode
Ausnutzung des Abnutzungsvorrats	maximal	gering	optimal
Restnutzungsdauer	keine	vorhanden	gering/keine
Definierte Überlebenswahrscheinlichkeit	keine	vorhanden	vorhanden
Verfügbarkeit	gering	hoch	hoch/sehr hoch
Fehlersuchzeit	hoch	keine	keine
Wartezeit auf Instandhaltung	hoch	keine	keine
Störung im Produktionsablauf	sehr hoch	gering	keine
Folgeschäden	hoch	gering	kaum
Gefahr gesundheitlicher Schäden	möglich	begrenzt möglich/sehr gering	begrenzt möglich/sehr gering
Kapazitätsausnutzung	gering	hoch	hoch
Lohnkosten	sehr hoch	hoch	gering
Planungsaufwand	kleiner	sehr hoch	gering
Technischer Aufwand	kleiner	gering	hoch
Lagerhaltungskosten	sehr hoch	relativ hoch, aber vertretbar	gering

2.1.4 Vorausschauende Instandhaltung

Eine weitere Strategie für die Instandhaltung ist die sogenannte vorausschauende Instandhaltung. Bei der vorausschauenden Instandhaltung werden durchgängig Messungen mithilfe von Sensoren an der Anlage durchgeführt und die gewonnenen Daten mittels Computeralgorithmen analysiert. Dadurch können Fehler frühzeitig erkannt und der perfekte Zeitpunkt für die Instandhaltungsmaßnahmen bestimmt werden. (Schuh, et al., 2020) Die Grundlage für die vorausschauende Instandhaltung bietet das *Condition Monitoring*. Hierbei werden Echtzeitdaten über die Anlagenteile gesammelt. Diese werden allerdings noch nicht für die Prog-

nose der Zuverlässigkeit verwendet. (Feldmann, et al., 2017) Ein Hauptvorteil dieser Strategie wird darin gesehen, die Anlagenverfügbarkeit zu steigern (Schuh, et al., 2020). Laut einer Umfrage der Roland Berger GmbH liegt für 79% der befragten Unternehmen der Nutzen der vorausschauenden Instandhaltung in der Leistungssteigerung. Darunter fallen eine erhöhte Anlagenverfügbarkeit, eine erhöhte Produkt-/Prozessqualität, verbesserte Planung von Servicezyklen, eine erhöhte Lebensdauer und erhöhte Betriebssicherheit und Nachhaltigkeit. 21% der befragten Unternehmen sehen den Nutzen dahingegen in einer Kostensenkung. (Feldmann, et al., 2017)

2.2 Zuverlässigkeit

In diesem Abschnitt wird zunächst der Begriff Zuverlässigkeit erläutert. Des Weiteren werden quantitative Zuverlässigkeitskenngrößen dargestellt sowie die qualitativen Zuverlässigkeitsmethoden Fehlermöglichkeits- und Einflussanalyse (Failure Mode and Effects Analysis (FMEA)) und die Fehlerbaumanalyse (Fault Tree Analysis (FTA)) kurz vorgestellt.

Die Zuverlässigkeit ist die „die Fähigkeit eines Objekts, eine geforderte Funktion unter gegebenen Bedingungen für eine gegebene Zeitspanne zu erfüllen“ (DIN EN 13306, 2017, S. 15). Somit lässt sich die Zuverlässigkeit im technischen Sinne mit der Funktionsfähigkeit gleichsetzen (Ryll & Freund, 2010). Um die Zuverlässigkeit darstellen zu können, gibt es quantitative und qualitative Methoden (Bertsche & Lechner, 2004). Für die quantitative Bestimmung der Zuverlässigkeit werden stochastische bzw. statistische Kenngrößen herangezogen. Zu den Kenngrößen zählen die Ausfallwahrscheinlichkeit, die Überlebenswahrscheinlichkeit, die Ausfallrate, die mittlere Lebensdauer und die Verfügbarkeit. (Ryll & Freund, 2010) Die verbreitetste qualitative Bestimmung der Zuverlässigkeit stellt die FMEA dar. Zudem zählt auch die FTA zu den qualitativen Methoden. (Bertsche & Lechner, 2004) Bei der Bewertung der Zuverlässigkeit von Anlagen kann in die Zustände funktionstüchtig und nicht funktionstüchtig unterteilt werden (Ryll & Freund, 2010). Des Weiteren kann eine Aufteilung zwischen einfachen und strukturierten Systemen getroffen werden. Bei einfachen Systemen hängt die Zuverlässigkeit nur von einer einzelnen Betrachtungseinheit ab, wohingegen bei strukturierten Systemen auch in Betracht gezogen werden muss, dass einzelne Komponenten miteinander verknüpft sind. (Ryll & Freund, 2010) Strukturierte Systeme können in Seriensysteme und Parallelsysteme unterteilt werden. Seriensysteme zeichnen sich dadurch aus, dass der Ausfall einer Komponente einen Ausfall des Systems hervorruft. Die Komponente mit der geringsten Lebensdauer ist also entscheidend. Parallelsysteme hingegen fallen erst nach Ausfall aller einbezogenen Komponenten aus. Bei den Komponenten werden zwischen Arbeits- und Reserveelementen unterschieden. Erst nach Ausfall aller Reserveelemente kommt es zum Systemausfall. (Ryll & Freund, 2010)

2.2.1 Quantitative Zuverlässigkeitskenngrößen

Im folgenden Abschnitt werden verschiedene quantitative Zuverlässigkeitskenngrößen vorgestellt und erläutert.

Ausfallwahrscheinlichkeit

$$F(t) = P(t \leq T) \quad (1)$$

Bei der Ausfallwahrscheinlichkeit handelt es sich um die Wahrscheinlichkeit eines Ausfalls einer Betrachtungseinheit in einem bestimmten Zeitintervall $(0, T)$. Sie beschreibt demnach, dass die Betrachtungseinheit bis zum Zeitpunkt t nicht ausfällt. Zudem stellt die Ausfallwahrscheinlichkeit dar, wie wahrscheinlich ein Ausfall in einem gewählten Intervall ist. (Strunz, 2012) Die Ausfallwahrscheinlichkeit experimentell wie folgt ermittelt werden (Strunz, 2012):

$$F(t) = \frac{m(t) - m(t + \Delta t)}{m} \quad (2)$$

$m(t)$	Anzahl der Funktionsfähigen BE zum Zeitpunkt t
$m(t + \Delta t)$	Anzahl der Funktionsfähigen BE zum Zeitpunkt $t + \Delta t$
M	Anfangsbestand der BE $m(t = 0) = N$

Überlebenswahrscheinlichkeit (Zuverlässigkeit)

$$R(t) = P(t > T) \quad (3)$$

Die Überlebenswahrscheinlichkeit stellt die Wahrscheinlichkeit dar, dass eine Betrachtungseinheit in einer bestimmten Betriebsdauer ohne Ausfall bleibt. Die Überlebenswahrscheinlichkeit beschreibt das Komplement der Ausfallwahrscheinlichkeit. (Strunz, 2012)

$$R(t) = 1 - F(t) \quad (4)$$

Ausfallrate

Die Ausfallrate beschreibt, wie wahrscheinlich ein Ausfall einer Anlage ist, wenn die Anlage bis zu einem bestimmten Zeitpunkt noch funktionsfähig ist (Ryll & Freund, 2010).

$$\lambda(t) = \frac{f(t)}{R(t)} \quad (5)$$

Die Ausfallrate lässt sich alternativ auch wie folgt darstellen (Strunz, 2012):

$$\lambda(t) = \frac{n(t) - n(t + \delta t)}{n * \delta t} \quad (6)$$

$n(t)$	Anzahl an intakten Betrachtungseinheiten zum Zeitpunkt t
$n(t + \delta t)$	Anzahl an intakten Betrachtungseinheiten zum Zeitpunkt $t + \delta t$
$n * \delta t$	Summe der noch intakten Einheiten zum Beginn des Intervalls

Ausfalldichte

Die Ausfalldichte gibt an, mit welcher Wahrscheinlichkeit eine Anlage in kleinen Zeitintervallen ausfallen. Sie lässt sich aus der Ausfallwahrscheinlichkeit ableiten. (Ryll & Freund, 2010)

$$f(t) = \frac{\delta F(t)}{\delta t} \quad (7)$$

Mittlere Lebensdauer

Die mittlere Lebensdauer beschreibt die erwartete Lebensdauer eines zufälligen Systems (Ryll & Freund, 2010).

Verfügbarkeit

Die Verfügbarkeit ist die „Fähigkeit eines Objektes, unter gegebenen Bedingungen und wenn erforderlich in einem Zustand zu sein, eine geforderte Funktion zu erfüllen, vorausgesetzt, dass die erforderlichen externen Ressourcen bereitgestellt sind“ (DIN 31051, 2019, S. 11). Sie lässt sich in theoretische (V_i) und technische Verfügbarkeit (V_t) unterteilen. Die technische Verfügbarkeit ist die theoretische Verfügbarkeit erweitert durch die geplanten Instandhaltungsmaßnahmen. (Pawellek, 2013) Die Verfügbarkeit lässt sich mathematisch wie folgt darstellen (Pawellek, 2013):

$$V_i = \frac{MTBF}{MTBF+MTTR} \quad (8)$$

$$V_t = \frac{MTBF}{MTBF+MTTR+MTTPM} \quad (9)$$

MTBF: Mean Time Between Failure
MTBR: Mean Time To Repair
MTTPM: Mean Time To Preventive Maintenance

2.2.2 Qualitative Zuverlässigkeitsmethoden

Nachfolgend werden die qualitativen Zuverlässigkeitsmethoden FMEA und FTA beschrieben.

FMEA

Bei der FMEA handelt es sich um ein Verfahren, bei dem alle möglichen Ausfallarten bestimmt werden und damit einhergehend auch alle Folgen und Ursachen des Ausfalls. Es handelt sich dabei um eine Risikoanalyse, die das Ziel verfolgt, ein Risiko in einem System frühzeitig zu erkennen und zu verbessern. (Bertsche & Lechner, 2004) Die FMEA läuft nach VDA FMEA-Formblatt in 5 Schritten ab. Zunächst muss eine hierarchische Struktur des Systems aus den Elementen erstellt werden. Darauf folgt die Beschreibung der Funktion. Als dritter Schritt erfolgt die Fehleranalyse einschließlich der möglichen Ursachen und Folgen. Infolgedessen geschieht die Risikobewertung und abschließend Optimierung des Systems zur Fehlervermeidung. (Bertsche, et al., 2009)

FTA

Die FTA ist eine Top-Bottom-Analyse, deren Ausgangspunkt ein Ausfall oder Schaden einer Betrachtungseinheit ist. Bei der FTA werden alle Ursachen oder Kombinationen von Ursachen eines Fehlers analysiert. (Bertsche, et al., 2009) Ziel ist es auch hier alle Ausfälle und deren Ursachen zu identifizieren sowie die Ausfallmechanismen zu dokumentieren. Beim Vorgehen wird zuerst ein unerwünschtes Ereignis betrachtet. Nachfolgend wird analysiert, welche Ausfälle auf der nächsten Ebene des Systems möglich sind und wie diese mit dem Systemausfall in Verbindung gebracht werden können. Dies wird so lange fortgesetzt, bis man an der untersten Systemebene angelangt ist. (Bertsche & Lechner, 2004)

2.3 Referenzmodelle

Das Ziel dieser Arbeit ist es mögliche Referenz- und Vorgehensmodelle zur Zuverlässigkeitsanalyse von Heizsystemen zu finden. Dazu muss zunächst definiert werden, was Referenzmodelle sind und wie sie angewendet werden können. Daher werden im nachfolgenden Abschnitt Definitionen für den Begriff Referenzmodell vorgestellt und ein Anwendungsbeispiel dargestellt.

Der Begriff Referenzmodell taucht in der Literatur erstmals Ende der 1980er Jahre für die Entwicklung von unternehmensspezifischen Modellen auf (Thomas, 2006). Bis heute gibt es keine allgemeingültige Definition des Begriffs. Es existieren viele Definitionen, die auch vom Anwendungsbereich abhängig sind. (Arora & Rabe, 2023a) Thomas (2006) beschreibt den Begriff Referenzmodell als ein konkreteres Informationsmodell, welches von seinen Nutzern als solches akzeptiert wird und zur Entwicklung von anderen Modellen dienen kann. Ähnlich definieren Rosemann & Schütte (1997) anhand des Kriteriums der Individualität, dass Referenz-Informationsmodelle eine Spezialisierung eines Informationsmodells sind. Rabe et al. (2006) definieren Referenzmodelle als einen konzeptionellen Rahmen, welcher das Erstellen von Modellen verbessern soll, indem die Effizienz bei der Erstellung gesteigert sowie die Qualität erhöht und die Wiederverwendbarkeit verbessert wird.

Van der Aalst (2006) verknüpft die Anwendung von Referenzmodellen zudem mit dem Paradigma *Design by Reuse*. Als Ziel von Referenzmodellen nennt er die Rationalisierung des Entwurfs spezieller Modelle durch das Bereitstellen von allgemeinen Lösungen, sodass der Modellierungs- und Konfigurationsprozess effizienter und schneller gestaltet werden kann. (van der Aalst, et al., 2006)

Trotz der Tatsache, dass es keine allgemeingültig festgelegte Definition des Begriffs Referenzmodell gibt, lassen sich dennoch Gemeinsamkeiten und bestimmte Charakteristiken feststellen. Arora et al. (2023a) stellt eine beispielhafte allgemeine Darstellung für die Anwendung eines Referenzmodells vor (s. Abbildung 4). Zu Beginn müssen parallel zum einen

die für das spezielle Modell benötigten Elemente aus dem Referenzmodell ausgewählt werden und zum anderen müssen die Elemente identifiziert werden, welche für das spezielle Modell benötigt werden. Beide Schritte laufen dabei iterativ ab. Die aus dem Referenzmodell gewählten Elemente werden anschließend in das zu erstellende Modell eingesetzt. Dabei kann es vorkommen, dass nicht alle identifizierten Elemente des zu erstellenden Modells vom Referenzmodell abgedeckt werden. Die fehlenden Elemente müssen folglich neu entwickelt werden. (Arora & Rabe, 2023a) Des Weiteren haben Arora & Rabe (2023a) bei einer Untersuchung von 17 relevanten Beiträgen 9 Eigenschaften mit insgesamt 41 Attributen herausgestellt. Bei den Eigenschaften für Referenzmodelle handelt es sich um *Wiederverwendbarkeit*, *Flexibilität*, *Zuverlässigkeit*, *Allgemeingültigkeit*, *Erforderlichkeit*, *Benutzerzentriertheit*, dass sie *umfangreich*, *systematisch* designt und *bildend* sind. Während die Charakteristiken der *Wiederverwendbarkeit*, *Allgemeingültigkeit*, *Benutzerzentriertheit*, *Bildung*, *Flexibilität* und *Umfangreiche* von einer Mehrheit der betrachteten Beiträge aufgenommen wurden, finden die Eigenschaften der *Zuverlässigkeit*, *Systematik* und *Erforderlichkeit* nur bei einer Minderheit Zustimmung. (Arora & Rabe, 2023a)

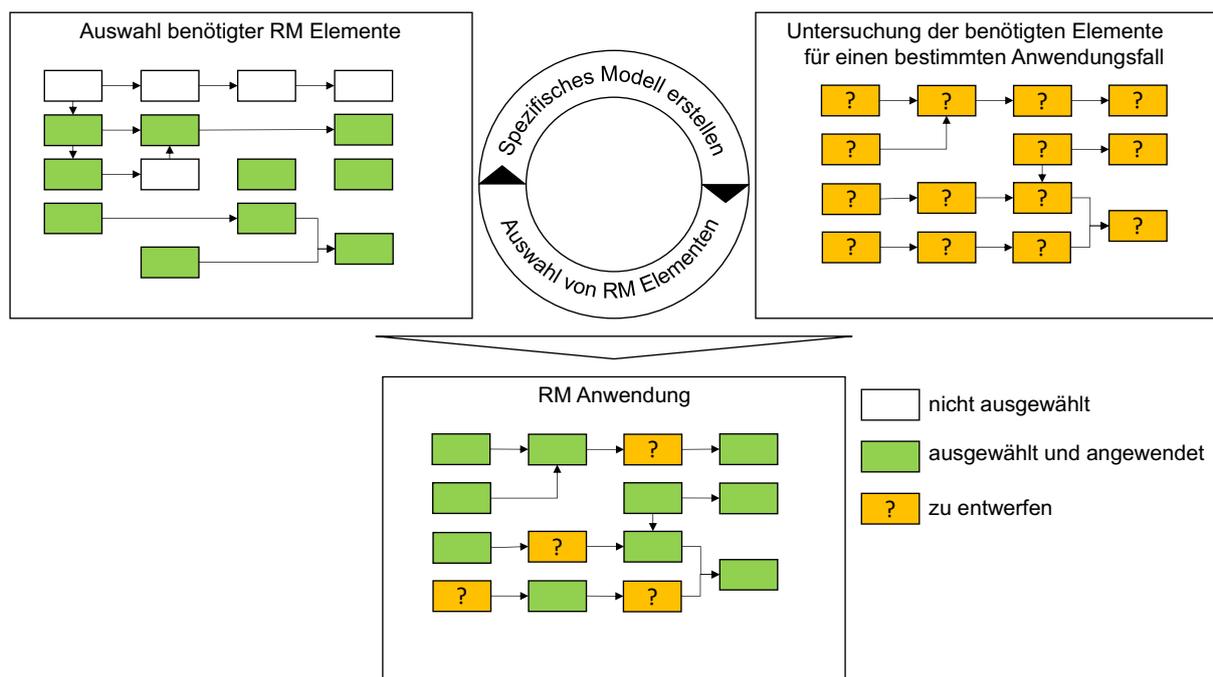


Abbildung 4: Anwendungsbeispiel für ein Referenzmodell (vgl. Arora & Rabe, 2023a, S. 3)

2.4 Wissensentdeckung in Datenbanken

Im folgenden Abschnitt werden zwei Vorgehensmodell für die Wissensentdeckung in Datenbanken beschrieben. Zum einen wird das *Knowledge Discovery in Databases (KDD)* - Vorgehensmodell nach Fayyad et al. (1996) vorgestellt und zum anderen das *Cross Industry Standard Process for Data Mining (CRISP-DM)* -Modell (Chapman, et al., 2000).

