



Fakultät II – Informatik, Wirtschafts- und Rechtswissenschaften  
Department für Informatik

# **Intelligente Systeme zur Verbesserung der Qualität von IEC 62559 Use Cases und SGAM Architekturmodellen**

Von der Fakultät für Informatik, Wirtschafts- und Rechtswissenschaften der Carl von  
Ossietzky Universität Oldenburg zur Erlangung des Grades und Titels

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

angenommene Dissertation

von Herrn Rene Andreas Kuchenbuch

geboren am 28. September 1996 in Koblenz

**Gutachter:**

Prof. Dr.-Ing. Jürgen Sauer

**Weitere Gutachter:**

Prof. Dr. Sebastian Lehnhoff

**Tag der Disputation:** 25. Februar 2026

---

## Zusammenfassung

Die IEC 62559 Use Case-Methodik und das Smart Grid Architecture Model (SGAM) Framework stellen etablierte Techniken des Requirements Engineering innerhalb energiebezogener Forschungs- und Entwicklungsprojekte dar und dienen zur Beschreibung des System-of-Concerns (SoCs). Bei der Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM-Modellen können jedoch Fehler aufgrund unterschiedlicher Interpretationen des System-of-Concerns durch heterogene Stakeholder und deren Hintergründe auftreten, was kostspielige Konsequenzen in späteren Projektphasen haben kann. Um diesem Problem entgegenzuwirken, untersucht diese Dissertation das Potenzial von intelligenten Systemen innerhalb der vorgestellten Requirements Engineering-Methoden. Ziel ist es, die Qualität von IEC 62559 Use Case-Beschreibungen und SGAM-Modellen zu steigern sowie deren Erstellung zu vereinfachen. Dies soll das gemeinsame Verständnis des SoC fördern.

Im Rahmen der Design Science Research Methodology werden zur Beantwortung der Forschungsfragen KI-Prototypen konzipiert, implementiert und evaluiert. Forschungsgegenstand ist die Untersuchung, wie KI-Systeme für Fachexperten entworfen werden können und inwiefern diese zur Qualitätsverbesserung von IEC 62559 Use Cases und SGAM Modellen beitragen. Hierfür werden Qualitätseigenschaften- und Metriken anhand der Grundsätze ordnungsmäßiger Modellierung (GoM) entworfen, die als Evaluierungsgrundlage für die zu konzipierenden KI-Systeme vorgesehen sind.

Die Ergebnisse dieser Forschung tragen dazu bei, die Qualität und Effizienz der angewendeten Requirements Engineering-Verfahren in der Entwicklung von Smart Grid-Lösungen zu verbessern und das gemeinsame Verständnis des SoC bei den beteiligten Stakeholdern zu fördern. Dadurch werden zuverlässigere und interoperablere Lösungen ermöglicht. Darüber hinaus zielt das Forschungsvorhaben darauf ab, die toolgestützte Erstellung von IEC 62559 Use Cases und SGAM Modellen auf eine innovative Ebene zu heben.



## Abstract

The IEC 62559 Use Case Methodology and the Smart Grid Architecture Framework (SGAM) represent established techniques of requirements engineering within energy-related research and development projects and are used to describe the system-of-concern. However, errors can occur due to different interpretations of the system-of-concern by heterogeneous stakeholders and their backgrounds, when creating use case descriptions according to IEC 62559 and SGAM models, leading to costly consequences in later project phases. To address this problem, this dissertation explores the potential of intelligent systems in specified requirement engineering methods to improve the quality of IEC 62559 use case descriptions and SGAM models in simplifying its creation to promote common understanding of the system-of-concern.

Within the framework of the Design Science Research Methodology, AI prototypes are designed, implemented, and evaluated to address the research questions. The research focuses on exploring how AI systems can be designed for domain experts and to what extent they contribute to the improvement of the quality of IEC 62559 Use Cases and SGAM models. Quality attributes and metrics are developed based on the principles of sound modeling practices (GoMs) to serve as evaluation criteria for the proposed AI systems.

The findings of this research contribute to enhancing the quality and efficiency of the applied Requirement Engineering techniques in the development of Smart Grid solutions and fostering a shared understanding of the System-of-Concern among stakeholders. This, in turn, enables the creation of more reliable and interoperable solutions. Moreover, the research endeavor aims to advance the tool-based creation of IEC 62559 Use Cases and SGAM models to an innovative level.



---

## Danksagung

Meine ersten offiziellen Worte zu meinem Promotionsvorhaben im OFFIS waren, dass es sich um einen Marathon handelt. Dieser Vergleich trifft es sehr gut, denn ebenso wie im Sport gilt auch in der Wissenschaft, dass der Weg in einer Gemeinschaft deutlich mehr Freude bereitet. So hatte ich im Verlauf meiner Promotion mit vielen Weggefährten zu tun und habe erlebt, wie Personen in meiner Forschungsgruppe kamen und gingen. Die prägendste Person, mit der ich den intensivsten fachlichen Austausch führte, war Johann Schütz. Mit Johann habe ich jederzeit - und gerne auch an längeren Abenden, begleitet von ein oder zwei ungesunden Kaltgetränken - über zahlreiche Themen und Theorien philosophiert, kritisch, gelegentlich zynisch, aber immer zielführend. Dieser Austausch hat mir maßgeblich geholfen, mein Thema zu schärfen und einzuschlagen. Dafür danke ich ihm herzlich.

Auch in meiner Forschungsgruppe „Standardized Systems Engineering and Assessment“ sowie am OFFIS durfte ich auf viele Weggefährten und Unterstützer zählen. Laura Niemann, Arlena Wellßow, Marcel Otte, Meng Lin, Mana Azamat, Dr. Saskia Rathjen, Dr. Carsten Krüger sowie mein Gruppenleiter Dr. Oliver Werth haben mich aktiv in der Gruppe unterstützt, sei es durch hilfreiche Ratschläge, das Teilen ihrer wissenschaftlichen Erfahrungen, methodische Unterstützung, inhaltliche Diskussionen, das Freihalten des Rückens oder einfach durch ein offenes Ohr und die Zeit, die sie sich für mich genommen haben. Ihr alle habt mir im Verlauf dieses Marathons immer wieder den nötigen Motivationsschub gegeben. Mein Dank gilt zudem Dr. Mathias Uslar, der mich als ehemaliger Gruppenleiter und als wissenschaftlicher Leiter durch die Bereitstellung der finanziellen Förderung unterstützt hat.

Während meiner Promotionszeit durfte ich zudem zahlreiche externe Projektpartner und themenverwandte Kolleginnen und Kollegen kennenlernen. Besonders hervorheben möchte ich das Projekt int:net, das einen äußerst bereichernden fachlichen Austausch rund um das Thema Interoperabilität ermöglichte. Mein Dank gilt hierbei insbesondere den Projektpartnern Carlos Ayon Mac Gregor, Joep van Genuchten, Laia Guitart, Ludwig Karg, Sebastian Kosslers, Dr. Gianluca Lipardi, Dr. Maider Santos Mugica, Beate Schmidt, Priv.-Doz. Dr. Thomas I. Strasser und Markus Stroot sowie allen weiteren Beteiligten. Eine besonders prägende Konferenzbekanntschaft, aus der sich ein toller Austausch entwickelte, war Dominik Vereno. Mit niemandem sonst konnte ich über institutionelle Grenzen hinweg so leidenschaftlich und ausführlich über das SGAM fachsimpeln wie mit Dominik.

Ich möchte mich außerdem herzlich bei meinem Erstgutachter Prof. Dr.-Ing. Jürgen Sauer für die hervorragende Betreuung, insbesondere für seine Geduld und Ruhe, bedanken. Der Austausch mit ihm war stets sehr entspannt und zugleich zielführend - gemeinsamer Kaffeekonsum inklusive - und fand immer auf Augenhöhe statt. Ebenso bin ich Prof. Dr. Sebastian Lehnhoff als meinem Zweitgutachter sehr dankbar für seine Unterstützung in den Themen der Energieinformatik sowie für die Möglichkeit und das entgegengebrachte Vertrauen durch einen Lehrauftrag an der Universität Oldenburg,

im Rahmen dessen ich wertvolle Erfahrungen in der Durchführung von Vorlesungen in den Interoperabilitätsthemen sammeln durfte. Mein herzlicher Dank gilt zudem Prof. Dr. Wolfram Wingerath für die Übernahme des Vorsitzes der Prüfungskommission sowie Dr. Marco Grawunder für seine Mitwirkung als wissenschaftlicher Mitarbeiter.

Ein besonderer Dank gilt Dave, der mich während der gesamten Promotionszeit in außergewöhnlichem Maße begleitet, unterstützt und mir stets Rückhalt gegeben hat. Seine beständige Präsenz, die vielen Gespräche und sein Verständnis haben diese Zeit für mich maßgeblich geprägt. Im privaten Freundes- und Bekanntenkreis möchte ich mich zudem bei David und Christian bedanken. Ihr habt mir auf unterschiedliche Weise geholfen, die Promotionszeit so angenehm wie möglich zu gestalten, mir Ausgleich zum wissenschaftlichen Alltag verschafft und mir stets ein offenes Ohr geboten.

*An letzter Stelle möchte ich diese Arbeit meinen Eltern, Ute Kuchenbuch († 2022) und Andreas Kuchenbuch († 2022), widmen, die leider nur den Anfang meiner Promotion miterleben konnten, mir dieses Vorhaben jedoch auch als erstem Akademikerkind der Familie zugetraut haben. Ihr habt mich im Leben stets nach besten Kräften unterstützt, mir Rückhalt gegeben und mir den Mut vermittelt, meinen eigenen Weg zu gehen. Ich weiß, dass ihr stolz auf mich wärt und diese Arbeit ohne eure Unterstützung nicht möglich gewesen wäre.*

# Inhalt

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Problemstellung und Zielsetzung . . . . .	3
1.2	Forschungskontext . . . . .	5
1.3	Verwandte Arbeiten . . . . .	8
1.4	Aufbau der Dissertation . . . . .	11
<b>2</b>	<b>Methodische Herangehensweise</b>	<b>13</b>
2.1	Anforderungen an das Vorgehensmodell . . . . .	13
2.2	Design Science . . . . .	14
2.3	Vorgehensmodell der Arbeit . . . . .	16
2.4	Verwendete Definition für Künstliche Intelligenz . . . . .	19
<b>3</b>	<b>Grundlagen im Smart Grid Kontext</b>	<b>23</b>
3.1	Smart Grid als System-of-Systems . . . . .	24
3.1.1	System-of-Systems Problematik . . . . .	24
3.1.2	Interoperabilität . . . . .	25
3.1.3	Architekturbeschreibung nach ISO/IEC/IEEE 42010 . . . . .	27
3.2	Anwendungsfallmethodik nach IEC 62559 (Use Case Methodology) . . . . .	29
3.2.1	IEC 62559 Use Cases . . . . .	30
3.2.2	Use Case XML-Serialisierung . . . . .	36
3.2.3	Use Case Repository (UCR) . . . . .	37
3.3	Smart Grid Architecture Model (SGAM)-Famework . . . . .	40
3.3.1	Komponenten . . . . .	41
3.3.2	SGAM Modelle . . . . .	44
3.3.3	Unterstützungswerkzeuge und Serialisierung . . . . .	46
<b>4</b>	<b>Grundlagen im KI-Kontext</b>	<b>49</b>
4.1	Machine Learning (ML) . . . . .	49
4.1.1	Überwachtes Lernen . . . . .	51
4.1.2	Künstliche neuronale Netze (KNNs) und Deep Learning . . . . .	53
4.2	Natural Language Processing (NLP) . . . . .	55
4.2.1	Textrepräsentation und -Einbettungstechniken . . . . .	58
4.2.2	Transformer . . . . .	60
4.3	(Large) Language Models . . . . .	63
4.3.1	Bidirectional Encoder Representations from Transformers (BERT) . . . . .	64
4.3.2	Robustly Optimized BERT Pretraining Approach (RoBERTa) . . . . .	66
4.3.3	Decoding-enhanced BERT with Disentangled Attention (DeBERTa) . . . . .	66

4.3.4	Generative Pretrained Transformer (GPT) . . . . .	66
4.3.5	FLAN-T5 . . . . .	67
4.3.6	Large Language Model Meta AI (LLaMA) . . . . .	68
4.4	Betrachtung der Anwendungsbereiche . . . . .	68
4.4.1	Anwendungsbereich: Textklassifikation . . . . .	69
4.4.2	Anwendungsbereich: Textregression . . . . .	71
4.4.3	Anwendungsbereich: Tokenklassifikation . . . . .	71
4.4.4	Anwendungsbereich: Textgenerierung . . . . .	73
4.5	Betrachtung der etablierten Techniken . . . . .	74
4.5.1	Transfer Learning (TL) . . . . .	75
4.5.2	Text Data Preprocessing . . . . .	75
4.5.3	Text Data Augmentation . . . . .	77
4.5.4	eXplainable Artificial Intelligence (XAI) . . . . .	78
4.5.5	Retrieval Augmented Generation (RAG) . . . . .	80
<b>5</b>	<b>Anforderungsanalyse von SGAAIRE</b>	<b>83</b>
5.1	Anforderungs- und Kontexterhebung: Teilstandardisierter Fragebogen . . . . .	84
5.1.1	Aufbau und Entwicklung der Umfrage . . . . .	84
5.1.2	Durchführung der Umfrage . . . . .	85
5.1.3	Datenaufbereitung- und Analyse . . . . .	86
5.2	Nicht-funktionale Anforderungen . . . . .	95
5.3	Funktionale Anforderungen . . . . .	97
5.3.1	Anforderungspriorisierung: Nutzwertanalyse (NWA) . . . . .	97
5.3.2	Anforderungsselektion . . . . .	106
<b>6</b>	<b>Konzeptionalisierung der Bewertungsmetriken</b>	<b>109</b>
6.1	Adaptierung des Qualitäts- und Fehlerbegriffes . . . . .	110
6.2	Anforderungen an IEC 62559-2 Use Case-Beschreibungen und SGAM Modelle . . . . .	111
6.3	Herleitung der (primären) Qualitätsdimensionen . . . . .	112
6.3.1	Bestehende Datenqualitätsmerkmale . . . . .	112
6.3.2	Herausforderungen bei Anwendung des Modellverständnisses . . . . .	115
6.3.3	Qualitätsdimensionen- und Metriken . . . . .	116
6.4	Bewertungsmetriken für Qualitätsdimensionen . . . . .	117
<b>7</b>	<b>Konzeptionalisierung von SGAAIRE</b>	<b>121</b>
7.1	Zugrundeliegender Prozess . . . . .	122
7.1.1	Betrachtung der Herausforderungen . . . . .	124
7.1.2	Einordnung in der Projektlandschaft . . . . .	126

---

7.2	Statisches Systemkonzept . . . . .	126
7.2.1	Systemkontext . . . . .	127
7.2.2	Systemarchitektur . . . . .	130
7.3	Dynamisches Systemkonzept . . . . .	131
7.3.1	Anwendungsfälle . . . . .	132
7.3.2	Interaktionen . . . . .	133
7.4	Konzeptualisierung der Anwendungsfälle . . . . .	136
7.4.1	KI-Modul für Klassifikationsprobleme . . . . .	136
7.4.2	KI-Modul für Fehlererkennungen . . . . .	137
7.4.3	KI-Modul für Generierungsaufgaben . . . . .	142
7.5	Konzeptionalisierung der KI-Modelle . . . . .	147
7.5.1	SGAM Domänen Klassifikator (M1) . . . . .	148
7.5.2	SGAM Zonen Klassifikator (M2) . . . . .	149
7.5.3	SGAM Interoperabilitätslayer Klassifikator (M3) . . . . .	149
7.5.4	Service Step Klassifikator (M4) . . . . .	150
7.5.5	SGAM Element Proposer (M5) . . . . .	150
7.5.6	Abstraktionsgrad Regressor (M6) . . . . .	152
7.5.7	Acronym Identification (M7a) . . . . .	152
7.5.8	Acronym Disambiguation (M7b) . . . . .	153
<b>8</b>	<b>Implementierung von SGAAIRE</b>	<b>155</b>
8.1	Systemumsetzung . . . . .	156
8.1.1	Programmiersprache . . . . .	156
8.1.2	Bibliotheken und Tools . . . . .	156
8.1.3	Laufzeitumgebung . . . . .	157
8.1.4	Pufferung- und Cachingverfahren für Inferenzen . . . . .	158
8.1.5	SGAAIRE Application Programming Interfaces (API) . . . . .	159
8.1.6	Integration im Use Case Management Repository (UCMR) . . . . .	161
8.2	Datenerhebung- und Überführung . . . . .	163
8.2.1	Datenerhebung . . . . .	164
8.2.2	Datenüberführung . . . . .	166
8.2.3	Datenbereitstellung . . . . .	169
8.2.4	Künstliche Datenanreicherung mittels Data Augmentation . . . . .	171
8.3	Feinabstimmung, Selektion und Integration von KI-Modellen . . . . .	173
8.3.1	Selektion der vortrainierten Sprachmodelle . . . . .	173
8.3.2	Fine-Tuning der vortrainierten Sprachmodelle . . . . .	173
8.3.3	Bewertungskriterien für die Selektion der KI-Modelle . . . . .	176
8.3.4	SGAM Domänen Klassifikator (M1) . . . . .	177
8.3.5	SGAM Zonen Klassifikator (M2) . . . . .	180

8.3.6	SGAM Interoperabilitätslayer Klassifikator (M3)	181
8.3.7	Service Step Klassifikator (M4)	182
8.3.8	SGAM Element Proposer (M5)	183
8.3.9	Abstraktionsgrad Regressor (M6)	186
8.3.10	Acronym Identification und -Disambiguation (M7)	188
<b>9</b>	<b>Evaluation und Diskussion</b>	<b>191</b>
9.1	Demonstration	192
9.1.1	Fallbeispiel A: Optimierung der Lebensdauer von (smarten) Transformatoren mithilfe von Smart Services	192
9.1.2	Fallbeispiel B: Verwendung von dynamischen Stromtarifen für Electric Vehicle (EV)-Ladestationen	197
9.2	Evaluation	201
9.2.1	Durchführung der randomisierten kontrollierten Studie	205
9.2.2	Einfluss von SGAAIRE auf die Identifikation eingebauter Fehler	208
9.2.3	Einfluss von SGAAIRE auf die Identifikation nicht-eingebauter Fehler	209
9.2.4	Einfluss von SGAAIRE auf die Qualitätsverbesserung	211
9.2.5	Einfluss der eingeschätzten Expertise	212
9.2.6	Systembewertung der User Experience	215
9.3	Diskussion	217
9.3.1	Einordnung der Studienergebnisse	218
9.3.2	Weiterführende Einordnung	219
9.3.3	Limitierungen	220
9.4	Beiträge aus der Arbeit	222
9.4.1	Wissenschaftliche Beiträge	222
9.4.2	Technische Beiträge	225
<b>10</b>	<b>Zusammenfassung und Schluss</b>	<b>229</b>
10.1	Fazit	229
10.2	Ausblick	232
<b>A</b>	<b>Anhang</b>	<b>235</b>
A.1	Erklärung zum Online-Anhang und GitLab Projekten	235
A.2	Anwendungsfallbeschreibung gemäß IEC 62559-2: #1 Optimierung der Lebensdauer von (smarten) Transformatoren mithilfe von Smart Services	236
A.3	Anwendungsfallbeschreibung gemäß IEC 62559-2: #2 Verwendung von dynamischen Stromtarifen für EV-Ladestationen	245
A.4	Anhang zur Nutzwertanalyse	253
A.4.1	Tabellen zur NWA	253

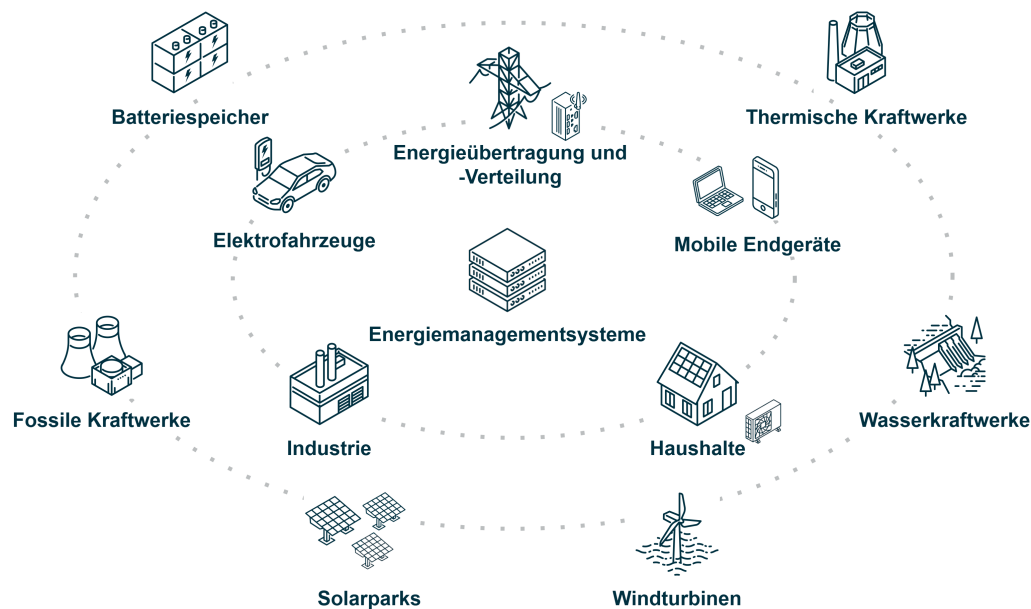
---

A.4.2	Berechnungsdatei zur Nutzwertanalyse . . . . .	255
A.5	Weitere Darstellungen aus den Demonstrationen . . . . .	257
A.6	Weitere Darstellungen aus der Evaluation . . . . .	258
A.7	Anwendungsfallbeschreibungen zu Smart Grid Assisting Artificial Intelligence in Requirements Engineering (SGAAIRE) . . . . .	259
<b>Abkürzungen</b>		<b>267</b>
<b>Abbildungen</b>		<b>271</b>
<b>Tabellen</b>		<b>274</b>
<b>Literatur</b>		<b>279</b>



# 1 Einleitung

Die gegenwärtige Ära der Klimakrise zwingt den Energiesektor dazu, sich verstärkt mit der Entwicklung von Smart Grids auseinanderzusetzen. Im Gegensatz zu traditionellen Energienetzen beruht die Grundidee hinter intelligenten Energienetzen auf Dezentralität und der Integration von Informations- und Kommunikationstechnologien (IKT). Das Ziel ist es, Energieverteilung und -übertragung effizienter zu gestalten und gleichzeitig Energieverluste zu minimieren [USD<sup>+</sup>13]. Zudem sollen Produzenten und Konsumenten in einer neuen Rolle als sogenannte „Prosumer“ in das Energienetz eingebunden werden [USD<sup>+</sup>13], was in Abbildung 1.1 verdeutlicht wird. Dies führt zu einer noch komplexeren Landschaft von Akteuren, Technologien und Geschäftsmodellen, die innerhalb eines System-of-Systems miteinander agieren [USD<sup>+</sup>13].



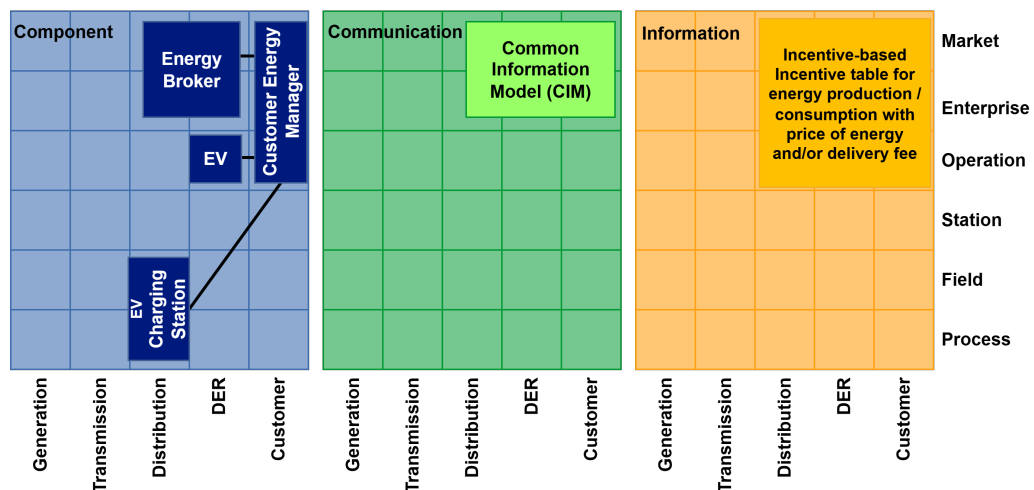
Quelle: Eigene Darstellung

Abbildung 1.1: Darstellung möglicher Akteure im Smart Grid

Die Gewährleistung der Interoperabilität zwischen diesen Systemen ist von entscheidender Bedeutung und erfordert zudem Standards und ein adäquates Requirements Engineering, um das gemeinsame Verständnis über das zu betrachtende System-of-Concern zu erhöhen [USD<sup>+</sup>13]. Wesentliche Vertreter in energiebezogenen Projekten stellen hierfür die IEC 62559 Use Case-Methodik (dt. Anwendungsfallmethodik) [USD<sup>+</sup>13] und das SGAM Framework [USD<sup>+</sup>13] dar, welche hochgradig zueinander stehen [FSBU20]. Die Anwendung der IEC 62559 Use Case-Methodik zur Erstellung von IEC 62559 Use

Case-Beschreibungen gemäß dem in IEC 62559-2 definierten Template [IEC15] sowie die Nutzung des SGAM Frameworks zur Erstellung von SGAM Modellen [CEN14] haben sich im Energiesektor etabliert und sind in einigen nationalen und europäischen Forschungs- und Innovationsprojekten im Energiebereich (z. B. H2020 und Horizon Europe) vertragliche Grundlagen [RSJ<sup>+</sup>23]. Bei der Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen können jedoch Fehler aufgrund unterschiedlicher Interpretationen des System-of-Concerns durch heterogene Stakeholder und deren Hintergründe [PKML22] sowie die Betrachtung unterschiedlicher Abstraktions- sowie Granularitätsebenen auftreten.

Ein Beispiel für ein fehlerhaftes SGAM Modell ist in Abbildung 1.2 dargestellt. Dort sind unter anderem Verbindungen auf der Komponentenebene falsch gesetzt, Elemente falsch auf Domänen und Zonen zugeordnet, Standards sowie Kommunikations- bzw. Informationsangaben auf der Kommunikations- und Informationsebene fehlerhaft oder unvollständig, textuelle Beschreibungen schwer interpretierbar und weitere Unstimmigkeiten erkennbar. Auch bei IEC 62559 Use Case-Beschreibungen können Akteure, Standards oder Referenzen fehlen, fehlerhaft oder redundant sein, Klassifikationsangaben inkorrekt erfolgen oder Texte undefinierte Begriffe enthalten. Diese Beispiele verdeutlichen lediglich einen potenziellen Ausschnitt des möglichen Fehlerrahms.



Quelle: Eigene Darstellung

Abbildung 1.2: Ausschnitt eines SGAM Modells mit potenziellen Fehlern

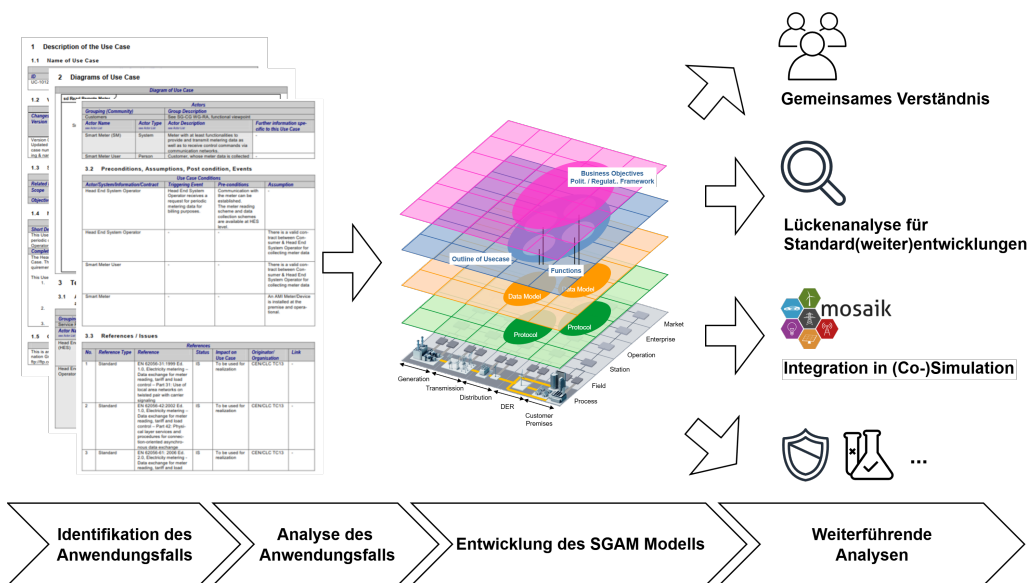
Fehler, die aus einem inkorrekten Verständnis der (System-)Anforderungen resultieren, können sich in späteren Projektphasen vervielfachen und ungeplante Mehrkosten sowie weitere negative Nebeneffekte verursachen [Fin94]. Werden solche Fehler nicht bereits in der Anforderungsphase, sondern erst im operativen Einsatz entdeckt, können die daraus resultierenden Kosten bis zu hundertmal höher liegen [HSD<sup>+</sup>04]. Neben wirtschaftlichen Folgen können zudem in kritischen Systemen technische Folgen, wie

Stromausfälle oder auch Blackouts aus Fehlern resultieren [PKT06]. Die exemplarisch benannten Fehler, insbesondere Unvollständigkeiten, Redundanzen und inkorrekte Angaben, können zu Fehlannahmen und unangemessenen Designentscheidungen führen und im schlimmsten Fall in nicht interoperablen Lösungen resultieren, die erhebliche Auswirkungen auf die Gesamtkosten haben.

### 1.1 Problemstellung und Zielsetzung

Die Gewährleistung einer hohen Qualität von IEC 62559 Use Cases und SGAM Modellen ist von zentraler Bedeutung, da diese dazu beiträgt, ein einheitliches Verständnis des System-of-Concern zu schaffen. Durch ein gemeinsames Verständnis können Missverständnisse und Fehlinterpretationen vermieden werden, was wiederum zu einer effektiveren Zusammenarbeit der beteiligten Stakeholder führt. Gleichzeitig verfolgen Fachexperten mit IEC 62559 Use Cases und SGAM Modellen unterschiedliche Ziele und Interessen (vgl. Abbildung 1.3), etwa die Analyse von Standardisierungslücken (insb. in den frühen Jahren) [CEN12], das Schaffen eines gemeinsamen Verständnisses [GDU17], die Integration in (Co-)Simulationen zur Abbildung dezentraler Akteure [VHN23] sowie weitere Anwendungen wie Sicherheitsanalysen [URS14] oder Testing [EBB<sup>+</sup>19].

Diese Dissertation basiert auf der Annahme, dass der Einsatz von intelligenten Systemen bei der Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen durch Fachexperten im Vergleich zur herkömmlichen Arbeitsweise zu einer Mini-



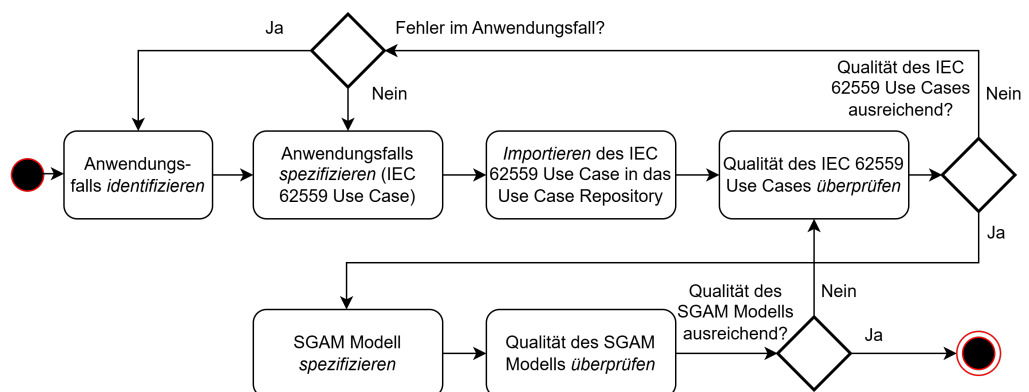
Quelle: Eigene Darstellung, angelehnt an [CEN12, GDU17]

Abbildung 1.3: Abhängigkeiten zwischen IEC 62559 Use Cases, SGAM Modellen und weiterführenden Tätigkeiten

mierung von Fehlern und Verbesserung der Erfüllungsgrade der Qualitätseigenschaften führen kann. In jüngerer Zeit hat sich insbesondere das Potenzial von Künstliche Intelligenz (KI) im Requirements Engineering deutlich gezeigt, wie [NBAKK17] hervorheben. Darüber hinaus zeichnet sich gegenwärtig eine starke Entwicklung in Bezug auf generative KI sowie Large Language Model ab, exemplarisch vertreten durch Technologien wie ChatGPT [AGA23]. Diese Entwicklungen unterstreichen die zunehmende Bedeutung und die vielfältigen Einsatzmöglichkeiten der KI in der Gestaltung und Verbesserung von Prozessen im Requirements Engineering, was neue Perspektiven für Forschung und Anwendung eröffnet. Exemplarische kontextbezogene Anwendungsfälle für den Einsatz von intelligenten Systemen könnten sein:

- Überführen von IEC 62559 Use Case-Beschreibungen in SGAM Modelle im Sinne eines Erstentwurfs (Arbeitsmodelle)
- Entscheidungsunterstützung bei der Klassifizierung von Komponenten oder Akteuren in SGAM Zonen und -Domänen
- Erkennen von fehlenden, redundanten sowie fehlerhaften Elementen (wie Informationsobjekte und Akteure) innerhalb der IEC 62559 Use Case-Beschreibungen

Zum Zeitpunkt dieser Dissertation besteht eine Forschungslücke darin, dass noch keine intelligenten Systeme für den spezifischen Kontext entwickelt und erprobt worden sind. Diese Arbeit zielt darauf ab, ein solches System zu konzipieren, zu entwickeln und praktisch zu evaluieren. Hierbei orientiert sich die Arbeit an dem in Abbildung 1.4 dargestellten Prozessmodell, das auf dem in [GDU17] beschriebenen, weit verbreiteten anwendungsfallorientierten Ansatz basiert. Es beschreibt die übergeordneten Phasen der Spezifikation und Qualitätssicherung von IEC 62559 Use Cases sowie den darauf aufbauenden SGAM Modellen. Innerhalb dieses Prozesses besteht insbesondere in der



Quelle: Übersetzt aus [KNN<sup>+</sup>24]

Abbildung 1.4: Erstellungsprozess von IEC 62559 Use Cases und SGAM Modellen

Spezifikations- und Qualitätssicherungsphase ein hohes Potenzial für eine gezielte Unterstützung durch intelligente Systeme. Um dies zu erreichen, werden zunächst relevante Anwendungsfälle durch systematische Methoden wie Literaturrecherche oder Befragungen von Anwendern ermittelt und für die Entwicklung eines prototypischen intelligenten Systems, namens *Smart Grid Assisting Artificial Intelligence in Requirements Engineering (SGAAIRE)* ausgewählt. Anschließend wird das konzipierte System SGAAIRE realisiert, das entworfene Artefakt evaluiert und in einem realen Umfeld integriert. Aus der dargelegten Problemstellung leiten sich die folgenden Zielsetzungen 1 und 2 ab.