2.4.1 Vorgehensmodell Knowledge Discovery in Databases (KDD)

Das KDD Vorgehensmodell nach Fayyad et al. (1996), beschreibt einen Prozess zur Wissensentdeckung in großen Datensätzen. Ziel des Prozesses ist es neue Informationen, Muster und Zusammenhänge bzw. allgemein gesehen Wissen aus den betrachteten Datenmengen zu generieren. Der KDD-Prozess lässt sich in 5 allgemeine Schritte gliedern (s. Abbildung 5). (Fayyad, et al., 1996)

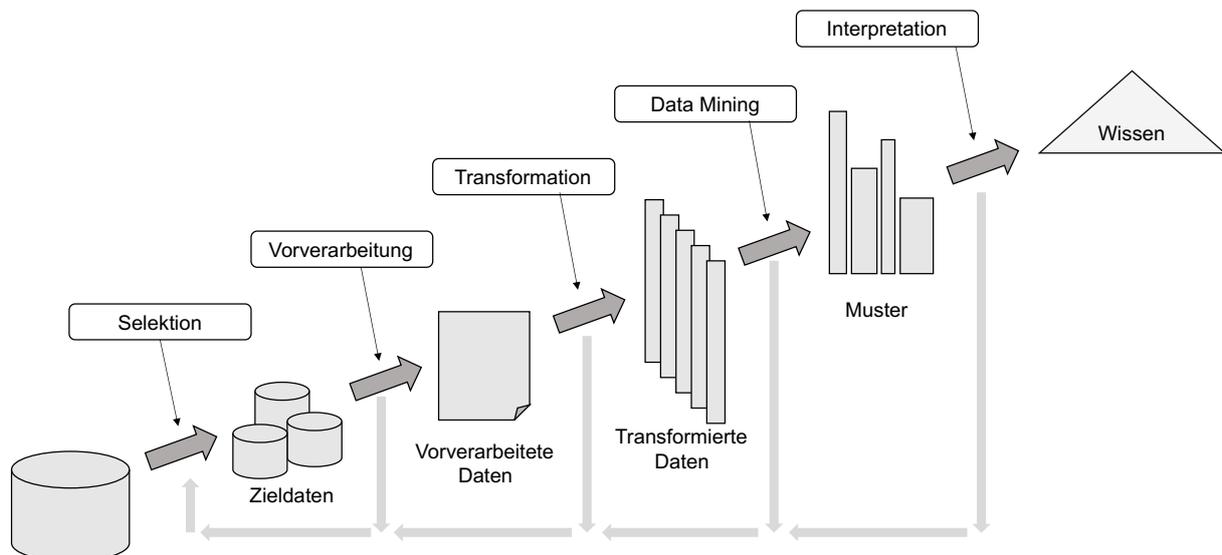


Abbildung 5: KDD-Prozess (vgl. Fayyad, et al., 1996, S. 41)

Die erste Phase des KDD-Prozesses nach Fayyad et al. (1996) ist die Datenselektion. Dazu muss der Analyst sich zuvor das gewünschte Domänenwissen aneignen und das Ziel des KDD-Prozesses aus Sicht des Anwenders definieren (Fayyad, et al., 1996). In dieser Phase werden je nach Ziel und vorhandenem Wissen in der jeweiligen Domäne die erforderlichen Daten für den weiteren KDD-Prozess ausgewählt (Sharafi, 2013). Neben den vorhandenen Daten des Anwenders können auch externe Daten gewählt werden, solange sie dem Anwender für das Ziel nützlich erscheinen (Cleve & Lämmel, 2020). Für das weitere Vorgehen kann es zudem nötig sein, die Daten in ein passendes Format zu bringen (Sharafi, 2013). Dabei kann es aus technischen und rechtlichen Gründen dazu kommen, dass die Daten nicht konvertierbar sind und die Selektion aus diesem Grund erneut durchgeführt werden

muss. Die Gründe können dabei Kapazitäts- oder Datentyp-Beschränkung sowie Beschränkungen des Zugriffsrechts sein. Eine Lösung dafür ist es, sich nur auf repräsentative Teilmengen von Daten zu konzentrieren. (Cleve & Lämmel, 2020)

In der zweiten Phase des KDD-Prozesses geht es um die Datenvorverarbeitung. Die vom Anwender selektierten Daten in Ihrer Form sind oft nicht für Data-Mining-Analysen geeignet sind. (Cleve & Lämmel, 2020) Daher werden in diesem Schritt das Rauschen sowie die Fehler der selektierten Daten bereinigt, da es das Ergebnis der Analyse sonst beeinflussen kann (Fayyad, et al., 1996). Dies geschieht mittels einer Fehleranalyse und unter Verwendung verschiedener Methoden. (Sharafi, 2013) Bei den Fehlern unterscheidet man verschiedene Arten. Zunächst gibt es die fehlenden Werte. Methoden zur Bereinigung dieser Fehlerart sind beispielsweise unter anderem das manuelle Einfügen der Werte, das Erstellen einer globalen Konstante oder das Bilden des Durchschnittswertes. Eine weitere Fehlerart sind Ausreißer, also Werte die stark vom restlichem Wertenniveau abweichen. Ein zusätzliches Ziel der Datenvorverarbeitung ist die Reduktion des Wertebereichs. Dafür können Werte eines Attributes verallgemeinert oder Wertebereiche zusammengefasst werden. (Cleve & Lämmel, 2020)

Im Anschluss an die Datenvorverarbeitung geschieht die Datentransformation in ein für das Data-Mining-System passendes Datenbankschema. Hier werden die Daten so angepasst, dass das Data-Mining-System sie verarbeiten kann. (Cleve & Lämmel, 2020) Zudem werden auf Grundlage der Ziele Attribute und Merkmale bestimmt, an die die Daten angepasst werden müssen (Sharafi, 2013).

Das Data-Mining stellt die vierte Phase des KDD-Prozesses nach Fayyad et al. (1996) dar. Ziel des Data-Minings ist es hauptsächlich Muster in den Daten zu finden und in einer anschaulichen Form darzustellen. Dazu zählen die Klassifizierung, die Regression oder das Clustern. (Fayyad, et al., 1996) Des Weiteren kann als Ziel angesehen werden, vorher aufgestellte Hypothesen zu prüfen (Sharafi, 2013). Als erster Schritt des Data-Mining muss anhand der Aufgabe (z. B. Klassifizierung) ein passendes Verfahren ausgewählt werden. Anschließend muss das Verfahren Anwendungs- bzw. Zielbezogen konfiguriert werden. Abschließend beginnt die Suche nach Mustern und es wird ein Modell als Grundlage für die Bewertung geschaffen. (Cleve & Lämmel, 2020)

Die letzte Phase ist die Interpretation und Evaluation der Muster, die beim Data-Mining-Prozess gefunden wurden, sowie die Visualisierung dieser (Fayyad, et al., 1996). Auch die Prüfung nach Gültigkeit, Neuartigkeit, Nützlichkeit und Verständlichkeit wird in diesem Schritt vorgenommen, da häufig viele gefundene Muster für den Anwender bedeutungslos sind (Cleve & Lämmel, 2020). Zudem steht die Verwendung der gewonnen Daten an. Entweder werden die Daten direkt genutzt oder zunächst dokumentiert und an andere weitergegeben. (Sharafi, 2013)

2.4.2 Cross Industry Standard Process for Data Mining (CRISP-DM)

Das CRISP-DM Modell beschreibt ein Vorgehensmodell zum Data-Mining, welches ein hierarchisches Referenzmodell und Anleitung der Aufgaben beinhaltet. Es wurde in Zusammenarbeit von verschiedenen Unternehmen (NVR Corporation, Daimler AG, SPSS und OHRA) entwickelt. (Chapman, et al., 2000)

Das CRISP-DM-Modell lässt sich in vier hierarchische Ebenen (s. Abbildung 6) aufgliedern, die sich von Ebene zu Ebene spezifizieren. Die erste Ebene besteht aus einem Phasenmodell mit sechs Phasen (s. Abbildung 7). Die zweite Ebene bezieht sich auf alle generischen Aufgaben. Sie sollen so allgemein wie möglich gehalten sein, um alle Data Mining Situationen abzudecken. Die dritte Ebene stellt die spezifischen Aufgaben dar. Hier wird beschrieben, wie mit Situationen aus der zweiten Ebene umgegangen werden soll. Auf der vierten und letzten Ebene befindet sich die Prozessinstanz. Auf dieser Ebene werden alle Handlungen, Entscheidungen und Resultate dokumentiert. Die Prozessinstanz muss sich nach den höheren Ebenen richten und stellt konkrete Handlungsschritte dar. (Chapman, et al., 2000)

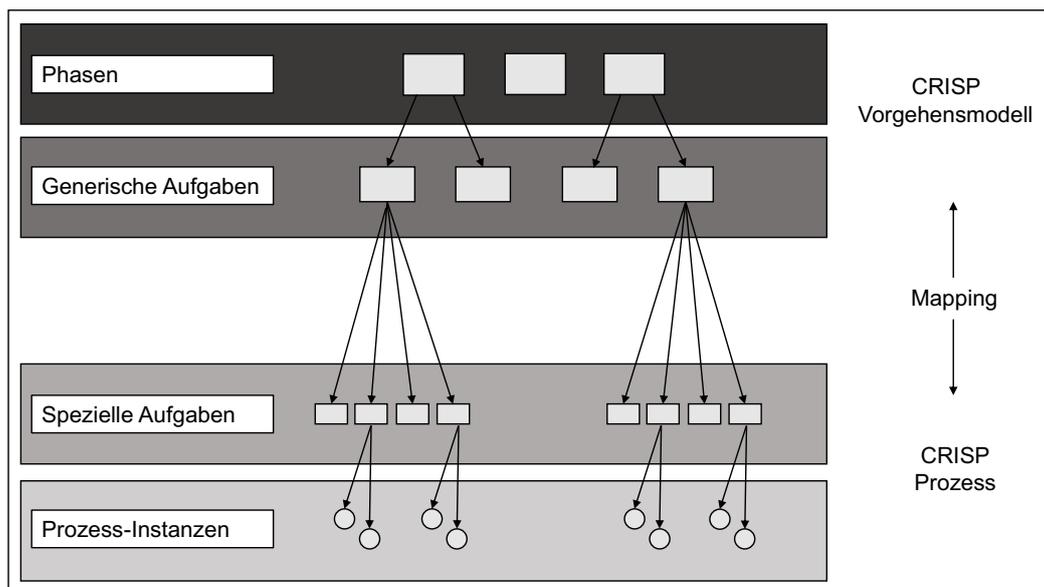


Abbildung 6: Ebenen des CRISP-DM Modells (vgl. Chapman, et al., 2000, S. 9)

Im Folgenden werden die sechs Schritte (Geschäftsverständnis, Datenverständnis, Datenverarbeitung, Modellierung, Bewertung und Einsatz) des Phasenmodell der ersten Ebene beschrieben (Chapman, et al., 2000):

Geschäftsverständnis

In der Phase *Geschäftsverständnis* sollen die Ziele und Anforderungen aus wirtschaftlicher Sicht verstanden werden. Dieses Verständnis wird dann in eine Problemstellung für das Data-Mining übertragen und es wird ein Plan entwickelt, wie die Ziele erreicht werden können.

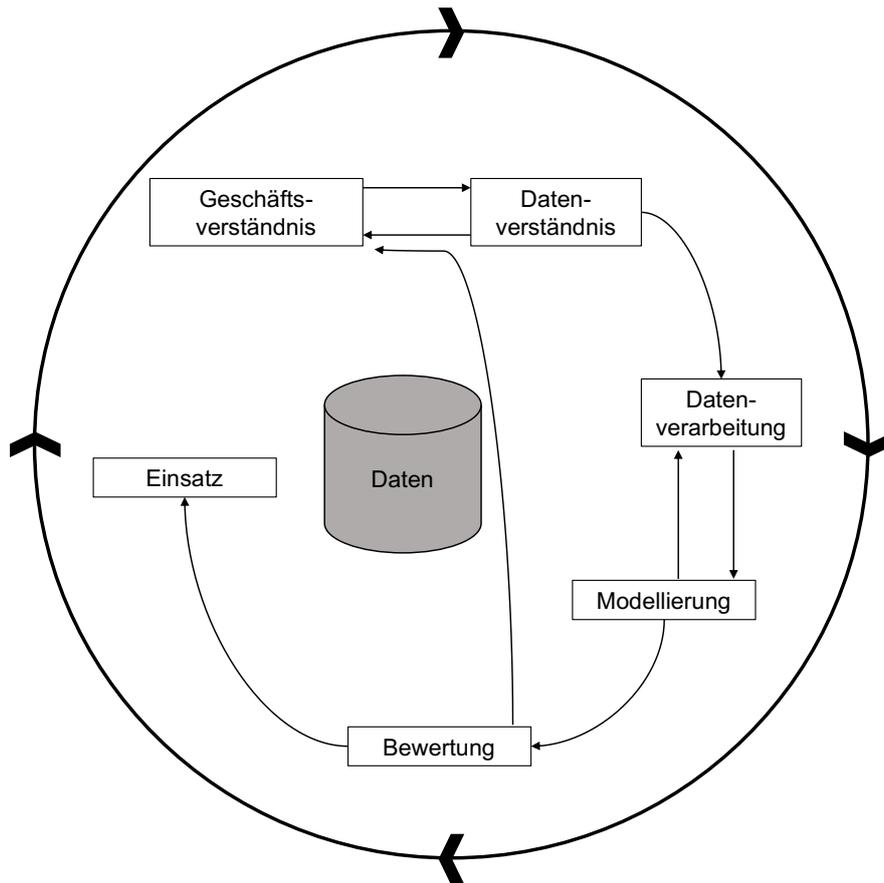


Abbildung 7: Phasen des CRISP-DM-Modells (vgl. Chapman, et al., 2000, S. 13)

Datenverständnis

Die zweite Phase *Datenverständnis* befasst sich konkret mit den Daten. Hier werden zunächst Daten gesammelt und es wird sich mit diesen vertraut gemacht. Dabei werden die Qualität der Daten betrachtet sowie erste Erkenntnisse gesammelt, die zu Erstellung von Hypothesen führen.

Datenverarbeitung

Im dritten Schritt werden die Daten vorbereitet, sodass sie in das Data-Mining Modell passen. Konkret werden in der Phase Datensätze und Attribute gewählt und die Daten bereinigt und transformiert.

Modellierung

In der *Modellierungs*-Phase werden verschiedene Modellierungstechniken ausgewählt und kalibriert. Da einige Techniken bestimmte Anforderungen haben, kann es nötig sein, zurück zum vorherigen Schritt zu gehen.

Bewertung

In der Phase der *Bewertung* wird das vorangegangene Vorgehen und die Qualität des Modells bewertet. Im konkreten wird betrachtet, ob das Vorgehen die Ziele aus der ersten Phase erreicht hat und ob sich weitere wichtige Fragestellungen ergeben haben. Zum Ende der Phase wird über den Gebrauch der gewonnenen Daten entschieden.

Einsatz

In der letzten Phase wird das gewonnene Wissen als Output zurückgegeben. Das Wissen wird in diesem Schritt folglich in einer beliebigen Form visualisiert und dargestellt und kann so für Entscheidungen genutzt werden. Je nachdem wofür die Daten genutzt werden sollen, kann die Phase einfach mit einem Bericht oder komplex mit einem wiederholbaren Data-Mining Prozess gestaltet werden.

2.5 Heizungssysteme

Im letzten Abschnitt des zweiten Kapitels wird die grundlegende Funktionsweise von Heizungssystemen erläutert.

Bei Heizungsanlagen kann man grundsätzlich zentrale und dezentrale Heizsysteme unterscheiden. Dezentrale Heizungsanlagen oder auch Einzelheizungen genannt werden Heizungen genannt, die die erforderliche Wärme zur Beheizung eines Raumes im selben Raum erzeugt. Mit dezentralen Heizungen können hohe Wirkungsgrade erzielt werden, da Wärmeverluste durch gesonderte Heizräume und Rohrleitungen entfallen. Außerdem kann der Betrieb einfach auf den gewünschten Raum beschränkt werden. Dementsprechend sind dezentrale Heizungen sparsam. (Laasch & Laasch, 2013) Als Einzelheizgeräte finden entweder Öfen, in denen Brennstoffe, Öl oder Gas verbrannt werden, oder elektrische Heizleiter Anwendung (Hainbach, 2018). Bei zentralen Heizungsanlagen hingegen können mehrere Räume mittels eines zentralen Wärmeerzeugers beheizt werden. Zentralheizungen beziehen ihren Namen nach den Wärmeträgern. (Laasch & Laasch, 2013) Als Gemeinsamkeit haben alle zentralen Heizungsanlagen den allgemeinen Aufbau bestehend aus einem Wärmeerzeuger, den Rohrleitungen und den Heizkörpern (Hainbach, 2018). Als Wärmeerzeuger können verschiedene Technologien verwendet werden. Dazu zählen Heizkessel, Wärmepumpen, Blockheizkraftwerke, Nah-/Fernwärme oder solarthermische Anlagen als Kombination mit anderen Technologien. (Lauckner & Krimmling, 2020)

Durch den Wärmeerzeuger wird ein Wärmeträger erhitzt und über die Rohrleitung zu den Heizkörpern transportiert. Die Heizkörper geben die Wärme in den Raum ab. Als Wärmeträger können Wasser, Luft oder Dampf zum Einsatz kommen (Lauckner & Krimmling, 2020). Als häufigster Energieträger für die Beheizung von Wohngebäuden in Deutschland kommt

Erdgas vor (BDWE, 2019). Gasheizkessel haben einen hohen Wirkungsgrad, relativ geringe Emissionen und eine komfortable Brennstoffzuführung (Arora & Rabe, 2023b).

Heizungsanlagen können zudem anhand von Heizkreisen strukturiert werden (s. Abbildung 8), wobei die Anzahl an Heizkreisen pro Anlage variieren kann. In Einfamilienhäusern kommen häufig drei Heizkreise zum Einsatz. Einer für die Heizkörper, einer für eine Fußbodenheizung und einer für die Warmwasserbereitung. (Lauckner & Krimmling, 2020)

Zentrale Heizungsanlagen haben einige Vorteile. Sie erwärmen ein gesamtes Gebäude mit allen Räumen gleichermaßen und stellen somit sicher, dass alle Räume über das ganze Jahr genutzt werden können. Zudem benötigen sowohl der Wärmeerzeuger als auch die Heizkörper wenig Platz. Des Weiteren erlangen die zentralen Heizungen eine gute Brennstoffausnutzung. Allerdings sind die Anschaffungskosten und Betriebskosten größer als bei Einzelheizungen. (Laasch & Laasch, 2013)

Am häufigsten finden heutzutage sogenannte Wassersystem mit Heizwasser als Wärmeträger Anwendung. Diese setzen sich aus einem Wärmeerzeuger, einer Umwälzpumpe, Raumregleinrichtungen, Heizkörpern, einem Ausdehnungsgefäß sowie Vor- und Rücklaufleitungen zusammen. Das Heizwasser wird durch den Wärmeerzeuger erwärmt. Durch eine Umwälzpumpe zirkuliert das Heizwasser kontinuierlich durch das System und das erwärmte Heizwasser gelangt zu den Heizkörpern. Die Wärme wird durch die Heizkörper schließlich an den Raum abgegeben. (Lauckner & Krimmling, 2020)

Der energieeffizienteste Wärmeerzeuger ist die elektrische Wärmepumpe (BDWE, 2019). Die Wärmepumpe ist ein Gerät, das den technischen Betrieb und die Umweltwärme aus einer Wärmequelle mit Hilfe eines Carnot-Prozesses kombiniert, um die Temperatur einer Wärmesenke zu erhöhen. Die Wärmepumpen sind umweltfreundlich, da sie drei Viertel der Heizenergie aus der Umwelt beziehen. Dabei sind die häufigsten Wärmequellen die Luft, das Erdreich und Grundwasser. Um diese Wärme folglich zu nutzen, wird nur wenig Strom benötigt. Die Technik, Anforderungen und Kosten unterscheidet sich je nach Wärmequelle. (Arora & Rabe, 2023b) Eine Wärmepumpen-Heizungsanlage besteht aus drei Komponenten: die Wärmequellenanlage, welche die Umweltwärme aufnimmt; die Wärmepumpe, welche die Umweltwärme nutzbar macht; und dem Wärmeverteilsystem, welches die Wärme verteilt oder speichert (BWP, 2022).

Im vorangegangenen Kapitel wurden alle wesentlichen Grundlagen zur Zuverlässigkeitsprognose von Heizsystemen mittels Datenanalyse erläutert sodass im Folgenden das Anforderungsprofil erstellt und anschließend die Literaturanalyse durchgeführt werden kann.

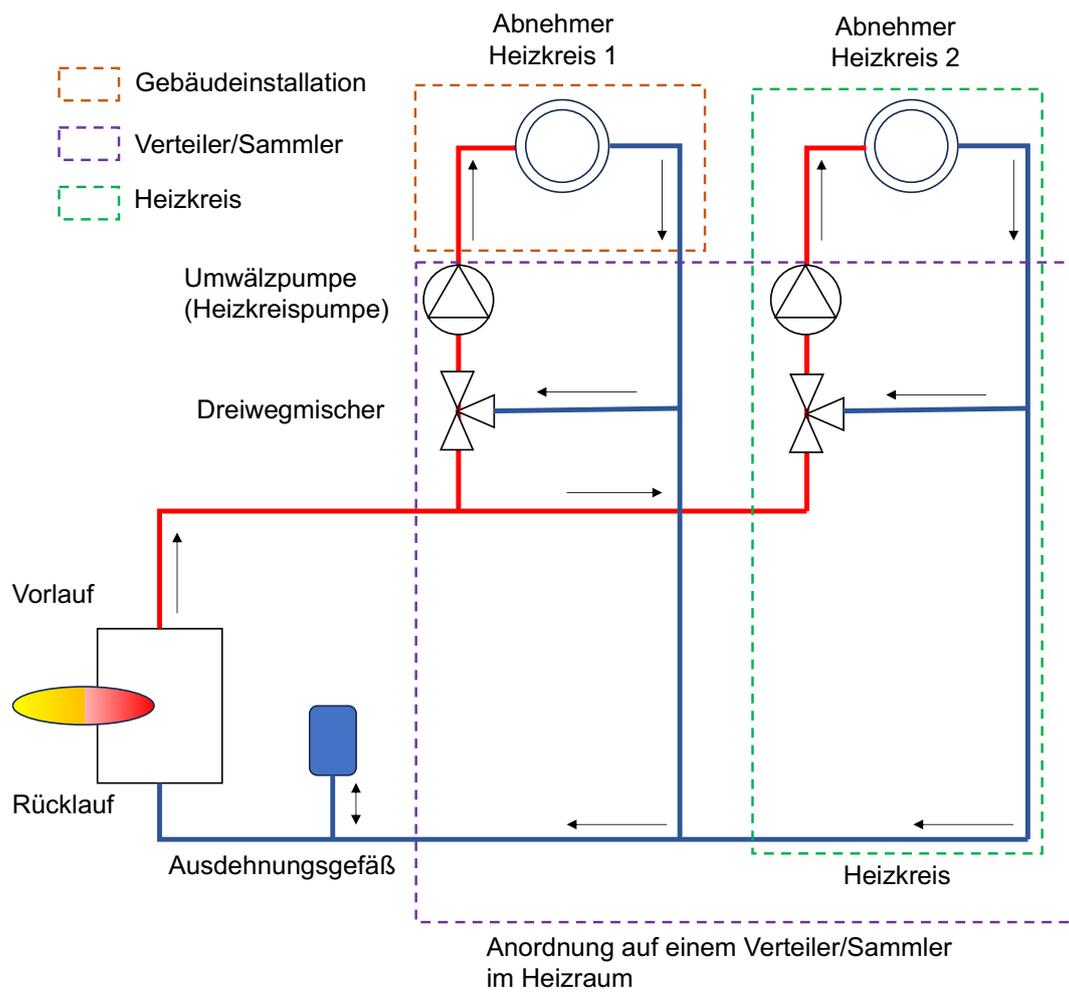


Abbildung 8: Beispielhafte Heizungsanlage mit zwei Heizkreisen (vgl. Lauckner & Krimmling, 2020, S. 99)

3 Anforderungsprofil zur Zuverlässigkeitsprognose von Heizungsanlagen

Im nachfolgenden Kapitel wird ein Anforderungsprofil für ein Modell zur Zuverlässigkeitsprognose von Heizsystem mittels Datenanalyse vorgestellt. Die Anforderungen werden in drei Bereiche unterteilt: allgemeine Anforderung an das Modell, Anforderungen an die Datenverarbeitung und -analyse und Anforderungen aus dem Heizsektor an das Zuverlässigkeitsprognosemodell.

Tabelle 2: Anforderungsprofil für das Zuverlässigkeitsprognose-Modell

Anforderungsprofil für ein Referenz- und Vorgehensmodell zur Zuverlässigkeitsprognose von Heizsystemen
Allgemeine Anforderungen
1. Modell ist wiederverwendbar, flexibel und allgemeingültig
2. Modell ist benutzerzentriert und lehrreich
3. Modell beinhaltet verschiedene Lösungselemente
4. Festlegung des Zwecks des Modells
Datenbezogene Anforderungen
5. Risikobewertung des Datenanalysemodells
6. Erfassen, Speichern und Verarbeiten von Daten
7. Test und Validierung des Datenanalysemodells
8. Beobachtung und Verbesserung/Anpassung des Datenanalysemodells
Heizungssektor bezogene Anforderungen
9. Einbindung aller Akteure des Heizsektors in die Zuverlässigkeitsprognose
10. Durchgängige Informationsbereitstellung über die Nutzungsphase der Heizsysteme

Das Ziel der vorliegenden Arbeit ist es, Ansätze für Vorgehens- und Referenzmodelle für die Zuverlässigkeitsprognose von Heizsystemen zu suchen. Aus diesem Grund müssen die gesuchten Ansätze die Anforderungen und Eigenschaften von Referenzmodellen erfüllen. Diese Anforderungen werden im Folgenden als *allgemeine Anforderungen* beschrieben. In einer Literaturanalyse haben Arora und Rabe (2023a) Merkmale und Eigenschaften von Referenzmodellen herausgestellt. Insgesamt ergeben sich aus der besagten Literaturanalyse 41 Merkmale, die in neun Eigenschaften eingeteilt sind. Die herausgestellten Eigenschaften sind: *wiederverwendbar*, *allgemeingültig*, *benutzerzentriert*, *lehrreich*, *flexibel*, *umfassend*, *zuverlässig*, *systematisch* und *notwendig*. Die Eigenschaften *wiederverwendbar*, *flexibel* und *allgemeingültig* haben in den Merkmalen Übereinstimmungen, weswegen sich diese Eigen-

schaften gegenseitig bestärken (Anforderung 1). Dies ist z. B. bei dem Merkmal *anpassbar* der Fall. Die Merkmale *zuverlässig*, *systematisch* und *notwendig* sind in der besagten Literaturanalyse in einer Minderheit (< 50%) der untersuchten Literaturen erwähnt worden und gelten daher als nicht allgemeingültig. Ein Modell ist dann *benutzerzentriert*, wenn es unter anderem einfach zu nutzen ist, gut dokumentiert ist oder Anwendungsbeispiele beinhaltet. Als *lehrreich* kann ein Modell beschrieben werden, wenn es z.B. Wissen oder Empfehlungen für die Anwendung beinhaltet. (Anforderung 2) Die vorgestellte Architektur von Referenzmodellen geht mit der Eigenschaft *umfassend* einher. Referenzmodelle bestehen unter anderem aus einer Sammlung von verschiedenen Leitfäden, Vorgehen, Methoden und Einheiten, die je nach Anwendungsfall gewählt werden. (Arora & Rabe, 2023a) Daher ist eine weitere Anforderung, dass die gefundenen Ansätze ebenfalls verschiedene Methoden, Vorgehensweisen, Leitfäden oder Einheiten beinhalten (Anforderung 3).

Eine weitere allgemeine Anforderung ergibt sich aus Schnieder (2023). Schnieder (2023) nennt als erste Phase der Entwicklung einer KI (Künstlichen Intelligenz) die Festlegung des genauen Zwecks der Verwendung (Anforderung 4).

Um das Versagen von Heizungsanlagen vorherzusagen, müssen eine Reihe an Daten aus dem Betrieb der Anlage analysiert werden. Um die Vielzahl an Daten zu analysieren, können z.B. sogenannte Machine-Learning (ML) Modelle eingesetzt werden. Bei ML-Modellen handelt es sich um Künstliche Intelligenz. Schnieder (2023) beschreibt auf Grundlage des Dokuments „*Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetze über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union (2021)*“ (Europäische Kommission, 2021) den Lebenszyklus inklusive der Entwicklungsphasen von KI-Systemen und somit vorgeschriebene Phase zur sicheren Entwicklung von KI-Systemen. (Schnieder, 2023) Anhand des Lebenszyklus lassen sich aus diesem Grund Anforderungen an ein Modell zur Zuverlässigkeitsprognose von Heizsystemen ableiten. Diese können sowohl für die Datenanalyse mit ML-Modelle angewendet werden als auch für generelle Datenanalyse.

Es ist notwendig eine Risikoermittlung und -analyse von bekannten und vorhersehbaren Risiken durchzuführen. Risiken, die bei zweckbestimmter Anwendung oder auch vorhersehbarer Fehlanwendung auftreten können, müssen abgeschätzt werden. Zudem werden Risiken aus Felderfahrungen bewertet. (Schnieder, 2023) Als Anforderung für das Datenanalyse-Modell ergibt sich folglich die Durchführung einer Risikobewertung (Anforderung 5). Mit Hilfe der Risikobewertung können im Folgenden die Systemanforderungen festgelegt werden (Schnieder, 2023).

Nachfolgend beschreibt Schnieder (2023) das Erfassen von Daten aus dem Zielsystem, in diesem Fall der Heizungsanlage, zum Aufstellen des KI-Systems. Das Erfassen, Speichern

und anschließende (Vor-)Verarbeiten von Daten stellt dementsprechend eine weitere Anforderung an das Datenanalyse-Modell dar (Anforderung 6).

Im nächsten Schritt wird das KI-System mit den zuvor gewonnenen Daten trainiert (Schnieder, 2023). Da das Trainieren sich auf KI-Systeme bzw. ML-Modelle bezieht und nicht allgemein für Datenanalyse-Modelle gilt, stellt dies keine Anforderung an das Modell zur Zuverlässigkeitsprognose dar.

Anschließend folgen das Testen und Validieren des Datenanalyse-Modells (Anforderung 7). Das entwickelte KI-System kann in der Folge in Betrieb genommen werden. (Schnieder, 2023)

Nachdem das Datenanalyse-Modell zur Verwendung freigegeben wurde, muss es weiter beobachtet werden. Dies hat den Zweck das Modell zu aktualisieren bzw. zu korrigieren (Anforderung 8). Bei KI-Systemen erfolgt dies dadurch, dass Bewertungsergebnisse aus dem KI-System zusammen mit den Quelldaten an den Hersteller übermittelt werden und der Hersteller die vorliegenden Daten auswertet und das System auf der Grundlage anpasst. (Schnieder, 2023)

Weitere Anforderungen an das Modell zur Zuverlässigkeitsprognose ergeben sich aus dem Bereich des Heizsektors sowie der Instandhaltung von Heizungsanlagen. Arora und Rabe (2023b) stellen unter anderem den Aufbau des Heizsektors dar sowie die Voraussetzungen und Probleme zur Implementierung von Lösungsansätzen der vorausschauenden Instandhaltung in diesem Bereich.

Der Heizsektor besteht zunächst aus den Lieferanten, die Elemente oder Subsysteme von Heizungen herstellen. Die Elemente werden anschließend an den Heizungshersteller geliefert. Der Hersteller baut folglich die Heizungen zusammen. Nachdem die Heizungen gebaut wurden, können sie verkauft werden. Manchmal besitzen die Heizungshersteller ihren eigenen Vertrieb. Meistens wird der Vertrieb, die Installation und Wartung allerdings von Serviceunternehmen durchgeführt. Diese Serviceunternehmen vertreiben häufig nicht nur die Heizungsanlagen eines Herstellers. In der Regel werden defekte Bauteile bei akuten Fehlern oder bei festgelegten Service-Terminen erneuert. Dies entspricht einem Mix aus reaktiver und proaktiv-zeitbezogener Instandhaltung. (Arora & Rabe, 2023b)

Der Aufbau des Heizsektors hat zur Folge, dass die informationsbezogene Distanz zwischen den Endkunden und den Herstellern groß ist. Der Hersteller hat keinen Einblick in den Betrieb der Anlagen und hat somit wenig Möglichkeiten die Heizungsanlagen zu optimieren. (Arora & Rabe, 2023b)

Um in Zukunft Lösungen zur vorausschauenden Instandhaltung integrieren zu können und die Zuverlässigkeit von Heizungen zu steigern, müssen einige Probleme gelöst werden. Zunächst benötigen die Hersteller und ggf. auch die Lieferanten, mehr Feedback von den Ser-

viceunternehmen und den Endkunden aus dem Betrieb der Anlagen, um die Zuverlässigkeit der Anlage weiter steigern zu können. (Anforderung 9)

Auch wenn bereits internetfähige Heizsysteme verkauft werden, müssen die Datenbereitstellung und -übermittlung in separaten Verträgen mit den Endkunden vereinbart werden. Das führt zu einer unstrukturierten und unvollständigen Datenbasis bezüglich Gebrauchs-, Fehler- und Systemdaten. (Arora & Rabe, 2023b) Für das Modell zur Zuverlässigkeitsprognose gilt: Je besser die dauerhafte Datenübermittlung der Heizsysteme, desto bessere Zuverlässigkeitsprognose-Modelle sind möglich. Dies hat den Grund, dass erst durch aktuelle und vor allem vergangene Zustands-, Fehler- und Ausfalldaten sowie Wartungs- und Reparaturinformationen, erneute Fehler bzw. Ausfälle vorausgesagt werden können. Als Anforderung ergibt sich folglich die durchgängige Bereitstellung von qualitativ und quantitativ hochwertigen Daten und Informationen aus den Heizsystemen (Anforderung 10).