**Zielsetzung 1**

Konzeption und Implementierung von unterstützenden intelligenten Systemen, die Fachexperten bei der Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen assistieren.

**Zielsetzung 2**

Steigerung der Qualität von IEC 62559 Use Case-Beschreibungen und SGAM Modellen durch den Einsatz entwickelter intelligenter Systeme im Vergleich zur traditionellen Herangehensweise.

Angesichts des Ziels, die Qualität sowohl von IEC 62559 Use Case-Beschreibungen als auch von SGAM Modellen zu verbessern, ist es ebenfalls notwendig, Bewertungsmetriken zu entwerfen. Diese Metriken sollen es ermöglichen, den Erfüllungsgrad von Qualitätsmerkmalen im Rahmen einer Evaluation zu messen. Dadurch können die erstellten Modelle und Beschreibungen systematisch überprüft und optimiert werden, um einen Vergleich zwischen der Verwendung von intelligenten Systemen und der konventionellen Arbeitsweise zu erreichen. Für den Entwurf dieser Bewertungsmetriken werden die Grundsätze ordnungsmäßiger Modellierung (GoM) nach [BRS95] zur Hand gezogen, welche die Ziele von Modellierungsprozessen generisch betrachtet und somit auch entscheidend für die Qualität der Beschreibungen und Modelle sind. Aus der genannten Problemstellung ergibt sich die Zielsetzung 3.

**Zielsetzung 3**

Identifizierung und Festlegung von Qualitätsmerkmalen und -Metriken für IEC 62559 Use Case-Beschreibungen und SGAM Modellen, gestützt auf die Grundsätze ordnungsmäßiger Modellierung.

## 1.2 Forschungskontext

In dem vorherigen Abschnitt wurde bereits auf die Problemstellung sowie auf die herausgehenden Zielsetzungen dieser Dissertation eingegangen. Auf Basis der definierten Zielsetzungen werden die Forschungsfragen sowie Hypothesen abgeleitet.

**Zielsetzung 1:** Wie im vorangehenden Abschnitt dargelegt wurde, besteht zum Zeitpunkt dieser Dissertation eine Lücke in der Forschung sowie der Praxis: Es existieren

keine intelligenten oder KI-basierten Systeme, die speziell für den Einsatz im Kontext der IEC 62559 Use Case-Methodik oder des SGAM Frameworks entwickelt wurden, um insbesondere die Erstellung von IEC 62559 Use Cases und SGAM Modellen zu unterstützen.

Angesichts der Komplexität und der erforderlichen Fachkenntnisse in Bezug auf die IEC 62559 Use Case-Methodik sowie das SGAM Framework und das jeweilige System-of-Concern, erfordert die Thematik ein tiefgehendes Verständnis des Nutzungskontextes. Daraus folgt die Notwendigkeit, spezifische Anforderungen zu identifizieren, die nicht nur die Qualität der Ergebnisse positiv beeinflussen, sondern auch den Modellierungsprozess verbessern. Auf der Grundlage dieser selektierten Anforderungen soll das intelligente System SGAAIRE entwickelt werden, unter Anwendung fortschrittlicher Methoden aus dem Bereich der KI. Dieser Ansatz zielt darauf ab, eine effektive Unterstützung für Fachexperten bereitzustellen und wird von Forschungsfrage 1 adressiert.

#### **Forschungsfrage 1**

Wie können **intelligente Systeme** für Fachexperten zur Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen konzipiert werden?

In Hinblick zur Beantwortung von Forschungsfrage 1 wird Hypothese 1 formuliert, welche die Anwendung von Techniken der Natural Language Processing (NLP) in den Vordergrund rückt. Angesichts der Tatsache, dass IEC 62559 Use Case Beschreibungen und SGAM Modelle vorrangig aus Textsequenzen bestehen, erweist sich NLP als eine fundamentale Technologie aus dem Spektrum der KI für diesen Kontext. Die Anwendung von NLP-Methoden zielt darauf ab, die Verarbeitung, Analyse und das Verständnis dieser Textdaten zu optimieren, was für die Effektivität der vorgeschlagenen intelligenten Systeme entscheidend ist.

#### **Hypothese 1**

Intelligente Systeme, die auf natürliche Sprachverarbeitung setzen, können dazu verwendet werden, Fachexperten bei der Erstellung von IEC 62559 Use Cases und SGAM Modellen zu unterstützen, indem sie relevante Informationen aus verschiedenen Quellen extrahieren und strukturieren.

**Zielsetzung 2:** Ein zentraler Bestandteil dieses Dissertationsvorhabens ist die Untersuchung, inwiefern der Einsatz von intelligenten Systemen eine Qualitätssteigerung im Vergleich zu konventionellen Techniken erzielen kann. Die Frage, ob eine derartige Steigerung tatsächlich realisiert werden kann, adressiert eine bestehende Forschungslücke und wird eingehend in Forschungsfrage 2 thematisiert.

#### **Forschungsfrage 2**

Führt der Einsatz von **intelligenten Systemen** bei der Erstellung von IEC 62559

Use Case-Beschreibungen und SGAM Modellen gegenüber der konventionellen Arbeitsweise zur **Qualitätsverbesserung**?

Hypothese 2, die direkt aus Zielsetzung 1 und Forschungsfrage 1 abgeleitet wird, konzentriert sich auf die Untersuchung des entwickelten Artefakts SGAAIRE. Diese Hypothese zielt darauf ab, empirisch zu überprüfen, ob und inwieweit das intelligente System SGAAIRE eine positive Auswirkung auf die Qualität von IEC 62559 Use Case-Beschreibungen und SGAM Modellen hat. Durch die Evaluation von SGAAIRE in praxisnahen Szenarien soll die Hypothese die potenziellen Verbesserungen in der Modellqualität, die durch den Einsatz von intelligenten Systemen erreicht werden können, aufzeigen.

### **Hypothese 2**

Der Einsatz von SGAAIRE bei der Erstellung von IEC 62559-2 Use Case-Beschreibungen und SGAM Modellen kann zu einer Verbesserung der Qualität beitragen.

**Zielsetzung 3:** Wie bereits im vorherigen Abschnitt dargelegt, existieren zum Zeitpunkt dieser Dissertation keine spezifischen Qualitätsmerkmale und Metriken für IEC 62559 Use Cases sowie SGAM Modelle. Die Entwicklung und Definition solcher Qualitätsmerkmale und Metriken sind jedoch essentiell, um die Forschungsfrage 2 adäquat zu adressieren. Angesichts der in Zielsetzung 3 formulierten Notwendigkeit, zunächst Qualitätsmerkmale zu untersuchen, werden innerhalb des Rahmens von Forschungsfrage 2 zwei spezifische Unterforschungsfragen definiert. Die erste Unterforschungsfrage 2.1 konzentriert sich auf die Identifikation von relevanten Qualitätsmerkmalen und zugehörigen Metriken, die den Grundsätzen ordnungsmäßiger Modellierung unterliegen.

### **Forschungsfrage 2.1**

Welche relevanten *Bewertungsmetriken für die Erfüllungsgrade von Qualitätsmerkmalen* können auf Basis der Grundsätze ordnungsmäßiger Modellierung (GoM) abgeleitet werden, um die Modellanforderungen von IEC 62559-2 Use Cases und SGAM Modellen zu erfüllen?

Auf Grundlage von Forschungsfrage 2.1 wird Hypothese 3 formuliert, die postuliert, dass die entwickelten Bewertungsmetriken effektiv eingesetzt werden können, um die Einhaltung der Anforderungen von IEC 62559 Use Cases und SGAM Modellen zu gewährleisten.

### **Hypothese 3**

Bewertungsmetriken können als Grundlage dienen, um die Qualität von erstellten IEC 62559-2 Use Case-Beschreibungen und SGAM Modellen zu bewerten und somit deren Konformität mit den vorgegebenen Modellanforderungen zu überprüfen.

Die zweite Unterforschungsfrage 2.2 fokussiert sich darauf, die Auswirkungen des Einsatzes intelligenter Systeme auf die Erfüllungsgrade der identifizierten Qualitätsmerkmale und -Metriken zu evaluieren. Durch Anwendung dieser Qualitätseigenschaften und -Metriken wird es ermöglicht, eine Bewertung von IEC 62559 Use Cases und SGAM Modellen durchzuführen. Zudem ist die Festlegung eines Referenzpunktes basierend auf konventionellen Methoden unerlässlich, um einen quantitativen Vergleich zu ermöglichen.

### **Forschungsfrage 2.2**

Welche *Auswirkungen* hat der Einsatz von *intelligenten Systemen* im Hinblick auf die Erfüllungsgrade der Qualitätsmerkmale hergeleitet aus den Grundsätzen ordnungsmäßiger Modellierung (GoM) gemäß Modell-Anforderungen von IEC 62559-2 Use Cases und SGAM Modellen?

Während Forschungsfrage 2 die allgemeine Verbesserung der Qualität durch den Einsatz intelligenter Systeme im Vergleich zu traditionellen Methoden thematisiert, zielt Forschungsfrage 2.2 spezifischer auf die Untersuchung der in Forschungsfrage 2.1 identifizierten Qualitätseigenschaften- sowie Metriken ab. Es wird analysiert, inwiefern sich die Erfüllungsgrade dieser Qualitätseigenschaften durch den Einsatz intelligenter Systeme im Vergleich zu konventionellen Techniken verändern. Des Weiteren basiert diese Forschungsfrage 2.2 zudem auch auf Hypothese 2, die darauf abzielt zu demonstrieren, dass der Einsatz intelligenter Systeme zu einer messbaren Verbesserung führt.

## **1.3 Verwandte Arbeiten**

Erste Arbeiten zu Qualitätsverbesserungen von IEC 62559 Use Case-Beschreibungen wurden im Rahmen von Konsistenzprüfungen von [GS15] durchgeführt. Der Schwerpunkt dieser Arbeit lag auf der Konzeption und Entwicklung eines UCMR, das Akteure, Informationsobjekte, Anforderungen und weitere Elemente als globale, Use Case-übergreifende Assets bereitstellt, die innerhalb der Use Case-Spezifikation verlinkt werden. Ziel war es, die Erstellung mehrerer Anwendungsfälle in unterschiedlichen Projekten zu unterstützen [GS15].

Unveröffentlichte Masterarbeiten von [Pet19] und [Ogu19] bauten auf dem UCMR auf und untersuchten mithilfe regelbasierter NLP-Techniken, wie sich die Konsistenz weiter verbessern lässt. [Ogu19] erstellte dazu eine manuell definierte NLP-Regelbasis, die von [Pet19] in Form eines Expertensystems sowie mit Regular Expressions (RegExs) in das UCMR integriert wurde. Der KI-Anteil war hierbei aufgrund der rein regelbasierten Ansätze begrenzt, was in der Praxis dazu führte, dass Modellierende aus unterschiedlichen Fachrichtungen ein hohes technisches Vorwissen benötigten. Zudem fokussierten sich diese Arbeiten ausschließlich auf Konsistenz als Qualitätsmerkmal, obwohl dies nur einen Teilaspekt der Modellierungsqualität darstellt.

Für den Einsatz von KI-Verfahren bei der Erstellung von SGAM Modellen existieren zum Zeitpunkt dieser Dissertation weder wissenschaftliche Arbeiten noch praktische Umsetzungen. Allerdings wird an zugehörigen Modellierungstools, wie der SGAM Toolbox [VNE<sup>+</sup>25] (basierend auf [NET<sup>+</sup>14]) weiterhin aktiv geforscht und entwickelt. Das SGAM sowie die IEC 62559 Use Case-Methodik stellen auch aktuell ein zentrales Instrument zur Adressierung von Interoperabilität dar [RSJ<sup>+</sup>23].

Im allgemeinen Requirements Engineering wurden zahlreiche Studien durchgeführt, in denen NLP-Techniken zur Verbesserung der Anforderungsqualität eingesetzt wurden. Eine systematische Literaturübersicht von [NBAKK17] vergleicht verschiedene Ansätze und zeigt, dass NLP insbesondere zur Beseitigung von Mehrdeutigkeiten, zur Erkennung von Inkonsistenzen und zur Priorisierung von Anforderungen beitragen kann. Seit Beginn dieser Dissertation im September 2022 hat sich das Forschungsfeld stark in Richtung Language Model (LM)/Large Language Model (LLM) (dt. Sprachmodelle bzw. große Sprachmodelle) entwickelt, ausgelöst durch die Veröffentlichung von OpenAI's ChatGPT [HJS22] im November 2022. Dieser Paradigmenwechsel hat auch die Verfahren des Software- und Systems Engineering in Forschung und Praxis beeinflusst. Eine aktuelle Literaturübersicht von [HZL<sup>+</sup>24] untersucht die Potenziale und Techniken von LLMs, unter anderem auch im Hinblick auf die Generierung von Use Case-Beschreibungen, wenngleich ohne Bezug zu standardisierten IEC 62559 Use Case Methodik [NBJ<sup>+</sup>24].

Im Vergleich zu den in Tabelle 1.1 dargestellten Arbeiten adressiert diese Dissertation erstmals den kombinierten Einsatz von intelligenten Systemen auf Basis von KI-Verfahren, insbesondere (L)LMs, zur Verbesserung der Qualität sowohl von IEC 62559 Use Cases als auch von SGAM Modellen. Daraus ergibt sich eine Forschungslücke, da beide Ansätze weiterhin aktiv weiterentwickelt und in der Praxis eingesetzt werden, um Interoperabilitätsprobleme methodisch zu identifizieren und zu adressieren [RSJ<sup>+</sup>23]. Beispielsweise erfolgt im Rahmen des int:net-Projektes eine Weiterentwicklung in Form einer sechsten SGAM-Interoperabilitätsebene, der sogenannten Framework Layer, sowie die Entwicklung einer methodischen Herangehensweise zur Abgrenzung und Fokussierung konkreter Anwendungsfälle [KMH24]. Zudem zeigen Ergebnisse aus der eigenen Studie [KNN<sup>+</sup>24], dass viele Fach- und Methodenexperten bei der Erstellung sowohl von IEC 62559 Use Cases als auch von SGAM Modellen mit Komplexitäten konfrontiert sind, die durch gezielte intelligente Unterstützungsmaßnahmen wirksam adressiert werden können. Während bestehende Arbeiten wie [HZL<sup>+</sup>24] und [NBJ<sup>+</sup>24] den Einsatz von LLMs im allgemeinen Requirements Engineering bzw. für generische Use Case-Beschreibungen untersuchen, fehlt dort der Bezug zu den standardisierten IEC 62559 Use Cases sowie gänzlich zu Referenzarchitekturmodellen, wie dem SGAM Framework. Arbeiten wie die SGAM Toolbox [NET<sup>+</sup>14, VNE<sup>+</sup>25] oder der SGAM Visualizer [GDU17] fokussieren dagegen ausschließlich auf die Weiterentwicklung von SGAM Modellierungs- und Anzeigewerkzeugen, ohne den Einsatz von KI-Methoden oder eine direkte Qualitätsverbesserung zu adressieren. Forschungen im Bereich der Konsistenzprüfung, beispielsweise [GS15], [Ogu19] und [Pet19], behandeln zwar IEC 62559 Use Cases, setzen jedoch primär auf regelbasierte NLP-Techniken und

Tabelle 1.1: Vergleich der wesentlichen verwandten Arbeiten

Referenz	IEC 62559 Use Case	Smart Grid Architecture Model	KI-Verfahren	Qualitätsverbesserung
Large Language Models for Software Engineering: A Systematic Literature Review [HZL <sup>+</sup> 24]	Generelles Requirements Engineering	Nein	LLM	Ja
Automating Software Documentation: Employing LLMs for Precise Use Case Description [NBJ <sup>+</sup> 24]	Generelle Use Cases	Nein	LLM	Ja
SGAM Toolbox Revisited: A Standards-Based Domain-Specific Modeling Language and Toolset [VNE <sup>+</sup> 25]	Nein	Ja	Nein	Nein
Towards Identifying an Approach for Consistency Checks to Smart Grid Descriptions [GS15]	Ja	Nein	Nein	Konsistenz (innerhalb eines Textes)
Konsistenzprüfung im Anforderungsmanagement mithilfe von Text Mining [Ogu19]	Ja	Nein	NLP	Konsistenz (innerhalb eines Textes)
Konzeption und Implementierung eines Expertensystems für die Unterstützung bei der Erstellung von Use Cases nach IEC 62559 [Pet19]	Ja	Nein	NLP	Konsistenz (innerhalb eines Textes)
<i>Diese Dissertation:</i> Intelligente Systeme zur Verbesserung der Qualität von IEC 62559 Use Cases und SGAM Architekturmodellen	Ja	Ja	Ja	Ja

beschränken sich auf Konsistenz als Qualitätsmerkmal innerhalb von Texten. Im Gegensatz dazu verfolgt diese Arbeit einen umfassenderen Qualitätsbegriff, der neben Konsistenz auch weitere Qualitätsmerkmale einbezieht und KI-gestützte Verfahren auf sowohl IEC 62559 Use Cases als auch SGAM Modelle angewendet werden können.

Zum Zeitpunkt der Dissertation liegen zudem keine Arbeiten vor, die sich systematisch mit den Herausforderungen in der praktischen Anwendung der Methoden selbst, sowohl aus Sicht von Fach- als auch von Methodenexperten, befassen. Ein Aspekt dieser Arbeit liegt daher in der Erfassung sowie der zielgerichteten Adressierung dieser Herausforderungen bei der Modellierung von IEC 62559 Use Cases und in SGAM Modellen. Darüber hinaus verfolgt die Arbeit das Ziel, den praktischen Einsatz der entwickelten Ansätze so weit wie möglich zu ermöglichen, insbesondere durch die Verwendung einer flexiblen, systemunabhängigen Architektur, um eine Beschränkung auf eine einzelne Implementierung zu vermeiden.

## 1.4 Aufbau der Dissertation

In diesem Kapitel wurde die Problemstellung der Arbeit motiviert sowie deren Zielsetzung und Forschungskontext erläutert. Darüber hinaus erfolgte eine Einordnung in den Stand der Forschung, um die bestehende Forschungslücke herauszuarbeiten.

Kapitel 2 beschreibt die methodische Herangehensweise und stellt das Vorgehensmodell der Dissertation vor. Der Schwerpunkt liegt hierbei auf dem Design Science Research Framework und dessen Konkretisierung im Rahmen der Arbeit. Kapitel 3 behandelt die theoretischen Grundlagen der verwendeten Methodiken und erläutert die relevanten Aspekte der IEC 62559 Use Case-Methodik sowie des SGAM Frameworks. Kapitel 4 widmet sich den Grundlagen des KI-Bereichs und stellt insbesondere die Techniken dar, die für die Entwicklung von SGAAIRE eingesetzt werden.

Da SGAAIRE als intelligentes KI-System ein Softwareartefakt darstellt, orientiert sich der weitere Aufbau der Arbeit an den Phasen nach [Som12]. Zunächst wird in Kapitel 5 die Anforderungs- und Kontextanalyse für die Entwicklung des KI-Prototyps SGAAIRE dargestellt. Dort erfolgt die Herleitung der nicht-funktionalen und funktionalen Anforderungen. Die in den Forschungsfragen relevanten Qualitätsmerkmale sowie deren Bewertungsmetriken sind Gegenstand von Kapitel 6. Die Konzeption von SGAAIRE wird in Kapitel 7 beschrieben und in Kapitel 8 in die konkrete Umsetzung überführt. Die Evaluation von SGAAIRE sowie die Herleitung und Analyse der Daten zur Beantwortung der Forschungsfragen erfolgen in Kapitel 9, das zudem eine kritische Diskussion der Ergebnisse enthält. Dort werden auch die im Dissertationsprozess entstandenen wissenschaftlichen und technischen Beiträge aufgeführt.

Abschließend fasst Kapitel 10 die Ergebnisse der Arbeit zusammen. Es beantwortet die Forschungsfragen im Hinblick auf die definierte Zielsetzung und zeigt im Ausblick potenzielle Ansätze für weiterführende Forschung auf.



## 2 Methodische Herangehensweise

In diesem Kapitel wird die methodische Herangehensweise dieser Dissertation detailliert vorgestellt. Zunächst erfolgt eine Erhebung der Anforderungen an das Vorgehensmodell, wobei der spezifische Forschungskontext sowie das zugrundeliegende Forschungsfeld berücksichtigt werden. Dies bildet die Grundlage für die Entwicklung eines methodischen Rahmens und der Auswahl geeigneter Forschungsmethoden. Auf dieser Basis wird der Forschungsansatz des Design Science vorgestellt, der die methodische Ausrichtung der Arbeit maßgeblich prägt. Im Anschluss wird der Aufbau der Dissertation erläutert, der das Vorgehensmodell, die eingesetzten Forschungsmethoden und deren Integration in die Gesamtstruktur der Arbeit umfasst. Zudem erfolgt eine präzise Eingrenzung des Begriffs „Künstliche Intelligenz“ in Bezug auf die Anwendung innerhalb dieser Arbeit.

### 2.1 Anforderungen an das Vorgehensmodell

Eine systematische Vorgehensweise als Forschungsmethodologie stellt eine wesentliche Basis für wissenschaftliche Arbeiten dar [Kot04]. Zunächst gilt es, das Forschungsfeld detaillierter zu betrachten und das adäquate Forschungsrahmenwerk zu identifizieren. Die vorliegende Arbeit ortet sich in der Energieinformatik (vgl. Definition 1) als Disziplin der „angewandten Informatik“ sowie auch der Wirtschaftsinformatik/Information Systems Research ein [WBC10]. Das Dissertationsvorhaben konzentriert sich nicht auf die direkte Entwicklung von Energiesystemen, sondern auf die weiterführende Forschung von Methoden aus der Energieinformatik, welche die Konzeption von Energiesystemen unterstützen.

**Definition 1** (Energieinformatik)

Die **Energieinformatik** befasst sich mit der Analyse, dem Entwurf und Implementierung von Systemen zur Steigerung der Effizienz von Energienachfrage- und Versorgungssystemen, was eine Sammlung und Analyse von Energiedatensätzen der Optimierung von Energieverteilungs- und Verbrauchsnetzen zur Unterstützung vorsieht. [WBC10]

Sowohl in der Wirtschaftsinformatik als auch in der anglo-amerikanischen Schwesterdisziplin der Information Systems Research lassen sich zwischen dem konstruktionswissenschaftlichen Paradigma („Design Science“) und dem verhaltenswissenschaftlichen Paradigma („Behavioral Science“) unterscheiden [WH07]. Nach [HMPR04] dient das aus den Naturwissenschaften hervorgegangene verhaltenswissenschaftliche Paradigma dazu, die Interaktion zwischen Menschen und Technologien zu untersuchen, wobei insbesondere das Verständnis für deren Nutzung und Auswirkungen im sozialen Kontext im Vordergrund steht. Im Gegensatz dazu wird das konstruktionswissenschaftliche Paradigma verwendet, um neue artefaktbasierte Lösungen zu entwickeln, wobei der Fokus auf der Schaffung von funktionalen und innovativen Designartefakten liegt, die als Grund-

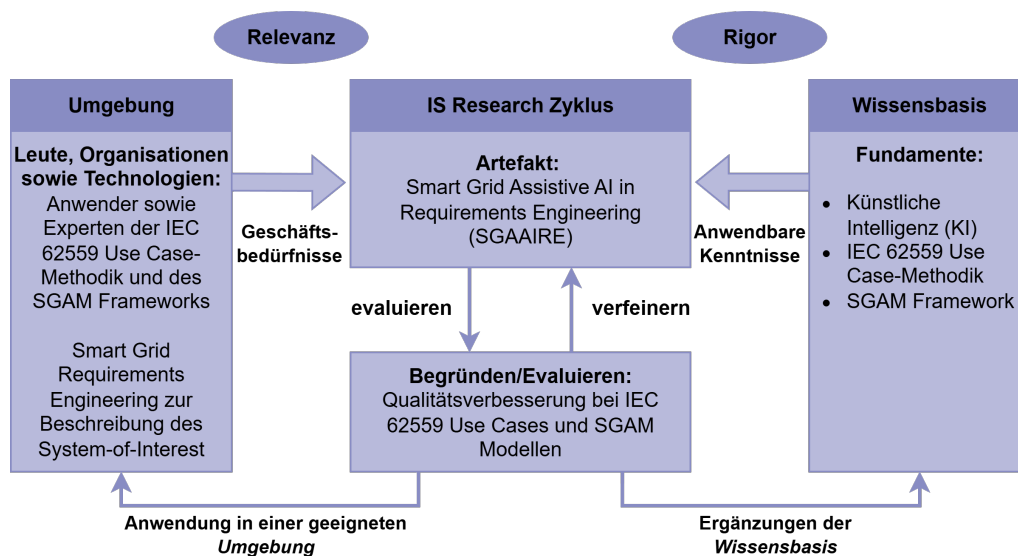
lage für die Lösung praktischer Probleme dient [HMPR04]. Aus der Zielsetzung heraus wurde folglich impliziert, dass das *konstruktionswissenschaftliche Paradigma* („*Design Science*“) im Rahmen des Forschungsvorhabens selektiert wurde.

Eine zentrale Anforderung an das Vorgehensmodell ist die Unterstützung von Prototyping im ganzheitlichen Sinne. Reine Softwareentwicklung stellt nicht zwangsläufig Forschung dar, sondern dient in dieser Dissertation als „Mittel zum Zweck“, um einen Prototyp (im Hinblick auf Forschungsfrage 1) zu entwickeln. Aus der Problemstellung und den daraus abgeleiteten Forschungsfragen ergibt sich die Notwendigkeit, mithilfe eines Prototypen ein Problem zu lösen, konkret die Reduktion von (bestenfalls schwerwiegenden) Fehlern, was das motivierte Forschungsproblem darstellt. Neben der Entwicklung des Artefakts und der Integration KI-basierter Ansätze soll das Vorgehen auch die Anforderungs- und Kontexterhebung umfassen. Dadurch soll Wissen darüber erzeugt werden, ob und in welcher Weise intelligente Systeme zur Verbesserung der Qualität von IEC 62559 Use Cases und SGAM Modellen beitragen können. Darüber hinaus muss das Vorgehensmodell eine Demonstration und Evaluierung des Prototyps beinhalten, um dessen Wirksamkeit empirisch zu untersuchen. Der in der Dissertation für die Evaluation gewählte Ansatz folgt der *deskriptiven Forschung*, insbesondere in der Messung von Phänomenen wie der Fehlerreduktion. Des Weiteren soll in dem Vorgehensmodell sowohl die Perspektive von Fach- als auch Methodenexpertise betrachtet werden. Dies bedeutet, dass sowohl die Perspektive der Anwender als auch jene von Forschung und Standardisierung eingenommen werden soll, um ein umfassendes Bild zu erhalten.

## 2.2 Design Science

Im Rahmen der Anforderungen an das Vorgehensmodell (Abschnitt 2.1) wurde festgestellt, dass die Dissertation die konstruktionswissenschaftlichen Paradigmen („*Design Science*“) fokussiert. Ein zentraler Aspekt der Design Science ist die Entwicklung und Evaluierung von IT-Artefakten zur Lösung praxisnaher Probleme [HMPR04]. Diese Artefakte können in unterschiedlichen Formen repräsentiert werden, von Software und (mathematischer) Modellierung bis hin zu informellen sprachlichen Beschreibungen [HMPR04]. Artefakte im Sinne von [MS95] sind Konstrukte, Modelle, Methoden sowie Instanziierungen. Prozesse im Sinne von [MS95] sind (Design-)Prozesse des Bauens und Evaluierens von Artefakten. Das etablierteste Rahmenwerk für Forschungsarbeiten im Rahmen von Design Science stellt das *Design Science Research Framework (DSRF)* dar.

Die Abbildung 2.1 veranschaulicht eine Konkretisierung des Dissertationsvorhabens basierend auf dem von [HMPR04] entwickelten DSRF-Framework. Der *Problemraum* (Umgebung) lässt sich durch die primäre Zielgruppe der Anwender sowie Experten der IEC 62559 Use Case-Methodik sowie des SGAM Frameworks eingrenzen. Diese nutzen die Methodiken zur Beschreibung des System-of-Interest innerhalb des energiebezogenen Anforderungsmanagement. Der *Relevanz-Zyklus* wird durch die Erreichung einer Qualitätssteigerung der IEC 62559-2 Use Cases und SGAM Modelle mittels des Arte-



Quelle: Eigene Darstellung, angelehnt an [HMPR04]

Abbildung 2.1: Visualisierung der Konkretisierung des DSRF für das Dissertationsvorhaben

fakts realisiert, was zu einer verbesserten Verständlichkeit des System-of-Interest führt. Dies hat wiederum womögliche positive Effekte auf die Interoperabilität und kann zur Minimierung von Folgefehlern im Produktivbetrieb beitragen.

Im *Lösungsraum* steht primär die Wissensbasis aus dem Bereich der künstlichen Intelligenz, IEC 62559 Use Case-Methodik und dem SGAM Framework. Daten, Standards, wissenschaftliche Methoden, Vorgehensweisen etc. spiegeln ebenfalls Teile der Wissensbasis wider. Erkenntnisse, die aus der Entwicklung und dem Einsatz des intelligenten Systems hinsichtlich ihrer Auswirkungen auf die Erfüllungsgrade der Qualitätsmerkmale gewonnen werden, fließen in den *Stringenz-Zyklus* zurück. Dies stellt das Wissen bezüglich der Bewertungsmetriken, deren Erfüllungsgrade sowie Kenntnisse zur Entwicklung und Anwendung von SGAAIRE dar.

Der Kern des Frameworks basiert auf dem *Information Systems Forschungszyklus*. Das zentrale *Artefakt* bildet SGAAIRE als Software-Prototyp, der die im Lösungsraum entwickelten Verfahren auf die im Problemraum definierte Fragestellung anwendet. Über die *Evaluierung* gilt es, die Ziele der Relevanz- und Stringenz-Zyklen zu validieren, wofür gängige Forschungsmethoden gemäß des Vorgehensmodells verwendet werden [HMPR04]. Dies wäre in dem Forschungsvorhaben die Qualitätsverbesserung von IEC 62559 Use Cases sowie SGAM Modellen. Zur Evaluierung werden die im Abschnitt 2.3 aufgeführten Forschungsmethoden verwendet.

## 2.3 Vorgehensmodell der Arbeit

Das in dieser Arbeit zugrunde gelegte Vorgehensmodell basiert auf dem Design Science Research Methodology (DSRM)-Prozess nach [PTRC07] und integriert die im Framework definierten Aktivitäten. Die spezifischen Maßnahmen (Aktionen) wurden aus dem im Abschnitt 1.2 hergeleiteten Forschungsgegenstand systematisch abgeleitet. Ein zentrales Artefakt des Dissertationsvorhabens ist SGAAIRE. Daher flossen für dessen Entwicklung Konzepte des Human-Centred Design (HCD) [DIN20] in die Phasen ein, die analog zum konventionellen Softwareentwicklungszyklus („Wasserfallmodell“) strukturiert sind. Sowohl DSRM als auch HCD fördern ein iteratives Vorgehen, indem sie das wiederholte Durchlaufen einzelner Phasen ermöglichen. Die Aktivitäten wurden wie folgt definiert und sind in Abbildung 2.2 dargestellt:

**Phase 1: Problemidentifikation** Zunächst erfolgt die Identifikation des Forschungsproblems, indem bestehende Herausforderungen bei der Qualitätssicherung von Use Case-Beschreibungen und SGAM Modellen analysiert werden. Eine zentrale Rolle spielt dabei die Definition von Fehlern und Qualität nach DIN EN ISO 9000:2015 [DIN15]. Auf Basis dieser Erkenntnisse werden vorläufige Hypothesen zur Entwicklung einer evaluierbaren Theorie induktiv abgeleitet, die anschließend durch deduktive Forschung überprüft werden.

**Phase 2: Zielsetzung einer Lösung definieren** Basierend auf der Problemidentifikation werden spezifische Qualitätsmetriken für die Evaluierung von IEC 62559 Use Cases und SGAM Modellen entwickelt. Diese Metriken leiten sich aus den GoMs sowie den Anforderungen an IEC 62559 Use Cases sowie SGAM Modelle ab und geschieht über eine *Literaturrecherche*. Die Berücksichtigung der Literatur soll dazu beitragen, die Anforderungen an geeignete IEC 62559 Use Cases und SGAM Modelle aus bestehenden Projekten sowie die bereits identifizierten Herausforderungen zu integrieren. Die Definition der *Anforderungen* an das zu entwickelnde KI-gestützte System erfolgt sowohl ebenfalls über die realisierte *Literaturrecherche* als auch durch die Einbindung von Methoden- und Fachexperten in Rahmen einer *Umfrage* mittels einem *teilstandardisierten Fragebogen*, welche auch den Nutzungskontext betrachtet. Die Umfrage mit dem teilstandardisierten Fragebogen ermöglicht eine fundierte Erhebung praxisrelevanter Anforderungen von den Anwendenden und Experten, da durch die Umfrage unterschiedliche Perspektiven erfasst und mithilfe des teilstandardisierten Fragebogens sowohl strukturierte als auch freie Antworten ermöglicht werden und eine höhere Erreichbarkeit (ggü. Experteninterviews) realisiert wird. Die Erhebung des *Nutzungskontexts* soll zudem in die Berücksichtigung der Entwicklung des KI-Systems einspielen und zudem für mehr Klarheit des Forschungskontexts sorgen. Die erhobenen funktionalen Anforderungen sollen in der Phase mittels einer Nutzwertanalyse für die Selektion priorisiert werden.

**Phase 3: Konzeption und Entwicklung** Auf Basis der ermittelten Anforderungen wird das intelligente System unter Anwendung eines *Prototyping*-Ansatzes konzipiert und implementiert. Die Wissensbasis des Systems fokussiert sich auf den KI-Bereich. Da-

bei kommen Methoden wie Natural Language Processing (NLP), Künstliche neuronale Netze (KNNs) insb. in Hinblick auf Language Model (LM)/Large Language Model (LLM) zum Einsatz, um relevante Informationen aus textbasierten Use Case-Beschreibungen und SGAM Modellen zu extrahieren und zu analysieren. Studien zeigen, dass solche Verfahren effektiv für Klassifikations- und Clustering-Probleme genutzt werden können [RN16, NBACK17, Cho20] und insbesondere (decoderbasierte) LM/LLM für generative Problemstellungen [NNN24].

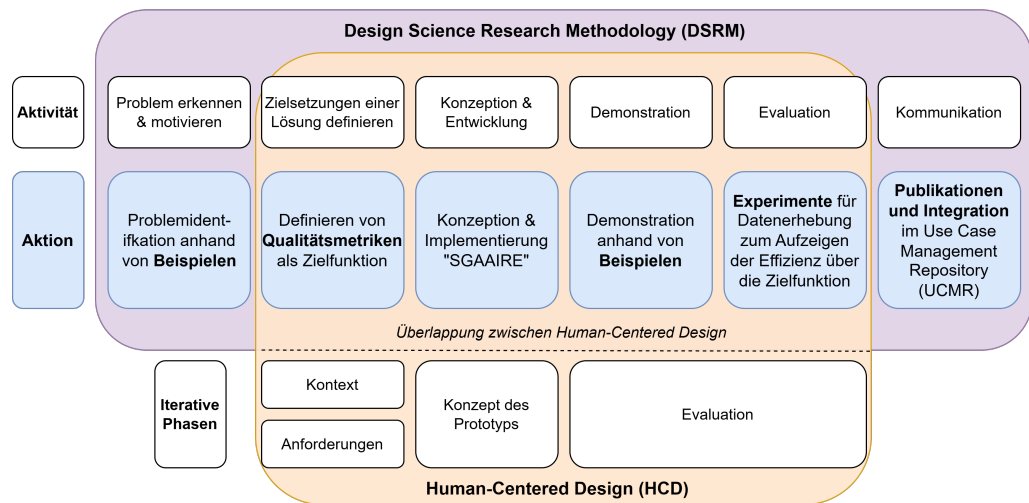
**Phase 4: Demonstration** Die Demonstration der entwickelten Lösung erfolgt anhand konkreter *Fallbeispiele*. Hierbei wird überprüft, ob und inwiefern das vorliegende intelligente System die Qualität der IEC 62559 Use Case-Beschreibungen und SGAM Modelle verbessert.