4 Methodik der systematischen Literaturrecherche

Im folgenden Kapitel wird die Methode der vorliegenden Arbeit erläutert. Dazu werden im ersten Abschnitt drei Konzepte einer systematischen Literaturrecherche vorgestellt. Anschließend wird anhand der vorgestellten Konzepte die Vorgehensweise der Literaturrecherche in dieser Arbeit geschildert und die Durchführung beschrieben.

4.1 Konzepte der systematischen Literaturrecherche

Es gibt eine Reihe an vorgeschlagenen Vorgehensweisen bei einer systematischen Literaturrecherche. Im Folgenden werden die drei Methoden nach Webster und Watson (2002), Kitchenham et al. (2004) und vom Brocke (2009) vorgestellt werden.

Nach dem Vorgehen von Webster und Watson (2002) ist der erste Schritt das Identifizieren von relevanter Literatur für das zu untersuchende Forschungsthema. Dabei soll das Thema bei der Suche möglichst vollumfänglich betrachtet werden. Zur Suche werden Suchbegriffe angewendet, sodass die Literatur auf das Forschungsthema limitiert wird. Um weitere Literaturen zu finden, beschreiben Webster und Watson (2002) das Durchführen einer Rückwärts-Suche sowie einer Vorwärts-Suche. Bei der Rückwärts-Suche werden die Zitationen aus den zuvor gefunden Artikeln untersucht und relevante Artikel mit einbezogen. Bei der Vorwärts-Suche werden Artikel gesucht, die den entsprechenden Artikel zitieren. Im nächsten Schritt des Vorgehens wird die gefundene Literatur strukturiert. Dazu schlagen Webster und Watson eine Konzeptmatrix vor, in der die Literatur sortiert wird (s. Abbildung 9). Als nächster Schritt wird die Ausarbeitung einer Forschungslücke als Ziel einer systematische Literaturanalyse genannt sowie das Aufstellen von Themen für die zukünftige Forschung. Als letzter Schritt nennen Webster und Watson die Evaluation der vorher erstellten Theorie. Dazu geben Kollegen eine Bewertung ab. (Webster & Watson, 2002)

Konzept Matrix					
Artikel	Konzept				
	A	B	C	D	...
1		x	x		x
2	x	x			
...			x	x	

Abbildung 9: Literatur-Konzeptmatrix (vgl. Webster & Watson, 2002, S. 17)

Kitchenham (2004) beschreibt ein Verfahren mit fünf Schritten zur Durchführung einer systematischen Literaturrecherche. Als ersten Schritt nennt Kitchenham (2004) das Ermitteln einer Suchstrategie sowie die anschließende Suche. Auf Grundlage des Forschungsthemas sollen Schlüsselwörter und Suchbegriffe sowie Synonyme oder ähnliche Begriffe aufgelistet werden. Dabei soll das Forschungsthema zudem in einzelne Bereiche gegliedert werden und die Begriffe in die verschiedenen Bereiche eingeordnet werden. Mit Suchoperatoren können auf Datenbanken die Schlüsselwörter bei der Suche zu einer Suchkette verknüpft werden. Als nächsten Schritt sieht Kitchenham (2004) das Festlegen von Ein- und Ausschlusskriterien vor. Dabei handelt es sich z. B. um einen zeitlichen Rahmen oder bestimmte Sprachen. Anschließend erfolgt eine Qualitätskontrolle anhand einer ausgewählten Bewertungsmethode. Der vorletzte Schritt des Vorgehens ist die Datenerfassung, bei der Informationen aus der Literatur gesammelt werden. Im letzten Schritt werden die erfassten Informationen analysiert und dargestellt. (Kitchenham, 2004)

Vom Brocke et al. (2009) beschreiben ebenfalls fünf Phasen einer systematischen Literaturanalyse. In der ersten Phase wird die Forschungsfrage und Forschungsumfang definiert. Dafür schlagen vom Brocke et al. (2009) z. B. die Taxonomie nach Cooper vor. In der zweiten Phase des Vorgehens wird ein Konzept des Themas bzw. der Forschungsfrage erstellt. Hier werden relevante Schlüsselwörter, Datenbanken sowie Ein- und Ausschlusskriterien festgelegt. In der dritten Phase findet die Literaturrecherche statt. Phase vier des Vorgehens beinhaltet die Literaturanalyse und -synthese der gefundenen Literatur. Abschließend werden in Phase 5 die Ergebnisse zusammengefasst und ein Ausblick für die Forschung erstellt. (vom Brocke, et al., 2009)

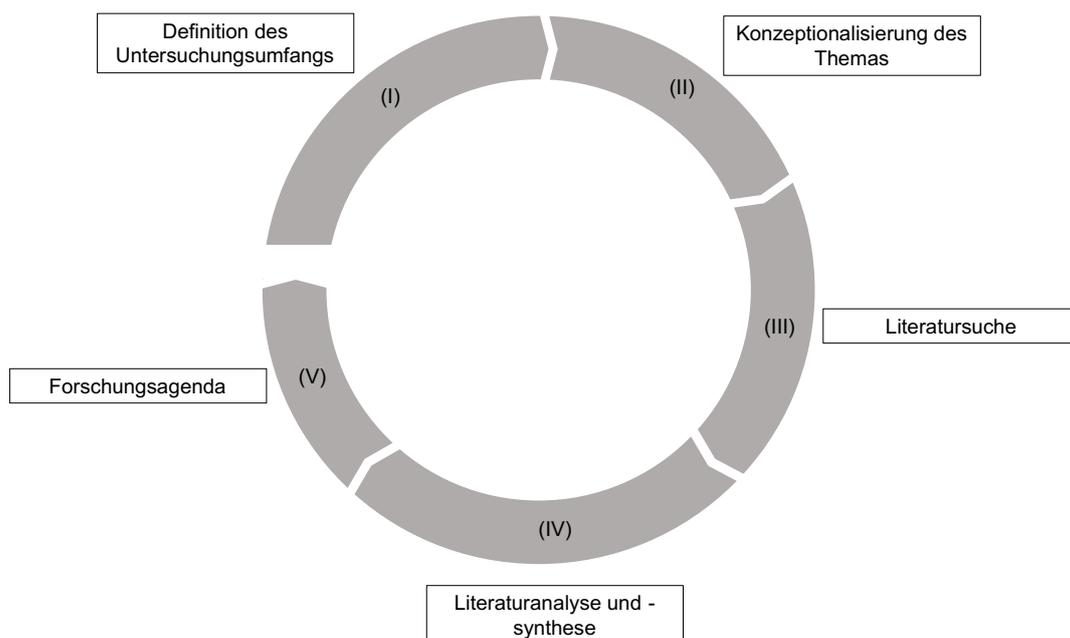


Abbildung 10: Vorgehen nach vom Brocke et al. (vgl. vom Brocke, et al., 2009, S. 7)

4.2 Durchführung der Literaturrecherche

Auf Grundlage der drei vorgestellten Vorgehensweisen wurde die Literaturrecherche in dieser Arbeit aufgebaut. Das Vorgehen der Recherche wird in diesem Abschnitt näher erläutert. Zunächst wird wie nach Kitchenham (2004) und vom Brocke et al. (2009) die Forschungsfrage und das Forschungsziel genau definiert. Im zweiten Schritt wird die Literaturrecherche organisiert, indem Datenbanken, Suchbegriffe und Ein-/Ausschlusskriterien festgelegt werden. Im Anschluss wird die Literatursuche auf den Datenbanken durchgeführt und dokumentiert. Als viertes folgt eine Qualitätsbewertung der Literatur in mehreren Schritten. Im fünften Schritt wird anhand der Literatur, die durch die Qualitätskontrolle für die weitere Analyse in Betracht gezogen wird, eine Rückwärts- sowie Vorwärts-Suche durchgeführt. Die so gesuchten Literaturen haben die gleichen Anforderungen wie bei der Literatursuche. Im letzten Schritt wird die final ausgewählte Literatur analysiert, ausgewertet und dargestellt.

Definition des Forschungsziels

Das Ziel dieser Arbeit ist es, eine Übersicht über bereits bestehende Modelle für die Zuverlässigkeitsprognose von Heizungsanlagen mittels Datenanalyse zu erstellen. Dabei wurden verschiedene Schwerpunkte der Literaturrecherche festgelegt. Die Schwerpunkte beinhalten Datenanalyse im allgemeinen sowie Zuverlässigkeitsprognose und vorausschauende Instandhaltung.

Organisation der Literatursuche

Anhand der Forschungsfrage und der Schwerpunkte wurde in der Folge eine erste Suche auf *Google Scholar* durchgeführt, um einen groben Überblick über das Thema zu generieren. Nachdem der Überblick geschaffen wurde, konnte die Vorbereitung für die Literatursuche getroffen werden. Dazu wurde zunächst eine Auswahl an Datenbanken getroffen. Insgesamt sind zehn Datenbanken ausgewählt worden, auf denen die Suche nach Literatur durchgeführt werden soll: *Google Scholar*, *Scopus*, *Web of Science*, *Engineering Village*, *Proquest*, *IEEEExplore*, *ScienceDirect*, *EBSCOhost*, *Universitätsbibliothek der TU Dortmund*, *SpringerLink*.

Des Weiteren wurden Suchbegriffe für die Suche nach Literatur aufgelistet. Dabei wurden zunächst englische und deutsche Suchbegriffe in Betracht gezogen. Diese wurden in vier Ebenen kategorisiert und aufgelistet (s. Tabelle 3 und Tabelle 4). Ebene 1 bezieht sich auf die Art der gesuchten Darstellung und beinhaltet Suchbegriffe wie *Modell* oder *Methode*. Auf Ebene 2 ist der allgemeine Themenbereich dargestellt, also z. B. die Suchbegriffe *Zuverlässigkeit* und *Instandhaltung*. Die Suchbegriffe zum Sektor Heizungssysteme sind in Ebene 3 aufgelistet. Ebene 4 beinhaltet abschließend Suchbegriffe zur Datenanalyse.

Tabelle 3: Vorläufige Liste der englischen Keywords für die Literaturrecherche

Ebene 1	Ebene 2	Ebene 3	Ebene 4
Model	Reliability	Heater	Analysis
Reference	Maintenance	Heating	Prediction
Method	Service	HVAC	Forecast
Concept	Availability	Heat pump	Predictive
Methology	Dependability	Heating installation	Data
Sample		Heating system	Time series
Procedure		Residential heating system	Mining
Process		Radiant heating	Machine learning
Procedure Model		District heating	Data analysis
Reference Model		Central heating	Smart Data
Process Model			Time-Series Data-analysis
			Data mining
			Big Data
			Assesment
			Prognosis
			Anomaly detection

Tabelle 4: Vorläufige Liste der deutschen Keywords für die Literaturrecherche

Ebene 1	Ebene 2	Ebene 3	Ebene 4
Modell	Zuverlässigkeit	Zentralheizung	Datenanalyse
Verfahren	Instandhaltung	Heizungsanlage	Zeitreihendatenanalyse
Methode	Service	Heizsystem	Prognose
Konzept	Wartung	Heizungssystem	Vorhersage
Vorgehen	Verfügbarkeit	Wärmepumpe	Analyse
Methodik	Verlässlichkeit		Anomaliedetektion
Referenzmodell			Machine Learning
Vorgehensmodell			Data Mining
Verfahrensmodell			Smart Data
Prozessmodell			

Als weitere Eingrenzung wurde der Zeitraum der zu suchenden Literatur auf die Jahre 2019 bis 2023 beschränkt, um Modelle zu finden, die einen möglichst aktuellen Stand der Technik beinhalten. Zudem sollen nur Literaturen gewählt werden, die in den Datenbanken offen zur Verfügung stehen und somit für die Suche direkt zugänglich sind. Des Weiteren wurden die Suchergebnisse nach Relevanz sortiert.

Nach einer ersten Suche mit den Suchbegriffen auf der Datenbank *Scopus* stellte sich heraus, dass bei der Suche mit deutschen Suchbegriffen keine hilfreichen Literaturen zu finden waren. Dies wurde auf den anderen Datenbanken ebenfalls geprüft und es wurden keine relevanten Literaturen gefunden. Daher beschränken sich die Suchergebnisse in der Arbeit auf die englische Sprache.

Mit den englischen Suchbegriffen wurde ebenfalls zunächst auf *Scopus* gesucht. Dabei wurden alle Suchbegriffe einer Ebene mit dem Suchoperator *OR* verknüpft und die vier Ebenen mit dem Suchoperator *AND* zu einer Suchkette verknüpft (s. Tabelle 5). Danach wurden experimentell einzelne Suchbegriffe aus der Suchkette entfernt, um deren Einfluss auf die Anzahl der Suchergebnisse zu bestimmen. Folglich konnten einige Suchbegriffe ausgeschlossen werden. Daraus ergibt sich die finale Liste der Suchbegriffe für die anschließende Literaturrecherche (s. Tabelle 6).

Tabelle 5: Beispielhafte Suchkette

Operator	Suchkette
	(model OR reference OR method OR concept OR methology OR sample OR procedure OR process OR "process model" OR "procedure model" OR "reference model" OR "process model")
AND	(reliability OR maintenance OR service OR availability OR dependability)
AND	(heating OR hvac OR heater OR "heat pump" OR "heating installation" OR "heating system" OR "residential heating system" OR "radiant heating" OR "district heating" OR "central heating")
AND	(analysis OR prediction OR forecast OR predictive OR "Machine Learning" OR data OR "time series" OR mining OR "data analysis" OR "time-series data analysis" OR "smart data" OR "data mining" OR "Big data" OR assessment OR prognosis OR "anomaly detection")

Literatursuche

Nachdem alle Kriterien für die Literatursuche festgelegt wurden, konnte die Literatursuche auf den Datenbanken durchgeführt werden. Die Suche wurde, wie in Tabelle 5 beispielhaft dargestellt, mit den finalen Suchbegriffen auf allen Datenbanken durchgeführt. Um die Ergebnisse zu reduzieren und präzisere Suchergebnisse zu erhalten, wurde die Suche falls möglich auf bestimmte, für jede Datenbank individuelle, Suchfelder begrenzt. Zum Beispiel

wurde bei der Universitätsbibliothek die Suche der Suchbegriffe auf die Zusammenfassung begrenzt. Eine vollständige Liste der Eingrenzungen in den einzelnen Datenbanken ist in Tabelle 7 dargestellt.

Tabelle 6: Finale Liste der Keywords

Ebene 1	Ebene 2	Ebene 3	Ebene 4
Model	Reliability	Heater	Analysis
Reference	Maintenance	Heating	Prediction
Method	Service	HVAC	Forecast
Concept	Availability	Heat pump	Predictive
Process		Heating installation	Data
		Heating system	Time series
			Mining
			Machine learning

Tabelle 7: Suchfelder auf den Datenbanken

Datenbank	Suchfelder	Beschreibung
Scopus	TITLE-ABS-KEY	Titel, Zusammenfassung, Keywords
IEEEExplore	All Metadata	Titel, Zusammenfassung, Keywords
Proquest	NOFT	Alle Felder außer Volltext
Engineering Village	Subject/Title/Abstract	Thema, Titel, Zusammenfassung
Web of Science	Topic (TS)	Titel, Zusammenfassung Keywords
EBSCOhost	Abstract (AB)	Zusammenfassung
Universitätsbibliothek	Abstract	Zusammenfassung
TU Dortmund		
SpringerLink	Keine Beschränkung	-
Google Scholar	Keine Beschränkung	-
ScienceDirect	TAK und T und Terms	Titel, Zusammenfassung, Keywords und Titel und kompletter Artikel

Bei der Eingrenzung der Keywords auf bestimmte Suchfelder gab es zwei Besonderheiten. Zum einen konnten auf den Datenbanken *SpringerLink* und *Google Scholar* keine Beschränkungen festgelegt werden. Aus diesem Grund wurden die Suchbegriffe auf diesen Datenbanken in allen Feldern gesucht. Und zum anderen stellte die Datenbank *ScienceDirect* eine Besonderheit dar, da aufgrund von einer Begrenzung an Suchoperatoren pro Feld mehrere Felder genutzt werden mussten. Dazu wurden die Suchbegriffe der Ebene 1 und 3 in dem Feld Titel, Zusammenfassung und Keywords gesucht, die Suchbegriffe der Ebene 2 im Titel und die Suchbegriffe der Ebene 4 im ganzen Artikel. Zudem mussten bei *ScienceDirect* zwei

Suchen durchgeführt werden, da die Anzahl der Keywords in den Ebenen 1 und 3 zu hoch war und wiederum die Anzahl der maximalen Anzahl an Suchoperatoren überschritten hat. Getrennt wurden die Suchen in der Ebene 3. In der ersten Suche wurden die Keywords *HVAC* und *heat pump* mit einbezogen und in der zweiten Suche die Keywords *heating system*, *heating*, *heater* und *heating installation*. Auf der Datenbank *Google Scholar* mussten ebenfalls mehrere (6) Suchen durchgeführt werden, da die Anzahl an Zeichen beschränkt ist. Von den Suchergebnissen wurden je Suche die ersten 30 Literaturen für die weitere Untersuchung in Betracht gezogen. Insgesamt ergaben sich dadurch 463 Literaturen, die für die weitere nachfolgende Bewertung in Betracht gezogen wurden. Eine Liste mit den jeweils ersten 30 Literaturen sowie den dazugehörigen Datenbanken und dazugehörigen Suchketten ist im Anhang in Mappe1 aufgeführt.

Qualitätsbewertung der Literatur

Die gefundenen Literaturen wurden in einem ersten Schritt nur anhand des Titels und der *Keywords* bewertet. Nach der ersten Bewertungsphase sind 61 Literaturen übriggeblieben. Literaturen, die auf einer Datenbank (*ScienceDirect* oder *Google Scholar*) bei mehreren Suchanfragen vorgekommen sind, sind bereits aussortiert worden und werden nicht mehrfach aufgeführt. Von den 61 Literaturen waren 20 doppelte und dementsprechend sind 41 Literaturen für die weitere Untersuchung betrachtet worden. In einer weiteren Bewertungsphase sind auf Basis der Zusammenfassungen weitere, für das Thema der Arbeit nicht passende Literaturen, aussortiert worden, sodass 25 Literaturen für die weitere Untersuchung in Betracht gezogen wurden. Diese wurden anschließend komplett gelesen und erneut nach ihrer Themenrelevanz bewertet. Letztendlich sind 18 Literaturen für die Arbeit relevant, wobei sich in drei Fällen jeweils zwei Literaturen mit dem gleichen Ansatz beschäftigt haben, sodass durch die Suche insgesamt 15 Ansätze relevante Ansätze gefunden wurden. Eine Liste mit den jeweiligen Literaturen aus den Bewertungsschritten ist im Anhang in Mappe2 dargestellt.

Rückwärts- und Vorwärts-Suche

Anhand der 15 relevanten Ansätze wurde anschließend eine Rückwärts- und Vorwärtssuche durchgeführt. Dabei wurden für die Literatur die gleichen Kriterien angewendet wie bei der ursprünglichen Literatursuche. Durch die Rückwärts- und Vorwärts-Suche wurden so vier weitere Ansätze gefunden.

Literaturanalyse und -auswertung

In dem letzten Schritt der systematischen Literaturrecherche wurden die finalen 19 gefundenen Ansätze (s. Tabelle 8) ausgewertet. Dazu wurden die Ansätze anhand des Anforderungsprofils für ein Referenz- und Vorgehensmodell zur Zuverlässigkeitsprognose von Heizsystemen analysiert und zusammenfassend dargestellt. Die Darstellung der Ansätze und die Gegenüberstellung dieser mit dem Anforderungsprofil erfolgt im folgenden Kapitel.

Tabelle 8: Literaturübersicht der gefundenen Ansätze

Autor (Jahr)	Titel	Art der Literatur
Fernandes et al. (2020)	Forecasting Appliances Failures: A Machine-Learning Approach to Predictive Maintenance	Zeitschriftenartikel
Bouabdallaoui et al. (2021)	Predictive maintenance in building facilities: A machine learning-based approach	Zeitschriftenartikel
Villa et al., (2022)	Machine Learning Framework for the Sustainable Maintenance of Building Facilities	Zeitschriftenartikel
Hosamo et al., (2022)	A Digital Twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics	Zeitschriftenartikel
Albayati et al., (2023)	Semi-Supervised Machine Learning for Fault Detection and Diagnosis of a Rooftop Unit	Zeitschriftenartikel
Galvez et al., (2021)	Fault detection and RUL estimation for railway HVAC systems using a hybrid model-based approach	Zeitschriftenartikel
Mansson et al., (2021)	A Fault Handling Process for Faults in District Heating Customer Installations	Zeitschriftenartikel
Chen et al., (2022)	Integration of FDD data to aid HVAC system maintenance	Konferenzprotokoll
Prabhu & Chaudhary, (2021)	Machine Learning enabled Condition Monitoring Models for Predictive Maintenance of Boilers	Konferenzprotokoll
Jayatilleka & Peters, (2021)	Reliability Modelling in HVAC & Other Industry Equipment Using Big Data	Konferenzprotokoll
Ciani et al., (2020)	Reliability evaluation of an HVAC ventilation system with FTA and RBD analysis	Konferenzprotokoll
Li et al., (2020)	A data-driven method for fault detection and isolation of the integrated energy-based district heating system	Zeitschriftenartikel
Yang et al., (2020)	Machine learning-based prognostics for central heating and cooling plant equipment health monitoring	Zeitschriftenartikel

Zhu et al., (2022)	An Effective Fault Detection Method for HVAC Systems Using the LSTM-SVDD Algorithm	Zeitschriftenartikel
Xie et al., (2023)	Digital twin enabled fault detection and diagnosis process for building HVAC systems	Zeitschriftenartikel
Ahern et al., (2023)	Implementation of the IDAIC framework on an air handling unit to transition to proactive maintenance	Zeitschriftenartikel
Borda et al., (2023)	Development of Anomaly Detectors for HVAC Systems Using Machine Learning	Zeitschriftenartikel
Martinez-Viol et al., (2022)	Semi-Supervised Transfer Learning Methodology for Fault Detection and Diagnosis in Air-Handling Units	Zeitschriftenartikel
Haruehansapong et al., (2022)	Deep Learning-Driven Automated Fault Detection and Diagnostics Based on a Contextual Environment: A Case Study of HVAC System	Zeitschriftenartikel

5 Literaturlauswertung und Darstellung relevanter Ansätze

Im folgenden Kapitel werden die relevanten Ansätze zur Zuverlässigkeitsprognose von Heizsystemen zusammengefasst und vorgestellt. Zudem werden die Ansätze im letzten Abschnitt des Kapitels anhand des Anforderungsprofils ausgewertet.

5.1 Machine-Learning Ansatz für vorausschauende Instandhaltung nach Fernandes et al. (2020)

Im folgenden Ansatz beschreiben Fernandes et al. (2020) eine Methode zur Identifikation von Fehlfunktionen in Heizungs-, Lüftungs- und Klimatisierungssystemen (HLK-Systeme) bevor diese tatsächlich auftreten. Dazu werden zunächst relevante Bausteine für Lösungen zur vorausschauenden Instandhaltung vorgestellt und in Verbindung gesetzt. Die Bausteine sind die Datenquelle, die Datenbeschaffung, die Datenspeicherung, die Datenverarbeitung, das Data Mining und die Datenvisualisierung. Für die Datenverarbeitung wird eine Smart Cloud of Things (SCoT) Architektur verwendet, die aus verschiedenen Softwareanwendungen besteht. Die vorgestellte Methode hat den Zweck, eine Fehlfunktion oder einen Ausfall eines Heizkessels bis zu 7 Tage bevor dieser Auftritt zu erkennen. (Fernandes, et al., 2020) Für die Studie wurden zunächst Daten aus 1000 Heizkesseln und 10 verschiedenen Gerätemodellen über 16 Monate gesammelt, wobei einige Geräte nur Daten über einen Monat bereitstellen konnten. Aus diesem Grund werden nur Daten gewählt, die aus einem Gerät und drei aufeinanderfolgenden Monaten stammen. Die Daten werden jeweils in einem Abstand von einer Millisekunde aufgenommen. Insgesamt kann der Zustand aus den Daten dadurch mit 91 Merkmalen beschrieben werden. Die erhaltenen Daten werden folglich vorverarbeitet indem fehlende Daten aufgefüllt werden, die Daten in nicht überschneidende 10-Minuten-Intervalle zusammengefasst werden und irrelevante Merkmale aussortiert werden. Die vorverarbeiteten Daten werden folglich mit einem der drei möglichen Zustände gekennzeichnet: (i) kein Fehler, (ii) leichter Fehler und (iii) schwerer Fehler. Anschließend werden die gekennzeichneten Daten in verschiedene Machine-Learning-Modelle (ML-Modelle) eingesetzt, mit denen die Fehler vorhergesagt werden sollen. Die ML-Modelle werden abschließend anhand ihrer Genauigkeit bewertet. (Fernandes, et al., 2020)

5.2 Machine-Learning Ansatz für vorausschauende Instandhaltung nach Bouabdallaoui et al. (2021)

In ihrem Ansatz beschreiben Bouabdallaoui et al. (2021) ein Vorgehen zur Implementierung von vorausschauender Instandhaltung bei Gebäudeinstallationen anhand eines HLK-Systems. Das Modell hat den Zweck, Fehler vorherzusagen. (Bouabdallaoui, et al., 2021)

Zunächst müssen verschiedene Daten aus vorher definierten und verfügbaren Datenquellen gesammelt werden. Diese werden anschließend gespeichert und verarbeitet. Zur Verarbeitung werden die Daten bereinigt und transformiert. Bei der Datenbereinigung werden irrelevante Einträge gestrichen, fehlende Werte ersetzt und Ausreißer behandelt. Zur Datentransformation werden zwei Methoden vorgeschlagen: Zum einen die Normalisierung von numerischen Merkmalen und zum anderen die Kodierung von kategorischen Merkmalen. Anschließend wird das ML-Modell erstellt. Dazu wird der vorher gewonnene Datensatz in einen Trainingsdatensatz zum Trainieren des Modells und einen Testdatensatz zur Validierung des Modells unterteilt. In der Studie wird ein Autoencoder bestehend aus mehreren *Long Short-Term Memory (LSTM)*-Ebenen verwendet. Dieser bestimmt anhand eines Anomalie-Werts und eines festgelegten Schwellenwerts, ob ein möglicher Fehler vorliegt. Nachdem das ML-Modell getestet und validiert wurde, kann es eingesetzt werden. Sobald der Anomalie-Wert höher ist als der Schwellenwert, wird eine Fehlerbenachrichtigung abgegeben. Abschließend werden während des Einsatzes des ML-Modells Rückmeldungen gesammelt. In den Rückmeldungen können falsche Fehlermeldungen oder unentdeckte Fehler an das ML-Modell gemeldet werden, um es nachfolgend zu verbessern. (Bouabdallaoui, et al., 2021)

5.3 Machine-Learning Framework für nachhaltige Instandhaltung nach Villa et al. (2022)

In dem Ansatz von Villa et al. (2022) wird ein Modell zur Vorhersage von Anomalien mit Hilfe von ML vorgeschlagen. Das Modell wird zudem in ein Gebäudeinformationsmodell (BIM: Building Information Model) integriert. Das vorgeschlagene Modell soll Anomalien klassifizieren und vorhersagen und in Realzeit auf einem Dashboard visualisieren. Als Anwendungsbeispiel wird in einer Studie ein Gebläsekonvektor eines HLK-Systems herangezogen. (Villa, et al., 2022)

An dem Gebläsekonvektor und im Raum wurden Internet of Things (IoT)-Sensoren angebracht, die verschiedene Daten kontinuierlich sammeln. Die Daten werden anschließend lokal sowie auf einer Cloud gespeichert. Die gesammelten Daten müssen vor dem Einsetzen in das ML-Modell vorverarbeitet werden. Dabei müssen die Daten bereinigt und integriert sowie die relevanten Merkmale skaliert, extrahiert und ausgewählt werden. Für die Vorverarbeitung können dazu verschiedene Methoden eingesetzt werden. Die vorverarbeiteten Daten

werden zudem mit dem Fehler, falls vorhanden, gekennzeichnet. Um anschließend das optimale ML-Modell für die Fehlererkennung zu identifizieren, werden die gekennzeichneten Daten in eine automatische ML-Methode eingesetzt. Das automatische ML-Modell trainiert und testet daraufhin verschiedene ML-Modelle. Die Modelle werden anschließend anhand von Messgrößen evaluiert und ausgewählt. (Villa, et al., 2022)

5.4 Digitaler Zwilling für die vorausschauende Instandhaltung nach Hosamo et al. (2022)

Im Folgenden wird der Ansatz aus dem Zeitschriftenartikel von Hosamo et al. (2022) vorgestellt. Der vorgeschlagene Vorgang nutzt einen digitalen Zwilling und ML zur Fehlerdetektion und -diagnose (FDD) von Gebäudekomponenten, sodass das Gebäudemanagement-Personal bessere Entscheidungen in der Instandhaltung treffen kann. Der Ansatz wurde in einer Studie anhand eines Klimageräts angewendet. (Hosamo , et al., 2022)

Das Modell beinhaltet drei Hauptschritte: Datenerfassung, vorausschauende Instandhaltung und ein BIM-Modell zur Visualisierung und Überwachung. Zunächst werden bei der Datenerfassung die Parameter identifiziert, die im Zusammenhang mit der vorausschauenden Instandhaltung stehen. Dazu werden die Daten in drei Gruppen klassifiziert: (1) Daten aus dem BIM-Modell, um Informationen über die Geräte zu gewinnen, (2) Sensor-Daten und (3) Daten zur Instandhaltung. Die Daten und Informationen aus den genannten Datenquellen werden auf einer Plattform gespeichert. Für die Überwachung während des Betriebs der Geräte wird ein Netzwerk aus IoT-Sensoren genutzt, das Daten aus den angebrachten Sensoren sammelt. Die Daten werden ebenfalls auf dem BIM-Modell gespeichert und visualisiert. Anschließend erfolgt der Prozess der vorausschauenden Instandhaltung. Zunächst werden die Daten bzw. Merkmale ausgewählt und vorverarbeitet. Die Vorverarbeitung wird in zwei Schritte unterteilt: Datenbereinigung und Datennormalisierung. Die Zustandsüberwachung erfolgt über das Erfassen und Interpretieren von Komponentenparametern. Anhand dieser wird bewertet, ob sich der Zustand einer Komponente im Laufe der Zeit im Vergleich zum Normalzustand verändert hat. Das Zustandsbewertungssystem wird mithilfe von sogenannten Expertenregel errichtet. Die Regeln sind von der Betriebsart abhängig und werden ausgelöst, sobald die Betriebsart bekannt ist. (Hosamo , et al., 2022)

Zur Prognose von Fehlern werden verschiedene ML-Modelle angewendet: *Artificial Neural Network* (ANN), *Support Vector Machine* (SVM) und *Decision Tree* Algorithmen. Den Modellen werden die Daten aus dem BIM-Modell, dem IoT Sensor-Netzwerk und der Instandhaltung zur Verfügung gestellt. Anhand der Daten werden die ML-Modell trainiert, validiert und getestet. Der Output des Prognose Prozesses beinhaltet die Fehler und eine Instandhal-

tungsanfrage, die in der Folge bewertet werden muss. Abschließend kann auf Grundlage des Ergebnisses eine Anpassung des Modells vorgenommen werden. (Hosamo , et al., 2022)

5.5 Machine-Learning Ansatz für die Fehlerdetektion und -diagnose nach Albayati et al. (2023)

Im folgenden Abschnitt wird der Ansatz nach Albayati et al. (2023) beschrieben. Der Zeitschriftenartikel beschäftigt sich mit ML-Modellen als Methode zur FDD. Die Studie wird anhand einer Dachanlage eines HLK-Systems aus einem Warenhaus durchgeführt. (Albayati, et al., 2023)

In der Studie werden aus einer Vielzahl von möglichen Fehlern in HLK-Systemen zunächst nur vier Fehler betrachtet: Unterversorgung von Kühlmittel, Überversorgung von Kühlmittel, verschmutzte Kondensatoren und verschmutzte Verdampfer. Für die vier Fehler werden jeweils ein Schwellenwert für die Fehlerintensität festgelegt. Zum Erfassen von Daten werden über 20 Sensoren an das System angebracht. Die Sensoren übertragen dabei in einem Intervall von einer Minute Daten zu dem Zustand und Verhalten des Systems. Anhand der Daten wird die Fehlerintensität der einzelnen Fehler berechnet und mit dem Schwellenwert verglichen. Die einzelnen Datenpunkte werden anschließend mit dem jeweiligen Fehler gekennzeichnet. Für die Studie wurden zwei Datensätze gesammelt. Der Unterschied der beiden Datensätze liegt darin, dass im ersten Datensatz mehrere Fehler möglich sind und in dem Zweiten angenommen wird, dass nur ein Fehler pro Beobachtung auftritt. Abschließend werden die für die Fehler relevanten Merkmale herausgesucht. (Albayati, et al., 2023)

In der Studie werden drei verschiedene ML-Anwendungen getestet und verglichen. Zum einen der SVM Algorithmus, zum anderen eine Kombination aus den Algorithmen SVM und dem *k-nearest Neighbor* (k-NN) und zuletzt eine Kombination aus SVM, Clustering und des k-NN Algorithmen. Um die Modelle zu trainieren, werden zufällig jeweils 25-30 Beobachtungen aus jeder Fehlerklasse in das Modell eingespielt. Die restlichen Beobachtungen werden als Testdaten genutzt und dienen zur Validierung des Modells. Die vorgestellten Modelle werden anschließend angewendet und die Genauigkeit der Fehlererkennung wird bestimmt, um einen Vergleich der Modelle zu ermöglichen. (Albayati, et al., 2023)