**Phase 5: Evaluation** Der zentrale Bestandteil der Evaluation ist der Vergleich zwischen der Nutzung des entwickelten KI-gestützten Systems und der konventionellen Arbeit bei IEC 62559 Use Case-Beschreibungen und SGAM Modellen. Dabei wird insbesondere die Reduktion von Fehlern als Qualitätskriterium herangezogen. Die Evaluation erfolgt quantitativ durch ein(e) *Randomized Controlled Trial (RCT)* (dt. randomisierte kontrollierte Studie) im Between-Subject Study Design, bei dem die beiden Verfahren hinsichtlich ihrer Qualitätsauswirkungen verglichen werden. Darüber hinaus erfolgt die Bewertung der Usability des Systems mithilfe des standardisierten System Usability Scale (SUS)-Verfahrens. Dieses methodische Vorgehen ermöglicht eine hohe interne Validität, da durch die zufällige Zuteilung von Teilnehmenden zu den Gruppen Störvariablen minimiert und Zusammenhänge zwischen den Herangehensweisen untersucht werden können [SNT04].

**Phase 6: Kommunikation** Die Ergebnisse dieser Dissertation werden sowohl wissenschaftlich als auch praktisch kommuniziert. Das entwickelte KI-System wird in das OFFIS Use Case Management Repository (UCMR) integriert und für den Einsatz in Projekten sowie zur Lehre zur Verfügung gestellt. Die Funktionalität des Systems wird durch begleitende Publikationen dokumentiert, um die wissenschaftliche und industrielle Relevanz der Forschung zu unterstreichen.

Die in diesem Vorgehensmodell *hervorgehobenen Methoden* bilden den Multimetodenansatz dieser Dissertation. Sie dienen dazu, die einzelnen Forschungsfragen (vgl. Abschnitt 1.2) gezielt zu adressieren:

- **Forschungsfrage 1 untersucht die Konzeption intelligenter Systeme zur Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen.** Zur Anforderungs- und Kontextanalyse wurden eine Umfrage sowie eine Literaturrecherche durchgeführt und mittels qualitativer Inhaltsanalyse ausgewertet. Die identifizierten Anforderungen wurden mithilfe einer Nutzwertanalyse priorisiert. Anschließend wurde ein prototypisches System entwickelt, das anhand zweier Fallbeispiele demonstriert wurde. Die zentralen Artefakte zur Beantwortung dieser Forschungsfrage sind ein Anforderungskatalog sowie ein KI-Prototyp, der belegt, dass sich intelligente Systeme für diesen Anwendungsfall entwickeln lassen.



Quelle: Eigene Darstellung, angelehnt an [PTRC07]

Abbildung 2.2: Vorgehensmodell der Arbeit

- **Forschungsfrage 2.1** befasst sich mit der **Ableitung relevanter Bewertungsmetriken zur Beurteilung der Erfüllungsgrade von Qualitätsmerkmalen in IEC 62559 Use Cases und SGAM Modellen**. Zur Identifikation dieser Qualitätsmerkmale wurde eine Literaturrecherche durchgeführt, die sowohl methodenbezogene als auch anwendungsorientierte Ansätze berücksichtigte.
- **Forschungsfrage 2.2** untersucht die **Auswirkungen intelligenter Systeme auf IEC 62559 Use Cases und SGAM Modelle**. Da Forschungsfrage 2 einen direkten Vergleich zwischen der Nutzung intelligenter Systeme und der konventionellen Herangehensweise erfordert, dienen die in Forschungsfrage 2.1 erarbeiteten Bewertungsmetriken als Grundlage zur Qualitätsbewertung des KI-Systems. Zur Untersuchung dieser Auswirkungen wird neben einer Demonstration ein Experiment in Form einer RCT (in Between-Subject Studiendesign) mit Probanden durchgeführt.

Forschungsfragen	Forschungsmethoden							Artefakt(e)
	Experimente	Umfrage	Qualitative Inhaltsanalyse	Literaturrecherche	Nutzwertanalyse	Demonstration	Prototyping	
RQ 1		●	●	●	●	●	●	Anforderungen + KI-Prototyp
RQ 2	RQ 2.1			●				Bewertungsmetriken
	RQ 2.2	●					●	Bewertung von KI-Systemen bei Qualitätsmerkmalen

Quelle: Eigene Darstellung

Abbildung 2.3: Zuordnung der Forschungsmethoden zu den Forschungsfragen

Das Vorgehensmodell sowie die verwendeten Forschungsmethoden sind in der Struktur der Dissertation verankert, wie in Tabelle 2.1 ersichtlich ist.

Tabelle 2.1: Einordnung des Vorgehensmodells in den Aufbau der Arbeit

Kapitel der Dissertation	Vorgehensmodell der Arbeit	
	Aktivität	Forschungsmethoden
1. Einleitung	Problemidentifikation	Literaturrecherche
2. Methodische Herangehensweise		
3. Grundlagen „Smart Grid-Kontext“		
4. Grundlagen „KI-Kontext“		
5. Anforderungsanalyse „SGAAIRE“	Zielsetzung definieren	Umfrage (teilstandardisierter Fragebogen), Literaturrecherche, Nutzwertanalyse
6. Konzept „Bewertungsmetriken“	Zielsetzung definieren	Literaturrecherche
7. Konzept „SGAAIRE“	Konzeption & Entwicklung	Prototyping
8. Implementierung „SGAAIRE“	Konzeption & Entwicklung	Prototyping
9. Evaluation und Diskussion	Demonstration & Evaluation & Kommunikation	Experimente, Demonstration
10. Schlusswort		

## 2.4 Verwendete Definition für Künstliche Intelligenz

Ein maßgeblicher Forschungskontext dieser Arbeit stellt der Themenkomplex „Künstliche Intelligenz“ dar. Da KI nach [RN16] ein interdisziplinäres Fundament aus der Informatik, Mathematik, Wirtschaft, Neurowissenschaften, Psychologie, Kontrolltheorie bzw. Kybernetik und Linguistik aufweist, können auf den Themenkomplex unterschiedliche Sichtweisen eingenommen werden. Zudem existiert keine allgemeingültige Definition von KI, weswegen im Forschungskontext und der Artefaktentwicklung eine Definition und deren Verwendung zu selektieren und festzulegen ist. Tabelle 2.2 zeigt eine Aufführung von etablierten KI-Definitionen chronologisch auf.

Für die Selektion der Definition gilt es zudem die Zielsetzung von KI festzulegen. Hierbei hat [Win92] Definitionen für sowohl technische, als auch wissenschaftliche Zielsetzungen von KI erarbeitet, welche in Tabelle 2.3 aufgeführt sind. Im Rahmen dieser Arbeit wird die **technische Zielsetzung** verfolgt, da der in Abschnitt 1.2 festgelegte Forschungskontext auf die Problemlösung innerhalb der Realität gemäß des Design Science Forschungsdesign abzielt. Insbesondere soll eine Software zur Entscheidungsfindung im Sinne eines (**intelligenten**) **wissensbasierten Systems** entwickelt werden, welche einen Themenbereich aus der KI darstellt [RN16]. Es gilt folglich eine etablierte Definition zu selektieren, welche die Entwicklung von solchen Systemen miteinbezieht, was schlussendlich die technische Zielsetzung der KI untermauert.

[RN16] unterscheidet bei den Definitionen die Kategorien zwischen „Menschlich Denken“, „Rational Denken“, „Menschlich Handeln“ und „Rational Handeln“. Im Hinblick auf das zu konzipierende, wissensbasierte System werden die Definitionen des **mensch-**

Tabelle 2.2: Aufführung beispielhafter Definitionen für Künstliche Intelligenz

Autoren und Quellenverweis	Definitionen
John McCarthy (1956) [MMRS06]	„Künstliche Intelligenz ist die Wissenschaft und Technik zur Herstellung intelligenter Maschinen.“ [MMRS06, übersetzt]
Patrick Henry Winston (1992) [Win92]	„Künstliche Intelligenz ist die Lehre von Berechnungen, die es ermöglichen, wahrzunehmen, zu denken und zu handeln.“ [Win92]
Raymond Kurzweil (1999) [Kur99]	„Künstliche Intelligenz ist die Fähigkeit, Maschinen zu erschaffen, die Funktionen ausführen, die normalerweise menschliche Intelligenz erfordern, wenn sie von Menschen durchgeführt werden.“ [Kur99]
EU Kommission (2020) [Eur20]	„KI-Systeme sind vom Menschen entwickelte Software- (und möglicherweise auch Hardware-) Systeme, die in Bezug auf ein komplexes Ziel auf physischer oder digitaler Ebene agieren, indem sie ihre Umgebung durch Datenerfassung wahrnehmen, die gesammelten strukturierten oder unstrukturierten Daten interpretieren, Schlussfolgerungen daraus ziehen oder die aus diesen Daten abgeleiteten Informationen verarbeiten und über die geeignete(n) Maßnahme(n) zur Erreichung des vorgegebenen Ziels entscheiden.“ [Eur20]

**lichen Handelns** betrachtet [AS10]. Diese Gruppe sieht vor, dass intelligente Systeme das menschliche Verhalten imitierten [RN16]. Für den Begriff der Intelligenz wurde der *Turing Test* (entwickelt im Jahr 1950) als ein Verfahren zur Bewertung der Fähigkeit eines Computers, menschliches Verhalten so überzeugend zu imitieren, dass er nicht von einem menschlichen Teilnehmer unterschieden werden kann, vorgesehen [RN16]. Hierfür sieht ein Computer zur Realisierung Techniken aus den Themenbereichen NLP, ML, aber auch Wissensdarstellung und automatisiertes Schließen vor [RN16]. Diese Themenbereiche stellen zum Zeitpunkt dieser Arbeit wesentliche Techniken im Hinblick auf das zu entwickelnde KI-System dar und werden im Kapitel 4 beleuchtet. Im Rahmen dieser Dissertation wird Künstliche Intelligenz basierend auf den Notwendig-

Tabelle 2.3: Zielsetzungen von KI nach [Win92]

Zielsetzung von KI	Definition
Technische Zielsetzung	„Das technische Ziel der künstlichen Intelligenz besteht darin, Probleme der realen Welt zu lösen, indem künstliche Intelligenz als ein Arsenal von Ideen zur Darstellung von Wissen, zur Nutzung von Wissen und zum Aufbau von Systemen eingesetzt wird.“ [Win92]
Wissenschaftliche Zielsetzung	„Das wissenschaftliche Ziel der künstlichen Intelligenz ist es, Konzepte zur Darstellung von Wissen, zur Nutzung von Wissen und zum Aufbau von Systemen um verschiedene Arten von Intelligenz zu erklären.“ [Win92]

keiten der aufgeführten Techniken und der technischen Zielsetzung zur Entwicklung des intelligenten Systems aus der Definition 2 von [Kur99] betrachtet.

**Definition 2** (Künstliche Intelligenz)

**Künstliche Intelligenz** ist die Fähigkeit, Maschinen zu erschaffen, die Funktionen ausführen, die normalerweise menschliche Intelligenz erfordern, wenn sie von Menschen durchgeführt werden. [Kur99]

Die vorliegende Definition bezieht sich auf den Begriff „Maschinen“, welcher im Rahmen dieser Dissertation jedoch auf ein „softwareseitiges System“ heruntergebrochen wird. Hierbei orientieren sich die Methoden im Kontext des Turing-Tests an dem vorherigen Absatz. Da zum Zeitpunkt der Arbeit keine Prototypen für starke KI existieren, sondern nur Konzepte, ordnet sich das intelligente System der schwachen KI zu, die zur Erfüllung konkreter Aufgaben entwickelt sind [Liu21]. Diese Arbeit gliedert sich zudem in den Bereich der angewandten KI (begründet durch die technische Zielsetzung) ein, weshalb die Grundlagenforschung von KI (über die wissenschaftliche Zielsetzung) nicht im Vordergrund steht. Die verwendete Definition spiegelt sich insbesondere in der Perspektive und Verwendung der KI-Themen wider, welche im Kapitel 4 genauer beleuchtet werden.

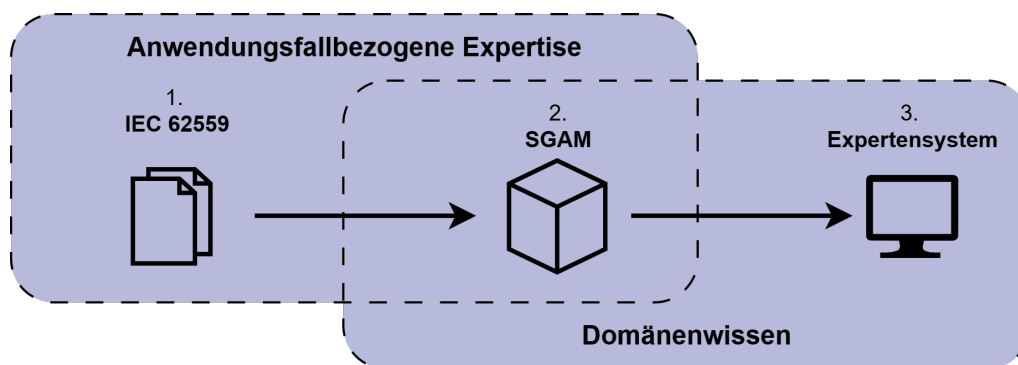


### 3 Grundlagen im Smart Grid Kontext

Dieses Kapitel gibt einen Überblick über die verwendeten Grundlagen im Rahmen des Smart Grid-Kontext. Zunächst wird im Abschnitt 3.1 die Problematik des Smart Grid als System-of-Systems aufgeführt, welche die Notwendigkeit von energiebezogenem Anforderungsmanagement aufführt, um so die Interoperabilität zwischen Systemen zu fördern. Dies umfasst die IEC 62559 Use Case-Methodik sowie das Smart Grid Architecture Model (SGAM) Framework als im Smart Grid-Sektor etablierte Requirements Engineering-Techniken [USD<sup>+</sup>13] [GDU17]. Diese werden in den Abschnitten 3.2 sowie 3.3 beleuchtet.

Die im Kapitel vorgestellten Methodiken benötigen in der Anwendung unterschiedlichste Expertise der Stakeholder. Hierbei unterscheidet [SCUG21] in Abbildung 3.1 zwischen anwendungsfallbezogene Expertise sowie Domänenexpertise. Die folgenden Begriffe werden im Rahmen der Dissertation verwendet und wie folgt betrachtet:

- Ein **Domänenexperte** verfügt über umfassendes Wissen in Bezug auf die Energiedomäne [SCUG21].
- **Technische Expertise** bzw. **anwendungsfallbezogene Expertise**: Diese Form der Expertise umfasst das Wissen, das für die inhaltliche Entwicklung eines Anwendungsfalls erforderlich ist. Technische Expertise kann auch als anwendungsfallbezogene Expertise betrachtet werden [SCUG21].
- Ein **Methodenexperte** besitzt das Wissen zur Anwendung der Methodik, wie in diesem Fall der IEC 62559 Use Case Methodik und des SGAM Frameworks.



Quelle: Basierend auf [SCUG21]

Abbildung 3.1: Unterschiedliche Arten der Expertise bei der Erstellung von IEC 62559 Use Cases sowie SGAM Modellen

### 3.1 Smart Grid als System-of-Systems

Die Transition von konventionellen Energiesystemen zum Smart Grid (dt. intelligentes Stromnetz) ist ein Schritt in Richtung einer nachhaltigeren Energiezukunft [USD<sup>+</sup>13]. Durch die Verwendung von Informations- und Kommunikationstechnologien (IKT) soll es durch das Smart Grid möglich sein, effizienter auf Schwankungen in der Energieerzeugung und -nachfrage zu reagieren [USD<sup>+</sup>13]. Hierbei soll eine bessere Integration erneuerbarer Energien möglich sein, was somit die Dezentralisierung der Energieerzeugung [USD<sup>+</sup>13] sowie die Dekarbonisierung als ein Teil des European Green Deal fördert [RSJ<sup>+</sup>23]. Durch die zunehmende Anzahl an Akteure, Geschäftsmodelle sowie Technologien [USD<sup>+</sup>13] steigt jedoch auch die Komplexität des Netzes, was die Notwendigkeit von Verfahren zum Fördern der Interoperabilität benötigt. Die folgenden Unterabschnitte beleuchten zum einen die Charakteristika des Smart Grids als System-of-Systems, zum anderen, wie Interoperabilität in der Energiedomäne betrachtet wird. Darüber hinaus wird der ISO/IEC/IEEE 42010-Standard „Software, systems and enterprise — Architecture description“ vorgestellt, der den Aufbau von Architekturbeschreibungen, unter anderem für das System-of-Interest innerhalb eines System-of-Systems, definiert [ISO11] und die Grundlage für die in dieser Dissertation behandelten Methodiken bildet.

#### 3.1.1 System-of-Systems Problematik

Das (intelligente) Energiesystem wird durch seine Charakteristiken als ein (operationales) „System-of-Systems“ (vgl. Definition 3) klassifiziert [USD<sup>+</sup>13, Alj18]. Charakteristiken stellen *betriebliche Unabhängigkeit* dar, wobei einzelne Komponenten auch dann autonom funktionieren können, wenn sie vom größeren Netzwerk getrennt sind, sowie die *verwaltungsbezogene Unabhängigkeit*, bei der jedes Element unabhängig arbeitet und dennoch kollektiv zur Gesamtleistung des Netzwerks beiträgt. Zudem sieht das Smart Grid eine *evolutionäre Entwicklung* vor, d.h. es werden Komponenten nahtlos integriert oder entfernt, ohne andere Teile des Systems zu beeinträchtigen [Alj18]. Es weißt zudem ein *emergentes Verhalten* auf, bei dem die Interaktion verschiedener Elemente zu einer effizienten, wirtschaftlichen und sauberen Energieverteilung führt, um unterschiedlichen Anforderungen gerecht zu werden [Alj18]. Zudem ist die *geografische Verteilung* gegeben, da die Elemente des (intelligenten) Energienetzes räumlich verteilt sind [Alj18].

**Definition 3** (System-of-Systems (SoS))

Das **System-of-Systems** (dt. System von Systemen) ist eine Menge von Systemen oder Systemelementen, die miteinander interagieren, um einen einzigartigen Leistungsumfang bereitzustellen, die keines der Teilsysteme alleine erreichen kann.

*Hinweis 1 zum Eintrag:* Systemelemente können erforderlich sein, um die Interaktion der Teilsysteme im System-of-Systems zu ermöglichen. (übersetzt aus [ISO19])

### 3.1.2 Interoperabilität

Durch den vermehrten Einsatz von IKT muss auch sichergestellt werden, dass verschiedene (Energie-)Systeme von verschiedenen Akteuren miteinander kommunizieren und sich verstehen können. Hierbei besteht die Herausforderung, Interoperabilität (vgl. Definition 4, verwendet im GridWise Architecture Council (GWAC) Interoperability Context-Setting Framework von [WHA<sup>+</sup>08]) zu schaffen.

**Definition 4** (Interoperabilität (IOP))

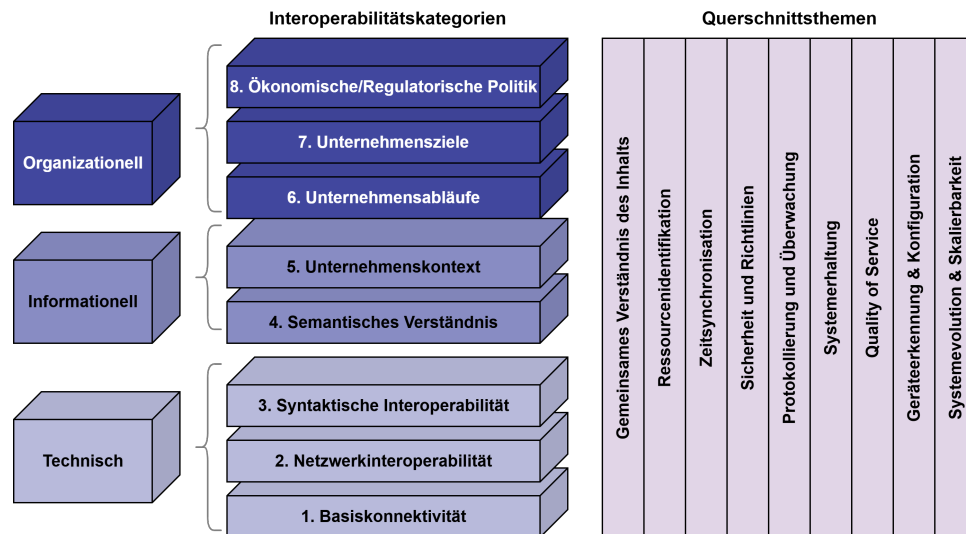
**Interoperabilität** ist die Fähigkeit von zwei oder mehr Netzwerken, Systemen, Geräten, Anwendungen oder Komponenten, Informationen untereinander auszutauschen und die ausgetauschten Informationen zu nutzen. (übersetzt aus [WHA<sup>+</sup>08])

Der US-amerikanische GWAC stellt mit dem Interoperability Context-Setting Framework ein Rahmenwerk dar um Interoperabilitätsbemühungen für den Energiesektor zu verbessern. Der GWAC [WHA<sup>+</sup>08] unterscheidet zwischen den folgenden drei Hauptkategorien von Interoperabilität, welche in Abbildung 3.2 dargestellt werden:

- **Syntaktische (technische) Interoperabilität** bezieht sich auf die Fähigkeit von Systemen, Datenformate und Kommunikationsprotokolle korrekt zu interpretieren, um den technischen Datenaustausch zu ermöglichen.
- **Semantische (informative) Interoperabilität** bezieht sich auf die Fähigkeit von Systemen, dass die ausgetauschten Daten in gleicher Weise interpretiert und verstanden werden. Hierunter fällt das gemeinsame semantische Verständnis innerhalb der Daten, aber auch der Geschäftskontext.
- **Pragmatische (organisatorische) Interoperabilität** bezieht sich auf die Fähigkeit von Systemen, ausgetauschte Informationen im spezifischen Nutzungskontext korrekt zu interpretieren und anzuwenden. Hierunter fallen eine Übereinstimmung der ökonomischen sowie regulatorischen Richtlinien, Unternehmensziele und Geschäftsprozesse.

Neben den Herausforderungen in den Interoperabilitätskategorien identifizierte GWAC Querschnittsaspekte (vgl. Abbildung 3.2), die zur Erreichung von Interoperabilität notwendig sind [WHA<sup>+</sup>08]. Das GWAC Interoperability Context-Setting Framework verdeutlicht, dass das Erreichen von Interoperabilität keine triviale Aufgabe ist und zu dem Verständnis auch spezifische Prozesse erforderlich sind. Solche Prozesskategorien umfassen unter anderem [RSJ<sup>+</sup>23]:

- **Standardisierung:** Die Standardisierung fördert in der Regel die Implementierung und Nutzung bewährter Verfahren für technische Lösungen und Systeme, was ein gemeinsames Verständnis sowie eine nahtlose Interoperabilität zwischen verschiedenen Anbietern unterstützt [RSJ<sup>+</sup>23]. Die Beteiligung an der Entwicklung neuer Standards sowie die Nutzung bestehender Standards können die Interoperabilität steigern.



Quelle: Basierend auf [WHA<sup>+</sup>08]

Abbildung 3.2: GridWise Architecture Council Interoperability Context-Setting Framework

*Beispiele:* IEC 61850 [IEC21a], Common Information Model (CIM) [IEC21b] und Common Grid Model Exchange Standard (CGMES) [IEC21c] definieren die Struktur und den Austausch von Daten insbesondere zwischen den Akteuren wie Transmission System Operator (TSO) (dt. Übertragungsnetzbetreiber) und Distribution System Operator (DSO) (dt. Verteilnetzbetreiber).

- **Anforderungsmanagement:** Methoden des Anforderungsmanagements fördern ein gemeinsames Verständnis der Stakeholder für das System-of-Interest und bieten Ansätze zur Bewältigung der Herausforderungen der Interoperabilitätsthemen [USD<sup>+</sup>13]. *Beispiele:* Neben Verfahren des klassischen Anforderungsmanagements (z. B. Nutzerbefragungen über Interviews, generelle Use Cases) werden im Energiesektor die IEC 62559 Use Case-Methodik [IEC19] oder das SGAM Framework [CEN14] als Verfahren zur Erfassung von Anforderungen an dem System-of-Interest in Forschungs- und Innovationsprojekten verwendet [USD<sup>+</sup>13] [RSJ<sup>+</sup>23].
- **Interoperabilitäts- und Konformenztests:** Konformitätstests sind Prozesse, die sicherstellen, dass die Implementierung spezifizierten Standards und Profilen entspricht, indem die Ausgaben bzw. Antworten der Implementierung gegen Muster und Regeln geprüft werden, während Interoperabilitätstests sich darauf konzentrieren, zu überprüfen, ob zwei oder mehr Systeme in der Lage sind, Informationen auszutauschen und diese Informationen gegenseitig zu nutzen [FGP19]. *Beispiele:* Das Projekt Integrating the Energy System (IES) bietet ein Rahmenwerk, wie Konformitäts- sowie Interoperabilitätstests ausgeführt werden können [GFF<sup>+</sup>18]. Über die IES/Energy Connectathons treffen sich Akteure wie TSOs, DSOs, Energy Service

Providers (ESPs) sowie Lieferanten mit ihren Systemen, um mit diesen Interoperabilitätstests für konkrete Use Cases durchzuführen. Hierbei wird das im Gesundheitssektor etablierte Tool IHE Gazelle für die Durchführung vorgesehen [GFF<sup>+</sup>18, SWK<sup>+</sup>24].

- **Tracking von Interoperabilitätsreife:** Die Verwendung von Reifegradmodellen unterstützt Organisationen und Communities dabei, ihre Interoperabilitätsziele und -bemühungen zu messen und zu verfolgen. Durch die Bewertung dieser Modelle erhalten sie Hilfestellungen, um die nächsthöheren Reifestufen zu erreichen. *Beispiele:* Smart Grid Interoperability Maturity Model [WLMD10] oder Evaluating the Maturity of Interoperability for the ENergy (EMINENT) [VGHK<sup>+</sup>25] stellen Reifegradmodelle für die Messung von Interoperabilitätsbemühungen dar.

### 3.1.3 Architekturbeschreibung nach ISO/IEC/IEEE 42010

Ein Weg die Interoperabilitätsziele zu erreichen ist der Einsatz von Verfahren des Systems Engineering für die Beschreibung des System-of-Interests (vgl. Definition 5), welcher im System-of-Systems Kontext ein Constituent System (vgl. Definition 6) darstellt. Das Rahmenwerk einer adäquaten Beschreibung des System-of-Interests wird über den ISO/IEC/IEEE 42010 Standard „Systems and software engineering — Architecture description“ definiert [ISO11].

#### **Definition 5** (System-of-Interest)

Das **System-of-Interest** (dt. System von Interesse) ist das System, dessen Lebenszyklus im Kontext dieses Dokuments betrachtet wird. (übersetzt aus [ISO19])

#### **Definition 6** (Constituent System (CS))

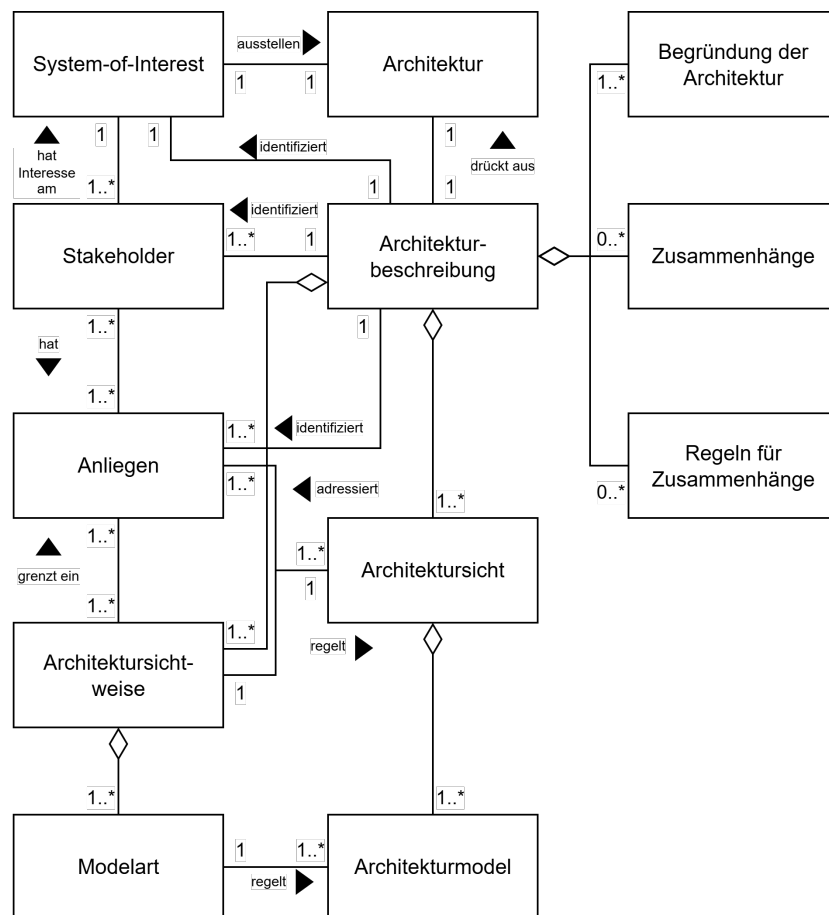
Ein **Constituent System** (dt. Konstituierendes System) ist ein selbstständiges System, das Teil eines System-of-Systems (vgl. Definition 3) ist.

*Anmerkung 1 zum Eintrag:* Constituent Systems können Teil eines oder mehrerer System-of-Systems (vgl. Definition 3) sein. Jedes Constituent System ist an sich ein nützliches System, das seine eigene Entwicklung, Verwaltung, Nutzung, Ziele und Ressourcen hat, interagiert jedoch innerhalb des System-of-Systems, um den einzigartigen Leistungsumfang bereitzustellen. (übersetzt aus [ISO19])

Abbildung 3.3 zeigt das konzeptionelle Modell der Architekturbeschreibung gemäß ISO/IEC/IEEE 42010. Eine *Architekturbeschreibung* soll die vorliegende *Architektur* eines Systems dokumentieren. Der Begriff *Architektur* wird von [ISO11] als „grundlegende Konzepte oder Eigenschaften eines Systems in seiner Umgebung, die sich in seinen Elementen, Beziehungen und den Grundsätzen seiner Gestaltung und Entwicklung widerspiegeln“ definiert. Für das betrachtete *System-of-Interest* existiert sowohl eine Architektur als auch eine entsprechende Architekturbeschreibung. Diese Architekturbeschreibung erfasst die *Stakeholder* des Systems, die ein Interesse an dem System-of-Interest

haben. Stakeholder haben verschiedene *Anliegen*, die in der Architekturbeschreibung identifiziert und adressiert werden müssen. [ISO11]

Eine Architekturbeschreibung besteht aus mehreren Komponenten. Hierbei kann eine Architekturbeschreibung eine oder mehrere *Architektursichtweisen* enthalten, welche die unterschiedlichen Anliegen der Stakeholder eingrenzt. Diese Architektursichtweisen regeln die *Architekturansichten*, welche ebenfalls als Teile der Architekturbeschreibung gesehen werden. Eine Architektursichtweise besteht aus einem oder mehreren *Modelltypen*, welche die Architekturmodelle regelt, welche wiederum Bestandteil der *Architekturansicht* sind. Zudem liegen einer Architekturbeschreibung eine oder mehrere *Begründungen für die Architektur* vor, welche Architekturentscheidungen erklärt, begründet oder rechtfertigt. Einer Architekturbeschreibung können *Zusammenhänge* und *Regeln für Zusammenhänge* vorliegen. [ISO11]



Quelle: Basierend auf [ISO11]

Abbildung 3.3: Konzeptionelles Modell von Architekturbeschreibungen gemäß ISO/IEC/IEEE 42010

Sowohl die im Abschnitt 3.2 behandelte IEC 62559 Use Case-Methodik als auch das in Abschnitt 3.3 erläuterte SGAM Framework basieren gemeinsam auf dem vorliegenden Modell als Metamodell, um das System-of-Interest adäquat und vollumfassend zu beschreiben und somit ISO/IEC/IEEE 42010-konforme Architekturbeschreibungen zu ermöglichen. Beide Methoden können eigenständig verwendet werden, sind jedoch nah verbunden und ergänzen sich gegenseitig. [URN<sup>+</sup>19]

### 3.2 Anwendungsfallmethodik nach IEC 62559 (Use Case Methodology)

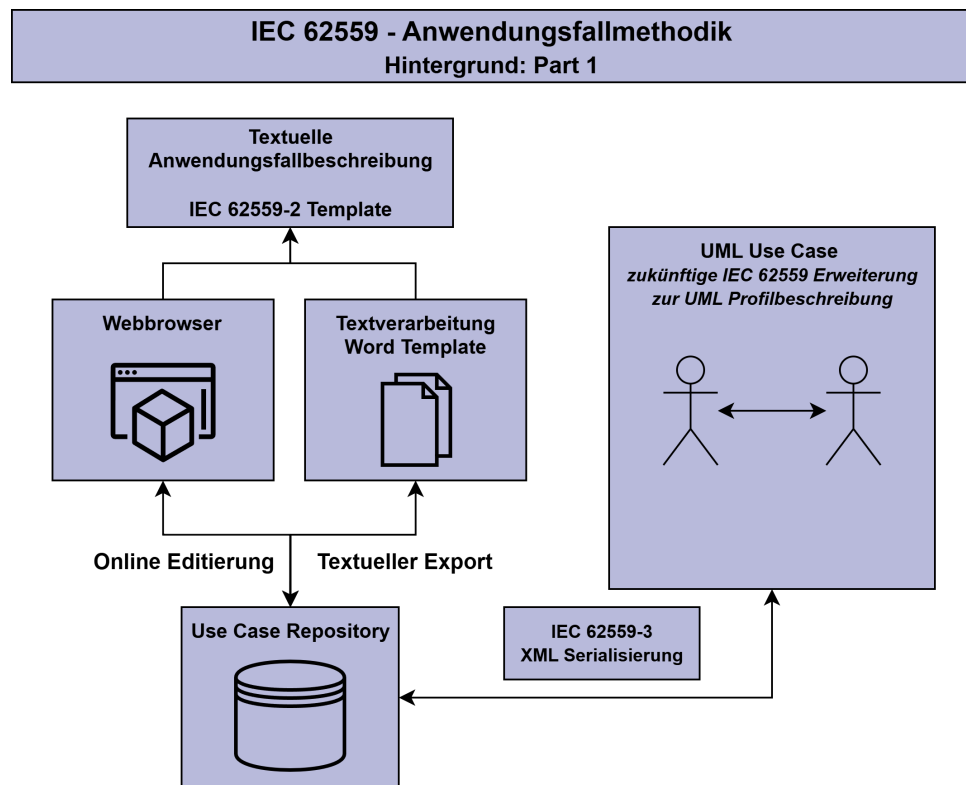
Die Use Case-Methodik (dt. Anwendungsfallmethodik) nach IEC 62559 stellt eine im Energiebereich etablierte, systematische Herangehensweise für die Erfassung von Use Cases für komplexe Systeme dar [IEC19]. Definition 7 repräsentiert eine im Standard IEC 62559-1 verwendete Definition eines Use Cases. Die Methodik wurde insbesondere von der IEC Technical Committee (TC) 8 entwickelt und dient als Basis für IEC Standardisierungsarbeiten [IEC19], aber auch für das Anforderungsmanagement für komplexe Systeme [GDU17].

#### **Definition 7** (Use Case)

Ein **Use Case** (dt. Anwendungsfall) ist eine Spezifikation einer Reihe von Aktionen, die von einem System ausgeführt werden, welche ein beobachtbares Ergebnis erzeugt, das typischerweise für einen oder mehrere Akteure oder andere Interessengruppen des Systems von Wert ist. (übersetzt aus [ISO12])

Der Aufbau der Standardserie zur Anwendungsfallmethodik ist in Abbildung 3.4 dargestellt. Die Motivation und Hintergründe zur Anwendungsfallmethodik sind im Teil IEC 62559-1 erläutert [IEC19]. Dieser sieht zudem vor, dass Use Cases als Grundlage für Standardisierungsarbeiten erhoben werden und beschreibt neben dem Aufbau der Standardserie auch die vorgesehenen Prozesse, Terminologien, Arten von Use Cases sowie das Tool Use Case Repository, welches zur Verwaltung und Verteilung von Anwendungsfällen verwendet wird [IEC19]. Anwendungsfälle werden in der Methodik über strukturierte Anwendungsfallbeschreibungen dokumentiert, deren Struktur im Teil IEC 62559-2 festgelegt ist und für Textverarbeitungsprogramme vorgesehen ist [IEC15]. Eine maschinenlesbare XML-Serialisierung einschließlich der Unified Modeling Language (UML) Klassendiagramme sind im Teil IEC 62559-3 spezifiziert [IEC18]. Der Teil IEC 62559-4 (nicht in der Abbildung enthalten) beschreibt Best Practices im Kontext der IEC-Normierungsprozesse sowie einige Beispiele außerhalb der Normierung [SyC18].

In dem Entwicklungsprozess von IEC 62559 Use Cases werden unterschiedliche Skills und Aufgaben realisiert, weshalb der Standard die folgenden Rollen vorsieht: [IEC19]



Quelle: Basierend auf [IEC19]

Abbildung 3.4: Aufbau der IEC 62559 Standard Serie

- Die **Anwendungsfallentwickler** sind entweder Domänen- oder technische Experten und sind mit dem IEC 62559-2 Use Case Template vertraut.
- Das **Qualitätsteam** entwickelt die Datenqualitätsziele und -regeln und stellt deren Einhaltung bei neuen und aktualisierten IEC 62559 Use Cases sicher.
- Das **Validierungsteam** gewährleistet die Konsistenz und Vollständigkeit der existierenden IEC 62559 Use Cases des gesamten UCR.
- Die **UC(M)R-Administratoren** sind für die technische Verwaltung des Tools verantwortlich.

### 3.2.1 IEC 62559 Use Cases

Im Rahmen der Dissertation werden die durch Anwendung der IEC 62559 Use Case-Methodik zur Beschreibung des System-of-Interest entwickelten Anwendungsfälle als IEC 62559 Use Cases bezeichnet. Hierbei definiert der Standard IEC 62559-2 „Use case methodology - Part 2: Definition of the templates for use cases, actor list and require-

ments list“ das Template, wie solche Use Cases aufgebaut sind. Folglich werden IEC 62559 Use Cases gemäß der Definition 8 spezifiziert.

**Definition 8** (IEC 62559 Use Case)

Ein **IEC 62559 Use Case** (oder auch IEC 62559(-2) Use Case-Beschreibung) ist ein spezifischer Anwendungsfall, der gemäß dem IEC 62559-2 Use Case-Template dokumentiert ist.

Der folgende Unterabschnitt erläutert den Aufbau des Templates sowie die verschiedenen Typen von Use Cases.

### 3.2.1.1 IEC 62559-2 Use Case Template

Das IEC 62559-2 Use Case Template bietet die Struktur zur Dokumentation von Anwendungsfällen gemäß der IEC 62559 Use Case-Methodik. Es existiert sowohl in einer Kurz- als auch in einer Langform.

Im Folgenden wird die Langform des Templates anhand der Struktur erläutert und anhand eines Beispiels aus Abschnitt A.3 verdeutlicht. Der Standard sieht hierbei die Nutzung eines textverarbeitungsprogrammgerichten Templates (z. B. in Microsoft Word) vor, ermöglicht aber auch die digitale Erfassung über das im Standard 62559-1 [IEC19] eingeführte Konzept des UCR [IEC15]. Da im Rahmen der Dissertation überwiegend das englische Template gemäß dem IEC 62559-2 Use Case-Template Standard betrachtet wird, werden die Bezeichnungen im Folgenden sowohl in deutscher Form (primär) als auch in englischer Form (sekundär) dargestellt, um eine eindeutige Zuordnung zu gewährleisten.

#### 1. Beschreibung des Anwendungsfalls (Description of the Use Case)

Die erste Sektion des Templates befasst sich mit der einleitenden Beschreibung des Anwendungsfalls, seiner Einordnung sowie der Aufführung von allgemeinen Metainformationen. Diese Sektion ist in acht Unterabschnitte gegliedert.

**1.1 Name des Anwendungsfalls (Name of Use Case)** Innerhalb dieses Abschnitts werden der Name des Anwendungsfalls sowie eine ID angegeben. Zudem wird der Anwendungsfall gemäß der vorliegenden Referenzarchitektur in Domänen und Zonen eingeordnet. Die im Standard als Beispiel referenzierte Referenzarchitektur für den Energiesektor stellt das SGAM dar. Hierbei wird der Use Case in die SGAM-Domänen und -Zonen eingeordnet [IEC15]. Abbildung A.21 zeigt, wie die Sektion „Use Case Identification“ des Templates für ein Beispiel ausgefüllt wurde.

**1.2 Versionsverwaltung (Version Management)** Die Versionierung der Anwendungsfallbeschreibung wird in diesem Abschnitt behandelt. Hierbei werden die Autoren aufgeführt, die an den verschiedenen Versionen mitgewirkt haben. Es werden sowohl das

Datum der Version als auch die vorgenommenen Änderungen sowie der Versionsstatus angegeben [IEC15]. Tabelle A.22 zeigt ein ausgefülltes Beispiel dieser Sektion.

**1.3 Umfang und Ziele des Anwendungsfalls (Scope and Objectives of Use Case)** In diesem Abschnitt wird der Umfang (Scope) des Projekts definiert, welcher den spezifischen Anwendungsfall eingrenzt. Die Ziele werden festgelegt und die damit verbundenen Geschäftsvorfälle beschrieben [IEC15]. Abbildung A.23 stellt ein Beispiel für das Ausfüllen dieser Informationen dar.

**1.4 Beschreibung des Anwendungsfalls (Narrative of the Use Case)** Der Abschnitt enthält sowohl eine kurze als auch eine vollständige Beschreibung des Anwendungsfalls. Während der Standard eine maximale Länge von 150 Wörtern für die kurze Beschreibung empfiehlt, kann die vollständige Beschreibung je nach Komplexität von einigen Sätzen bis hin zu mehreren Seiten reichen [IEC15]. Tabelle A.24 zeigt ein Beispiel für die ausgefüllte Kurz- sowie Langbeschreibung.

**1.5 Leistungskennzahlen (Key Performance Indicators, KPI)** Die für den Anwendungsfall relevanten Key Performance Indicators (KPIs) werden in diesem Abschnitt aufgeführt. Hierbei wird jeder KPI mit einer ID, einem Namen und einer Beschreibung versehen. Zudem werden die KPIs in Bezug auf die Zielsetzungen referenziert [IEC15]. Abbildung A.25 zeigt ein Beispiel für eine ausgefüllte Sektion, die eine KI beinhaltet.

**1.6 Bedingungen des Anwendungsfalls (Use Case Conditions)** Dieser Abschnitt sieht die Beschreibung von vorgenommenen Annahmen des Anwendungsfalls, aber auch die notwendigen Vorbedingungen vor [IEC15]. Tabelle A.26 zeigt ein Beispiel für eine ausgefüllte Sektion, die Annahmen und Vorbedingungen aufführt.

**1.7 Weitere Informationen zur Klassifikation / Zuordnung des Anwendungsfalls (Further Information to the Use Case for Classification / Mapping)** Der Abschnitt enthält zusätzliche Meta- sowie Klassifikationsinformationen. Hierzu gehören Referenzen zu anderen Anwendungsfällen, der Grad der Detailtiefe, die Priorisierung (relevant für die Betrachtung des gesamten Use Case-Clusters), der räumliche Bezug des Anwendungsfalls, sowie die Art des Anwendungsfalls (Aufführung unterschiedlicher Arten vgl. Abschnitt 3.2.1.2) sowie eine Auflistung von Schlüsselwörtern (Keywords) [IEC15]. Tabelle A.27 zeigt ein ausgefülltes Beispiel für diese Sektion.

**1.8 Allgemeine Anmerkungen (General Remarks)** Dieser Abschnitt besteht aus Kommentaren, welche nirgendwo anders in der Anwendungsfallbeschreibung benannt sind [IEC15]. Ein Beispiel für solche Kommentare ist in Tabelle A.28 enthalten.

## 2. Diagramme des Anwendungsfalls (Diagrams of Use Case)

In diesem Abschnitt werden Abbildungen wie generelle Grafiken sowie UML-Diagramme (z. B. Anwendungsfalldiagramme, Sequenzdiagramme und Aktivitätsdiagramme) zur Veranschaulichung des Anwendungsfalls eingesetzt. Der Standard sieht vor, dass Diagramme Interaktionen darstellen sollten [IEC15]. Tabelle A.29 zeigt ein Beispiel für den Diagrammteil.

## 3. Technische Details (Technical Details)

Diese Sektion befasst sich mit der Aufführung der Akteure (im Abschnitt *3.1 Akteure / Actors*) sowie der verwendeten Referenzen (im Abschnitt *3.2 Referenzen / References*).

**3.1 Akteure (Actors)** Dieser Abschnitt befasst sich mit der Darstellung der Akteure, die für den Anwendungsfall relevant sind. Hierbei können Akteure innerhalb zugehöriger Gruppen eingeordnet werden. Eine Gruppe weist einen Namen sowie eine Beschreibung auf. Akteure werden ebenfalls mit einem Namen und einer Beschreibung versehen und einem Akteurstypen zugeordnet. Zudem werden spezifische Informationen für diesen Anwendungsfall angegeben. Dies kann gemäß Standard beispielsweise eine Einordnung in das SGAM-Framework darstellen [IEC15]. Tabelle A.30 stellt eine beispielhafte Übersicht der Akteure dar.

**3.2 Referenzen (References)** Der Anwendungsfall sowie die Anwendungsfallbeschreibung verwenden Dokumente im Sinne von Standards, Deliverables, technischen Spezifikationen, Handbüchern, Publikationen etc. Dieser Abschnitt soll eine Übersicht der verwendeten Literatur (im Sinne von Referenzen) aufführen. Das Template sieht hierbei eine Identifikation über eine Nummer, eine Einordnung des Referenztyps, den Titel der Referenz, den Reifestatus des Dokuments, die Einschätzung der Auswirkungen auf den Anwendungsfall, die dahinterstehende bzw. veröffentlichende Organisation/den Urheber sowie einen Link vor [IEC15, GDU17]. Tabelle A.32 zeigt ein Beispiel für die Aufführung der Referenzen.

## 4. Schrittweise Analyse des Anwendungsfalls (Step by Step Analysis of Use Case)

Ein Kernaspekt des Use Case-Templates ist die detaillierte Szenariobeschreibung des Anwendungsfalls. Diese soll so gestaltet sein, dass ein klarer Zusammenhang zur Anwendungsfallbeschreibung (Narrativ) erkennbar ist. Die Sektion ist in den Abschnitt *4.1 Übersicht der Szenarien / Overview of Scenarios* sowie *4.2 Schritte – Szenarien / Steps – Scenarios* untergliedert [IEC15].

**4.1 Übersicht der Szenarien (Overview of Scenarios)** Die Sektion beginnt mit einer Übersicht der Szenarien. Jedes Szenario ist mit einer Nummer (sequenziell stei-

gend), einem Szenarionamen, einer Beschreibung, einer Auflistung der primären Akteure, dem auslösenden Ereignis sowie Vor- und Nachbedingungen versehen [IEC15]. Eine beispielhafte Szenarioübersicht ist in Tabelle A.33 dargestellt.

**4.2 Schritte – Szenarien (Steps – Scenarios)** Der Folgeabschnitt befasst sich mit der detaillierten Beschreibung der einzelnen Schritte der Szenarien. Hierbei wird für jedes Szenario ein Abschnitt angelegt, der die Beschreibung der vorgesehenen Schritte enthält. Jeder Schritt ist einer sequenziell steigenden Nummer (alphanumerisch) zugewiesen und umfasst einen Service, der den Informationsaustausch mit HTTP-ähnlichen Verben beschreibt. Zudem werden die Akteure, die Informationen produzieren und konsumieren, die IDs der ausgetauschten Informationen sowie die IDs der zugehörigen Anforderungen angegeben [IEC15]. Tabelle A.34 zeigt einen beispielhaften Abschnitt zur Beschreibung der Schritte eines Szenarios.

#### 5. Ausgetauschte Informationen (Information Exchanged)

Dieser Abschnitt listet die ausgetauschten Informationen des Anwendungsfalls auf. Dabei werden die definierten Informationsaustausche jeweils mit einer ID, einem Namen, einer Beschreibung sowie ggf. einer Referenz zu den zugehörigen Anforderungen versehen [IEC15]. Tabelle A.35 zeigt ein Beispiel für die Auflistung der ausgetauschten Informationen.

#### 6. Anforderungen (Requirements)

Ein optionaler Abschnitt umfasst die Auflistung der nicht-funktionalen Anforderungen. Diese Anforderungen können innerhalb der Sektion „Schrittweise Analyse“ verlinkt werden. Anforderungen, die jeweils mit einer ID, einem Namen sowie einer Beschreibung aufgeführt sind, können im Dokument über zugehörige Kategorien kategorisiert werden, die ebenfalls eine ID, einen Namen und eine Beschreibung aufweisen [IEC15]. Ein Beispielabschnitt ist in Tabelle A.36 aufgeführt.

#### 7. Fachbegriffe und Definitionen (Common Terms and Definitions)

In den Anwendungsfällen werden fachbezogene Begriffe und Akronyme verwendet, die es zu definieren gilt [IEC15]. Ein Beispiel für einen solchen Abschnitt ist in Tabelle A.37 dargestellt.

#### 8. Zusätzliche Informationen (Custom Information)

Zusätzliche Informationen, die nicht in die vorgegebenen Bereiche des Templates passen, können in einer Key-Value-Darstellung mit Verweis auf die zugehörige Sektion angegeben werden [IEC15]. Tabelle A.38 zeigt ein Beispiel für einen solchen (optionalen) Abschnitt.

### 3.2.1.2 Arten von Use Cases

Nicht nur auf IEC 62559 Use Cases, sondern auch auf allgemeine Use Cases bezogen, ist je nach Kontext die Verwendung von bestimmten Arten von Use Cases erforderlich [Coc01]. Das IEC 62559-2 Use Case Template (vgl. Abschnitt 3.2.1.1) sieht im Bereich *1. Description of the use case* Einordnungen der Use Case-Typen vor. Des Weiteren werden die Wahl der Akteure, technische Spezifikationen und die Szenariobeschreibung maßgeblich vom jeweiligen Use Case-Typ beeinflusst [IEC15]. Der Grad der Spezifikation kann dabei zwischen *High-Level* und *Low-Level Use Cases* variieren:

- Ein **High-Level Use Case (HLUC)** ist ein Anwendungsfall, der eine Idee oder ein Konzept unabhängig von einer spezifischen technischen Realisierung wie einer architektonischen Lösung beschreibt. [IEC15]
- Ein **Low-Level Use Case (LLUC)** bzw. ein **Specialized Use Case** ist ein Anwendungsfall, der eine Idee oder ein Konzept konkret spezifisch auf technische Realisierung beschreibt. [IEC15]

Der Betrachtungswinkel, ob der Use Case aus technischer Implementierungssicht oder auf Geschäftsprozessebene betrachtet wird, unterscheidet sich in *System/Technical Use Cases* und *Business Use Cases*:

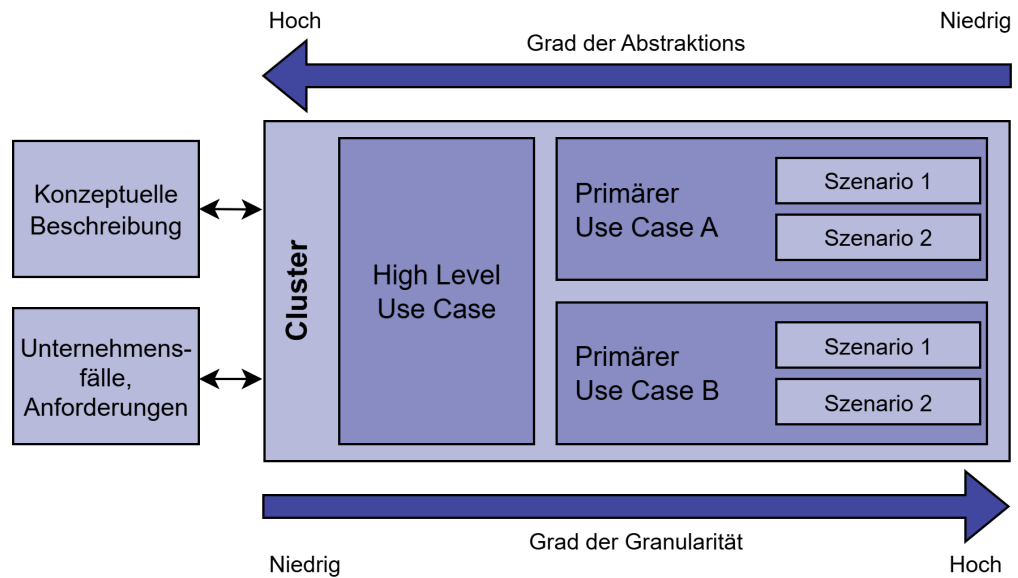
- Bei einem **System Use Case (SUC) oder Technical Use Case (TUC)** liegt der Schwerpunkt auf dem (typischerweise) zu entwickelnden Computersystem. [Coc01]
- Bei einem **Business Use Case (BUC)** liegt der Schwerpunkt auf dem geschäftlichen Betrieb und nicht auf dem Betrieb des Computersystems. [Coc01]

Use Cases, die konkret in einem System implementiert werden sollen, können in Beziehungen zueinander stehen. Hierbei wird zwischen *Primary Use Cases* und *Secondary Use Cases* unterschieden:

- Ein **Primary Use Case** ist ein Anwendungsfall, der in einem bestimmten System implementiert wird und durch eine definierte Grenze gekennzeichnet ist (d.h. er kann auf eine definierte Architektur abgebildet werden). [IEC19]
- Ein **Secondary Use Case** ist ein Anwendungsfall, um Kernfunktionalitäten zu beschreiben, die von mehreren PUCs genutzt werden. [IEC19]

Use Cases können entweder *generisch* für die breite Masse oder *spezifisch* auf ein konkretes Projekt, Unternehmen oder eine Organisation zugeschnitten sein:

- Ein **Generic Use Case** ist ein Anwendungsfall, der für die Standardisierung breit akzeptiert ist und in der Regel verschiedene individuelle Anwendungsfälle sammelt und harmonisiert, ohne auf einer projekt- oder technologiespezifischen Lösung zu basieren. [IEC15]



Quelle: Basierend auf [IEC19]

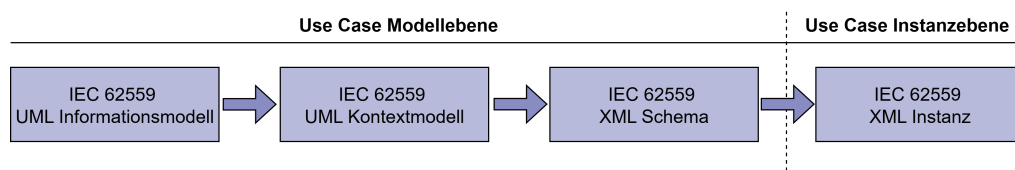
Abbildung 3.5: Zusammenfassung von Use Cases in einem Cluster

- Ein **Individual Use Case** ist ein Anwendungsfall, der spezifisch für ein Projekt oder innerhalb eines Unternehmens/einer Organisation verwendet wird. [IEC15]

Im Rahmen der IEC 62559 Methodik lassen sich mit Abbildung 3.5 einige benannte Arten von Use Cases innerhalb eines Clusters darstellen. Ein Cluster stellt hierbei eine Ansammlung von zugehörigen IEC 62559 Use Cases dar (z. B. Smart EV Charging) [IEC19]. Die Art des Use Cases bestimmt auch die Granularität bzw. den Abstraktionsgrad. Ein High-Level Use Case weist einen hohen Abstraktionsgrad und eine niedrige Granularität auf, während ein Primary Use Case eher einen geringeren Abstraktionsgrad und eine höhere Granularität besitzt [IEC18].

### 3.2.2 Use Case XML-Serialisierung

Das IEC 62559-2 Template sieht durch die Struktur eine Verwendung von Textverarbeitungsprogrammen wie Microsoft Word vor [IEC15]. Da Textdaten im docx-Format unstrukturierte Daten darstellen, sind sie für die maschinelle Weiterverarbeitung in z. B. UML-Modellierungstools oder im Use Case Repository nicht geeignet. Der Standard IEC 62559-3 definiert eine XML-Serialisierung, die eine maschinenlesbare Darstellung von Use Cases ermöglicht und somit Export- und Importmöglichkeiten in (Third-Party)-Tools sowie dem Use Case Repository bietet [IEC18]. Dabei legt der Standard das Rahmenwerk gemäß Abbildung 3.6 fest.



Quelle: Basierend auf [IEC18]

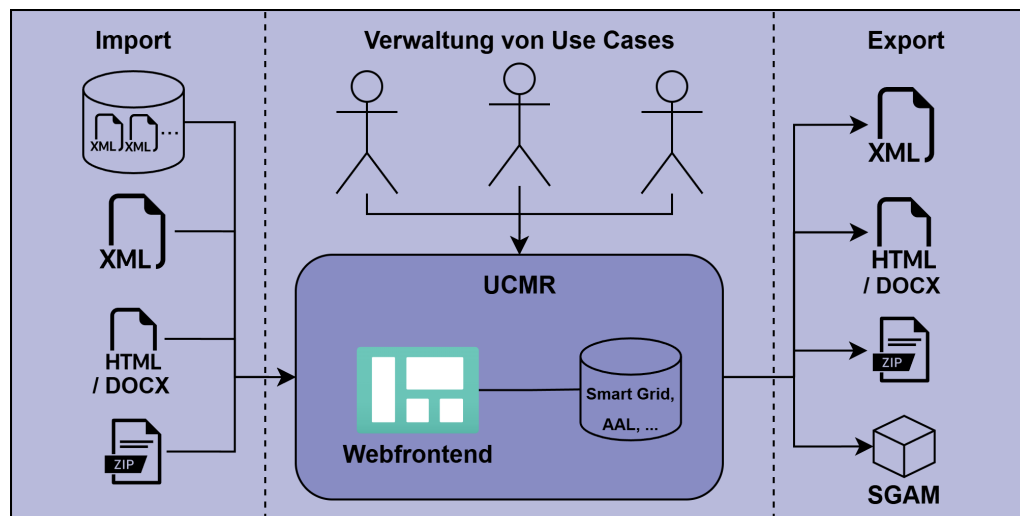
Abbildung 3.6: Rahmenwerk zur Definition des XML-basierten Datenaustauschformats des IEC 62559-3 Standards

Der Standard IEC 62559-3 unterscheidet zwischen der *Use Case Modellebene*, die das Metamodell definiert, und der *Use Case Instanzebene*, die einen spezifischen IEC 62559 Use Case auf Grundlage dieses Modells beschreibt. Das UML-Informationsmodell basiert auf den Konzepten des IEC 62559-2 Standards. Während das Informationsmodell weitestgehend generisch nutzbar ist, beschreibt das UML-Kontextmodell ein restriktiveres Modell, das den Austausch von Use Cases definiert. Das zugrunde liegende XML-Schema ist konform zum kontextuellen Modell und wird im XML Schema Definition (XSD)-Format beschrieben, welches die Struktur für konkrete XML-Instanzdateien vorgibt. [IEC18]

### 3.2.3 Use Case Repository (UCR)

Ein im Standard IEC 62559 verankertes Tool zur Verwaltung und Bereitstellung stellt das Use Case Repository dar [IEC19], welches sowohl im Rahmen der Literatur als auch in dieser Dissertation als Use Case Management Repository (UCMR) bezeichnet wird [GDU17]. Durch die Verwendung globaler Akteure und Anforderungen sowie die Harmonisierung der angelegten Use Cases sollen hierbei Unstimmigkeiten in der Namensgebung, Unterschiede in der Granularität und Unvollständigkeit reduziert werden [IEC19]. Der Standard spezifiziert die Anforderungen, die das Tool erfüllen muss. Ein Aspekt ist die Bereitstellung vordefinierter Einträge, um die Wiederverwendung von Elementen zu motivieren [IEC19]. Des Weiteren soll das UCR die Integration anderer (Standardisierungs-)Prozesse und Toolumgebungen ermöglichen [IEC19]. Der Standard identifiziert auch Stakeholder, die am UCMR beteiligt sind [IEC19]:

- **Editoren** erstellen IEC 62559 Use Cases.
- **Reviewer** überprüfen IEC 62559 Use Cases und geben als Qualitätssicherungsinstanz den Editoren Feedback, haben aber keine Schreibrechte.
- **Leser** lesen IEC 62559 Use Cases.
- **Administratoren** verwalten das Use Case Repository und sind für das Nutzermanagement verantwortlich.



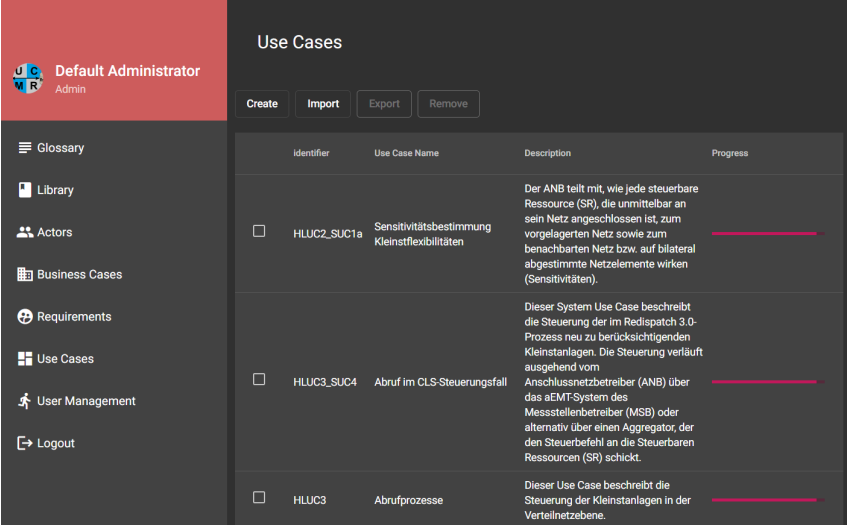
Quelle: Basierend auf [GDU17]

Abbildung 3.7: Generelles Design des Use Case Management Repository

Abbildung 3.7 zeigt das generelle Design des Use Case Management Repository. Das Tool soll als webbasierte Anwendung zur Verfügung gestellt werden und nicht nur die zentrale Speicherung von Use Cases im Smart Grid-Kontext, sondern auch deren Nutzung in anderen Kontexten und Domänen unterstützen. Das Konzept umfasst Import- und Exportmöglichkeiten von IEC 62559 Use Cases mittels IEC 62559-3 Use Case XML-Serialization sowie die Unterstützung weiterer Formate wie HTML/DOCX und ZIP-Archive. Zusätzlich ist eine zukünftige Integration zur Generierung von SGAM Modellen auf Basis der IEC 62559 Use Case-Beschreibungen geplant. [GDU17]

Zum Zeitpunkt dieser Dissertation liegt das OFFIS UCMR in der Version 3.0 vor und bietet eine leichtgewichtige, standardorientierte Implementierung für die Verwaltung und Bereitstellung von IEC 62559 Use Cases. Es basiert auf dem UML-Informationsmodell des IEC 62559-3 Standards (vgl. Abschnitt 3.2.2) und verwendet globale Elemente (Assets). Zudem ermöglicht es den Import und Export von XML- sowie DOCX-Dateien. Abbildung 3.8 zeigt einen Screenshot der OFFIS UCMR-Instanz für das Redispatch 3.0 Projekt ([KOH<sup>+</sup>23]), der die Übersicht der angelegten Anwendungsfälle und Menüpunkte veranschaulicht. Diese Version wird in den Projekten Redispatch 3.0 sowie IDiNA eingesetzt.

Das öffentlich zugängliche und von der EU Kommission geförderte BRIDGE Use Case Repository basiert ebenfalls auf dem IEC 62559 Standard und nutzt für die Verwaltung sowie der Bereitstellung von Markdown-Dateien sowie ein öffentliches GitHub Repository [DAZ<sup>+</sup>23a] anstelle einer dedizierten Web-Applikation. Abbildung 3.9 zeigt die Übersichtsseite des BRIDGE Repository.

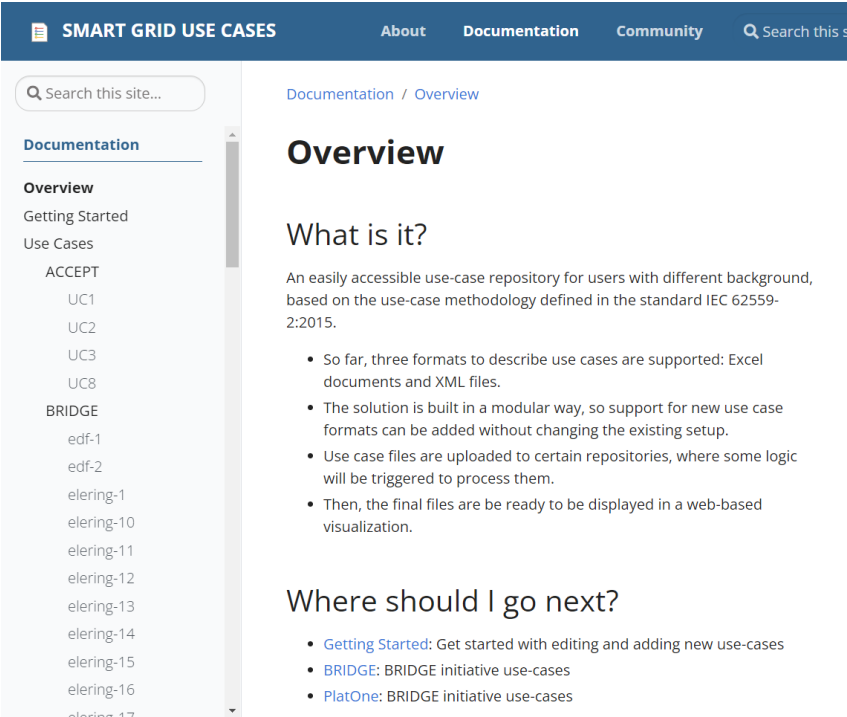


The screenshot shows the 'Use Cases' management interface. At the top left, it identifies the user as 'Default Administrator Admin'. Below this is a navigation menu with options like 'Glossary', 'Library', 'Actors', 'Business Cases', 'Requirements', 'Use Cases', 'User Management', and 'Logout'. The main area contains a table of use cases with columns for 'Identifier', 'Use Case Name', 'Description', and 'Progress'. There are buttons for 'Create', 'Import', 'Export', and 'Remove' at the top of the table.

Identifier	Use Case Name	Description	Progress
<input type="checkbox"/> HLUC2_SUC1a	Sensitivitätsbestimmung Kleinflexibilitäten	Der ANB teilt mit, wie jede steuerbare Ressource (SR), die unmittelbar an sein Netz angeschlossen ist, zum vorgelagerten Netz sowie zum benachbarten Netz bzw. auf bilateral abgestimmte Netzelemente wirken (Sensitivitäten).	Progress bar
<input type="checkbox"/> HLUC3_SUC4	Abruf im CLS-Steuerungsfall	Dieser System Use Case beschreibt die Steuerung der im Redispatch 3.0-Prozess neu zu berücksichtigenden Kleinanlagen. Die Steuerung verläuft ausgehend vom Anschlussnetzbetreiber (ANB) über das aEMTS-System des Messstellenbetreiber (MSB) oder alternativ über einen Aggregator, der den Steuerbefehl an die Steuerbaren Ressourcen (SR) schickt.	Progress bar
<input type="checkbox"/> HLUC3	Abrufprozesse	Dieser Use Case beschreibt die Steuerung der Kleinanlagen in der Verteilnetzebene.	Progress bar

Quelle: Eigene Darstellung, Redispatch 3.0 UCMR Instanz [KOH<sup>+</sup>23]

Abbildung 3.8: Screenshot des OFFIS Use Case Management Repository



The screenshot shows the 'SMART GRID USE CASES' documentation page. The top navigation bar includes 'About', 'Documentation', and 'Community', along with a search bar. The left sidebar shows a 'Documentation' menu with 'Overview' selected. The main content area is titled 'Overview' and includes a section 'What is it?' which describes the repository as an easily accessible use-case repository based on the IEC 62559-2:2015 standard. It lists supported formats (Excel, XML), modular architecture, upload process, and visualization. A 'Where should I go next?' section provides links to 'Getting Started', 'BRIDGE', and 'PlatOne'.

**SMART GRID USE CASES** About Documentation Community Search this site...

Search this site...

Documentation / Overview

## Overview

### What is it?

An easily accessible use-case repository for users with different background, based on the use-case methodology defined in the standard IEC 62559-2:2015.

- So far, three formats to describe use cases are supported: Excel documents and XML files.
- The solution is built in a modular way, so support for new use case formats can be added without changing the existing setup.
- Use case files are uploaded to certain repositories, where some logic will be triggered to process them.
- Then, the final files are ready to be displayed in a web-based visualization.