5.6 Hybrider modellbasierter Ansatz für die Fehlererkennung nach Galvez et al., (2021)

Galvez et al., (2021) stellen einen Ansatz eines hybriden Modells, bestehend aus einem physikbasierten Modell und einem datenbasierten Modell, vor. Das Modell soll zur FDD in einem HLK-System dienen. In einer Studie wird der Ansatz anhand eines HLK-Systems in einem Zug angewendet. (Galvez, et al., 2021)

Das physikbasierte Modell basiert auf dem realen System und generiert Daten zum Zustand auf Basis von modellierten Betriebsmodi. Das Modell hat reale und virtuelle Sensoren, welche auf den gemessenen Daten der realen Sensoren beruhen. Die Daten der realen und virtuellen Sensoren müssen in das physikbasierte Modell gespielt werden, um eine Antwort auf ein Eingangssignal zu simulieren. Die von den Sensoren gemessenen Daten müssen zunächst vorverarbeitet werden. Zuerst müssen die Daten bereinigt werden, um Fehler bei der Simulation und der Extraktion der Merkmale zu vermeiden. Im nächsten Schritt werden die Daten analysiert, um Schlüsselmerkmal zu identifizieren und zu klassifizieren. Anschließend werden die Daten nach ihrem Betriebsmodus und dem Fehlertyp klassifiziert. Das physikbasierte Modell simuliert folglich die gleiche Antwort wie das reale System. Die Reaktion des Systems wird anhand der Merkmale aus den realen und virtuellen Sensoren definiert. Nachdem das physikbasierte Modell die Daten generiert hat, wird jede Simulation mit dem jeweiligen Fehlercode gekennzeichnet. Die Merkmale werden dann aus den Daten extrahiert. Jedes Merkmal, welches in Relation zu einem Fehlercode steht, wird genutzt, um das datenbasierte Modell zu trainieren, zu validieren und zu testen. Nachdem die Daten vorverarbeitet wurden, wird das datenbasierte Modell erstellt. In der Studie wird ein überwachtetes Klassifizierungs-Modell verwendet. Es werden verschiedene Ansätze des überwachten Lernens getestet und anhand ihrer Genauigkeit evaluiert. (Galvez, et al., 2021)

5.7 Prozess zur Fehlerhandhabung nach Mansson et al., (2021)

Nachfolgend wird der Ansatz von Mansson et al. (2021) dargestellt. Sie schlagen ein Vorgehen zur Fehlerbehandlung in Fernwärme-Anlagen vor. Zunächst werden Daten aus den Anlagen der Kunden gewonnen. Da die Daten für das Analysetool genutzt werden sollen, müssen diese von hoher Qualität und vorverarbeitet sein, damit das Datenformat zu dem Algorithmus der Fehlererkennung passt. Im nächsten Schritt müssen die Daten analysiert werden, um ein abweichendes Verhalten zu erkennen. Dazu bedarf es einem Datenanalysetool, das eine große Menge an Daten verarbeiten kann. Dieses soll in der Lage sein, Fehler zu erkennen und zu diagnostizieren. Sobald also ein abweichendes Muster in den Daten auftritt, soll es anzeigen, dass ein Fehler in der Installation herrscht. Zudem wird beschrieben, dass trotz der automatisch ablaufenden Fehlererkennung ein Datenspezialist zur Verfügung stehen sollte. Auch andere Indikatoren der Fehlererkennung werden in Betracht gezogen, z.B. die Meldung eines Kunden über spürbare Änderung im Wärmekomfort der Anlage. (Mansson, et al., 2021)

Der nächste Schritt ist die Entscheidung, wie mit der Fehlermeldung umgegangen wird. Dazu wird in dem Modell ein Service-Admin integriert, der mit den Charakteristiken des Fernwärme-Systems vertraut ist. Dieser soll die Fehler einstufen und priorisieren, da es zu möglichen

Fehlermeldungen kommen kann. Beim Auftreten einer Fehlermeldung sammelt der Admin in der Folge zusätzliche Information zur Schwere des Fehlers und entscheidet, ob die Anlage besucht und der Fehler behoben werden muss. Anschließend muss dokumentiert werden, falls keine weitere Untersuchung und Bearbeitung des Fehlers veranlagt wird. Dies ist wichtig, falls in Zukunft weitere Fehler in der Anlage auftreten. Sollte eine Fehler weiter untersucht werden müssen, wird daraufhin eine Arbeitsanordnung erstellt. Diese muss alle wichtigen Informationen zu der Anlage und dem Fehler beinhalten. Im Anschluss wird die Anlage von einem Techniker untersucht. Während des Besuchs erstellt der Techniker ein Serviceprotokoll, das alle Informationen zur Anlage und zur Fehlerbehebung beinhaltet. Nach dem Besuch wird noch das Erstellen eines Berichts vorgesehen, der auf dem Protokoll beruht. Im letzten Schritt des Vorgehens wird der Bericht an verschiedene Bereiche des Betreibers versendet, unter anderem auch an das Datenanalysetool. Dies hat den Zweck das Datenanalysetool in der Folge zu verbessern. (Mansson, et al., 2021)

5.8 Fehlerdetektion und -diagnose für die Instandhaltung nach Chen et al. (2022)

Chen et al. (2022) schlagen einen Ansatz zur Integration eines FDD-Tools in ein computergestütztes Instandhaltungsmanagementsystem (Computerized Maintenance Management System (CMMS)). In einer Studie wird der Ansatz anhand von 2162 Dachanlagen von HLK-Systemen angewendet. (Chen, et al., 2022)

Zu Anfang werden Daten aus verschiedenen Quellen, z. B. Sensoren, für das FDD Tool erfasst. Diese werden vorverarbeitet und in einen Algorithmus zur FDD gegeben. Der Algorithmus erkennt und diagnostiziert Fehler im System und gibt die klassifizierten Fehler zum einen an das Betriebs- und Instandhaltungsteam weiter und zum anderen werden die Daten an das CMMS übermittelt. Dort fließen die FDD-Daten gemeinsam mit weiteren instandhaltungsbezogenen Daten in einen weiteren Algorithmus, der eine Entscheidung bezüglich der Instandhaltungsmaßnahmen trifft. Die Entscheidung wird ebenso an das Betriebs- und Instandhaltungsteam übermittelt. (Chen, et al., 2022)

Des Weiteren wird ein FDD-Datenmodell für das FDD-Tool erstellt. Das Datenmodell beinhaltet vier Ebenen an Informationen: 1) Zeit des Fehlerauftretens, 2) die Gebäude und System ID, 3) eine Fehler ID und 4) zusätzliche Informationen zum Fehler. Es wird beschrieben, dass die vierte Ebene sowohl bereits vom FDD-Tool erstellt werden kann als auch vom CMMS. Die Fehler ID aus der dritten Ebene muss zuvor beschrieben werden, d. h. die Fehler müssen zuvor bekannt sein. Da es also zu falschen Meldungen kommen kann, müssen die Meldungen vom Personal validiert anschließend werden. (Chen, et al., 2022)

5.9 Machine-Learning Ansatz für die Zustandsüberwachung nach Prabhu & Chaudhary (2021)

Prabhu & Chaudhary (2021) stellen einen Ansatz zur vorausschauenden Instandhaltung für Heizkessel vor, bei dem Schlüsselpersonal-Indikatoren erstellt werden und mit ML-Modellen die restliche Zeit bis zum Eintreten eines kritischen Zustands bestimmt wird. In der Studie werden zwei konkrete Anwendungsmodelle für Heizkessel aufgestellt. Die vorgestellten Modelle sind ein Ölheizungs-Performance-Modell und ein Wasserpumpen-Performance-Modell. (Prabhu & Chaudhary, 2021)

Der Vorgang zur Erstellung der Modelle stellt sich wie folgt dar. Die Daten für die Modelle werden aus IoT-Sensoren gewonnen, die sich in den Räumen des Heizkessels befinden. Die IoT-Sensoren übertragen die Daten in einer Frequenz von einer Sekunde an einen Web-Server. Die autorisierten Benutzer können die Daten durch ein Benutzerportal einsehen. Der Web-Server übermittelt die Daten in der Folge an eine Cloud, die die Daten speichert. Über die Cloud werden die Daten durch ein ML-Modell modelliert und die Ergebnisse werden daraufhin in der Cloud gespeichert und im Benutzerportal visualisiert. (Prabhu & Chaudhary, 2021)

Die in der Studie verwendeten Daten bestehen aus einer Vielzahl an analogen und digitalen Daten. Digitale Daten sind z.B. die Zeit, die der Heizkessel ein-/ausgeschaltet ist. Ein analoger Parameter ist z.B. die Öltemperatur oder die Wasserflussrate. Auf Grundlage der Parameter wurden für die beiden Modelle jeweils ein Performance-Index eingeführt (Öl-Performance-Index und Wasserpumpen-Verschlechterungsfaktor). Die Daten, die für das ML-Modell bezogen werden, stellen sich aus den aktuellen kontinuierlich erfassten Daten und historischen Daten aus 6 (Ölheizung) bzw. 2 (Wasserpumpe) Monaten zusammen. Die Daten werden in der Folge durch verschiedene ML-Regressions-Modelle modelliert und das beste Modell wird ausgewählt. Mit den herausgestellten Indizes, den Daten und den ML-Modellen wird schließlich die restliche Zeit bis zum Eintreten eines kritischen Zustandes der jeweiligen Systeme bestimmt. (Prabhu & Chaudhary, 2021)

5.10 Zuverlässigkeitsmodellierung nach Jayatilleka & Peters (2021)

Jayatilleka und Peters (2021) beschreiben einen Ansatz mit Bezug zu HLK-Systemen, um die Zuverlässigkeit von Systemen mit Hilfe von Big Data zu bestimmen und zu verbessern. Zunächst wird die Lebensdauer der Systeme bestimmt. Diese wird anhand des Benutzerprofils und Lastenprofils bemessen. Die Lebensdauer wird als 90% der Benutzerrate bestimmt. Das bedeutet 90% der Kunden haben eine Nutzungsdauer, die über der Lebensdauer liegt und 10%, die unter der Lebensdauer liegt. (Jayatilleka & Peters, 2021)

Es werden drei Entwicklungslevel von Komponenten beschrieben: unterentwickelte, akkurat entwickelte und überentwickelten Komponenten. Die Entwicklungslevel werden anhand der Lebensdauer bemessen. In Zuverlässigkeitsprogrammen sollen unterentwickelte Komponenten im Entwicklungsprozess aufgedeckt werden. Wird dies aber nicht effektiv gestaltet, so kommt es dennoch zu unterentwickelten Komponenten. Big Data kann dabei helfen unterentwickelte Komponenten frühzeitig in der Feldanwendung zu erkennen, sodass diese in der Produktion behoben werden können. (Jayatilleka & Peters, 2021)

Um die Zuverlässigkeit zu beschreiben, wird anhand der Badewannenkurve die Änderung in der Fehlerrate beschrieben. Zu Beginn der Lebensdauer kommt es zur *Kindersterblichkeit* (Infant Mortality (IM)) von Komponenten oder Systemen. Diese Fehler treten in der Regel aufgrund von Herstellungsfehlern auf. Um diese proaktiv zu verhindern, kann die Prozess-Fehlermodus- und -Einfluss-Analyse (Process Failure Mode and Effect Analysis (PFMEA)) angewendet werden. Die Methode kann dazu führen die Fehlerrate zu verringern und die Dauer der Zeit der IM zu verkürzen. (Jayatilleka & Peters, 2021)

Darüber hinaus können unterentwickelte Teile zu Fehlern während der Lebenszeit führen. Der Grund für unterentwickelte Teile ist meist eine ungenaue Aufdeckung von Fehlermodi in der Entwicklung. Als proaktive Methoden kann die Entwurfs-Fehlermodus- und -Einfluss-Analyse (Design Failure Mode and Effect Analysis (DFMEA)) eingesetzt werden. Bei der DFMEA werden funktionale Anforderungen an ein System oder eine Komponente entworfen und physischen Konfigurationen zugeordnet. Mittels Designkontrollen werden die funktionalen Anforderungen kontrolliert. Ein Fehler kann als ein Vorfall definiert werden, bei dem die Anforderung nicht erfüllt wird. Wenn Anforderungen fehlen, kann die DFMEA sie nicht berücksichtigen. Die fehlenden Anforderungen werden als Fehlermodus in der Praxis entdeckt. Durch im Feld befindliche Systeme können fehlende Anforderungen hinzugefügt und minimiert werden. (Jayatilleka & Peters, 2021)

Zudem wird erwähnt, dass sogenannte Frühindikatoren einen großen Einfluss auf das Erkennen von Feldfehlern haben und dass Big Data-Systeme dabei helfen können die Verschlechterung des Zustands anhand von gesetzten Schwellenwerten zu beobachten und Fehler vorzeitig zu erkennen. Die Autoren beschreiben zudem, dass ML-Modelle helfen können diese Frühindikatoren zu erkennen. Ereignisse, die vor einem Fehler auftreten, können helfen die Ursache zu analysieren. (Jayatilleka & Peters, 2021)

5.11 Zuverlässigkeitsbewertung nach Ciani et al., (2020)

Ciani et al. (2020) beschrieben zwei Verfahren zur Bestimmung der Zuverlässigkeit eines Systems anhand des Ventilationssystems eines HLK-Systems. Die Verfahren sind die

Fehlerbaumanalyse (Fault Tree Analysis (FTA)) und das Zuverlässigkeitsblockdiagramm (Reliability Block Diagram (RBD)).

Bei der FTA wird zwischen Basis-Ereignissen, welche spontan auftreten und Ereignissen, welche durch einen oder mehrere andere Ereignisse hervorgerufen werden, unterschieden. An oberster Stelle steht ein Top-Event, also ein Systemausfall. Dieses wird analysiert, um die Fehler in den Subsystemen zu modellieren. Zudem bestehen die FTA-Diagramme noch aus sogenannte Gates. Die Gates beschreiben die Verbreitung der Fehler. Es wird zwischen AND und OR Gates unterschieden. Ein *AND* Gate sagt aus, dass mehrere Ereignisse zu einem übergeordneten Fehler führen. Ein *OR* Gate zeigt an, dass ein einzelner Fehler zu einem übergeordneten Fehler führen kann. Die Berechnung der Wahrscheinlichkeit das Auftreten eines *AND* Gates ist das Produkt der Wahrscheinlichkeit aller Ereignisse, die durch das *AND* Gate betrachtet werden. Die Wahrscheinlichkeit eines *OR* Gates ist die Summe der Wahrscheinlichkeiten der Ereignisse, die bei einem *OR* Gate in Betracht gezogen werden.

Die FTA kann also genutzt werden, um Ursachen von verschiedenen Ereignissen zu identifizieren und Informationen über die Wahrscheinlichkeit des Top-Events geben. Wenn alle Fehlerraten bzw. -wahrscheinlichkeiten bekannt sind, kann die Gesamtsystemzuverlässigkeit berechnet werden. (Ciani, et al., 2020)

Als zweites Verfahren wird das RBD beschrieben. Das RBD besteht aus Blöcken, die jeweils eine Komponente des Systems darstellen. Jeder der Blöcke beinhaltet die Fehlerrate der Komponente. Die Blöcke werden zudem mit den anderen Blöcken, mit denen die Komponenten in Verbindung stehen, verbunden, um so das System darzustellen. Die Blöcke können parallel oder in Serie geschaltet sein. Die Wahrscheinlichkeiten von zwei Blöcken, die in Reihe oder parallel geschaltet sind, lassen sich folglich berechnen und zu einer Fehlerrate zusammenfassen. (Ciani, et al., 2020)

Anhand des Ventilationssystems eines Zuges werden die Verfahren getestet. Für beide Methoden wird eine Gesamtdistanz des Zuges von 5000000km berücksichtigt anhand derer die Zuverlässigkeit beschrieben wird. Beide Methoden geben sehr ähnliche Ergebnisse und beschreiben eine Systemzuverlässigkeit von 0,44 nach den 5000000km. (Ciani, et al., 2020)

5.12 Datenbasierter Ansatz zur Fehlererkennung und -isolation nach Li et al., (2020)

In dem Zeitschriftenartikel beschreiben Li et al. (2020) eine Methode zur Fehlererkennung und -isolation (FDI). Die Methode wird anhand eines Fernwärmesystemen angewendet. (Li, et al., 2020)

Zunächst wird ein Modell zur Beschreibung des Systems erarbeitet. Dazu werden die einzelnen Systeme und Subsysteme mathematisch modelliert. Nachfolgend werden Fehlerszena-

rien klassifiziert. Für Kontrollsysteme werden drei Fehlertypen definiert: Sensor-Fehler, Antriebsfehler und Komponentenfehler. Sollte ein Fehler auftreten, ändert sich der Wert der entsprechenden Parameter, die aus den Daten erfasst werden. Die Änderung des Wertes kann wiederum in drei Unterfehlertypen unterteilt werden: Verzerrung, Drift und komplette Fehler. Anschließend wird die Zwei-Level Methode zur Fehlerdetektion und -isolation (FDI) vorgestellt. Die Methode besteht aus zwei Teilen, der Fehlerdetektion und der Fehlerisolation. Die Fehlerdetektion soll Fehler korrekt erkennen. Die Fehlerisolation soll die fehlerhaften Komponenten identifizieren und die Unterfehlerart bestimmen. Für die Fehlerdetektion werden Schwellenwerte für die Parameter gesetzt, um falsche Fehlmeldungen weitestgehend zu vermeiden. Bei Fehlerzuständen treten Residuen in den Parametern auf. Wenn die Residuen größer sind als der Schwellenwert, liegt ein Fehler vor. (Li, et al., 2020)

Die Methode zur Fehlerisolation beinhaltet zwei Klassifizierer. Der übergeordnete Klassifizierer soll Systemfehler lokalisieren und der untergeordnete Klassifizierer die Unterfehlerart bestimmen. In der Studie werden beide Klassifizierer auf Basis eines *Convolutional Neural Network* (CNN) aufgebaut. Dem CNN-Modell werden Daten hinzugefügt, die zuvor vorverarbeitet wurden und anschließend zum Trainieren des Systems genutzt werden. Anschließend wird das CNN-Modell anhand von Evaluationsmesswerten bewertet. Diese wurden auf Grundlage von beobachteten wahren positiven, falschen positiven, wahren negativen und falschen negativen Meldungen erstellt. (Li, et al., 2020)

5.13 Machine-Learning Ansatz zur Zustandsprognose nach Yang et al., 2020

Yang et al. (2020) stellen eine Methode zur Prognose von Fehlern von HLK-Systemen auf Basis von ML vor, mit der ebenfalls die Zeit bis zum Auftreten eines Fehlers bestimmt werden soll.