### Where should I go next?

- [Getting Started](#): Get started with editing and adding new use-cases
- [BRIDGE](#): BRIDGE initiative use-cases
- [PlatOne](#): BRIDGE initiative use-cases

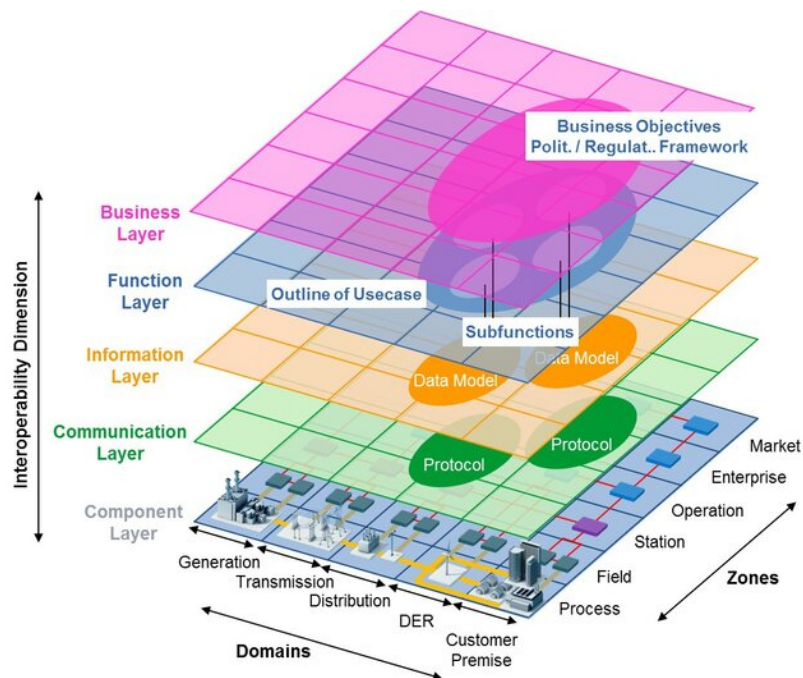
Quelle: [DAZ<sup>+</sup>23a]

Abbildung 3.9: Screenshot des BRIDGE Use Case Repository

### 3.3 Smart Grid Architecture Model (SGAM)-Framework

Ein aus dem europäischen M/490 Mandat entstandenes Rahmenwerk stellt das SGAM-Framework dar. Hierbei soll das Framework die Interoperabilitäts Herausforderungen in der Systementwicklung von Smart Grid-bezogenen Anwendungsfällen adressieren [CEN12]. Das SGAM Framework nimmt eine architektonische und technologie- sowie lösungsunabhängige Perspektive auf das System-of-Systems oder auch dem System-of-Interest ein [CEN12]. Hierbei wird gemäß [CEN14] das SGAM Framework angewendet, um:

- eine strukturierte Analyse von Smart-Grid-Anwendungsfällen zu ermöglichen,
- verschiedene Ansätze zu Smart-Grid-Architekturen, Paradigmen, Roadmaps und Perspektiven zu visualisieren und zu vergleichen,
- einen Leitfaden zur Analyse potenzieller Implementierungsszenarien bereitzustellen,
- ein gemeinsames Verständnis zwischen verschiedenen Stakeholder sicherzustellen,
- Standards und Standardisierungslücken zu identifizieren,
- den Umfang von Smart-Grid-Projekten zu visualisieren,
- und zusammengefasst, um mit der Komplexität von Smart Grids umzugehen.



Quelle: [CEN14]

Abbildung 3.10: Smart Grid Architecture Model (SGAM)

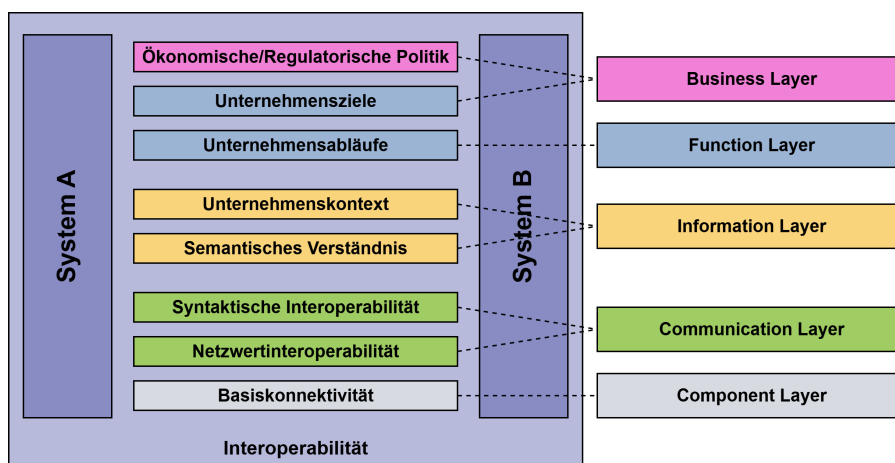
### 3.3.1 Komponenten

Das Framework sieht vor, System(komponenten) und Anwendungsfälle innerhalb einer einheitlichen Visualisierung darzustellen und diese auf die verschiedenen Interoperabilitätsebenen abzubilden [CEN14]. Abbildung 3.10 zeigt eine Visualisierung des SGAM Frameworks im Rahmen eines generischen SGAM Modells auf. Dieses Modell besteht aus drei Dimensionen, die in den Folgeabschnitten erläutert werden. Diese Dimensionen sind [CEN12]:

- Interoperabilitätsebenen, basierend auf den GWAC Interoperability Categories
- Domänen, basierend auf der Energiewertschöpfungskette
- Zonen, basierend auf den hierarchischen Stufen des Energiesystemmanagements

#### 3.3.1.1 SGAM Interoperability Layers

Ein zentrales Ziel des SGAM Frameworks besteht in der Bewältigung der Komplexität der Systementwicklung, insb. hinsichtlich der Interoperabilitäts Herausforderungen [CEN12]. Die im Abschnitt 3.1.2 dargestellten Interoperabilitätskategorien des GWAC Interoperability Context-Setting Frameworks dienen als Grundlage für die Auswahl der Interoperabilitätsebenen im SGAM Framework [CEN12]. Dabei wurde im SGAM Framework bewusst die Entscheidung getroffen, die Anzahl der Interoperabilitätskategorien zu reduzieren, wobei jedoch die Verwendung zusätzlicher Kategorien bei Bedarf ermöglicht wird. Abbildung 3.11 zeigt eine Gegenüberstellung der im GWAC Framework verwendeten Kategorien zu den im SGAM Framework verwendeten Kategorien. Diese werden wie folgt nach [CEN12] betrachtet:



Quelle: Basierend auf [CEN12]

Abbildung 3.11: Mapping der Interoperabilitätskategorien aus dem SGAM Framework und GWAC Interoperability Context-Setting Framework

### Business Layer

Der Business Layer adressiert Geschäftsprozesse, -modelle, regulatorische Rahmenbedingungen sowie marktbezogene Strukturen, die an Smart Grid-Anwendungsfällen beteiligt sind.

### Function Layer

Der Function Layer fokussiert sich auf die funktionalen Aspekte der Smart Grid-Anwendungsfälle. Hier werden die spezifischen Funktionen sowie Services modelliert, die für die Umsetzung der Geschäftsprozesse notwendig sind.

### Information Layer

Der Information Layer befasst sich mit den vorliegenden Informationsaustauschen und deren Datenmodellierung. Er beschreibt die relevanten Datenstrukturen und Informationsflüsse zwischen den Funktionen, einschließlich der Datenformate und semantischen Modelle, die für die Interoperabilität erforderlich sind.

### Communication Layer

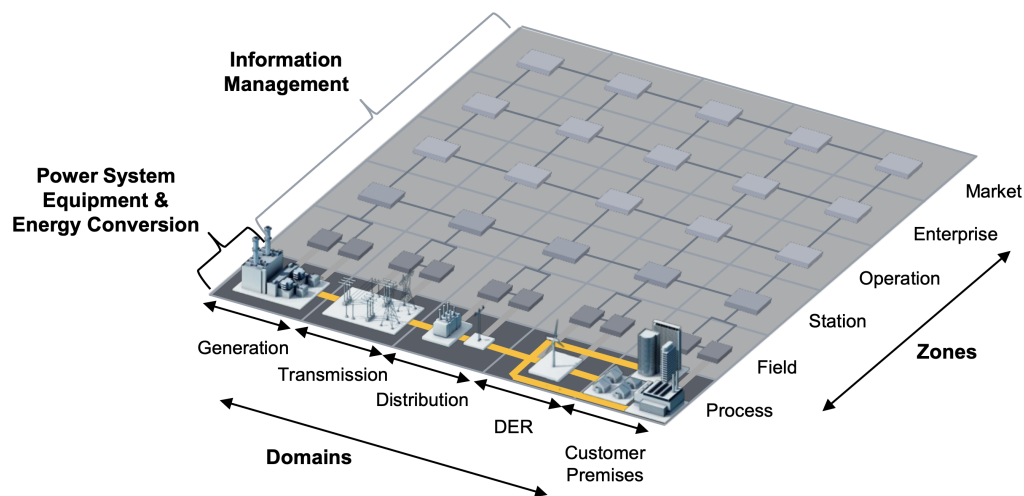
Der Communication Layer adressiert die Kommunikationsaspekte der Interoperabilität. Er umfasst die Kommunikationsprotokolle und -technologien, die für den Austausch von Informationen zwischen den Systemkomponenten erforderlich sind, sowie die Netzwerkinfrastruktur und Sicherheitsmechanismen.

### Component Layer

Der Component Layer bildet die physischen Geräte und Systeme ab, die an den Smart Grid-Anwendungsfällen beteiligt sind. Diese Ebene beschreibt die Hardware-Komponenten, deren technische Spezifikationen und Schnittstellen zur Realisierung der Funktionalitäten notwendig sind.

#### 3.3.1.2 SGAM Smart Grid Plane

Die beiden anderen Dimensionen des SGAM Frameworks sind die Domänen- und Zonenachsen, die für jede Interoperabilitätsebene in einer 2D-Matrix gespannt sind, die als (SGAM) Smart Grid Plane definiert wird (vgl. Abbildung 3.12) [CEN14]. Ziel der Smart Grid Plane ist es, die gesamte elektrische Energieumwandlungskette abzudecken – von der Erzeugung der Energie bis hin zum Endverbraucher oder -erzeuger. Zudem soll mithilfe einer Funktionssegmentierung die Rahmenbedingungen dieser Domänen und Zonen, wie beispielsweise der Fokus auf Echtzeitoperationen in der Field/Station Zone, geordnet werden [CEN14].



Quelle: [CEN14]

Abbildung 3.12: SGAM Grid Plane: Domänen und Zonen

### SGAM Domains

Das SGAM Framework berücksichtigt die im National Institute of Standards and Technology (NIST) Smart Grid Framework definierten Domänen und ergänzt diese mit Distributed Energy Resources (DER) [CEN12]. Diese Ergänzung wurde vorgenommen, da sie die aktuelle Situation in der gegenwärtigen Transition widerspiegelt und zusätzliche Klassen von Use Cases erforderlich macht [CEN12]. Die verwendeten Domänen welche der Wertschöpfungskette des Energiesektors entsprechen [CEN14] umfassen die folgenden:

- **(Bulk) Generation:** Erzeugung von elektrischer Energie in großen Mengen und am Hochspannungsnetz angebunden. Beispiele hierfür sind die Erzeuger aus fossilen Brennstoffen (z. B. Kohle oder Gas), hydroelektrische Kraftwerke, Atomkraftwerke und Off-Shore-Windfarmen.
- **Transmission:** Akteure, welche Elektrizität über längere Distanzen transportieren.
- **Distribution:** Akteure, welche Elektrizität an Kunden verteilen.
- **Distributed Energy Resources (DER):** Verteilte Energieressourcen (klein-skalierte Energieproduzenten sowie -Konsumenten), welche am öffentlichen Verteilnetz angeschlossen sind. Diese sind am Niedrig- sowie Mittelspannungsnetz angebunden.
- **Customer Premises:** Endkunden, bestehend aus industriellen, kommerziellen und privaten Einrichtungen sowie Haushalten, die Energie konsumieren und produzieren oder die Rolle von Prosumern innehaben.

### SGAM Zones

Die SGAM-Zonen repräsentieren die hierarchischen Ebenen des Energiemanagements und spiegeln ein Modell wider, das auf Aggregation und funktionaler Trennung in der Energieversorgung basiert [CEN14]. Hierbei werden die folgenden beiden Arten von Aggregation unterschieden: Datenaggregation bezeichnet die Konsolidierung von Daten aus verschiedenen Zonen, um die Datenmenge zu reduzieren, räumliche Aggregation umfasst die Zusammenfassung von Daten oder Systemen aus einzelnen Standorten zu größeren Einheiten [CEN14]. Die folgenden Zonen werden im SGAM Framework berücksichtigt: [CEN14]

- **Process:** Berücksichtigung der physikalischen, chemischen oder räumlichen Umwandlung von Energieformen sowie der direkt beteiligten physischen Ausrüstungen wie z. B. Generatoren, Transformatoren, Kabeln.
- **Field:** Berücksichtigung der Ausrüstungen zum Schutz, zur Steuerung und Überwachung des Energiesystems, wie beispielsweise Schutzrelais, Schaltfeldsteuerungen und intelligente elektronische Geräte, die Prozessdaten aus dem Energiesystem erfassen und nutzen.
- **Station:** Berücksichtigung der räumlichen Aggregationsebene, beispielsweise für Datenkonzentration, funktionale Aggregation, Automatisierung von Umspannwerken, lokale SCADA-Systeme und Anlagentechnik.
- **Operation:** Berücksichtigung der Durchführung von Steuerungsoperationen im Energiesystem in den jeweiligen Bereichen, wie z. B. Verteilungsmanagementsysteme, Energiemanagementsysteme für Erzeugungs- und Übertragungssysteme, Mikronetzmanagementsysteme, Virtual Power Plants (die mehrere dezentrale Energieerzeugungsanlagen aggregieren) und Managementsysteme für das Laden von EVs.
- **Enterprise:** Berücksichtigung der kommerziellen und organisatorischen Prozesse, Dienstleistungen und Infrastrukturen für Unternehmen, wie z. B. Asset-Management, Logistik, Abrechnung und Beschaffung.
- **Market:** Berücksichtigung der Marktoperationen entlang der Energieumwandlungskette, wie z. B. Energiehandel und Einzelhandelsmärkte.

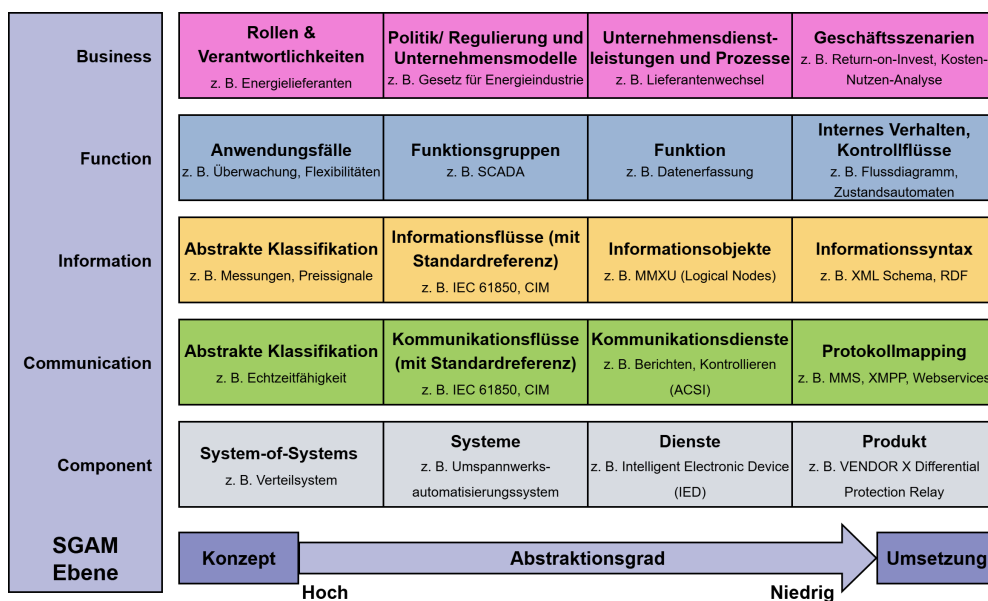
### 3.3.2 SGAM Modelle

Analog zu den Use Cases gemäß IEC 62559 können SGAM Modelle verschiedene Ausprägungen aufweisen. Eine relevante Dimension dieser Ausprägungen ist der Abstraktionsgrad, der bestimmt, ob der zugrunde liegende Anwendungsfall eher konzeptioneller oder implementierungsnaher Natur ist. Dieser Aspekt ist eng mit dem Typ des Use Cases verknüpft (vgl. Abschnitt 3.2.1.2) [CEN14]. Abbildung 3.13 veranschaulicht die unterschiedlichen Abstraktionsarten für jede Interoperabilitätsebene und zeigt entsprechende Beispiele der Komponenten für die einzelnen Kategorien auf. Dabei wird in [CEN14]

darauf hingewiesen, dass die Konsistenz des Abstraktionsgrads über die verschiedenen Ebenen hinweg nicht stets gegeben sein muss und es zu Abweichungen kommen kann.

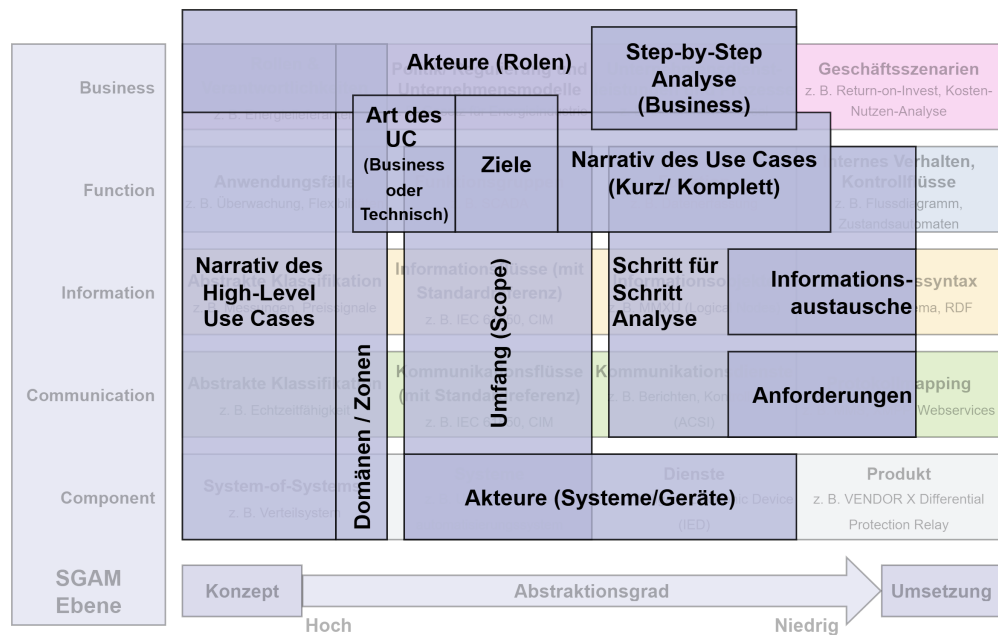
Die Erzeugung von SGAM Modellen basiert auf einem vorliegenden Anwendungsfall. Das SGAM Framework verweist auf die in Abschnitt 3.2 behandelte Use Case-Methodik gemäß ISO/IEC 62559 und das IEC 62559-2 Use Case-Template [CEN12, CEN14]. Durch die standardisierte Struktur des IEC 62559-2 UC-Templates können SGAM Modelle auf Basis der Informationen aus dem IEC 62559 Use Case erzeugt werden [GDU17, CEN14, FSBU20]. Diese Möglichkeit ist im Benutzerhandbuch des SGAM Frameworks beschrieben [CEN14]. Abbildung 3.14 zeigt eine angepasste Darstellung von Abbildung 3.13 und veranschaulicht das im SGAM Handbuch dargestellte Mapping zwischen den Sektionen des IEC 62559 Use Case-Templates und den Abstraktionsstufen der Interoperabilitätsebenen.

Der Prozess der manuellen Erzeugung von SGAM Modellen anhand von IEC 62559 Use Cases kann je nach Vorgehensweise als Top-Down, Bottom-Up oder Mixed-Up Herangehensweise durchgeführt werden [GDU17]. Beim Top-Down-Ansatz wird zunächst der Business Layer erstellt und zuletzt der Component Layer. Beim Bottom-Up-Ansatz hingegen werden zunächst die Akteure und Komponenten des Component Layer aufgeführt, während der Business Layer zuletzt erstellt wird. Bei der Mixed-Up-Herangehensweise kann je nach Informationslage an einer beliebigen Interoperabilitätsebene begonnen und beendet werden. In allen Herangehensweisen stellt die generelle Analyse



Quelle: Basierend auf [CEN14]

Abbildung 3.13: Kategorisierung der verschiedenen Abstraktionsgrade über die SGAM Interoperabilitätsebenen



Quelle: Basierend auf [CEN14]

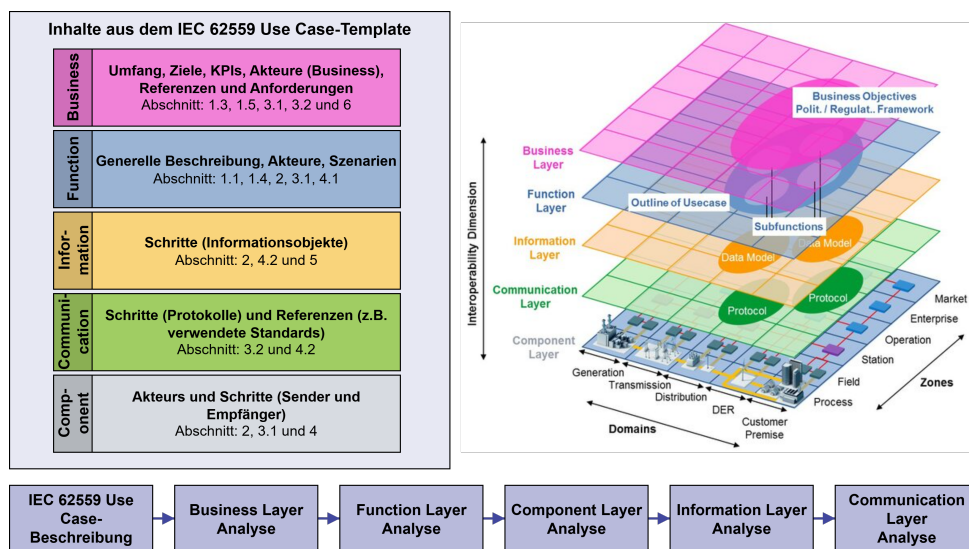
Abbildung 3.14: Beispielhaftes Mapping des IEC 62559 Use Case-Template und dem SGAM Framework mit Berücksichtigung der Abstraktionsgrade

des Anwendungsfalls jedoch den ersten Schritt dar. Ein Beispiel für einen möglichen Mixed-Up-Prozess sowie ein Mapping von IEC 62559 Use Cases zu SGAM Modellen (nach [FSBU20], basierend auf [GDU17]) ist in Abbildung 3.15 dargestellt.

### 3.3.3 Unterstützungswerkzeuge und Serialisierung

Für die Arbeit mit dem SGAM Framework stehen Werkzeuge zur Verfügung, die die Erstellung, Visualisierung und Serialisierung von SGAM Modellen unterstützen. In diesem Abschnitt werden die relevantesten Werkzeuge und ihre spezifischen Funktionen erläutert.

Ein verbreitetes Werkzeug zur Erstellung und Darstellung von SGAM Modellen ist die Nutzung von Microsoft PowerPoint-Vorlagen, wie sie im Rahmen der Projekte TDX-Assist [LCR<sup>+</sup>19] und CoordiNet [GMIFM<sup>+</sup>23] eingesetzt wurden. Dabei wird jede Interoperabilitätsebene des SGAM durch eine separate Folie repräsentiert, während das vollständige SGAM Modell mithilfe von Makros generiert wird. Obwohl das unstrukturierte Format keine Möglichkeit zur Serialisierung bietet, analog zur Word-Vorlage der IEC 62559 Use Case-Methodik, vereinfacht die Verwendung von PowerPoint die Modellierung, da es ein weit verbreitetes und akzeptiertes Tool darstellt.



Quelle: Basierend auf [FSBU20]

Abbildung 3.15: Modellierung von SGAM Modelle anhand von IEC 62559-2 Use Cases

Ein weiteres, etwas restriktiveres Werkzeug ist das Microsoft Visio-Template aus dem DISCERN-Projekt. Dieses Addon ermöglicht die Erstellung von SGAM Modellen durch die Verwendung vordefinierter Komponententypen, was eine strukturierte und konsistente Modellierung fördert. Das Visio-Template bietet somit eine gezielte Unterstützung bei der Erstellung von SGAM Modellen.

Die vom Josef Ressel Center for User-Centric Smart Grid Privacy entwickelte SGAM Toolbox stellt eine Erweiterung von Enterprise Architect dar und erlaubt es SGAM Modelle über die Paradigmen der Modellgetriebene Softwareentwicklung (MDD) zu erstellen und nutzt als Basis die standardisierte Modellierungssprache Systems Modeling Language (SysML) [NET<sup>+</sup>14]. Hierdurch können einzelne Elemente des SGAM Modells im Detail z. B. über UML-Diagramme im Detail modelliert werden [NET<sup>+</sup>14].

Für die reine Visualisierung von SGAM Modellen bietet der SGAM Visualizer [GDU17] eine Lösung. Dieses Tool ermöglicht eine browserbasierte 3D-Darstellung von SGAM Modellen und unterstützt den Import von Modellen, die mittels der im Folgenden beschriebenen Serialisierungstechniken erstellt wurden.

Zum Zeitpunkt dieser Dissertation existiert kein offizielles Serialisierungsformat für SGAM Modelle, wie es etwa bei IEC 62559 Use Cases mit der IEC 62559-3 Use Case XML-Serialisierung [IEC18] der Fall ist. Allerdings haben sowohl das DISCERN-Projekt [SUG<sup>+</sup>16] als auch die SGAM Toolbox [NET<sup>+</sup>14] ein semantisches Format unter Verwendung von XML-Schemata entwickelt, das zur Strukturierung und zum Export/Import von SGAM Modellen in ihren jeweiligen Tools verwendet wird. Diese Ansätze stellen derzeit die praktikabelsten Lösungen zur Serialisierung von SGAM Modellen dar.



## 4 Grundlagen im KI-Kontext

Dieses Kapitel widmet sich den methodischen Grundlagen aus dem Feld der Künstlichen Intelligenz, die dieser Arbeit zugrunde liegen. Die Forschungsfragen dieser Arbeit beziehen sich auf „intelligente Systeme“, die wesentlich auf KI-Methoden basieren und somit das zentrale Forschungsobjekt darstellen [RF08]. Konkret wird ein KI-basiertes intelligentes System betrachtet, dessen Konzeption, Implementierung und Evaluierung eine eingehende Betrachtung der zugrunde liegenden Technologien erfordert. Dieses Kapitel bietet eine strukturierte Einführung in die relevanten Methoden und Konzepte, die für die Bearbeitung der Forschungsfragen notwendig sind.

Eine eindeutige, allgemein akzeptierte Definition von Künstlicher Intelligenz (KI) ist bisher nicht etabliert, was bereits im Abschnitt 2.4 dargelegt wurde. Für den Kontext dieser Dissertation wird jedoch die Definition nach Kurzweil herangezogen, die im Abschnitt 2.4 hergeleitet wurde. Diese definiert KI als die Fähigkeit, „*Maschinen zu erschaffen, die Funktionen ausführen, die normalerweise menschliche Intelligenz erfordern, wenn sie von Menschen durchgeführt werden*“ [Kur99]. Die Herausforderung, Künstliche Intelligenz einem konkreten Themengebiet zuzuordnen, bleibt jedoch bestehen. Dies wird auch durch die AI Watch-Initiative der Europäischen Kommission verdeutlicht, die sich primär mit der Regulierung von KI in der EU befasst [SLCD<sup>+</sup>21]. Hierbei untersuchte AI Watch Definitionen, Taxonomien und Anwendungsdomänen von KI, um eine systematische Einordnung des Forschungsfeldes zu ermöglichen.

Die Gliederung dieses Kapitels orientiert sich an der Relevanz der behandelten KI-Themengebiete für die vorliegende Dissertation und sind gemäß [SLCD<sup>+</sup>21] der Disziplin der KI angesehen. Der Abschnitt 4.1 behandelt das Thema Machine Learning (ML), welches als fundamentaler Bestandteil moderner KI-Systeme gilt. Dabei werden grundlegende Konzepte, Methoden und die für diese Arbeit relevanten Ansätze des maschinellen Lernens erläutert. Da der Forschungskontext dieser Arbeit die Verarbeitung natürlicher Sprache umfasst, werden im Abschnitt 4.2 die wesentlichen Themenfelder der NLP betrachtet. Hierzu zählen unter anderem die Analyse, Modellierung und Generierung von Sprache sowie deren Anwendung in KI-Systemen. In Abschnitt 4.3 werden die in dieser Dissertation eingesetzten großen Sprachmodelle (L)LMs vorgestellt, die insbesondere in Abschnitt 4.4 für verschiedene, im Rahmen der Arbeit relevante Anwendungsbereiche genutzt werden. Ergänzend werden in Abschnitt 4.5 weitere zentrale Themen und Techniken behandelt, darunter der Einsatz von eXplainable Artificial Intelligence (XAI), Methoden des Text-Data-Preprocessings sowie Verfahren zur Textaugmenta-tion.

### 4.1 Machine Learning (ML)

In den 2010er-Jahren gewann das Thema Big Data zunehmend an Bedeutung und unterstrich die Relevanz von Machine Learning-Verfahren, die in der Lage sind, große und komplexe Datenmengen effizient auszuwerten [Mur14]. Charakterisiert wird Big Data

typischerweise durch fünf zentrale Merkmale – die 5 V: *Volume* (Datenmenge), *Velocity* (Verarbeitungsgeschwindigkeit), *Variety* (Vielfalt an Datenquellen und -formaten), *Veracity* (Zuverlässigkeit und Qualität der Daten) und *Value* (potenzieller Erkenntnisgewinn) [Zik12]. Das Themenfeld Machine Learning ist eines der zentralen Teilgebiete von KI und befasst sich mit der Entwicklung von Algorithmen und statistischen Modellen, die auf der Grundlage von Daten Muster erkennen, Vorhersagen treffen oder Entscheidungen ableiten können, ohne explizit für jede einzelne Aufgabe programmiert zu sein. [Mur14] Machine Learning ist somit ein Sammelbegriff verschiedenster Methoden und Algorithmen, welcher sich auf die Arten von Problemen spezifizieren lassen [Bis06]:

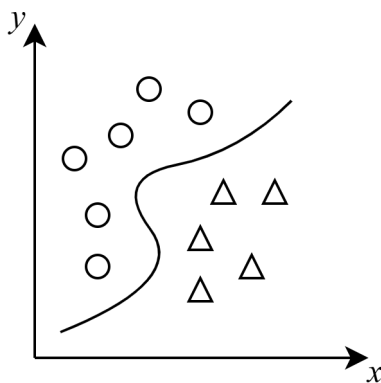
- **Überwachtes Lernen (Supervised Learning):** Beim überwachten Lernen besteht das Ziel darin, eine Abbildungsfunktion  $f : \mathcal{X} \rightarrow \mathcal{Y}$  zu erlernen, die Eingabevektoren  $x \in \mathcal{X}$  auf Ausgabewerte  $y \in \mathcal{Y}$  abbildet. Grundlage dafür ist ein Trainingsdatensatz  $D = \{(x_i, y_i)\}_{i=1}^N$ , bestehend aus  $N$  beobachteten Input/Output-Paaren. Das Modell wird anhand dieses Datensatzes trainiert, um eine Funktion  $f$  zu approximieren, die auch für bisher unbekannte Eingaben  $x$  eine möglichst genaue Vorhersage des zugehörigen Ausgabewerts  $y$  ermöglicht. Dies kann im Rahmen von Klassifikation oder auch in Regression stattfinden. Im Kontext des Smart Grid umfasst dies unter anderem die *Lastprognose* [AMA21], die *Vorhersage der Einspeisung aus Photovoltaik- oder Windkraftanlagen* [BMN09] sowie die *Fehlererkennung in smarten Transformatoren* [KAS23]. Häufig eingesetzte Verfahren sind Regressionsmodelle (z. B. lineare Regression, Random Forest), Support Vector Machines (SVM), Künstliche neuronale Netze (KNNs) (z. B. LSTM für Zeitreihen) und Gradient Boosting (z. B. XGBoost). [Mur14, Bis06]
- **Unüberwachtes Lernen (Unsupervised Learning):** Beim unüberwachten Lernen liegt kein Trainingsdatensatz in Form expliziter Input/Output-Paare vor. Stattdessen besteht die gegebene Datenmenge aus reinen Inputdaten  $D = \{x_i\}_{i=1}^N$ , wobei  $x_i \in \mathcal{X}$  und  $N$  die Anzahl der Beobachtungen bezeichnet. Ziel ist es, Strukturen, Muster oder Zusammenhänge innerhalb der Daten zu identifizieren, ohne dass explizite Zielwerte  $y$  vorliegen. Im Bereich intelligenter Energiesysteme findet unüberwachtes Lernen beispielsweise Anwendung in der *Anomalieerkennung in Netzbetriebsdaten* [RSAM21] oder der *Topologiebestimmung in Verteilnetzen* [DGH24]. Gängige Verfahren umfassen  $k$ -Means-Clustering, hierarchisches Clustering, DBSCAN, Principal Component Analysis (PCA) sowie Autoencoder zur Erzeugung latenter Repräsentationen. [Mur14, Bis06]
- **Bestärkendes Lernen (Reinforcement Learning):** Beim bestärkenden Lernen interagiert ein Agent mit einer Umgebung  $\mathcal{E}$  in diskreten Zeitschritten  $t \in \mathbb{N}$ . Zu jedem Zeitpunkt  $t$  beobachtet der Agent einen Zustand  $s_t \in \mathcal{S}$ , wählt eine Aktion  $a_t \in \mathcal{A}$  gemäß einer Strategie  $\pi : \mathcal{S} \rightarrow \mathcal{A}$  und erhält daraufhin eine Rückmeldung in Form einer Belohnung  $r_t \in \mathbb{R}$  sowie einen Folgezustand  $s_{t+1}$ . Ziel ist es, eine optimale Strategie  $\pi^*$  zu erlernen, welche die erwartete kumulierte Belohnung maximiert, also  $\mathbb{E}[\sum_{t=0}^{\infty} r_t]$ . In Smart Grid-Anwendungen wird bestärkendes Lernen unter anderem innerhalb der *Co-Simulation* [VHN23] oder zur *Lernen von Angriffen*

vektoren von *Misuse Cases* [VWU23] eingesetzt. Verwendete Methoden reichen von klassischem Q-Learning über Deep Q-Networks (DQN) bis hin zu fortgeschrittenen Deep-Reinforcement-Learning-Verfahren wie Markov Decision Process (MDP) oder Proximal Policy Optimization (PPO). [Mur14, Bis06]

Aufgrund der in Kapitel 7 beschriebenen Problemstellungen liegt der Fokus dieses Abschnitts auf dem überwachten Lernen (Supervised Learning). Die dargestellten Grundlagen bilden das Fundament für spätere zentrale Aspekte der Arbeit sowie für die daraus abgeleiteten architektonischen Entscheidungen im Bereich der LMs und LLMs. Im folgenden Abschnitt 4.1.1 werden die wichtigsten Kernkonzepte des überwachten Lernens erläutert. Die in dieser Arbeit eingesetzten (Large) Language Models (LLMs) basieren auf KNNs. (L)LMs sind dem Themenfeld des Deep Learning (DL) zuzuordnen [BHA<sup>+</sup>22]. Eine vertiefte Betrachtung erfolgt in Abschnitt 4.1.2.

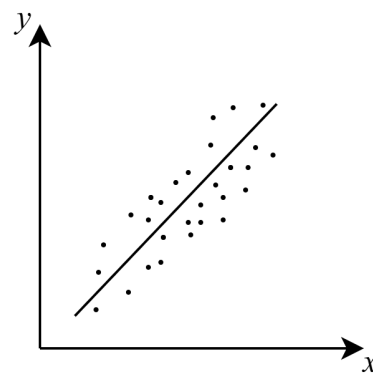
### 4.1.1 Überwachtes Lernen

Im Rahmen des Supervised Learning lassen sich zwei fundamentale Problemklassen unterscheiden: *Klassifikation* und *Regression*. Beide basieren auf der Modellierung einer Abbildungsfunktion  $f : X \rightarrow Y$ , wobei sich die Art der Zielvariablen  $Y$  unterscheidet. Handelt es sich bei  $Y$  um eine diskrete, kategoriale Variable, spricht man von einer *Klassifikationsaufgabe* (vgl. Abbildung 4.1); ist  $Y$  hingegen eine kontinuierliche Größe, liegt eine *Regression* vor [Bis06] (vgl. Abbildung 4.2). Klassifikation wird angewandt zur Entscheidung zwischen Klassen - beispielsweise bei der Fehlererkennung in Systemen - während Regressionsmodelle zur Vorhersage numerischer Kenngrößen wie etwa Energieverbrauch oder Lastprognosen eingesetzt werden.



Quelle: Eigene Darstellung

Abbildung 4.1: (Binäres) Klassifikationsproblem



Quelle: Eigene Darstellung

Abbildung 4.2: Regressionsproblem

Bei Klassifikations- sowie Regressionsproblemen wird der Datensatz typischerweise in disjunkte Mengen aufgeteilt: den Trainingsdatensatz  $\mathcal{D}_{\text{train}} = \{(x_i, y_i)\}_{i=1}^n$  zum Lernen der Modellparameter und den Testdatensatz  $\mathcal{D}_{\text{test}}$  zur Evaluierung der Modellleistung. Während das Modell auf  $\mathcal{D}_{\text{train}}$  durch Minimierung einer Loss-Funktion  $\mathcal{L}(f(x_i), y_i)$  lernt, bleibt  $\mathcal{D}_{\text{test}}$  während des Trainings unberührt, um eine objektive Abschätzung des Fehlers zu ermöglichen. Ein optionaler Validierungsdatensatz kann zur Auswahl von Hyperparametern oder zur Vermeidung von Overfitting eingesetzt werden, insbesondere bei komplexeren Modellklassen. [Mur14, Bis06]

In überwachten Lernverfahren wird der Datensatz typischerweise in Eingabemerkmale (Features) und Zielwerte (Labels) differenziert. Hierbei sind Features  $x \in \mathbb{R}^d$ , die Merkmalsdimensionen eines Objekts oder Zustands und haben zugehörige Labels  $y$ , die die zu lernende Zielgröße darstellen. Die Form von  $y$  bestimmt dabei die Art der Lernaufgabe: Bei der *binären Klassifikation* nimmt  $y$  genau zwei diskrete Werte an, d. h.  $y \in \{0, 1\}$ ; in der *Multi-Class Klassifikation* gilt  $y \in \{1, \dots, K\}$  für  $K > 2$  Klassen. Im Gegensatz dazu steht die *Multi Label-Klassifikation*, bei der jedem Beispiel mehrere Labels gleichzeitig zugeordnet werden können, was typischerweise durch einen Vektor  $y \in \{0, 1\}^K$  mit unabhängigen binären Indikatoren realisiert wird. [Bis06].

Das Ziel von Machine Learning-Modellen ist das adäquate Abbilden des Eingaberaums und das Generalisieren von neuen, unbekanntem Daten. Jedoch besteht die Problematik, dass ein KI-Modell entweder zum Overfitting oder Underfitting neigen kann: Underfitting tritt auf, wenn das Modell zu einfach ist, um die zugrunde liegende Struktur in den Daten zu erfassen [Bis06]. Overfitting hingegen beschreibt die Überanpassung an die Trainingsdaten, bei der auch Rauschen und zufällige Schwankungen modelliert werden, sprich das Auswendiglernen von den Trainingsdaten [Bis06]. Dies äußert sich in einer hohen Varianz und geringer Generalisierung auf Testdaten. Der sogenannte Bias-Variance Trade-off beschreibt das Spannungsfeld zwischen diesen beiden Fehleranteilen (Bias repräsentierend durch das Underfitting; Varianz durch das Overfitting). Mathematisch lässt sich der erwartete quadratische Fehler eines Modells  $f$  auf einem Punkt  $x$  zerlegen in:

$$\mathbb{E}[(f(x) - y)^2] = \underbrace{(\mathbb{E}[f(x)] - y)^2}_{\text{Bias}} + \underbrace{\mathbb{E}[(f(x) - \mathbb{E}[f(x)])^2]}_{\text{Varianz}} + \underbrace{\sigma^2}_{\text{Irreduzibler Fehler}},$$

wobei  $\sigma^2$  die inhärente Unschärfe in den Daten darstellt [GBD92, Bis06]. Die Modellwahl zielt darauf ab, ein Gleichgewicht zwischen Bias und Varianz zu erreichen, um eine möglichst geringe Gesamtfehlerrate zu erzielen.

Stellvertretende überwachte Machine Learning-Verfahren, die in der Vergangenheit häufig in der Praxis eingesetzt wurden, umfassen Decision Trees, SVMs und k-Nearest Neighbors (k-NN). Decision Trees (dt. Entscheidungsbäume) teilen Daten in hierarchische Strukturen, die leicht interpretierbar sind, während SVMs für Klassifizierungsaufgaben hochdimensionaler Daten optimal sind, indem sie eine optimale Trennlinie zwischen den Klassen finden. k-NN (nicht zu verwechseln mit Künstliche neuronale Netze (KNNs)) hingegen basiert auf der Annahme, dass ähnliche Datenpunkte in der Nähe voneinander liegen, und klassifiziert neue Instanzen basierend auf den häufigsten Klas-

sen ihrer Nachbarn. Im Rahmen dieser Dissertation werden KNNs als leistungsstarke Modelle betrachtet, die im Abschnitt 4.1.2 detailliert erläutert werden, da sie die Grundlage für die in der Arbeit verwendeten (L)LMs bilden.

### 4.1.2 Künstliche neuronale Netze (KNNs) und Deep Learning

Künstliche neuronale Netze (KNNs) sind Modelle, die sich an der Architektur und Funktionsweise biologischer Nervenzellen orientieren und als Grundbaustein moderner DL / Transformer-basierte Modelle darstellen [LBH15, VSP<sup>+</sup>17]. Analog zu biologischen Neuronen, die über Synapsen Signale empfangen, verarbeiten und weiterleiten, bestehen KNNs aus künstlichen Neuronen, die numerische Eingaben gewichten, durch eine Aktivierungsfunktion transformieren und als Ausgang weitergeben [Hay09]. Neuronen werden in einer Input Layer, mehreren Hidden Layer sowie einer Output Layer angeordnet werden. Diese Struktur erlaubt es dem Netzwerk, komplexe nichtlineare Zusammenhänge zu modellieren insbesondere durch den Einsatz von mehreren Hidden Layer [Hay09]. Künstliche neuronale Netze (KNNs) werden in vielfältigen Anwendungsfeldern eingesetzt, darunter Bild- und Spracherkennung, autonome Systeme, medizinische Diagnostik sowie Prognosemodelle in Energiesystemen [LBH15, Sch15]. Besonders im Bereich der NLP ermöglichen Deep Neural Network (DNN) wie Transformer-basierte-Modelle [VSP<sup>+</sup>17] Verfahren zur maschinellen Übersetzung, Textklassifikation und kontextsensitiven Generierung (vgl. Abschnitt 4.2).

Ein KNN kann als Funktion  $f(\mathbf{x}; \boldsymbol{\theta})$  aufgefasst werden, wobei  $\mathbf{x} = (x_1, \dots, x_n)$  die Eingabeneuronen und  $\boldsymbol{\theta}$  die Menge aller Gewichtungen und Biases umfasst. Im *Forward Propagation* wird die Aktivierung eines Neurons  $a_j^{(l)}$  in Schicht  $l$  durch

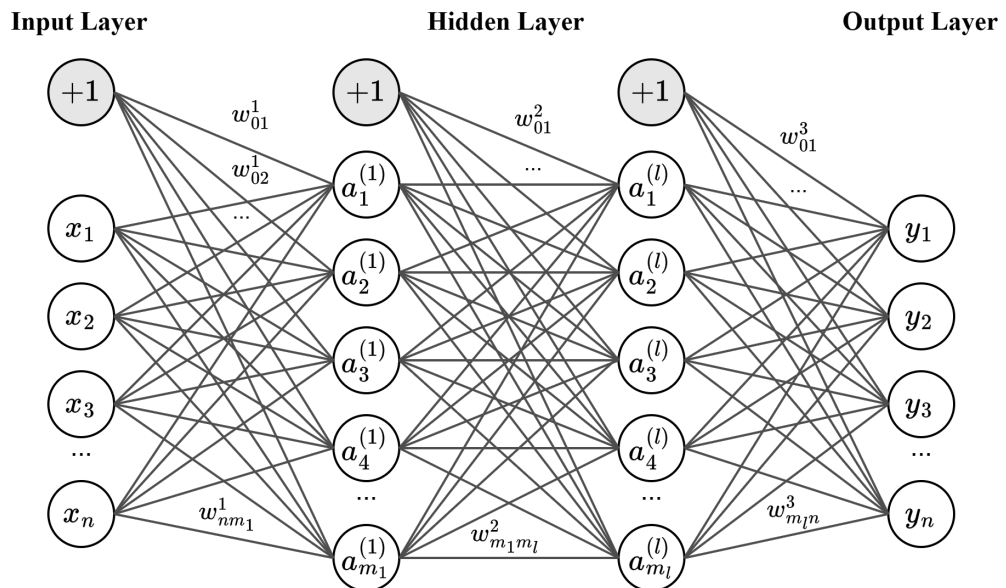
$$a_j^{(l)} = \phi^{(l)} \left( \sum_{i=1}^{m_{l-1}} w_{ji}^{(l)} \cdot a_i^{(l-1)} + b_j^{(l)} \right)$$

berechnet, wobei  $\phi^{(l)}(\cdot)$  eine Aktivierungsfunktion (z. B. ReLU, Sigmoid) ist,  $w_{ji}^{(l)}$  das Gewicht zwischen Neuron  $i$  der vorherigen Schicht  $l - 1$  und Neuron  $j$  in Schicht  $l$  darstellt, und  $b_j^{(l)}$  der Bias-Term ist. Die Eingabeschicht bildet den Vektor  $\mathbf{x} = (x_1, \dots, x_n)$ , und die Ausgabeschicht ergibt die Output-Neuronen  $\mathbf{y} = (y_1, \dots, y_k)$ . Eine abstrakte Darstellung eines solchen KNN ist in Abbildung 4.3 illustriert.

Zur Optimierung des Netzwerks wird eine differenzierbare Verlustfunktion  $\mathcal{L}(\hat{\mathbf{y}}, \mathbf{y})$  minimiert, wobei  $\hat{\mathbf{y}}$  die vom Netz berechneten Ausgaben sind. Für ein Klassifikationsmodell kann dies die Cross-Entropy darstellen, für regressionsbasierte Probleme der Mean Squared Error (MSE) (dt. mittlere quadratische Abweichung) oder Mean Absolute Error (MAE) (dt. mittlere absolute Abweichung). Die Gradienten  $\partial \mathcal{L} / \partial w_{ji}^{(l)}$  werden mittels der *Backpropagation* unter Anwendung der Kettenregel effizient berechnet. Diese fließen in ein Optimierungsverfahren wie den *Stochastic Gradient Descent (SGD)* ein, der die Gewichte gemäß

$$w_{ji}^{(l)} \leftarrow w_{ji}^{(l)} - \eta \cdot \frac{\partial \mathcal{L}}{\partial w_{ji}^{(l)}}$$

aktualisiert, wobei  $\eta$  die Lernrate bezeichnet [Hay09]. Auf diese Weise passt das Netzwerk seine Parameter iterativ an, um eine optimale Approximation der Zielausgabe  $y$  zu erreichen [Hay09]. Bei der Verwendung von SGD wird folglich die Ableitung der für den Layer  $l$  vorliegenden Aktivierungsfunktion  $\phi^{(l)}(\cdot)$  verwendet. Moderne Optimierungsverfahren wie Adam, RMSProp oder AdaGrad verbessern dabei Konvergenzgeschwindigkeit und Trainingsstabilität insbesondere in tiefen Architekturen [ALN23]. Abbildung 4.4 stellt die sigmoide, Tanh- und ReLU-Funktionen als etablierte Aktivierungsfunktionen dar. Während Sigmoid (—) und Tanh (---)-Funktionen aufgrund ihrer Sättigungseigenschaften genutzt werden, werden ReLU (—) in modernen Architekturen durch ihre Effizienz und ihre Fähigkeit, das vanishing-gradient-Problem (zu klein werdende Gradienten) abzumildern eingesetzt [Hay09].



Quelle: Eigene Abbildung

Abbildung 4.3: Vereinfachte Abbildung eines (tiefen) KNN

Mit dem Fortschritt in den DL-Architekturen haben sich insbesondere im Bereich der natürlichen Sprachverarbeitung (NLP) leistungsfähige Sprachmodelle etabliert. Basierend auf Trainingsdaten in Milliardenhöhe und großen Modellparametern erfassen Sprachmodelle semantische und syntaktische Muster, die Aufgaben wie maschinelle Übersetzung, Fragebeantwortung und Textgenerierung ermöglichen. Den entscheidenden Durchbruch markierte die Einführung des Transformer-Modells [VSP<sup>+</sup>17], das aufgrund seines selbstaufmerksamkeitsbasierten Rechenparadigma sowohl parallele Ver-

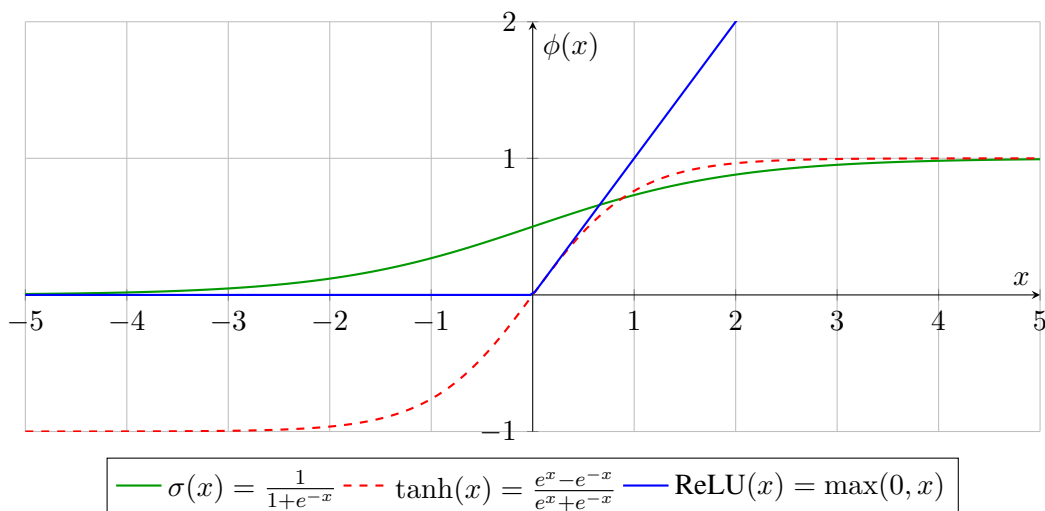


Abbildung 4.4: Aktivierungsfunktionen: Sigmoid, Tanh und ReLU

arbeitung als auch langfristige Abhängigkeitsmodellierung erlaubt. Darauf aufbauend entstanden (Large) Language Models wie BERT [DCLT19] und GPT [OAA<sup>+</sup>24], die durch vortrainierte Repräsentationen in zahlreichen NLP-Aufgaben den Stand der Technik in den Anfang der 2020er Jahre neu definierten. Dies wird im Rahmen des folgenden Abschnitts 4.2 detaillierter erläutert.

## 4.2 Natural Language Processing (NLP)

Das Thema der Natural Language Processing (NLP) (dt. natürliche Sprachenverarbeitung) stellt einen Teilbereich des Themenkomplexes der Künstlichen Intelligenz und Linguistik dar und beschäftigt sich als Forschungszweig, sprachliche Ausdrücke und Wörter für Computer zu verstehen, und zu verarbeiten um bestimmte Aufgaben zu erfüllen [KKKS23, Cho20]. Hierbei besteht das Themenfeld der NLP aus zwei Komponenten, nämlich **Natural Language Understanding (NLU)** und **Natural Language Generation (NLG)**, welches sich mit dem Verständnis und Generieren von Texten befasst [KKKS23, Cho20].

**Natural Language Understanding** dient dazu, natürliche Sprache für Maschinen verständlich zu machen und aus dieser Konzepte, Emotionen, Schlüsselwörter usw. zu extrahieren. Es wird bei NLU auf Inhalte der Linguistik zurückgegriffen. Linguistik ist die Wissenschaft, die sich mit der Bedeutung der Sprache, dem Sprachkontext und den verschiedenen Formen der Sprache beschäftigt [KKKS23]. Hierbei werden in NLU die folgenden Konzepte aus der Linguistik betrachtet [KKKS23]:

- **Phonologie:** Phonologie befasst sich mit den Lautsystemen und -mustern einer Sprache. Sie untersucht, wie Laute Bedeutung unterscheiden und in der Sprachstruktur organisiert sind. Im NLP spielt Phonologie eine Rolle bei der Spracherkennung und -synthese. [KKKS23]
- **Morphologie:** Morphologie untersucht die Struktur und Bildung von Wörtern in einer Sprache. Sie analysiert, wie Wörter aus kleineren Einheiten, den Morphemen, zusammengesetzt sind. Im NLP ermöglicht die morphologische Analyse die Erkennung von Wortstämmen und Flexionsformen. Diese Erkenntnisse unterstützen Aufgaben wie die Lemmatisierung und verbessern das Textverständnis. [KKKS23]
- **Lexik:** Lexik befasst sich mit dem Wortschatz einer Sprache und umfasst alle Wörter und deren Bedeutungen, die in einem bestimmten sprachlichen Kontext verwendet werden. Jedes vorgesehene Wort gilt es für Menschen aber auch NLP-Systeme zu interpretieren.
- **Syntax:** Syntax fokussiert sich auf die Struktur und die Anordnung von Wörtern in Sätzen, um deren grammatische Beziehungen zu identifizieren und zu verstehen. Durch den Einsatz von Parsing-Techniken und Part-of-Speech-Tagging können NLP-Systeme die Rolle einzelner Wörter im Satz (wie Subjekte, Verben, Objekte) bestimmen und die syntaktische Struktur von Sätzen analysieren. [KKKS23]
- **Semantik:** Das wesentliche Ziel der Semantik stellt die richtige Interpretation eines Satzes dar. Dies stellt insb. für Maschinen eine Herausforderung dar, da diese nicht das menschliche Verständnis der Sprache oder des Konzeptes besitzen und folglich besondere Techniken notwendig sind. [KKKS23]
- **Diskurs:** Diskurs betrachtet, wie Sätze in Texten und Gesprächen zusammenhängen und Bedeutung erzeugen. Sie betrachtet die Struktur und Kohärenz von längeren Texteinheiten. Im Rahmen von NLP unterstützt die Berücksichtigung des Diskurses den Kontext und die Absicht hinter Kommunikation zu verstehen. [KKKS23]
- **Pragmatik:** Pragmatik befasst sich mit der Nutzung von Sprache im Kontext und wie dieser die Bedeutung beeinflusst. Es wird dabei untersucht, wie Sprecher Absichten vermitteln und Empfänger Interpretationen vornehmen. Pragmatik unterstützt in NLP ein tieferes Verständnis von Ironie, Anspielungen und indirekten Äußerungen zu entwickeln. [KKKS23]

**Natural Language Generation** stellt den Gegensatz zur Natural Language Understanding dar und befasst sich mit der Generierung von Formulierungen, Sätzen und Absätzen auf Basis der internen Repräsentation [KKKS23]. Hierbei werden in der Literatur folgende Konzepte betrachtet [KKKS23]:

- **Lautsprecher und Generator:** Für die Textgenerierung werden ein Lautsprecher (z. B. in Form einer Applikation) und ein Generator (z. B. in Form eines Programms) benötigt. Dabei wandelt der Generator die Absichten des Lautsprechers in flüssige, der Situation angemessene Texte um. [KKKS23]

- **Komponenten und Ebenen der Repräsentation:** Dieser Bereich umfasst die systematische Bearbeitung von Aufgaben, die sich mit der Auswahl relevanter Inhalte, der grammatikalischen Strukturierung sowie der textuellen Gliederung befassen. Ebenso werden die Auswahl und der Einsatz sprachlicher Mittel sowie die Realisierung in Form eines Textes oder einer Sprachausgabe betrachtet. [KKKS23]
- **Applikation oder Lautsprecher:** Die Aufgabe besteht darin, das Modell der Situation aufrechtzuerhalten, wobei der Sprecher den Vorgang lediglich startet, aber nicht an der Erstellung der Sprache beteiligt ist. Es geht darum, die Historie zu bewahren, den Inhalt zu strukturieren und eine Repräsentation des Bekannten zu nutzen, um aus den Vorschlägen des Sprechers eine Auswahl zu treffen, wobei das Verständnis der Situation durch den Sprecher essentiell ist. [KKKS23]

Es hat sich gezeigt, dass NLP-basierende Systeme (in Bezug auf die Komponenten NLU sowie NLG) in verschiedene Anwendungsbereiche eingesetzt werden. Diese umfassen nach [Cho20] insbesondere:

- Textklassifikation
- Textgenerierung
- Fragebeantwortungssysteme
- Textübersetzung
- Informationsextraktion
- Informationsabfrage
- Erschließung von Wissen
- Indizierung und Textsuche

Aus der Entwicklungsgeschichte heraus lassen sich in der Praxis hauptsächlich drei Arten von NLP-Systemen unterscheiden: heuristikbasiertes NLP, Machine Learning-basiertes NLP und Deep Learning-basiertes NLP [SMGS20]. Insbesondere die frühen NLP-Systeme stellen *heuristikbasierte NLP-Systeme* dar und verließen sich auf das Formulieren von Regeln über einen längeren Zeitraum, wobei Methoden wie Regular Expression (RegEx), kontextfreie Grammatiken oder Ontologien zur Unterstützung der Regelextraktion eingesetzt wurden [SMGS20]. Ein wesentlicher Vorteil von heuristikbasierten NLP-Systemen besteht darin, dass sie aufgrund ihrer Nichtabhängigkeit von statistischen Modellen, noch heute in sicherheitskritischen Bereichen, insbesondere im Gesundheitswesen, eingesetzt werden [SMGS20]. *Machine Learning-basiertes NLP* nutzt Verfahren aus dem unüberwachten (Unsupervised), überwachten (Supervised) und verstärkenden Lernen (Reinforcement Learning) [SMGS20]. Zu den Methoden gehören beispielsweise Naive-Bayes-Klassifikatoren, Support Vector Machines oder Hidden Markov Modelle [SMGS20]. *Deep Learning-basiertes NLP* stützt sich auf Deep Neural Networks, die besonders in den letzten Jahrzehnten an Bedeutung gewonnen haben. Während zu Beginn der Deep Learning-Ära in NLP vor allem Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs) und Convolutional Neural Networks (CNNs) verwendet wurden [KKKS23], dominieren gegenwärtig Transformerbasierte Modelle [KKKS23, SMGS20].

Die nachfolgenden Unterabschnitte dieser Sektion widmen sich einerseits den relevanten NLP-Anwendungsbereichen und andererseits den relevanten spezifischen Algorithmen, Techniken und Methoden, die im Kontext des NLP innerhalb dieser Dissertation zur Anwendung kommen. Der Fokus wird auf jüngste sowie etablierte State-of-the-Art Verfahren aus dem Bereich *Machine Learning-basiertes NLP* sowie *Deep Learning-basiertes NLP* gesetzt, welche zum Zeitpunkt der Dissertation insbesondere vortrainierte LM/LLMs darstellen.

#### 4.2.1 Textrepräsentation und -Einbettungstechniken

Im vorherigen Abschnitt wurde ein Einblick in das Themenfeld der natürlichen Sprachverarbeitung vorgestellt. Ein wichtiges Themenfeld ist die Überführung von textuellen Daten in einer maschinenverarbeitbaren Darstellung, um anschließend rechnerische Operationen zur Sprachverarbeitung aufzuführen [KKKS23]. Ziel ist es, Textelemente, typischerweise Wörter, Phrasen oder ganze Dokumente, in einem Vektorraum  $\mathbb{R}^d$  so abzubilden, dass semantische und syntaktische Relationen durch geometrische Strukturen modelliert werden können [KKKS23].

Die Erzeugung semantisch bedeutungstragender Vektoren erfolgt durch Embedding Methoden, welche textuelle Elemente in einem Vektorraum  $\mathbb{R}^d$  abbilden. Eine rudimentäre Methode ist die *One-Hot-Encoding*, bei der jedes Token  $t_i$  durch einen Vektor  $\vec{v}_i \in \mathbb{R}^V$  mit exakt einer 1 an der positionsspezifischen Stelle dargestellt wird, z. B.  $\vec{v}_{\text{Netz}} = (0, 0, 1, 0, \dots, 0)^T$  [MRS08]. Ein weiteres einfaches Modell zur Textrepräsentation ist das *Bag-of-Words*, bei dem ein Dokument durch die Häufigkeiten seiner enthaltenen Wörter beschrieben wird [JMNR09]. Dabei geht jedoch jegliche Wortreihenfolge und syntaktische Information verloren, wodurch semantische Beziehungen unberücksichtigt bleiben [JMNR09]. Aufgrund der fehlenden semantischen Dichte und hohen Dimensionalität (da jedes Wort des Korpus eine Dimension darstellen kann) gilt es weitere Methoden zu betrachten [MRS08].

*Word2Vec* erzeugt semantisch strukturierte Vektoren, indem es Kontexte basierend auf Wort-Kookkurrenzen probabilistisch modelliert. In der *Skip-Gram*-Variante von *Word2Vec* wird die Wahrscheinlichkeit  $P(t_j | t_i)$  für Kontextwörter  $t_j$  gegeben ein zentrales Wort  $t_i$  maximiert, wobei  $\vec{v}_i, \vec{v}_j \in \mathbb{R}^d$ :

$$P(t_j | t_i) = \frac{\exp(\vec{v}_i^T \vec{v}_j)}{\sum_{k \in V} \exp(\vec{v}_i^T \vec{v}_k)}.$$

Die daraus resultierenden Einbettungen erfassen semantische Ähnlichkeiten durch geometrische Nähe im Vektorraum [MCCD13].

Das auf globalen Kookkurrenzen basierende *GloVe*-Modell minimiert hingegen den Unterschied zwischen Skalarprodukten eingebetteter Wortpaare und dem logarithmierten Kookkurrenzmaß:

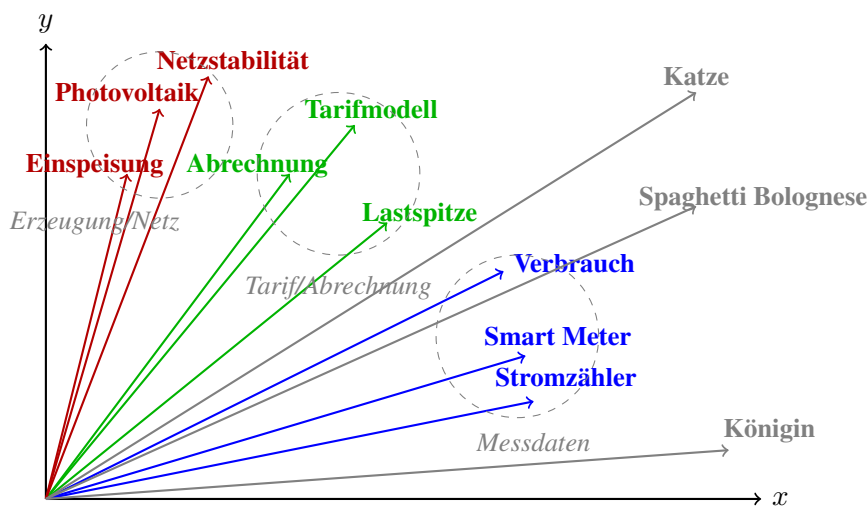
$$J = \sum_{i,j} f(X_{ij}) (\vec{v}_i^T \vec{v}_j + b_i + b_j - \log X_{ij})^2,$$

wobei  $X_{ij}$  die Häufigkeit des gemeinsamen Auftretens von  $t_i$  und  $t_j$  beschreibt und  $f(X_{ij})$  eine Gewichtungsfunktion darstellt [PSM14].

Während klassische Repräsentationen wie One-Hot-Vektoren oder Bag-of-Words hohe Dimensionalität bei gleichzeitigem Verlust semantischer Information aufweisen, erlauben modernere Embedding Techniken, insbesondere auf Basis verteilter Repräsentationen, hingegen eine dichte, kontinuierliche Abbildung sprachlicher Elemente (Wörter, Sätze, Texte) in einem hochdimensionalen Vektorraum  $\mathbb{R}^d$  [PBG23]. Hierbei wird jedes Token  $t_i \in T$  als Vektor  $\vec{v}_i \in \mathbb{R}^d$  modelliert, wobei  $\vec{v}_i = f(t_i)$  die Einbettung durch eine trainierte Funktion  $f$  bezeichnet, sodass semantisch ähnliche Tokens  $t_j \in T$  im Vektorraum proximale Lagen einnehmen, d. h.  $\text{sim}(\vec{v}_i, \vec{v}_j) \approx 1$  für  $t_i \sim t_j$  [PBG23]. In diesem Kontext bezeichnet die Funktion  $\text{sim}(\vec{v}_i, \vec{v}_j)$  typischerweise eine Ähnlichkeitsmetrik im Vektorraum, wobei insbesondere die Cosine Similarity (dt. Kosinosähnlichkeit) verwendet wird [MRS08]:

$$\text{sim}(\vec{v}_i, \vec{v}_j) = \frac{\vec{v}_i \cdot \vec{v}_j}{\|\vec{v}_i\| \cdot \|\vec{v}_j\|} \in [-1, 1].$$

Ein hoher Ähnlichkeitswert nahe 1 impliziert, dass die Vektoren semantisch ähnliche oder funktional austauschbare Bedeutungen kodieren, wobei die Ähnlichkeit je nach Modell und Kontext verschiedene linguistische Ebenen (vgl. Abschnitt 4.2) widerspiegeln kann, wie etwa thematische Nähe, taxonomische Relationen oder kontextuelle Abhängigkeiten [MRS08]. Ein energiebezogenes Beispiel ist in Abbildung 4.5 dargestellt und zeigt, wie bei Anwendung von Dimensionsreduktionsverfahren (z. B. Hauptkomponentenanalyse oder t-SNE) die semantische Nähe und Distanz von Begriffen in einem



Quelle: Eigene Darstellung

Abbildung 4.5: Beispielpflichte Darstellung von semantischer Nähe und Distanz im eingebetteten Vektorraum

2D-Vektorraum visualisiert werden kann. Es ist erkennbar, dass ähnliche Begriffe nahe beieinander liegen, was für Clustering- und Klassifikationsaufgaben hilfreich sein kann [JMNR09].

## 4.2.2 Transformer

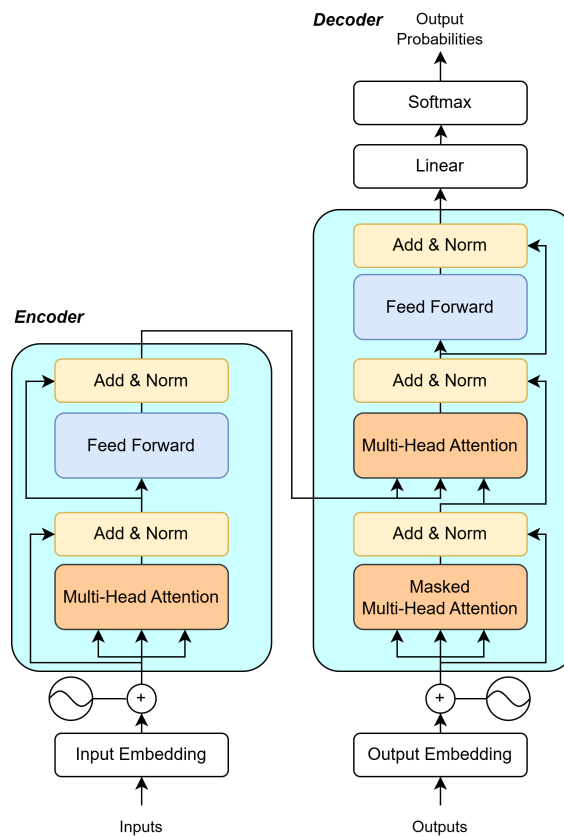
Eine Technik, die zum Zeitpunkt der Dissertation den State-of-the-Art definiert stellt die Transformer-Architektur aus dem Jahr 2017 dar und wurde im Rahmen des Papers „Attention Is All You Need“ von [VSP<sup>+</sup>17] veröffentlicht und dient als Basis in gegenwärtige (L)LMs [DCLT19, HGC23, OAA<sup>+</sup>24, TLI<sup>+</sup>23]. Im Gegensatz zu früheren rekurrenten Architekturen wie das Rekurrentes neuronales Netz (RNN) oder Long Short-Term Memory (LSTM) Netze basiert der Transformer auf der „Self-Attention“-Technik, die es dem Modell ermöglicht, globale Abhängigkeiten innerhalb einer Eingabesequenz effizient zu erfassen und dem Kontext besser zuzuordnen [VSP<sup>+</sup>17]. Obwohl Self-Attention im Transformer mit einer zeitlichen und speichertechnischen Komplexität von  $\mathcal{O}(n^2 \cdot d)$  aufwendiger ist als die sequentiellen Verfahren von RNNs oder LSTMs mit  $\mathcal{O}(n \cdot d^2)$ , erlaubt die vollständige Parallelisierbarkeit der Berechnung eine wesentlich effizientere Nutzung moderner Hardware [VSP<sup>+</sup>17]. Dies wird mit dem Blick auf die maximale Pfadlänge im Transformer deutlich, da diese lediglich bei  $\mathcal{O}(1)$  liegt, da jedes Token direkt mit jedem anderen interagieren kann, während bei rekurrenten Netzen eine Pfadlänge von  $\mathcal{O}(n)$  notwendig ist, was den effizienten Umgang mit langen Kontexten zusätzlich erschwert [VSP<sup>+</sup>17].