Zunächst wird beschrieben, dass in Gebäudesystemen zwei Arten von Fehlern auftreten können: harte und weiche Fehler. Harte Fehler sind physische Fehler, die im System auftreten. Weiche Fehler sind Fehler, die durch unpassende Steuerungsprogramme auftreten. Das Prognosemodell soll helfen auf Grundlage von historischen und aktuellen Betriebs- und Instandhaltungsdaten vorrauschauende Modelle zu entwerfen. Die Methode läuft in vier Schritten ab: Datenvorbereitung, Datenverarbeitung, Modellieren und Modellbewertung. (Yang, et al., 2020)

Bei der Datenvorbereitung wird zunächst festgelegt, dass es sich bei dem Prognoseproblem zum einen um ein binäres Klassifikations-Problem handelt. Entweder ist der Wert positiv oder negativ. Zum anderen beinhaltet es eine Regressionsaufgabe, um die Zeit bis zum Auftreten eines Fehlers zu bestimmen. Um folglich einen Datensatz zu erstellen der zur Modellierung dient, müssen Fehlerereignisse sowie die jeweils passenden Betriebsdaten heraus-

gestellt werden. Diese werden in der Folge mit dem Zustand gekennzeichnet. Die Qualität der Daten ist entscheidend für die Performance des Prognosemodells. Daher werden im nächsten Schritt die Daten vorverarbeitet. Es werden Merkmale aus dem ursprünglichen Datensatz generiert bzw. extrahiert, um die Datendarstellung zu verbessern. Aus den generierten und extrahierten Merkmalen werden mithilfe eines Algorithmus korrelierte oder irrelevante Merkmale aussortiert. Für das Klassifikations- und Regressionsproblem stehen verschiedene Algorithmen zur Verfügung. In der Studie werden einfache Algorithmen gewählt, da sie schneller angewendet werden können und einen Output generieren, der für Menschen interpretierbar ist. Im letzten Schritt werden die ML-Modelle evaluiert. Dazu können verschiedene Messwerte verwendet werden. Um die Zeit bis zum Auftreten eines Fehlers ebenfalls zu berücksichtigen, wurde zudem ein zusätzlicher score entwickelt. (Yang, et al., 2020)

5.14 Methode zur Fehlerdetektion nach Zhu et al., (2022)

Zhu et al. (2022) schlagen einen Ansatz zur Fehlererkennung in HLK-Systemen vor. Das Modell besteht aus zwei Hauptkomponenten: dem Simulationsmodell und System der Fehlererkennung. Zunächst werden während des Betriebs Daten aus dem realen System gesammelt. Die gesammelten Daten dienen dazu, das Simulationssystem zu kalibrieren, so dass sich das Simulationssystem der Funktion des realen Systems annähert. Anschließend werden Fehler in das Simulationsmodell eingespeist, um Daten über den fehlerhaften Betrieb des Systems zu erhalten. Schlussendlich extrahiert und identifiziert ein Algorithmus Merkmale aus den simulierten Daten, um Fehler zu erkennen und eine Fehlerwarnung in Echtzeit zu geben. Der Algorithmus muss zuvor mit Trainingsdaten aus dem Simulationsmodell trainiert werden. (Zhu, et al., 2022)

Die Fehlererkennung des Algorithmus läuft in vier Schritten ab. Im ersten Schritt werden die Daten erfasst. In Schritt zwei werden die Daten verarbeitet und normalisiert, um die Differenzen in den Merkmalen zu eliminieren. Anschließend werden die verarbeiteten Daten als Input für ein LSTM-Modell genutzt. Dieses wurde so entworfen, dass es die Summe der Anzahl an Verlustfunktionen minimiert. Im letzten Schritt wird der Residual als Eingang für das unterstützende Vektordaten Beschreibungs-Modell (Support Vector Data Description (SVDD)) genutzt. Der Radius R des SVDD wird als Schwellenwert für einen Fehler genutzt. Abschließend wird das Modell bewertet. Dies wird anhand der Genauigkeit, der Präzision, des Rückrufwertes und des F1-Scores gemacht. (Zhu, et al., 2022)

5.15 Prozess zur Fehlerdetektion und -diagnose mit Digitalem Zwilling nach Xie et al., (2023)

Xie et al. (2023) beschreiben ein FDD-Verfahren für HLK-Systeme, das um einen Digitalen Zwilling erweitert wurde. Zunächst werden Fachleute damit beauftragt die häufigsten Fehler des HLK-Systems zu benennen. Aus diesen Fehlern wird anschließend eine Fehlerüberwachungsliste erstellt. Damit nicht alle Daten aus dem HLK-Datensatz für den FDD-Prozess verwendet werden müssen, wird eine *Tasche mit Wörtern* (BoW: Bag of Words) basierte Methode zur Merkmalsextraktion und -auswahl angewendet. Dies wird angewendet, um die relevanten Sensordaten für die jeweiligen Fehler zu beschränken. Dies geschieht auf Basis von gekennzeichneten normalen und fehlerhaften Daten. In einem semantischen Modell werden auf Grundlage der Fehler und der relevanten Sensordaten sogenannte Fehler-Tags erstellt und als Ad-hoc Wissen in den digitalen Zwilling integriert. Dadurch werden die Fehler-Tags in einem Brick-Schema mit den relevanten Sensoren verknüpft. (Xie, et al., 2023)

Durch das Erstellen von semantischen Fehler-Tags zur Kennzeichnung der relevanten Sensoren wird ermöglicht, dass der FDD-Prozess die kontinuierlichen Echtzeitdaten ohne menschliches Eingreifen verarbeiten kann und die richtigen Sensordaten aus der Fehlerüberwachungsliste abrufen kann. Die Sensordaten werden folglich dazu verwendet die Fehler zu erkennen. Ein Fehleralarm wird ausgelöst, wenn die Daten vom Normalzustand abweichen. (Xie, et al., 2023)

5.16 Anwendung des IDAIC-Modells nach Ahern et al., (2023)

Ahern et al. (2023) stellen den Zyklus zur Verbesserung der industriellen Datenanalyse (Industrial Data Analysis Improvement Cycle (IDAIC)) vor, der zur proaktiven Instandhaltung beitragen soll. Das IDAIC-Modell wird in der Studie anhand eines Klimageräts angewendet. Das IDAIC-Modell basiert auf dem CRISP-DM-Vorgehensmodell und ist darauf angepasst, einen Übergang zur proaktiven Instandhaltung zu ermöglichen. Das Modell besteht aus einem inneren Kreis, der sich auf die Lösung von sichtbaren bzw. bekannten Fehlern konzentriert, und dem äußeren Kreis, der unsichtbare Problem, meist Vorläufer der sichtbaren Probleme, behandelt. (Ahern, et al., 2023)

Der innere Kreis besteht aus sechs Schritten. Der erste Schritt ist das Domänenverständnis. Dabei werden alle notwendigen Informationen und Daten des Systems gesammelt, um dieses vollständig zu verstehen. Im nächsten Schritt werden die Informationen und Daten genutzt, um das Problem in einen Kontext zu bringen. Anschließend werden die Daten auf Grundlage ihrer Qualität bewertet. Je nachdem wie hoch die Qualität der Daten ist, werden diese in der Folge bereinigt und vorverarbeitet. Im fünften Schritt des Vorgehens wird der Betrieb des Systems bewertet bzw. die sichtbaren Fehler werden behoben. Da datenbasierte

Modell für diesen Schritt zu aufwendig sind, werden wissensbasierte Modelle, die auf Fachwissen von Experten beruhen, eingesetzt. Der letzte Schritt des sichtbaren Kreises stellt den Übergang zu dem nicht sichtbaren Kreis dar. In diesem Schritt werden die Probleme, die in der Datenbewertung und der Betriebsbewertung festgestellt wurden, angegangen. Zudem wird das System in diesem Schritt in Betrieb genommen. (Ahern, et al., 2023)

Der unsichtbare Kreis läuft in weiteren vier Schritten ab. Im ersten Schritt wird untersucht, ob das Verständnis des Problems angemessen ist. Dazu wird geprüft, ob das System effizient und stabil funktioniert. In diesem Schritt wird die mögliche Anwendung von datengestützten Analysen für einzelne Bereiche geprüft. Anschließend werden diese Bereiche anhand ihrer Daten geprüft. Auf Basis dieser wird die Durchführbarkeit eines datenbasierten Ansatzes festgestellt. Es werden Analyseverfahren angewendet, um unerwartete Muster oder Anomalien während des Betriebes zu erkennen. In der Folge wird das beste Verfahren ausgewählt. Im letzten Schritt des nicht sichtbaren Kreises werden die Analyseergebnisse bewertet. (Ahern, et al., 2023)

5.17 Anomalie-Detektion mit Machine-Learning nach Borda et al., (2023)

Borda et al. (2023) beschreiben eine Methode zur Anomalie-Detektion in HLK-Systemen. Die Methode basiert auf ML und wird in einer Studie anhand eines Klimageräts eines HLK-Systems in einem Bürogebäude angewendet. (Borda, et al., 2023)

Zunächst werden das Gebäude sowie das Klimagerät und deren Spezifikationen vorgestellt. Die benötigten Daten für die Anomalie Detektion werden aus Sensoren und Messgeräten gewonnen, die an dem Klimagerät angebracht sind. Nachfolgend werden mit einem Algorithmus die relevanten Merkmale für die FDD ausgewählt. Die Daten werden jeweils in einem Intervall von 15 Minuten erfasst. Anschließend werden die Daten vorverarbeitet. Wenn weniger als 16 Datenpunkte eines Merkmals fehlen, werden die fehlenden Daten stückweise linear interpoliert. Bei mehr als 16 Datenpunkten in Folge wird eine Datenbereinigung vorgenommen. Zudem werden die Merkmale durch Skalierung auf eine Einheitsvarianz standardisiert. (Borda, et al., 2023)

Es werden zwei Arten von ML-Modelle getestet, zum einen überwachte Modelle und zum anderen halb-überwachte Modelle. Für das überwachte Modell werden insgesamt sechs verschiedene Modellvarianten angewendet. Die halb-überwachten Modelle werden mit Hilfe einer Selbstlernstrategie, die ein iterativen Trainingsprozess darstellt, entwickelt. Die überwachten und halb-überwachten Modelle werden abschließend jeweils getestet und evaluiert. (Borda, et al., 2023)

5.18 Transfer-Lern-Methode für Fehlerdetektion und -diagnose nach Martinez-Viol et al. (2022)

Martinez-Viol et al. (2022) schlagen einen Ansatz für Transfer-Lernen zur FDD in HLK-Systeme vor, welches durch ein *Neural Network* (NN)-Modell unterstützt wird. Die Methode wird anhand eines Klimagerätes in einer Studie getestet. Der Ansatz besteht aus vier Hauptteilen: (1) Konstruktion eines Referenzmodells, (2) Datenerfassung und Etikettierung des Ziel-Systems, (3) Erkennen von Unähnlichkeiten und (4) Training des Quellklassifizierers. (Martinez-Viol, et al., 2022)

Im ersten Schritt muss ein Referenzmodell erstellt werden. Das Modell muss dabei mit historischen Daten oder Simulationsmodellen anderer Systeme trainiert und getestet werden. Die Daten aus dem Referenzmodell sollten bereits bereinigt sein oder nur leicht formatiert werden müssen. Nachfolgend müssen Daten aus dem Zielsystem erfasst und gekennzeichnet werden. Dafür kann das Wartungsteam mit seinem Wissen und den Instandhaltungsaufzeichnung Stichproben kennzeichnen oder Fehler experimentell erzwingen, um Beispiele für die Fehler zu sammeln. Die Daten müssen anschließend vorverarbeitet werden. Im dritten Schritt müssen die Unähnlichkeiten zwischen Referenzmodell und Zielsystem verringert werden, da das Transfer-Lernen voraussetzt, dass beide Bereiche eine Ähnlichkeit aufweisen. Da dies bei HLK-Systemen nicht gewährleistet werden kann, wird vorgeschlagen zusätzlich zu der Transfer-Lern-Methode die Daten vorher zu filtern. Dabei kommt ein Verfahren zum Einsatz, bei dem die zu kennzeichnenden/zur verwendenden Daten priorisiert werden, sodass sie den größten Einfluss auf das Training des überwachten Modells haben. Im letzten Schritt der Methode wird das Lernen auf das Zielmodell übertragen. Dabei müssen die Daten aus dem Zielsystems erfasst und auf die gleiche Weise ausgeglichen werden wie bei dem Referenzmodell. (Martinez-Viol, et al., 2022)

5.19 Deep-Learning Ansatz zur Fehlerdetektion und -diagnose nach Haruehansapong et al. (2022)

Haruehansapong, et al. (2022) stellen eine Methode zur automatischen Fehlerdetektion und -diagnose (Automatic Fault Detection and Diagnosis (AFDD)) in HLK-Systemen unter der Verwendung von Deep-Learning (DL) vor. Zunächst werden verschiedene Störungsereignisse definiert, die durch defekte HLK-Systeme hervorgerufen werden. Um die Fehler zu erkennen, werden sogenannte Kontextfaktoren verwendet. Dazu werden relevante Umweltfaktoren betrachtet und aus diesen ermittelt, wie die HLK-Systeme Fehler verursachen können. Die AFDD soll mit Hilfe der Kontextfaktoren Fehlerereignisse in den HLK-Systemen vorhersagen. Zur Fehlerdiagnose müssen die Faktoren zunächst in digitale Signale umgewandelt werden. (Haruehansapong, et al., 2022)

Nachfolgend wird ein DL-Modell erstellt, das aus zwei Teilen besteht: der Umgebung und der Systemverarbeitung. Mit Hilfe von Sensoren können die oben beschriebenen Kontextfaktoren kontinuierlich gemessen und in digitale Form umgewandelt werden. Die Systemverarbeitung auf Basis eines ML-Modells läuft anschließend in vier Schritten ab: (1) Signalerfassung, (2) Signalvorverarbeitung, (3) Modelllernen und vorhersagen und (4) Modellbewertung und -visualisierung. Zudem lässt sich die Systemverarbeitung in zwei parallele Teilprozesse unterteilen, den Lernprozess des Modells und den Anwendungsprozess. Als ML-Modell wird ein ANN vorgeschlagen. Um das ML-Modell zu bewerten, wird die Genauigkeit gemessen, die auf Grundlage des Verhältnisses richtiger Vorhersagen (positiv oder negativ) zu allen Vorhersagen (richtig positiv und negativ sowie falsch positiv und negativ) berechnet wird. (Haruehansapong, et al., 2022)

5.20 Gegenüberstellung der relevanten Ansätze mit dem Anforderungsprofil

Im nachfolgenden Abschnitt werden die Ansätze dem Anforderungsprofil aus Kapitel 3 gegenübergestellt. Tabelle 9 stellt dar, welche Anforderungen von den jeweiligen Ansätzen erfüllt sind. Zusätzlich kann der Tabelle die Häufigkeit des Auftretens der einzelnen Anforderungen entnommen werden.