Abbildung 4.6 zeigt die Struktur der Transformer-Architektur und unterscheidet zwischen dem Encoder-Stack (linke Seite) und dem Decoder-Stack (rechte Seite). Die groß dargestellten blauen Boxen repräsentieren die Layer der Encoder- bzw. Decoder-Stacks, welche im ursprünglichen Paper jeweils  $n = 6$ -mal gestapelt werden [VSP<sup>+</sup>17]. Die wesentlichen Bestandteile des Modells lassen sich wie folgt aufzählen: [VSP<sup>+</sup>17]

- **Input- und Output-Embedding:** Die Eingabesequenz  $w_1, w_2, \dots, w_n$  wird über eine lernbare Embedding-Matrix  $E \in \mathbb{R}^{|V| \times d}$  in kontinuierliche Vektoren überführt:  $x_i = E[w_i]$  für jedes Token  $w_i$ . Dabei ist  $|V|$  die Vokabulargröße und  $d$  die Dimension der Einbettungen. Im Decoder wird dieselbe Matrix zur Projektion der Ausgabevektoren auf den Vokabelraum verwendet.
- **Positional Encoding:** Um Positionsinformationen ohne Rekurrenz für die parallele Verarbeitung bereitzustellen, werden zu den Embeddings  $x_i$  explizite Positional Encodings  $p_i \in \mathbb{R}^d$  addiert:  $z_i = x_i + p_i$ . Die Vektoren  $p_i$  basieren auf folgenden Funktionen:

$$p_{i,2k} = \sin\left(\frac{i}{10000^{2k/d}}\right), \quad p_{i,2k+1} = \cos\left(\frac{i}{10000^{2k/d}}\right)$$

- **(Masked) Multi-Head Self-Attention:** Dieser Mechanismus bildet das Herzstück der Architektur und ermöglicht es dem Modell, kontextuelle Abhängigkeiten zwischen Tokens zu erfassen – unabhängig von deren relativer Position. Dabei wird für



Quelle: Eigene Abbildung nach [VSP<sup>+</sup>17]

Abbildung 4.6: Architektur eines Transformers

jedes Token eine gewichtete Aggregation der Repräsentationen aller anderen Tokens berechnet, wobei die Gewichtung durch deren relative Relevanz im gegebenen Kontext bestimmt wird.

Technisch geschieht dies durch das Heranziehen dreier Matrizen: *Query* ( $Q$ ), *Key* ( $K$ ) und *Value* ( $V$ ), welche aus der Eingabesequenz  $X \in \mathbb{R}^{n \times d}$  durch lineare Transformationen hervorgehen:

$$Q = XW^Q, \quad K = XW^K, \quad V = XW^V, \quad W^Q, W^K, W^V \in \mathbb{R}^{d \times d_k}$$

Für jedes Token berechnet das Modell die Attention (dt. Aufmerksamkeit) zu allen anderen Tokens in der Sequenz durch das Skalarprodukt zwischen den zugehörigen Query- und Key-Vektoren. Die resultierende Attention-Matrix wird skaliert, um numerische Stabilität zu gewährleisten, und anschließend durch eine Softmax-Funktion normalisiert:

$$\text{Attention}(Q, K, V) = \text{softmax} \left( \frac{QK^\top}{\sqrt{d_k}} \right) V$$

Diese sogenannte *Scaled Dot-Product Attention* erlaubt es dem Modell, selektiv relevante Informationen aus der Sequenz zu extrahieren. Jeder Eintrag in der resultierenden Matrix gibt an, wie stark ein bestimmtes Token (via Query) die Inhalte eines anderen Tokens (via Value) gewichtet.

Der *Multi-Head*-Ansatz repliziert diesen Mechanismus  $h$ -fach parallel mit unterschiedlichen Gewichtungen, sodass verschiedene semantische oder syntaktische Relationen erfasst werden können. Die resultierenden Teilräume werden anschließend konkateniert und gemeinsam weiterverarbeitet:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O$$

Im Decoder wird zusätzlich eine sogenannte *Masked Multi-Head Attention* eingesetzt, um die Autoregressivität sicherzustellen. Dabei handelt es sich um die Eigenschaft, dass jedes Token nur auf vorherige Tokens zugreifen darf, um sicherzustellen, dass die Ausgabe Schritt für Schritt generiert wird und zukünftige Informationen nicht vorzeitig einfließen. Die Maskierung verhindert, dass ein Token auf spätere Positionen in der Sequenz zugreift. Dies wird durch eine additive Maskenmatrix realisiert, in der alle unzulässigen Positionen mit  $-\infty$  belegt werden.

- **Add & Norm:** Nach jedem im Transformer vorgesehenem Submodul wird die Ausgabe mit dem ursprünglichen Eingang durch eine Residual-Verbindung kombiniert und anschließend normalisiert:

$$\text{LayerNorm}(x + \text{Sublayer}(x))$$

Dies hilft, den Gradientenfluss während des Trainings zu stabilisieren und erleichtert das Lernen von tiefen KNNs. Die Normalisierung sorgt zusätzlich dafür, dass die Verteilung der Aktivierungen kontrolliert bleibt, was zu schnellerer Konvergenz führt.

- **Feed-Forward-Netzwerk:** Dieses besteht aus zwei linearen Transformationen mit nichtlinearer Aktivierung (z. B. ReLU) dazwischen:

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2$$

Dabei sind  $W_1, W_2$  und  $b_1, b_2$  trainierbare Parameter. Dieses Netzwerk wird positionsunabhängig auf jedes Token separat angewendet.

- **Finale lineare Schicht und Softmax:** Die finale Ausgabe wird mittels einer linearen Projektion  $W^O \in \mathbb{R}^{d \times |V|}$  in den Vokabelraum transformiert, worauf eine Softmax-Funktion Wahrscheinlichkeiten über mögliche Folgetokens liefert:

$$y = \text{softmax}(zW^O + b)$$

Ein zentraler Unterschied zwischen der Encoder- und der Decoder-Seite liegt in ihrer jeweiligen Funktion innerhalb der Modellarchitektur: Während der **Encoder** die gesamte Eingabesequenz parallel verarbeitet und eine kontextuelle Repräsentation generiert, nutzt der **Decoder** diese Repräsentation sequenziell, um Schritt für Schritt eine

Ausgabe zu erzeugen [VSP<sup>+</sup>17]. Dabei ist insbesondere die Verwendung der *Masked Multi-Head Attention* im Decoder hervorzuheben, welche sicherstellt, dass bei der Generierung eines Tokens nur Informationen aus vorangehenden Positionen berücksichtigt werden [VSP<sup>+</sup>17]. Zusätzlich kommt im Decoder eine *Encoder-Decoder Attention* zum Einsatz, die es dem Modell erlaubt, gezielt Informationen aus der Encoder-Ausgabe abzurufen [VSP<sup>+</sup>17].

Neben den eingangs dargestellten Vorteilen hinsichtlich der vollständigen Parallelisierbarkeit sowie der geringeren maximalen Pfadlänge, zeigt die Transformer-Architektur eine signifikant bessere Leistungsfähigkeit im Vergleich zu klassischen rekurrenten Modellen wie RNNs und LSTMs. Insbesondere in Aufgaben der natürlichen Sprachverarbeitung (darunter maschinelle Übersetzung, Textklassifikation und Fragebeantwortung) übertrifft der Transformer frühere Ansätze sowohl in Bezug auf Genauigkeit als auch auf Trainingszeit [VSP<sup>+</sup>17, DCLT19, RWC<sup>+</sup>19]. Dies lässt sich vor allem auf die Fähigkeit zurückführen, globale Kontextinformationen effizient zu modellieren sowie verschiedene semantische Relationen simultan über mehrere Attention-Köpfe hinweg zu erfassen.

Heutige Sprachmodelle, die auf der Transformer-Architektur basieren, implementieren nicht zwangsläufig sowohl den Encoder- als auch den Decoder-Stack. Stattdessen greifen viele Modelle selektiv auf einzelne Komponenten zurück, etwa ausschließlich auf den Encoder (z.,B. BERT [DCLT19]) oder ausschließlich auf den Decoder (z. B. GPT [RWC<sup>+</sup>19]). Eine vertiefende Darstellung dieser Varianten erfolgt im Abschnitt 4.3, wobei der Transformer weiterhin die konzeptionelle und architektonische Grundlage bildet [VSP<sup>+</sup>17].

### 4.3 (Large) Language Models

Language Models (dt. Sprachmodelle) sind zentrale Bausteine der natürlichen Sprachverarbeitung (NLP) [JMNR09]. Ihr Ziel besteht darin, Wahrscheinlichkeiten für Wortfolgen zu modellieren, um beispielsweise vorherzusagen, welches Wort mit hoher Wahrscheinlichkeit als Nächstes in einem gegebenen Kontext auftreten wird [ZZL<sup>+</sup>25]. Hierbei lassen sich nach [ZZL<sup>+</sup>25] vier Generationen großer Sprachmodelle identifizieren, die sich durch technologische Fortschritte in Architektur, Trainingsdaten und Anwendungsmöglichkeiten unterscheiden:

1. Frühere Ansätze der *ersten Generation* („*statistische LM*“) aus den 90er Jahren basierten überwiegend auf  $n$ -Gramm-Modellen [ZZL<sup>+</sup>25], die Wortwahrscheinlichkeiten auf Basis fester, meist kurzer Kontextfenster berechnen. Sie wurden zumeist für klar abgegrenzte sprachliche Aufgaben eingesetzt.
2. Mit dem Aufkommen neuronaler Netze entwickelte sich die *zweite Generation von Sprachmodellen* („*neuronale LM*“) in den Anfängen der 2010er Jahre, insbesondere durch den Einsatz von RNNs [Med00] sowie durch dichte Vektorraumdarstellungen wie Word2Vec [MCCD13]. Diese Modelle konnten bereits semantische Beziehungen zwischen Wörtern abbilden, waren jedoch noch in ihrer Kontexttiefe begrenzt.

3. Die *dritte Generation von Sprachmodellen* („*vortrainierte LM*“) basiert auf der Transformer Architektur [VSP<sup>+</sup>17] von 2017 und stellt die Basis der aktuell stärker verwendeten Sprachmodelle. Sie ermöglicht eine parallele Verarbeitung längerer Kontextfenster und damit eine deutlich gesteigerte Generalisierungsfähigkeit, was sich in der praktischen Leistungsfähigkeit moderner NLP-Anwendungen widerspiegelt [ZZL<sup>+</sup>25]. Repräsentative Kandidaten stellen z. B. BERT [DCLT19], GPT [RWC<sup>+</sup>19] oder Embeddings from Language Model (ELMo) [PNI<sup>+</sup>18] dar.
4. Zum Zeitpunkt dieser Dissertation hat sich die *vierte Generation* („*Large Language Model (LLM)*“) zu Beginn der 2020er Jahre etabliert. Diese basiert auf skalierbaren Transformer-Architekturen mit Milliarden von Parametern und nutzt umfangreiche, vortrainierte Textkorpora. Charakteristisch für diese Modelle ist ihre Fähigkeit zur kontextsensitiven Verarbeitung über große Sequenzlängen sowie zur Durchführung komplexer Aufgaben über Few- oder Zero-Shot [ZZL<sup>+</sup>25]. Zudem basiert ihre Anwendung auf dem gezielten Einsatz von Prompts, mit denen Aufgaben in natürlicher Sprache beschrieben und durch das Modell ausgeführt werden können. Repräsentative Modelle dieser vierten Generation sind GPT-4 [OAA<sup>+</sup>24], Large Language Model Meta AI (LLaMA) [TLI<sup>+</sup>23] oder Fine-tuned Language Net Text-to-Text Transfer Transformer (FLAN-T5) [CHL<sup>+</sup>22].

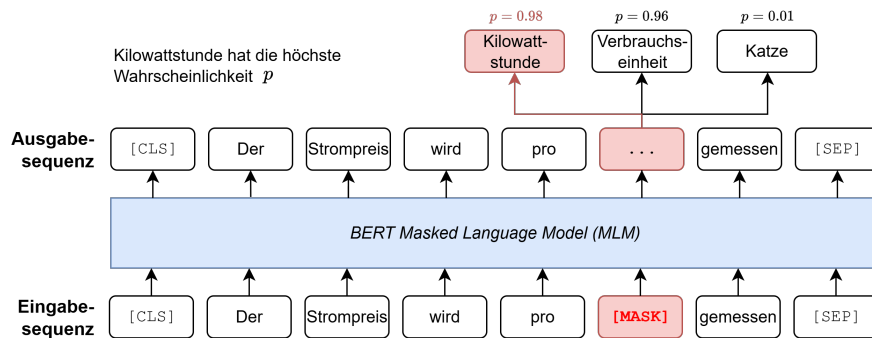
Die folgenden Unterabschnitte stellen die in dieser Dissertation verwendeten LMs vor. Diese lassen sich gemäß [ZZL<sup>+</sup>25] überwiegend der dritten sowie teilweise der vierten Generation von Sprachmodellen zuordnen.

#### 4.3.1 Bidirectional Encoder Representations from Transformers (BERT)

Eines der ersten populärsten vortrainierten LM, das auf der Transformer-Architektur basiert, ist Bidirectional Encoder Representations from Transformers (BERT), das 2018 vorgestellt wurde [DCLT19]. BERT nutzt ausschließlich die Encoder-Komponente der Transformer-Architektur und zeichnet sich insbesondere durch die Fähigkeit aus, Sprachkontexte tiefgreifend und bidirektional zu erfassen. Im Gegensatz zu früheren Ansätzen, die Texte lediglich unidirektional (von links nach rechts oder umgekehrt) verarbeiteten, erlaubt BERT durch das Masked Language Modeling (MLM) eine gleichzeitige Einbettung von linker und rechter Kontextinformation auf allen Modellschichten [DCLT19].

Abbildung 4.7 illustriert die Funktionsweise des MLM, das im Rahmen des Vortrainings von BERT eingesetzt wird. Dabei werden einzelne Tokens in einem Eingabetext zufällig maskiert, d.h. durch ein spezielles [MASK]-Token ersetzt. Das Modell lernt mittels Backpropagation, die maskierten Tokens allein auf Basis des bidirektionalen Kontexts korrekt vorherzusagen. [DCLT19]

Das Vortraining von BERT umfasst zwei Aufgaben: Zum einen das Masked Language Modeling (MLM), zum anderen Next Sentence Prediction (NSP). Bei der NSP-Aufgabe erhält das Modell zwei Sätze und muss vorhersagen, ob der zweite Satz im ursprünglichen Text unmittelbar auf den ersten folgt („IsNext“) oder ob er zufällig aus



Quelle: Eigene Abbildung nach [SM21]

Abbildung 4.7: Beispiel für Masked Language Modeling (MLM) mit BERT als Modell

dem Korpus ausgewählt wurde („NotNext“). Ziel dieser Aufgabe ist es, das Modell für Satzbeziehungen zu sensibilisieren, wie sie beispielsweise in Frage-Antwort-Systemen oder Textinferenz-Aufgaben auftreten. Für das Vortraining wurden sowohl das Books-Corpus (mit ca. 800 Millionen Wörtern) als auch die englischsprachige Wikipedia (rund 2,5 Milliarden Wörter) verwendet [DCLT19].

Das ursprüngliche BERT-Modell wurde in zwei Open Source Varianten veröffentlicht: BERT<sub>Base</sub> mit 110 Millionen und BERT<sub>Large</sub> mit 340 Millionen Parametern [DCLT19]. BERT<sub>Base</sub> umfasst 12 Layer, BERT<sub>Large</sub> 24 Layer, jeweils mit entsprechend größerer Modelltiefe und -breite. DistilBERT ist eine komprimierte Variante von BERT<sub>Base</sub>. Es nutzt nur 6 Layer, verzichtet auf die NSP-Aufgabe beim Vortraining und reduziert die Parameteranzahl auf etwa 66 Millionen und steht ebenfalls Open Source zur Verfügung [SDCW20]. Trotz geringerer Modellgröße erreicht DistilBERT etwa 97 % der Leistung von BERT<sub>Base</sub> bei deutlich höherer Effizienz, was es für ressourcenbeschränkte Szenarien besonders attraktiv macht.

BERT findet breite Anwendung in den Aufgabenfeldern rund um Textklassifikation, Named Entity Recognition, Fragebeantwortung sowie Textähnlichkeits- und Inferenzaufgaben [GDM<sup>+</sup>25], was im Abschnitt 4.4 verdeutlicht wird. Allerdings zeigt sich im Vergleich zu aktuelleren autoregressiven Sprachmodellen wie GPT-3 oder T5, dass BERT vor allem auf Verständnisaufgaben optimiert ist, während generative Aufgaben nicht vollumfänglich abgedeckt werden [Zay24]. Zudem ist das BERT-Modell im Vortraining auf die begrenzten Aufgaben MLM und NSP beschränkt, wohingegen moderne LLMs auf vielseitigeren, häufig auch multimodalen und dialogorientierten Daten trainiert werden [Zay24]. Spätere Studien zeigten jedoch, dass die NSP-Aufgabe die Modellleistung kaum verbessert und daher in Nachfolgemodellen weggelassen wurde (vgl. RoBERTa, Abschnitt 4.3.2).

### 4.3.2 Robustly Optimized BERT Pretraining Approach (RoBERTa)

Robustly Optimized BERT Pretraining Approach (RoBERTa) wurde 2019 als Replikationsstudie von BERT entwickelt, um den Einfluss von Hyperparameter- und Datensatz-Variationen auf die Modellleistung systematisch zu untersuchen [LOG<sup>+</sup>19]. Die Studie zeigt, dass BERT ursprünglich untertrainiert war, weshalb in RoBERTa folgende Anpassungen realisiert: längeres Training über mehr Daten, größere Batchgrößen, dynamisches Masking sowie der Verzicht auf die NSP-Aufgabe. Beim dynamischen Masking erfolgt die Maskierung bei jedem Trainingsepochendurchlauf neu, wodurch das Modell eine größere Vielfalt an Kontexten lernt [LOG<sup>+</sup>19]. RoBERTa verwendet dieselbe Encoder-Komponente der Transformer-Architektur wie BERT, erreicht jedoch dank dieser Optimierungen auf Benchmarks wie GLUE, SQuAD und RACE höhere Ergebnisse [LOG<sup>+</sup>19]. RoBERTa wurde in mehreren Varianten als Open Source Modelle veröffentlicht, darunter RoBERTa<sub>Base</sub> mit 125 Millionen und RoBERTa<sub>Large</sub> mit 355 Millionen Parametern [LOG<sup>+</sup>19].

### 4.3.3 Decoding-enhanced BERT with Disentangled Attention (DeBERTa)

Decoding-enhanced BERT with Disentangled Attention (DeBERTa) ist ein ebenfalls Encoder-basiertes LM, das von Microsoft 2020 als Weiterentwicklung von BERT und RoBERTa vorgestellt wurde und zwei architektonische Neuerungen einführt: Zum einen trennt das Modell in der sogenannten disentangled Attention die Repräsentation von Wortinhalt und Positionsinformation, um diese unabhängig voneinander zu verarbeiten [HGC23]. Dies verbessert insbesondere die Generalisierungsfähigkeit bei komplexen syntaktischen Strukturen [HGC23]. Zum anderen erweitert DeBERTa das MLM um einen sogenannten Enhanced Mask Decoder, der absolute Positionsinformationen gezielt in die Vorhersage maskierter Tokens einbezieht [HGC23]. Diese Anpassungen ermöglichen eine differenziertere Kontextmodellierung, was sich in besseren Ergebnissen auf Benchmarks wie MNLI, SQuAD v2.0 und RACE im Vergleich zu RoBERTa-Large zeigt [HGC23].

DeBERTa wurde über die letzten Jahre kontinuierlich aktualisiert: DeBERTa v2 (2021) integriert zusätzliche architektonische Optimierungen und DeBERTa v3 (2021) erweitert das Vortraining durch ein entropiebasiertes Maskierungsverfahren sowie Replaced Token Detection [HGC23], wodurch die Modellleistung weiter gesteigert werden konnte. DeBERTa wurde ebenfalls in mehreren Open Source Varianten veröffentlicht, darunter DeBERTa<sub>Base</sub> mit ~ 86 Millionen und DeBERTa<sub>Large</sub> mit ~ 304 Millionen Parametern [HGC23].

### 4.3.4 Generative Pretrained Transformer (GPT)

Generative Pretrained Transformer (GPT) ist ein Decoder-basiertes LM, das von OpenAI erstmals 2018 vorgestellt und seither in mehreren Versionen weiterentwickelt wur-

de [RWC<sup>+</sup>19]. Im Gegensatz zu Encoder-basierten Modellen wie BERT oder DeBERTa nutzt GPT eine unidirektionale Architektur auf Basis eines Transformer-Decoders, um Tokens autoregressiv vorherzusagen [RWC<sup>+</sup>19]. Durch autoregressives Training lernt das Modell, jedes Wort basierend auf dem vorherigen Kontext zu generieren, was insbesondere für Aufgaben der natürlichen Sprachgenerierung (z. B. Textvervollständigung, Dialogsysteme) von Vorteil ist [RWC<sup>+</sup>19].

GPT-2 wurde 2019 von OpenAI veröffentlicht und basiert auf einem rein autoregressiven Transformer-Decoder ohne bidirektionale Kontextverarbeitung [RWC<sup>+</sup>19]. Es wurde in vier Open Source-Varianten trainiert, die sich hinsichtlich der Modellgröße unterscheiden: GPT-2<sub>Small</sub> mit 117 Millionen, GPT-2<sub>Medium</sub> mit 345 Millionen, GPT-2<sub>Large</sub> mit 762 Millionen und GPT-2<sub>XL</sub> mit 1,5 Milliarden Parametern [RWC<sup>+</sup>19]. Alle Varianten verwenden denselben architektonischen Aufbau, unterscheiden sich jedoch in Layeranzahl, Hidden Size und Anzahl der Attention-Heads [RWC<sup>+</sup>19].

Mit GPT-3(.5), GPT-4(o) sowie GPT-4o mini wurden Modellgröße und Umfang der Trainingsdaten massiv erweitert, was zu einer deutlichen Steigerung der Zero-Shot- und Few-Shot-Leistung führte [OAA<sup>+</sup>24]. Diese Modelle zählen im Gegensatz zu GPT-2 zur Klasse der LLMs [ZZL<sup>+</sup>25]. GPT-Modelle verzichten vollständig auf eine explizite Aufgabenanpassung und demonstrieren stattdessen durch In-Context Learning eine bemerkenswerte Flexibilität bei einer Vielzahl sprachverarbeitender Aufgaben [OAA<sup>+</sup>24]. Diese Modelle sind nicht Open Source verfügbar, können jedoch über dedizierte APIs genutzt und mit eigenen Daten zentralisiert feinjustiert werden [OAA<sup>+</sup>24]. OpenAI bietet bspw. GPT-3.5, GPT-4 und GPT-4o mini über seine Cloud-API an, bei denen Entwickelnde ihre eigenen Datensätze hochladen und Modelle wie GPT-4o mini für eigene Zwecke nachtrainieren können [OAA<sup>+</sup>24].

Im Vergleich zu Encoder-basierten Modellen wie BERT oder DeBERTa zeigen autoregressive GPT-Modelle Schwächen bei rein verständnisorientierten Aufgaben, da ihnen eine vollständige bidirektionale Kontextverarbeitung fehlt [BMR<sup>+</sup>20]. Zudem neigen sie bei längeren Textgenerierungen stärker zu Halluzinationen und können faktisch falsche Aussagen erzeugen, da sie primär auf Wahrscheinlichkeiten statt auf echtes Wissen zurückgreifen [OAA<sup>+</sup>24, JLF<sup>+</sup>24]. Ihre Stärke liegt dagegen in der Sprachproduktion, nicht in der präzisen Analyse oder Bewertung von Textinhalten [JLF<sup>+</sup>24].

### 4.3.5 FLAN-T5

Fine-tuned Language Net Text-to-Text Transfer Transformer (FLAN-T5) ist ein in 2022 entwickeltes Encoder-Decoder-basiertes LLM, das auf dem von Google entwickelten T5 [RSR<sup>+</sup>23] aufbaut und durch eine gezielte Weiterentwicklung mittels Instruction-Tuning optimiert wurde [CHL<sup>+</sup>22]. Instruction Tuning sieht vor, dass das Modell auf viele verschiedene Aufgabenformate trainiert wird, bei denen die Aufgabenstellung explizit im Prompt enthalten ist [CHL<sup>+</sup>22]. Im Gegensatz zu autoregressiven Modellen wie GPT, die ausschließlich die Decoder-Komponente der Transformer-Architektur verwenden, oder bidirektionalen Ansätzen wie BERT, die lediglich den Encoder nutzen,

basiert FLAN-T5 auf der vollständigen Transformer-Architektur mit Encoder und Decoder. Dadurch können sowohl Eingabe- als auch Ausgabeinformationen explizit verarbeitet werden, was eine flexible Text-zu-Text-Transformation für unterschiedlichste Aufgabenformate ermöglicht [RSR<sup>+</sup>23].

FLAN-T5 wurde ebenfalls in mehreren Open Source-Varianten veröffentlicht, darunter FLAN-T5<sub>Small</sub> mit  $\sim 77$  Millionen, FLAN-T5<sub>Base</sub> mit  $\sim 248$  Millionen, FLAN-T5<sub>Large</sub> mit  $\sim 783$  Millionen, FLAN-T5<sub>XL</sub> mit  $\sim 2,85$  Milliarden und FLAN-T5<sub>XXL</sub> mit  $\sim 11,3$  Milliarden Parametern [CHL<sup>+</sup>22].

Obwohl T5-Modelle wie FLAN-T5 durch ihre Encoder-Decoder-Architektur für generative Aufgaben leistungsfähiger sind, benötigen sie häufig mehr Rechenressourcen und sind langsamer als autoregressive Modelle wie GPT [RSR<sup>+</sup>23]. Dies liegt an der doppelten Rechenlast durch die Nutzung von Encoder und Decoder sowie an der ineffizienteren Decodierung bei längeren Texten [RSR<sup>+</sup>23].

#### 4.3.6 Large Language Model Meta AI (LLaMA)

Large Language Model Meta AI (LLaMA) ist ein Decoder-basiertes LLM, das von Meta AI 2023 entwickelt wurde und sich durch eine offene, effiziente und skalierbare Architektur auszeichnet [TLI<sup>+</sup>23]. Die neueste Version, LLaMA 3, wurde 2024 veröffentlicht [TLI<sup>+</sup>23]. Im Gegensatz zu früheren Decoder-Modellen wie GPT-2 nutzt LLaMA moderne Trainingsstrategien, darunter skalierbares Pretraining auf hochwertigen, mehrsprachigen Datensätzen, effizientere Tokenizer sowie Architekturverbesserungen wie Grouped Query Attention, die insbesondere bei größeren Modellen eine schnellere und speichereffizientere Inferenz ermöglichen [TLI<sup>+</sup>23]. Grouped Query Attention reduziert den Rechenaufwand bei der Selbstaufmerksamkeit, indem mehrere Query-Gruppen gemeinsam mit geteilten Key- und Value-Projektionen verarbeitet werden, was insbesondere bei großen Modellen die Effizienz und Skalierbarkeit verbessert [TLI<sup>+</sup>23].

Die Modelle wurden in mehreren Größen veröffentlicht, darunter LLaMA 3<sub>8B</sub> mit 8 Milliarden und LLaMA 3<sub>70B</sub> mit 70 Milliarden Parametern, wobei beide Varianten als Open Source-Modelle zugänglich sind [TLI<sup>+</sup>23].

LLaMA bietet durch die Open-Source-Verfügbarkeit und Effizienz eine Alternative zu LLMs wie GPT-4, erreichen in Benchmarks vergleichbare Ergebnisse und erlauben eine flexible Anpassung [TLI<sup>+</sup>23]. Gegenüber Modellen wie FLAN-T5 oder BERT-basierten Varianten sind LLaMA-Modelle deutlich stärker in generativen Aufgaben, zeigen aber Schwächen bei strukturierten NLU-Aufgaben ohne zusätzliches Finetuning [ZZL<sup>+</sup>25].

### 4.4 Betrachtung der Anwendungsbereiche

Die vorherigen Abschnitte behandelten die Grundlagen und Auswahl etablierter Sprachmodelle (vgl. Abschnitt 4.3) sowie zentrale Verfahren des überwachten Lernens (vgl. Abschnitt 4.1), die für diese Arbeit von zentraler Bedeutung sind. Der folgende Ab-

schnitt widmet sich konkreten Aufgabenstellungen, in denen Sprachmodelle im Kontext des NLP gezielt eingesetzt werden. Im Fokus stehen dabei insbesondere jene Anwendungsbereiche, die für die Konzeption und Entwicklung der KI-Modelle relevant sind (vgl. Abschnitt 7.5).

#### 4.4.1 Anwendungsbereich: Textklassifikation

Textklassifikation stellt eine der größeren Aufgaben in Rahmen des NLP [LPL<sup>+</sup>22], bei der Texte oder Textsegmente automatisch vordefinierten Klassen oder Labels zugeordnet werden. Wie in Abschnitt 4.1.1 zum überwachten Lernen für den (allgemeinen) Klassifikationsteil beschrieben, kann es sich insb. um binäre Klassifikationen, Multi-Class Klassifikationen oder Multi-Label Klassifikationen handeln. Klassische Ansätze basieren auf Kombinationen von Textrepräsentationsstrategien wie Bag-of-Words oder TF-IDF-Vektoren (vgl. Abschnitt 4.2.1) in Kombination mit maschinellen Lernverfahren wie Support Vector Machines, Naive Bayes, k-Nearest Neighbours etc. [LPL<sup>+</sup>22]. Mit der Einführung tiefer neuronaler Netze [Sch15], insbesondere von Transformatorarchitekturen [VSP<sup>+</sup>17] wie BERT [DCLT19] oder GPT [RWC<sup>+</sup>19], hat sich die Textklassifikation in der Praxis verändert [LPL<sup>+</sup>22]: Sprachmodelle erfassen semantische und syntaktische Zusammenhänge kontextsensitiv und ermöglichen dadurch eine robustere Klassifizierung auch in komplexen Anwendungsdomänen [LPL<sup>+</sup>22]. Die Leistungsfähigkeit dieser Modelle hängt jedoch von der Qualität und Repräsentativität der Trainingsdaten ab [LPL<sup>+</sup>22], wodurch Aspekte wie Datenvorverarbeitung sowie das Fine-Tuning der Modelle für spezifische Klassifikationsaufgaben entscheidend für die Güte der Modellqualität sind [KJH<sup>+</sup>19, LPL<sup>+</sup>22].

Für Textklassifikation bei Sprachmodellen wird ein Eingabetext  $x = (t_1, t_2, \dots, t_n)$  zunächst tokenisiert und über eine Einbettungsfunktion  $E : \mathcal{V} \rightarrow \mathbb{R}^d$  in einen kontinuierlichen Vektorraum abgebildet [LPL<sup>+</sup>22]. Die resultierende Sequenz von Einbettungen  $(x_1, \dots, x_n)$  wird anschließend durch mehrere Transformer-Schichten propagiert [VSP<sup>+</sup>17], wobei sich die Modelle insbesondere in ihrer Kontextmodellierung unterscheiden [LPL<sup>+</sup>22]:

- *BERT* nutzt eine vollständig bidirektionale Self-Attention, wodurch jeder Token die volle Kontextinformation über alle Positionen hinweg erhält [DCLT19]. Für Klassifikationsaufgaben wird die Endrepräsentation des speziellen [CLS]-Tokens verwendet, typischerweise durch Anwendung eines linearen Klassifikators mit Softmax-Ausgabe:

$$\hat{y} = \text{softmax}(Wh_{[\text{CLS}]} + b), \quad W \in \mathbb{R}^{C \times d}, \quad b \in \mathbb{R}^C.$$

Dabei ist  $h_{[\text{CLS}]} \in \mathbb{R}^d$  der finale Hidden-State dieses Tokens aus der letzten Encoder-Schicht, und  $C$  bezeichnet die Anzahl der Klassen [DCLT19].

- *GPT* verwendet eine autoregressive, unidirektionale Architektur mit kausal maskierter Self-Attention. Dadurch wird die Repräsentation jedes Tokens nur auf vorhergehende Positionen konditioniert, d.h. für Position  $i$  gilt  $\alpha_{ij} = 0$  für  $j > i$  [RWC<sup>+</sup>19].

Bei Klassifikationsaufgaben wird häufig die Repräsentation des letzten Tokens der Eingabesequenz als Entscheidungsgrundlage verwendet. Die Ausgabe erfolgt analog über einen linearen Projektionslayer [RWC<sup>+</sup>19].

Die im Abschnitt 4.1.1 eingeführten Verlustfunktionen finden auch im Kontext der Textklassifikation Anwendung. Bei Multi-Class-Klassifikationsaufgaben wird typischerweise die kategoriale Kreuzentropie in Kombination mit einer Softmax-Aktivierung über den Klassenausgaben eingesetzt, während bei Multi-Label-Klassifikation aufgrund der Nicht-Exklusivität der Zielklassen die binäre Kreuzentropie mit Sigmoid-Aktivierung pro Ausgabeknoten zur Anwendung kommt [GBC16]. Die Auswahl der jeweils geeigneten Verlustfunktion erfolgt in der Regel implizit durch die Architektur der verwendeten Sprachmodelle sowie durch die Konfiguration entsprechender Deep-Learning-Frameworks und Bibliotheken [Jai22].

Die Bewertung der Modellgüte bei Textklassifikationsaufgaben erfolgt anhand von Metriken, deren Auswahl sich nach der Art der Klassifikation sowie den Eigenschaften des Datensatzes richtet [LPL<sup>+</sup>22]. Für binäre und Multi-Class-Klassifikation werden *Accuracy*, *Precision*, *Recall* und *F1-Score* eingesetzt [LPL<sup>+</sup>22]. Accuracy misst den Anteil korrekt klassifizierter Instanzen an allen Vorhersagen, ist jedoch bei unausgeglichene Klassenverteilungen nur bedingt aussagekräftig. Precision (Präzision) definiert sich als Anteil korrekt positiver Vorhersagen an allen positiven Vorhersagen, Recall (Sensitivität) den Anteil korrekt erkannter positiver Instanzen an allen tatsächlich positiven Instanzen beschreibt, während der F1-Score Precision und Recall als harmonisches Mittel vereint und beide Größen gleich gewichtet [LPL<sup>+</sup>22]:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad \text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Bei Multi-Class- oder Multi-Label-Klassifikation werden häufig aggregierte Varianten des F1-Scores eingesetzt [LPL<sup>+</sup>22]:

- *Macro-F1* berechnet den F1-Score für jede Klasse einzeln und bildet anschließend den ungewichteten Mittelwert. Dadurch werden alle Klassen gleich behandelt, unabhängig von ihrer Häufigkeit.
- *Micro-F1* aggregiert die Vorhersagen über alle Klassen hinweg und berechnet dann einen globalen F1-Score. Diese Variante ist stärker durch häufige Klassen beeinflusst.

Zur Evaluierung der Trennschärfe eines Modells, insbesondere bei schwach balancierten Daten, werden zusätzlich Receiver Operating Characteristic (ROC)- oder Precision-Recall-Kurven sowie die jeweilige Area Under the Curve (AUC) herangezogen [DG06].

Für eine Einschätzung der Generalisierungsfähigkeit eines Modells ist die Anwendung von klassischen Trainings-/Test(-/Validierungs)aufteilungen, Cross Validation (dt. Kreuzvalidation) oder Benchmark-Vergleichen (z. B. GLUE [WSM<sup>+</sup>19]) üblich. Beim Einsatz von *k*-fold Cross-Validation wird das Datenset in *k* disjunkte Teilmengen (sog.

Folds) unterteilt [RTL09]. In jedem der  $k$  Durchläufe wird jeweils ein Fold zur Validierung verwendet, während die verbleibenden  $k - 1$  Folds dem Training dienen. Durch dieses Verfahren wird jeder Datenpunkt genau einmal zur Validierung und  $(k-1)$ -mal zum Training herangezogen, was insb. bei kleineren Datensätzen zu einer effizienteren Nutzung der verfügbaren Daten führt [RTL09]. Übliche Werte für  $k$  sind 5 oder 10, da sie einen guten Kompromiss zwischen Rechenaufwand und Varianz der Schätzung darstellen [RTL09].