Bei Betrachtung der allgemeinen Anforderungen lässt sich sagen, dass die Anforderung an das Modell wiederverwendbar, allgemeingültig und flexibel (Anforderung 1) zu sein von 95% der relevanten Ansätze erfüllt wird. In einem Ansatz wird diese Anforderung nicht erfüllt. Des Weiteren erfüllen 17 von 19 (89%) Ansätzen die Anforderung benutzerzentriert und lehrreich zu sein (Anforderung 2). Lediglich 5 (26%) der Ansätze bieten verschiedene Lösungselemente (Anforderung 3). Dahingegen erfüllen alle Ansätze die Anforderung der Festlegung des Zwecks (Anforderung 4). Bei den datenbezogenen Anforderungen ist das Erfassen, Speichern und Verarbeiten von Daten in allen Ansätzen erfüllt (Anforderung 6). In 12 von 19 (63%) Ansätzen wird das Modell zudem getestet und validiert (Anforderung 7). In 3 Ansätzen wird die Anforderung der Verbesserung/Anpassung des Datenanalysemodells erfüllt (Anforderung 8). Eine Risikoanalyse des Datenanalysemodell wird in keinem der Ansätze durchgeführt (Anforderung 5). Die Heizungssektor bezogene Anforderung der Einbindung aller Akteure des Heizsektors ist nur ein einem Ansatz gegeben (Anforderung 9). Dahingegen ist die durchgängige Informationsbereitstellung über die Nutzungsdauer des Heizsystems in 17 (89%) der Ansätze erfüllt (Anforderung 10). Zusammenfassend lässt sich feststellen, dass die Anforderungen 1, 2, 4, 6, 7 und 10 in einer Mehrheit der Ansätze bereits berücksichtigt werden. Anforderungen 4 und 6 sind dabei in allen Ansätzen erfüllt. Anforderungen 3, 5, 8 und 9 sind nur in einer Minderheit in den Ansätzen gegeben. Anforderung 5 ist dabei von keinem Ansatz berücksichtigt worden und Anforderung 1 nur von einem Ansatz.

Aus Tabelle 9 sich zudem erkennen, dass die Ansätze zwischen vier und sieben Anforderungen erfüllen. Sechs Ansätze erfüllen sieben der zehn Anforderungen und stellen somit bereits gute Modell für die Zuverlässigkeitsprognose von Heizsystemen dar. Acht weitere Ansätze erfüllen jeweils sechs Anforderungen. In drei Ansätzen werden fünf Anforderungen erfüllt und in zwei Ansätzen lediglich vier. Dementsprechend lässt sich festhalten, dass 14 der relevanten Ansätze bereits eine Mehrzahl (<50%) der beschriebenen Anforderungen erfüllen.

Tabelle 9: Gegenüberstellung der Ansätze mit dem Anforderungsprofil

Anforderung										
Ansatz nach Autor	1	2	3	4	5	6	7	8	9	10
Fernandes, et al. (2020)	•	•	•	•		•	•			
Bouabdallaoui, et al. (2021)	•	•		•		•	•	•		•
Villa, et al., (2022)	•	•		•		•	•			•
Hosamo , et al., (2022)	•	•		•		•	•	•		•
Albayati, et al., (2023)	•	•	•	•		•	•			•
Galvez, et al., (2021)	•	•		•		•	•			•
Mansson, et al., 2021		•		•		•		•		•
Chen, et al., 2022	•			•		•				•
Prabhu & Chaudhary, 2021	•	•		•		•	•			•
Jayatileka & Peters, 2021	•		•	•		•	•		•	•
Ciani, et al., 2020	•	•		•		•				
Li, et al., 2020	•	•		•		•	•			•
Yang, et al., 2020	•	•	•	•		•	•			•
Zhu, et al., 2022	•	•		•		•	•			•
Xie, et al., 2023	•	•		•		•				•
Ahern, et al., 2023	•	•		•		•				•
Borda, et al., 2023	•	•	•	•		•	•			•
Martinez-Viol, et al., 2022	•	•		•		•	•			•
Haruehansapong, et al., 2022	•	•		•		•	•			•
Auftreten (%)	95	89	26	100	0	100	63	16	5	89

6 Zusammenfassung und Ausblick

Das Ziel der vorangegangenen Arbeit war es mittels einer systematischen Literaturrecherche Vorgehens- und Referenzmodell zur Zuverlässigkeitsprognose von Heizsystemen zu finden. Dabei wurden Schwerpunkte auf die Themenfelder Zuverlässigkeitsprognose, vorausschauende Instandhaltung und Datenanalyse gelegt. Zur Erreichung des Ziels wurden in Kapitel 2 zunächst die Grundlagen zu den Themenbereichen der Arbeit erläutert. Dazu wurden in Abschnitt 2.1 Grundlagen der Instandhaltung sowie Instandhaltungsstrategien beschrieben. In Abschnitt 2.2 wurden anschließend Grundlagen der Zuverlässigkeit sowie Zuverlässigkeitskenngrößen vorgestellt. Anschließend wurden in Abschnitt 2.3 verschiedene Definitionen zu Referenzmodellen vorgestellt. In Abschnitt 2.4 wurden zwei Modelle zur Wissensentdeckung in Datenbanken dargestellt. Abschließend wurden in Abschnitt 2.5 Grundlagen zu Heizungsanlagen dargelegt. Die Themenbereiche waren die Instandhaltung, Zuverlässigkeit, Referenzmodelle, Modelle zur Wissensentdeckung in Datenbanken und Heizungsanlagen. In Kapitel 3 wurde in der Folge ein Anforderungsprofil für ein Zuverlässigkeitsprognosemodell erstellt. Anschließend wurde in Kapitel 4 die Methodik der systematischen Literaturanalyse der Arbeit vorgestellt und das Vorgehen konkret beschrieben. Die Literaturrecherche wurde dabei mithilfe von Suchbegriffen auf zehn verschiedenen Datenbanken durchgeführt und die gefundene Literatur in verschiedenen Bewertungsschritten anhand der Relevanz für das Thema beurteilt. Bei der Literaturrecherche wurden 19 relevante Ansätze gefunden, die anschließend in Kapitel 5 dargestellt und anhand des Anforderungsprofils aus Kapitel 3 ausgewertet wurden. Am Ende der Arbeit wurden die Ansätze im Abschnitt 5.20 dem Anforderungsprofil gegenübergestellt und in einer Matrixform dargestellt sowie ausgewertet.

Durch die vorliegende Arbeit ist erkennbar geworden, dass anhand der Bewertung mit dem erstellten Anforderungsprofil bereits gute Ansätze zur Zuverlässigkeitsprognose von Heizsystemen vorliegen. Es fällt allerdings auf das einzelne Anforderungen nur in einer geringen Anzahl der relevanten Ansätze erfüllt werden. So wird in keinem der gefundenen Ansätze eine Risikobewertung des Datenanalysemodells vorgenommen und nur in wenigen Ansätzen wird das Datenanalysemodell nach Inbetriebnahme verbessert oder angepasst. Zudem werden die Akteure des Heizsektors kaum mit in die Zuverlässigkeitsprognose mit eingebunden. Des Weiteren bieten nur wenige Ansätze verschiedene Lösungselemente.

Um ein vollumfängliches Modell zur Zuverlässigkeitsprognose zu erhalten, müssen die Ansätze um die fehlenden Anforderungen ergänzt werden.

7 Literaturverzeichnis

AG Energiebilanzen, 2022. *Umweltbundesamt*. [Online] Available at:

<https://www.umweltbundesamt.de/daten/energie/energieverbrauch-nach-energietraegern-sektoren#allgemeine-entwicklung-und-einflussfaktoren>

[Zugriff am 16 Juni 2023].

AG Energiebilanzen, 2023. *Umweltbundesamt*. [Online] Available at:

<https://www.umweltbundesamt.de/daten/private-haushalte-konsum/wohnen/energieverbrauch-privater-haushalte#hochster-anteil-am-energieverbrauch-zum-heizen>

[Zugriff am 16 Juni 2023].

Ahern, M., O'Sullivan, D. T. J. & Bruton, K., 2023. Implementation of the IDAIC Framework on an Air Handling Unit to Transition to Proactive Maintenance. *Energy & Buildings*, Februar, Band 284.

Albayati, M. G. et al., 2023. Semi-Supervised Machine Learning for Fault Detection and Diagnosis of a Rooftop Unit. *Big Data Mining and Analytics*, Juni, 6(2), S. 170-184.

Arora, S.-J. & Rabe, M., 2023a. Approach to Reference Models for Building Performance Simulation: Establishing Common Understanding. *SN Computer Science*, März, 4.

Arora, S.-J. & Rabe, M., 2023b. Predictive Maintenance: Assessment of Potentials for Residential Heating. *International Journal of Computer Integrated Manufacturing*, April.

Bauernhansl, T., Krüger, J., Reinhart, G. & Schuh, G., 2016. WGP-Standpunkt Industrie 4.0. Wissenschaftliche Gesellschaft für Produktionstechnik e. V..

Bertsche, B. et al., 2009. Zuverlässigkeit mechatronischer Systeme. Grundlagen und Bertung in frühen Entwicklungsphasen. Berlin, Heidelberg: Springer.

Bertsche, B. & Lechner, G., 2004. Zuverlässigkeit im Fahrzeug- und Maschinenbau. Ermittlung von Bauteil- und Systemzuverlässigkeit. 3. Auflage Hrsg. Berlin, Heidelberg, New York: Springer.

Borda, D. et al., 2023. Development of Anomaly Detectors for HVAC Systems Using Machine Learning. *Processes*, Februar, 11(2).

Bouabdallaoui, Y. et al., 2021. Predictive Maintenance in Building Facilities: A Machine Learning-Based Approach. *Sensors*, Februar, 21(4).

- Bundesverband der Energie- und Wasserwirtschaft, 2019. *BDEW Studie "Wie heizt Deutschland?"*. [Online] Available at: <https://www.bdew.de/energie/studie-wie-heizt-deutschland/> [Zugriff am 19 Juni 2023].
- Bundesverband Wärmepumpe e. V., 2022. *Wie funktioniert die Wärmepumpe?*. [Online] Available at: <https://www.waermepumpe.de/waermepumpe/funktion-waermequellen/> [Zugriff am 19 Juni 2023].
- Chapman, P. et al., 2000. CRISP-DM 1.0. Step-by-step data mining guide.
- Chen, Y., Crowe, E., Granderson, J. & Lin, G., 2022. Intergration of FDD Data to Aid HVAC System Maintenance. New York, USA, ACM.
- Ciani, L., Guidi, G. & Galar, D., 2020. Reliability evaluation of an HVAC ventilation system with FTA and RBD analysis. Wien, Österreich, IEEE.
- Cleve, J. & Lämmel, U., 2020. Data Mining. 3. Auflage Hrsg. Berlin, Boston: De Gruyter.
- DIN 31051, 2019. Instandhaltung. Berlin: Beuth Verlag.
- DIN EN 13306, 2017. Instandhaltung - Begriffe der Instandhaltung. Berlin: Beuth Verlag.
- Europäische Kommission, 2021. Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetze über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union
- Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P., 1996. From Data Mining to Knowledge Discovery in Databases. 17(3), S. 37-54.
- Feldmann, S., Rauen, H., Herweg, O. & Synek, P.-M., 2017. Predictive Maintenance. Service der Zukunft - und wo er wirklich steht. München: Roland Berger GmbH.
- Fernandes, S. et al., 2020. Forecasting Appliances Failures: A Machine-Learning Approach to Predictive Maintenance. *Information*, April, 11(4).
- Galvez, A., Diez-Olivan, A., Seneviratne, D. & Galar, D., 2021. Fault Detection and Estimation for Railway HVAC Systems Using a Hybrid Model-Based Approach. *Sustainability*, Juni, 13(12).
- Hainbach, C., 2018. Systeme und Bauteile der Heizungstechnik. In: K. Grote, B. Bender & D. Göhlich, Hrsg. *Dubbel. Taschenbuch für den Maschinenbau*. 25. Auflage Hrsg. Berlin, Heidelberg: Springer Vieweg, S. 942-952.

- Haruehansapong, K. et al., 2022. Deep Learning-Driven Automated Fault Detection and Diagnostics Based on a Contextual Environment: A Case Study of HVAC System. *Buildings*, Dezember, 13(1).
- Hodapp, W., 2018. Die Bedeutung einer zustandsorientierten Instandhaltung. In: J. Reichel, G. Müller & J. Haeffs, Hrsg. *Betriebliche Instandhaltung*. 2. Auflage Hrsg. Berlin: Springer Vieweg, S. 135-152.
- Hosamo, H. H. et al., 2022. A Digital Twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics. *Energy & Buildings*, April, Band 261.
- Jayatilleka, S. & Peters, J., 2021. Reliability Modelling in HVAC & Other Industry Equipment Using Big Data. Orlando, USA, IEEE.
- Karuppiyah, K., Sankaranarayanan, B. & Ali, S. M., 2020. On Sustainable Predictive Maintenance: Exploration of Key Barriers Using an Intergrated Approach. *Sustainable Production and Consumption*, Juli, 27(3), S. 1537-1553.
- Kitchenham, B., 2004. Procedures for Performing Systematic Reviews. *Procedures for Performing Systematic Reviews*, August, S. 1-33.
- Laasch, T. & Laasch, E., 2013. Haustechnik. Grundlagen Planung Ausführung. 13. Auflage Hrsg. Wiesbaden: Springer Vieweg.
- Lauckner, G. & Krimmling, J., 2020. Raum- und Gebäudeautomation für Architekten und Ingenieure. Grundlagen-Orientierungshilfen-Beispiele. Wiesbaden: Springer Vieweg.
- Li, M. et al., 2020. A Data-Driven Method for Fault Detection and Isolation of the Integrated Energy-Based District Heating System. *IEEE Access*, Januar, 8.
- Mansson, S., Thern, M., Johansson Kallioniemi, P.-O. & Sernhed, K., 2021. A Fault Handling Process for Faults in District Heating Customer Installations. *Energies*, Mai, 14(11).
- Martinez-Viol, V. et al., 2022. Semi-Supervised Transfer Learning Methodology for Fault Detection and Diagnosis in Air-Handling Units. *Applied Sciences*, September, 12(17).
- Pawellek, G., 2013. Integrierte Instandhaltung und Ersatzteillogistik. Vorgehensweisen, Methoden, Tools. Berlin, Heidelberg: Springer Vieweg.
- Prabhu, V. & Chaudhary, D., 2021. Machine Learning enabled Condition Monitoring Models for Predictive Maintenance of Boilers. Noide, India, IEEE.

- Rabe, M., Jaekel, F.-W. & Weinaug, H., 2006. Reference Models for Supply Chain Design and Configuration. Monterey, USA, IEEE.
- Rosemann, M. & Schütte, R., 1997. Grundsätze ordnungsmäßiger Referenzmodellierung. In: J. Becker, M. Rosemann & R. Schütte, Hrsg. *Entwicklungsstand und Entwicklungsperspektiven der Referenzmodellierung: Proceedings zur Veranstaltung vom 10. März 1997*. Münster: Institut für Wirtschaftsinformatik.
- Ryll, D.-I. F. & Freund, D.-I. C., 2010. Grundlagen der Instandhaltung. In: M. Schenk, Hrsg. *Instandhaltung technischer Systeme*. Berlin, Heidelberg: Springer, S. 23-101.
- Schnieder, L., 2023. Qualitative Methodik zur Absicherung von KI-Systemen für den bodengebundenen Verkehr. Nürtingen, VDI, S. 271-282.
- Schuh, G. et al., 2020. Predictive Maintenance. Entwicklung vorausschauender Wartungssysteme für Werkzeugbaubetriebe und Serienproduzenten. Aachen: Werkzeugmaschinellabor RWTH Aachen.
- Sharafi, A., 2013. Knowledge Discovery in Databases. Eine Analyse des Änderungsmanagements in der Produktentwicklung. Wiesbaden: Springer Gabler.
- Strunz, M., 2012. Instandhaltung. Grundlagen - Strategien - Werkstätten. Berlin, Heidelberg: Springer.
- Thomas, O., 2006. Understanding the Term Reference Model in Information Systems Research: History, Literature Analysis and Explanation. In: C. Bussler & A. Haller, Hrsg. *Business Process Management Workshops*. Berlin, Heidelberg: Springer, S. 484-496.
- Van der Aalst, W. et al., 2006. Configurable Process Models as a Basis for Reference Modeling. In: C. Bussler & A. Haller, Hrsg. *Business Process Management Workshops*. Berlin, Heidelberg: Springer, S. 512-518.
- VDI 2884, 2005. Beschaffung, Betrieb und Instandhaltung von Produktionsmitteln unter Anwendung von Life Cycle Costing (LCC). Düsseldorf: Verein Deutscher Ingenieure.
- Villa, V. et al., 2022. Machine Learning Framework for the Sustainable Maintenance of Building Facilities. *Sustainability*, Januar, 14(2).
- Vom Brocke, J. et al., 2009. Reconstructing the Giant: on the Importance of Rigour in Documenting the Literature Search Process. *ECIS 2009 Proceedings*, Januar.

- Webster, J. & Watson, R. T., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, Juni, 26(2), S. 13-23.
- Xie, X. et al., 2023. Digital Twin Enabled Fault Detection and Diagnosis Process for Building HVAC Systems. *Automation in Construction*, Februar, 146.
- Yang, C., Gunay, B., Shi, Z. & Shen, W., 2020. Machine Learning-Based Prognostics for Central Heating and Cooling Plant Equipment Health Monitoring. *IEEE Transactions on Automation Science and Engineering*, Juni, 18(1), S. 346-355.
- Zhu, H., Yang Wen, Li, S. & Pang, A., 2022. An Effective Fault Detection Method for HVAC Systems Using the LSTM-SVDD Algorithm. *Buildings*, Februar, 12(2).

Anhang

Anhang A: Mappe1

Anhang B: Mappe2