#### 4.4.2 Anwendungsbereich: Textregression

Ein der Textklassifikation (vgl. Abschnitt 4.4.1) strukturell verwandter Anwendungsfall in Sprachmodellen ist die Textregression, bei der aus einem Eingabetext ein kontinuierlicher numerischer Zielwert vorhergesagt wird [Tah24]. Die (allgemeine) Regression wurde bereits im Abschnitt zum überwachten Lernen (vgl. Abschnitt 4.1.1) eingeführt. Klassische Verfahren nutzen (wie bei der Textklassifikation) eine Kombination aus textuellen Repräsentationen wie TF-IDF-Vektoren oder Bag-of-Words-Kodierungen in Kombination mit Regressionsmodellen wie Linearer Regression, Support Vector Machines, Random Forest Regressoren oder Gradient Boosting [Tah24]. Der Einsatz transformerbasierter Sprachmodelle erlaubt es, kontextsensitive semantische Strukturen im Text tiefergehend zu erfassen und verbessert in vielen Fällen die Regressionsleistung gegenüber traditionellen Methoden [Tah24].

In architektonischer Hinsicht unterscheidet sich die Textregression von der Textklassifikation im Wesentlichen in der Ausgestaltung des finalen Ausgabekopfs des Sprachmodells. Während bei der Klassifikation üblicherweise ein vollvernetzter Projektionslayer mit anschließender Softmax bzw. Sigmoid-Aktivierung verwendet wird, um eine Wahrscheinlichkeitsverteilung über diskrete Klassen  $\mathcal{Y} = \{1, \dots, C\}$  zu modellieren [DCLT19], besteht der Regressionskopf lediglich aus einer linearen Abbildung:

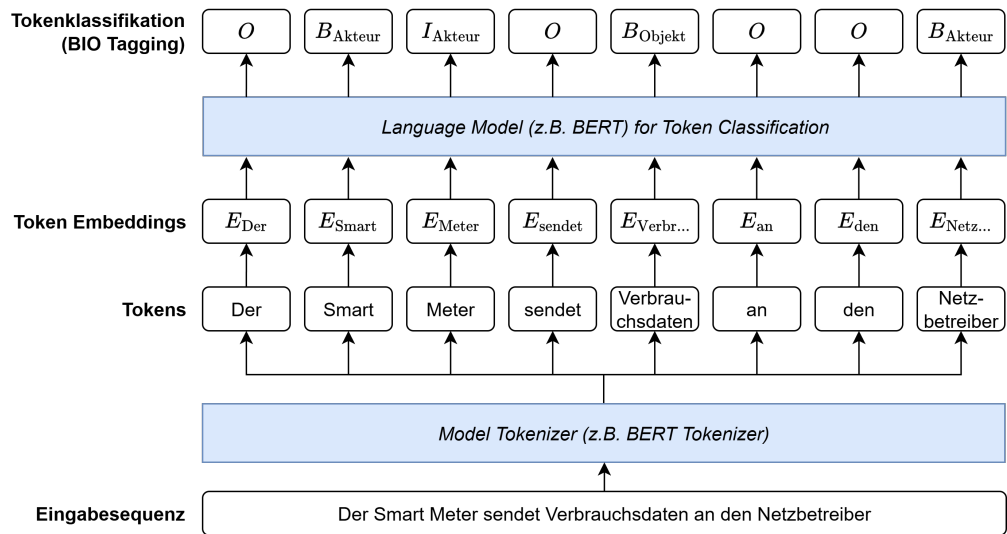
$$\hat{y} = w^\top h + b, \quad w \in \mathbb{R}^d, \quad b \in \mathbb{R},$$

wobei  $h \in \mathbb{R}^d$  die extrahierte Textrepräsentation ist (z. B. der [CLS]-Vektor bei BERT oder die finale Hidden-State-Repräsentation eines speziellen Tokens bei GPT) [Tah24].

Auch die Verlustfunktion unterscheidet sich: Statt der Kreuzentropie wird typischerweise der MSE oder MAE als Optimierungsziel verwendet [Tah24]. Damit unterscheidet sich die Textregression im Kern durch die Ausrichtung des Modelloutputs auf den kontinuierlichen Wertebereich  $\mathbb{R}$ , während die anderen Komponenten des Sprachmodells (wie Tokenisierung, Embedding und kontextuelle Repräsentation) unverändert bleiben [Tah24].

#### 4.4.3 Anwendungsbereich: Tokenklassifikation

Tokenklassifikation ist die Aufgabe im NLP, bei der jedem einzelnen Token eines Textes eine spezifische Kategorie oder Rolle zugewiesen wird. Im Unterschied zur Textklassi-



Quelle: Eigene Darstellung

Abbildung 4.8: Beispiel für Tokenklassifikation mit BIO-Tagging

fikation (vgl. Abschnitt 4.4.1) erfolgt hier die Vorhersage nicht auf Textebene, sondern sequentiell auf der Ebene einzelner Token [EHP<sup>+</sup>24]. Ein Anwendungsfall für Tokenklassifikation stellt die Named Entity Recognition (NER) dar, bei der Eigennamen bzw. benannte Entitäten (z. B. Akteure, Organisationen oder Orte) in einem Text identifiziert und typisiert werden [EHP<sup>+</sup>24]. Eine strukturierte Annotation in NER-Aufgaben stellt das Beginning-Inside-Outside (BIO)-Tagging-Schema dar [DTSOD23, Jai22]. Hierbei wird jedem Token ein Label der Form  $B-X$ ,  $I-X$  oder  $O$  zugewiesen, wobei  $X$  den Typ der Entität bezeichnet (z. B. Akteur) [DTSOD23]. Das  $B$ -Label markiert den Beginn einer Entität,  $I$  kennzeichnet die Fortsetzung derselben Entität, und  $O$  steht für Token außerhalb jeglicher benannter Entität [DTSOD23]. Dieses sequenzielle Labelling erlaubt es dem Modell, zusammenhängende Entitätsspannen zu erkennen und korrekt zu segmentieren [DTSOD23].

Für die Tokenklassifikation wird analog zur allgemeinen Textklassifikation pro Token eine kontextabhängige Repräsentation  $h_i \in \mathbb{R}^d$  erzeugt und über einen linearen Klassifikator mit Softmax-Aktivierung in ein Label überführt. Die Optimierung erfolgt sequenzbasiert mittels Kreuzentropie, aggregiert über alle Tokens der Eingabe [EHP<sup>+</sup>24].

Ein Beispiel für eine Tokenklassifikation ist in Abbildung 4.8 dargestellt. Die Textsequenz „Der Smart Meter sendet Verbrauchsdaten an den Netzbetreiber“ wurde dabei zunächst tokenisiert und anschließend mithilfe eines Sprachmodells gemäß dem BIO-Tagging annotiert. Ziel der Klassifikation ist die Extraktion relevanter Entitäten aus dem Text, insbesondere von Akteuren und (Informations-)Objekten.

#### 4.4.4 Anwendungsbereich: Textgenerierung

Insbesondere durch KI-Plattformen wie ChatGPT [OAA<sup>+</sup>24] geprägt, gewinnt das Thema der Textgenerierung zunehmend an Bedeutung. Ein zentraler Anwendungsbereich großer Sprachmodelle ist die Textgenerierung, bei der auf Basis eines gegebenen Kontexts ein kohärenter und semantisch sinnvoller Folgetext erzeugt wird. Die nachfolgend dargestellten Modellierungsansätze stellen etablierte Paradigmen für die automatische Textgenerierung sowie deren Trainingsstrategien dar [LTZ<sup>+</sup>24]:

- **Causal Language Modelling (CLM):** Bei diesem autoregressiven Ansatz wird die gemeinsame Wahrscheinlichkeit einer Textsequenz  $x = (x_1, \dots, x_n)$  durch die Zerlegung in bedingte Wahrscheinlichkeiten gemäß der Kettenregel modelliert:

$$p(x) = \prod_{t=1}^n p(x_t \mid x_1, \dots, x_{t-1}).$$

Die Generierung erfolgt sequenziell, wobei jedes Token auf Basis der vorangegangenen Tokens vorhergesagt wird. Aufgrund dieser Eigenschaft eignen sich CLM-Modelle insbesondere für die freie Textproduktion sowie dialogbasierte Systeme (z. B. GPT). [RWC<sup>+</sup>19]

- **Masked Language Modeling (MLM):** Dieses bidirektionale Vorhersageverfahren basiert auf der Rekonstruktion maskierter Tokens innerhalb einer Sequenz. Die mathematische Formulierung und weitere Details sind in Abschnitt 4.3.1 beschrieben (z. B. BERT). [DCLT19]
- **Encoder-Decoder-Architekturen (Seq2Seq):** Diese Modelle bestehen aus zwei Komponenten: Einem Encoder, der eine Eingabesequenz  $x = (x_1, \dots, x_n)$  in eine latente Repräsentation  $z$  überführt, und einem Decoder, der daraus eine Zielsequenz  $y = (y_1, \dots, y_m)$  generiert. Die Modellierung erfolgt typischerweise als bedingte Wahrscheinlichkeitsverteilung

$$p(y \mid x) = \prod_{t=1}^m p(y_t \mid y_1, \dots, y_{t-1}, z),$$

wobei der Decoder autoregressiv auf die bisher erzeugten Tokens sowie die kodierte Eingabe zugreift. Diese Architektur eignet sich insbesondere für Aufgaben wie maschinelle Übersetzung oder Textzusammenfassung (z. B. FLAN-T5). [RSR<sup>+</sup>23]

Typische Anwendungsfelder textgenerativer Modelle umfassen neben der offenen Textproduktion auch maschinelle Übersetzung, automatische Textzusammenfassungen, Frage-Antwort-Systeme, dialogbasierte Assistenten sowie die Generierung natürlichsprachlicher Beschreibungen strukturierter Daten. [LTZ<sup>+</sup>24]

Während sich Klassifikations- und Regressionsmodelle anhand etablierter Metriken wie Accuracy, Präzision oder F1-Score klar und objektiv vergleichen lassen [LPL<sup>+</sup>22],

gestaltet sich die Bewertung generativer Modelle komplexer. Dies liegt insbesondere daran, dass sowohl syntaktische als auch semantische Unterschiede zwischen Texten nicht eindeutig messbar sind und je nach Anwendungsfall unterschiedliche Anforderungen an Kohärenz, Korrektheit oder Diversität gestellt werden [LTZ<sup>+</sup>24]. Ein zentraler Bestandteil der automatischen Textgenerierung ist daher die Bewertung der Qualität der erzeugten Inhalte [LTZ<sup>+</sup>24]. Da generative Aufgaben in der Regel keine eindeutig richtige Zielsequenz besitzen, bedarf es spezialisierter Metriken zur Beurteilung textueller Eigenschaften, weshalb sich gemäß [LTZ<sup>+</sup>24] vier grundlegende Kategorien von Bewertungsverfahren etabliert haben:

- **Metriken basierend auf  $n$ -Gramm-Überlappung:** Diese Verfahren messen die Übereinstimmung zwischen generiertem und Referenztext auf Basis von  $n$ -Gramm-Vorkommen. Beispiele sind *BLEU*, das die Präzision der Übereinstimmungen berücksichtigt, *ROUGE*, das häufig zur Bewertung von Zusammenfassungen verwendet wird, sowie *METEOR*, das zusätzlich linguistische Aspekte wie Stemming und Synonyme einbezieht. [CCG21, LTZ<sup>+</sup>24]
- **Semantische Ähnlichkeitsmetriken:** Diese Verfahren basieren auf Repräsentationen aus vortrainierten Sprachmodellen und vergleichen generierte und Referenztexte im semantischen Vektorraum. *BERTScore* nutzt kontextuelle Embeddings, um Ähnlichkeiten zwischen Tokens zu erfassen. Weitere Verfahren wie *BLEURT*, *MoverScore* oder *COMET* kombinieren semantische Modellierung mit überwachenden Ansätzen und korrelieren gut mit menschlichen Bewertungen. [CCG21, LTZ<sup>+</sup>24]
- **Wahrscheinlichkeitsbasierte Metriken:** Diese Metriken bewerten, wie gut die Wahrscheinlichkeit einer Testsequenz von einem Sprachmodell abgeschätzt werden. Ein Maß ist die *Perplexität (PPL)*, die aus der Entropie der Verteilung abgeleitet wird. Eine geringe Perplexität deutet auf eine hohe Vorhersagegüte hin und damit auf eine hohe statistische Kohärenz des generierten Textes. [CCG21, LTZ<sup>+</sup>24]
- **Diversitätsmetriken:** Diese Metriken bewerten die lexikalische Vielfalt innerhalb der generierten Texte. *Distinct- $n$*  misst den Anteil einzigartiger  $n$ -Gramme und ist besonders in kreativen Anwendungsbereichen wie der Dialog- oder Geschichten-Generierung von Bedeutung. [CCG21, LTZ<sup>+</sup>24]

## 4.5 Betrachtung der etablierten Techniken

Die folgenden Abschnitte geben einen Überblick über wesentliche Techniken, die für die Entwicklung, Anpassung und den Betrieb moderner KI-Systeme, insb. im Bereich des NLP mit Fokus auf (L)LMs von Bedeutung sind. Dabei liegt der Schwerpunkt nicht auf der Erklärung einzelner Aspekte der Techniken sowie Paradigmen, sondern auf einer systematischen Darstellung des verfügbaren Spektrums an Verfahren. Die konkrete Relevanz und Implementierung einzelner Techniken wird im Kapitel 8 vertieft behandelt.

### 4.5.1 Transfer Learning (TL)

Ein Aspekt, der sich insbesondere in der Technik von LM/LLMs widerspiegelt ist das Transfer Learning (TL). Transfer Learning bezeichnet den Ansatz, Wissen aus einer Quellaufgabe  $\mathcal{T}_S$  auf eine Zielaufgabe  $\mathcal{T}_T$  zu übertragen. Insbesondere im Bereich (L)LMs besteht dieser Ansatz darin, ein Modell zunächst auf einer großen allgemeinen Textmenge  $\mathcal{D}_S$  auf Basis von weitreichenden Datenbasen, bspw. Wikipedia, Fachbüchern oder Papern vorzutrainieren (Pre-Training) und anschließend auf einer aufgabenspezifischen Datenmenge  $\mathcal{D}_T$  weiterzuentwickeln (Fine-Tuning). Diese zweistufige Vorgehensweise erlaubt eine effiziente Nutzung von Rechenressourcen und ermöglicht es, Wissen über Sprachstrukturen und Weltwissen in nachgelagerte Aufgaben zu übertragen. [ZQD<sup>+</sup>21]

Es werden zwischen verschiedenen TL-Strategien unterschieden, u.a. bei den Sprachmodellen [ZQD<sup>+</sup>21]. Beim klassischen Fine-Tuning werden alle Modellparameter weitertrainiert, alternativ können nur Teile des Modells angepasst oder zusätzliche Module eingebaut werden [ZQD<sup>+</sup>21]. Ansätze die sich insb. in LLMs vorzufinden sind, bei denen keine Parameteränderung nötig ist, sondern die Aufgabe direkt z. B. über Prompts oder RAG (vgl. Abschnitt 4.5.5) beschrieben werden, stellt sog. In-Context Learning dar [BMR<sup>+</sup>20].

### 4.5.2 Text Data Preprocessing

Ein wesentlicher Aspekt für das Fine-Tuning von (L)LMs ist eine qualitative sowie quantitativ adäquate Datenbasis. Textdaten, insb. IEC 62559 Use Case-Beschreibungen, stellen unstrukturierte Daten dar und können stark in ihrer syntaktischen Struktur variieren, etwa hinsichtlich der Verwendung von Füllwörtern, Satzzeichen, Groß- und Kleinschreibung oder typografischer Konventionen. Diese Aspekte beeinflussen das Trainingsverhalten und können die Performanz des Modells negativ beeinträchtigen [KJH<sup>+</sup>19, KA18].

Data Preprocessing (dt. Datenvorverarbeitung) bezeichnet den methodischen Prozess der Aufbereitung, Bereinigung und Transformation von Rohdaten in eine für das Modelltraining geeignete Struktur. Reale Datensätze sind häufig unvollständig, inkonsistent, redundant oder durch Rauschen verfälscht. Durch gezielte Vorverarbeitungsschritte wird die Qualität der Daten gesteigert, wodurch sich die Modellqualität verbessert. Im Wesentlichen werden die folgenden Aspekte betrachtet [HKP12]:

- **Datenbereinigung (Data Cleaning):** Fehlende oder fehlerhafte Werte werden durch geeignete Verfahren wie Löschung, Substitution oder statistische Schätzung behandelt. Messrauschen wird durch den Einsatz geeigneter Glättungstechniken reduziert, während die Vereinheitlichung von Formaten die Vergleichbarkeit der Daten gewährleistet. Ziel dieser Maßnahmen ist die Herstellung eines konsistenten Datensatzes.
- **Datenintegration (Data Integration):** Die Zusammenführung heterogener Datenquellen erfolgt unter Beachtung der Auflösung von Schema-, Struktur- und Namens-

konflikten. Redundante Einträge werden identifiziert und entfernt, um Speicher- und Verarbeitungsaufwand zu reduzieren. Durch die Sicherstellung semantischer Konsistenz entsteht ein konsolidierter und widerspruchsfreier Gesamtdatensatz.

- **Datenreduktion (Data Reduction):** Die Verringerung von Attributanzahl oder Datenvolumen erfolgt unter Erhalt der relevanten Informationen. Verfahren und Techniken wie die Principle Component Analysis, Feature Engineering, Sampling und Aggregation können angewandt werden, um Speicher- und Rechenaufwand ohne Qualitätsverlust zu reduzieren.

Zur Bewältigung der identifizierten Herausforderungen werden spezifische Techniken des Text-Data-Preprocessings eingesetzt, die die zuvor beschriebenen allgemeinen Schritte der Datenvorverarbeitung gezielt auf textuelle Daten übertragen. Ziel ist es, die syntaktische Qualität sowie die syntaktische Homogenität der Textdaten zu verbessern, wodurch die Generalisierungsfähigkeit des Modells gesteigert werden kann [KJH<sup>+</sup>19]. Typische Preprocessing-Schritte für Textvorverarbeitung gemäß [AHN19] umfassen u. a. die folgenden Aspekte:

- **Tokenisierung (Unitize & Tokenize):** Zerlegung des Textes in analysierbare Einheiten (Tokens), typischerweise Wörter oder Phrasen. Optional kann eine Repräsentation durch  $n$ -Gramme (z. B. Bigrams, Trigrams) erfolgen, um die lokale Kontextinformation zu bewahren. [AHN19]
- **Standardisierung und Bereinigung:** Vereinheitlichung der Texte durch Konvertierung in Kleinschreibung sowie Entfernung von Zahlen, Satzzeichen, Sonderzeichen und überflüssigen Leerzeichen, um lexikalische Varianz zu reduzieren. [AHN19]
- **Stopwort-Entfernung:** Entfernung häufig vorkommender, inhaltsarmer Wörter (z. B. *und, ist, the, of*), die für viele Analyseaufgaben wenig semantischen Mehrwert bieten. Neben generellen Standardlisten, z. B. System for the Mechanical Analysis and Retrieval of Text (SMART), können domänenspezifische Stopwortlisten verwendet werden. [AHN19]
- **Stemming und Lemmatisierung:** Reduktion von Wörtern auf ihren Wortstamm (Stemming) oder ihre Grundform unter Berücksichtigung grammatischer Merkmale (Lemmatisierung), um die Vokabulargröße zu verringern (z. B. *train, trains, training* → *train*) [AHN19]
- **Part-of-Speech (POS) Tagging:** Annotation von Tokens mit ihrer Wortart (z. B. Substantiv, Verb, Adjektiv), um semantische Bedeutungen besser zu erfassen und die Genauigkeit von Folgeprozessen wie Lemmatisierung oder Entity Recognition zu erhöhen. [AHN19]
- **Benutzerdefinierte Lexika:** Einsatz domänenspezifischer Wörterbücher zur Entfernung irrelevanter Tokens oder zur gezielten Extraktion relevanter Begriffe (z. B. Keywords, Named Entities), angepasst an den jeweiligen Analysekontext. [AHN19]

In modernen Sprachmodellen sind einige der Schritte des Text Preprocessings bereits im Tokenizer oder der Modellarchitektur integriert. Beispielsweise verwendet BERT den WordPiece-Tokenizer, der automatisch Tokenisierung, Normalisierung (z. B. Kleinschreibung im Fall von BERT) sowie das Handling unbekannter Wörter übernimmt [DCLT19]. GPT-2 und GPT-3 setzen auf Byte-Pair Encoding, wodurch selbst seltene oder fehlerhafte Eingaben in sublexikalische Einheiten zerlegt und modellierbar werden [RWC<sup>+</sup>19]. In FLAN-T5 wird ebenfalls ein eigener SentencePiece-Tokenizer genutzt, der auf Unicode-Normalisierung basiert und explizite Trennung von Wörtern nicht zwingend voraussetzt [RSR<sup>+</sup>23]. Sprachmodelle wie BERT oder GPT integrieren bereits einige klassische Preprocessing-Schritte implizit, da ihre Tokenizer und tiefen kontextuellen Repräsentationen semantische und syntaktische Informationen erfassen [DCLT19, RWC<sup>+</sup>19]. Dadurch entfällt oft die Notwendigkeit separater Verfahren wie Stemming, Lemmatisierung oder das Entfernen von Stopwörtern.

### 4.5.3 Text Data Augmentation

(Text) Data Augmentation bezeichnet die systematische Erzeugung zusätzlicher Trainingsdaten durch kontrollierte Transformation vorhandener Daten bzw. Texte. Ziel ist es, die Varianz im Trainingskorpus zu erhöhen, um die Generalisierungsfähigkeit von Modellen zu verbessern, Overfitting zu reduzieren und die Robustheit gegenüber Rauschen zu steigern. Die Notwendigkeit solcher Verfahren ergibt sich insbesondere in Szenarien mit begrenzter Datenverfügbarkeit, unbalancierten Klassenverteilungen oder domänenspezifischen Korpora. Folglich können Data Augmentation-Techniken insb. bei kleineren Datensätzen angewandt werden, um so eine Performanzsteigerung zu erzielen. [LHC22, FGW<sup>+</sup>21]

Im Gegensatz zum Data Preprocessing, das primär auf die Bereinigung, Standardisierung und strukturelle Harmonisierung von Rohdaten abzielt (vgl. Abschnitt 4.5.2), verfolgt (Text) Data Augmentation das Ziel, den Umfang und die Vielfalt der Trainingsdaten zu vergrößern [LHC22, FGW<sup>+</sup>21]. Während Preprocessing darauf bedacht ist, Inkonsistenzen und Rauschen zu eliminieren [HKP12], nutzt Augmentation gezielte Transformationen, um neue Varianten existierender Daten zu generieren [LHC22, FGW<sup>+</sup>21]. Die folgenden Herangehensweisen (kategorisiert über die Metastudie von [FGW<sup>+</sup>21]) können dabei u.a. betrachtet werden:

- **Regelbasierte Techniken:** Regelbasierte Transformationen, die auf deterministische, reproduzierbare Verfahren basieren. Hierzu zählen tokenbasierte Operationen wie das Einfügen, Löschen oder Vertauschen von Wörtern, syntaktische Umstrukturierungen sowie konsistenzbasierte Verfahren. Diese Methoden sind mit geringem Implementierungsaufwand verbunden und ermöglichen eine gezielte Erhöhung der Dateninstanzen für unterrepräsentierte Klassen in Klassifikationsaufgaben.
- **Beispielinterpolationstechniken:** Verfahren, bei denen zwei oder mehr existierende Beispiele kombiniert oder interpoliert werden, um neue Trainingsinstanzen zu erzeugen. Durch die resultierende Glättung der Entscheidungsgrenzen im Dimensionsraum

tragen sie nachweislich zur Verbesserung der Generalisierungsleistung und zur Verringerung von Over Fitting bei.

- **Modellbasierte Techniken:** Methoden, die (L)LMs zur Generierung synthetischer Dateninstanzen nutzen. Durch den Einsatz von gezielten Prompts zur Generierung von synthetischen Daten lassen sich auf Basis etablierter Modelle (z. B. ChatGPT [OAA<sup>+</sup>24] oder LLaMA [TLI<sup>+</sup>23]) domänen- und aufgabenspezifische Datensätze erzeugen.

Der Einsatz von Data-Augmentation-Strategien kann insbesondere dazu beitragen, die Datenbasis für das Fine-Tuning von Sprachmodellen zu erweitern und die Leistungsfähigkeit der Modelle zu steigern. Klassische, nicht-große Sprachmodelle (z. B. BERT [DCLT19], RoBERTa [LOG<sup>+</sup>19], GPT-2 [RWC<sup>+</sup>19]) sind in der Regel stärker auf umfangreichere Datensätze und größere Few-Shot-Mengen angewiesen, um adäquate Vorhersagen zu erzielen [ATCG<sup>+</sup>19]. Im Gegensatz dazu benötigen große Sprachmodelle (z. B. GPT-3 und aufwärts, FLAN-T5, LLaMA) in Few-Shot-Szenarien weniger zusätzliche Daten und können zudem im No-Shot-Setting Vorhersagen ohne explizite Trainingsexemplare generieren [ZLX<sup>+</sup>23].

#### 4.5.4 eXplainable Artificial Intelligence (XAI)

KI hat in den letzten Jahren in einer Vielzahl von Anwendungsbereichen an Bedeutung gewonnen, weshalb die Anforderungslage sich weiterentwickelt hat. Mit der zunehmenden Leistungsfähigkeit komplexer Modelle, insbesondere bei DL, geht jedoch eine wachsende Intransparenz ihrer Entscheidungsprozesse einher. Dabei handelt es sich um Black-Box-Modelle, deren Entscheidungsprozesse nur schwer nachzuvollziehen sind. Im Kontext sicherheitskritischer Anwendungen, gegebenen regulatorischen Anforderungen oder das Ziel in der Wissenschaft, Nachvollziehbarkeit zu realisieren, ist jedoch nicht allein das Ergebnis, sondern auch die Begründung der Entscheidung von Relevanz. An dieser Stelle setzt das Forschungsfeld der eXplainable Artificial Intelligence (XAI) an, das Methoden und Werkzeuge bereitstellt, um die internen Funktionsweisen von KI-Systemen für den Menschen verständlich zu machen. [SWM17]

Unter Erklärbarkeit wird in der Literatur häufig zwischen *transparenten Modellen* (intrinsisch erklärbar; vereinzelt auch als ante-hoc bezeichnet) und *post-hoc-Erklärbarkeit* unterschieden [RAS<sup>+</sup>24, BADRDS<sup>+</sup>20, XUD<sup>+</sup>19]. Transparente Modelle sind von vornherein so konstruiert, dass ihre Funktionsweise für Menschen unmittelbar nachvollziehbar ist. Die Erklärungen entstehen direkt aus der Modellstruktur, sodass keine nachträgliche Übersetzung erforderlich ist. Typische Beispiele sind Entscheidungsbäume, lineare Modelle, regelbasierte Systeme oder Generalized Additive Models (GAM) [RAS<sup>+</sup>24]. Post-hoc-Erklärbarkeit hingegen umfasst Methoden, die darauf abzielen, die Entscheidungen komplexer, oft als „Black Box“ bezeichneter Modelle, etwa tiefer KNNs, im Nachhinein zu interpretieren. Dies geschieht beispielsweise durch modellagnostische Ansätze wie LIME, SHAP, Feature-Importance-Analysen, Gegenfaktische Erklärungen

oder visuelle Verfahren wie Aktivierungs-Heatmaps [GMR<sup>+</sup>18]. Es zeigt sich, dass insbesondere transparente Modelle zwar eine höhere Erklärbarkeit aufweisen, ihre Vorhersagegenauigkeit jedoch in der Regel geringer ist als die komplexerer, post-hoc interpretierter Modelle [XUD<sup>+</sup>19].

Darüber hinaus wird zwischen *globaler und lokaler Erklärbarkeit* in der Granularität unterschieden [Mol22, GMR<sup>+</sup>18, BADRDS<sup>+</sup>20]. Globale Erklärbarkeit beschreibt Ansätze, die ein umfassendes Verständnis des gesamten Modells ermöglichen. Ziel ist es, allgemeine Zusammenhänge zwischen Eingabevariablen und Modellvorhersagen zu identifizieren und so das generelle Entscheidungsverhalten des Modells nachvollziehbar zu machen. Typische Beispiele sind globale Feature-Importance-Analysen, Partial Dependence Plots oder Surrogatmodelle, die ein komplexes Modell durch ein transparentes Modell approximieren [Mol22]. Lokale Erklärbarkeit fokussiert sich auf die Interpretation einzelner Modellvorhersagen. Hierbei wird untersucht, warum ein Modell für eine spezifische Eingabe eine bestimmte Ausgabe liefert. Beispiele hierfür sind LIME, SHAP, Individual Conditional Expectation Plots oder gegenfaktische Erklärungen [GMR<sup>+</sup>18, BADRDS<sup>+</sup>20].

Globale Methoden tragen zu einem übergreifenden Verständnis des Modellverhaltens bei, während lokale Methoden eine detaillierte Analyse einzelner Entscheidungen ermöglichen, was insbesondere in sicherheitskritischen Anwendungsfällen relevant ist. Lokale Ansätze können insbesondere in sicherheitskritischen Anwendungsfällen relevant sein, in denen die Nachvollziehbarkeit einzelner Entscheidungen erforderlich ist [Mol22].

LLMs zählen aufgrund ihrer komplexen, auf KNNs basierenden Architektur zu den typischen Black-Box-Modellen. Um ihre Vorhersagen und generierten Inhalte nachvollziehbar zu machen, kommen sowohl globale als auch lokale Erklärbarkeitsansätze zum Einsatz. Globale Methoden für LLMs zielen bspw. darauf ab, die Rolle einzelner Layer oder Attention-Muster im Gesamtmodell zu analysieren, während lokale Verfahren wie Attention-Visualisierung, Integrated Gradients oder Gegenfaktische Analysen die Erklärung einzelner Modelloutputs unterstützen. [BP21]

Ein zentraler Aspekt aktueller XAI-Forschung ist der Menschen-zentrierte Ansatz, bei dem die Gestaltung von Erklärungen an den Bedürfnissen und kognitiven Fähigkeiten der Nutzer ausgerichtet wird. Ziel ist es, nicht nur die technische Nachvollziehbarkeit von Modellentscheidungen zu verbessern, sondern auch sicherzustellen, dass diese Erklärungen in der jeweiligen Anwendungssituation korrekt interpretiert und angemessen genutzt werden z. B. innerhalb einer Systemumgebung. [KLZ24]

In diesem Kontext können innerhalb intelligenter Systeme sogenannte Cognitive Forcing Functions (CFFs) gezielt eingesetzt werden, um die Qualität menschlicher Entscheidungen zu unterstützen und kognitive Verzerrungen zu reduzieren. Ursprünglich aus der medizinischen Diagnostik stammend, dienen CFFs dazu, Entscheidungsträger gezielt zu einer kritischen Überprüfung ihrer ersten Schlussfolgerungen zu veranlassen und kognitive Verzerrungen wie den Automation Bias zu reduzieren. Übertragen auf den Bereich der menschen-zentrierten XAI können CFFs in Benutzerschnittstellen integriert

werden, um Anwender explizit dazu anzuregen, alternative Hypothesen zu prüfen, die Erklärungen des Systems kritisch zu hinterfragen und so die Qualität der menschlichen Entscheidungsfindung in KI-gestützten Prozessen zu erhöhen. [BMG21]

#### 4.5.5 Retrieval Augmented Generation (RAG)

LLMs wie GPT-4 [OAA<sup>+</sup>24] oder LLaMA [TLI<sup>+</sup>23] basieren auf dem Wissen, das sie während des Pre-Trainings aus einem umfangreichen Textkorpus extrahieren. Dieses Wissen ist jedoch statisch und auf den Stand des Trainingszeitpunkts beschränkt. Infolgedessen können solche Modelle bei Anfragen, die über den Trainingskorpus hinausgehen oder aktuelles bzw. domänenspezifisches Wissen erfordern, zu fehlerhaften oder haluzinierten Ausgaben neigen. Retrieval Augmented Generation (RAG) adressiert diese Limitationen, indem sie externe Wissensquellen zur Laufzeit (d.h. zur Inferenz) einbindet [GXG<sup>+</sup>24]. Folglich wird das Sprachmodell mit kontextrelevanter, aktualisierbarer Information angereichert, ohne dass ein vollständiges Fine-Tuning des Modells erforderlich ist [GXG<sup>+</sup>24]. Zudem können Quellen mit in der Ausgabe aufgeführt werden, was die Transparenz steigern kann [GXG<sup>+</sup>24]. RAG besteht typischerweise aus zwei zentralen Komponenten:

- **Retriever**  $\mathcal{R}$ : Gegeben einer Nutzereingabe  $q \in \mathcal{Q}$ , extrahiert der Retriever aus einer externen Wissensbasis  $\mathcal{D}$  eine Menge relevanter Dokumente  $D_q = \{d_1, \dots, d_k\} \subseteq \mathcal{D}$ , die semantisch zum Input  $q$  passen. Die Abfrage erfolgt über eine semantische Ähnlichkeitsfunktion im Vektorraum:

$$D_q = \mathcal{R}(q; \theta_R),$$

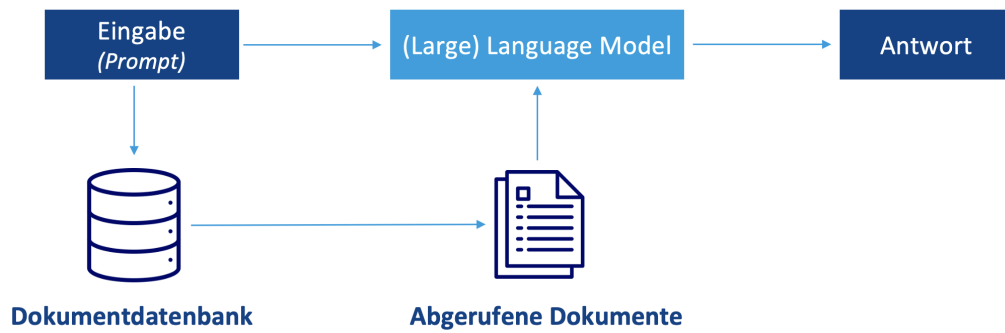
wobei  $\theta_R$  die Parameter des Retrievalmoduls (z. B. ein Embedding-Modell) bezeichnet. [GXG<sup>+</sup>24]

- **Generator**  $\mathcal{G}$ : Der Generator erhält als Input sowohl die ursprüngliche Anfrage  $q$  als auch die vom Retriever zurückgelieferten kontextrelevanten Dokumente  $D_q$ . Auf Basis dieser Informationen wird eine Antwort  $a \in \mathcal{A}$  erzeugt:

$$a = \mathcal{G}(q, D_q; \theta_G),$$

wobei  $\theta_G$  die Parameter des gen. Sprachmodells beschreibt (z. B. GPT-Architektur). [GXG<sup>+</sup>24]

Abbildung 4.9 illustriert den generischen Ablauf der RAG-Architektur. Der *Retriever*  $\mathcal{R}$  greift auf eine externe Dokumentendatenbank zu und selektiert kontextrelevante Dokumente, welche anschließend in den Prompt des Generators  $\mathcal{G}$  eingebettet werden. Die zugrundeliegende Dokumentenquelle kann aus eigenen Datenbanklösungen bestehen (z. B. GraphQL, PostgreSQL oder vektorbasierte Datenbanken), aber auch Quellen



Quelle: Eigene Abbildung nach [RML24]

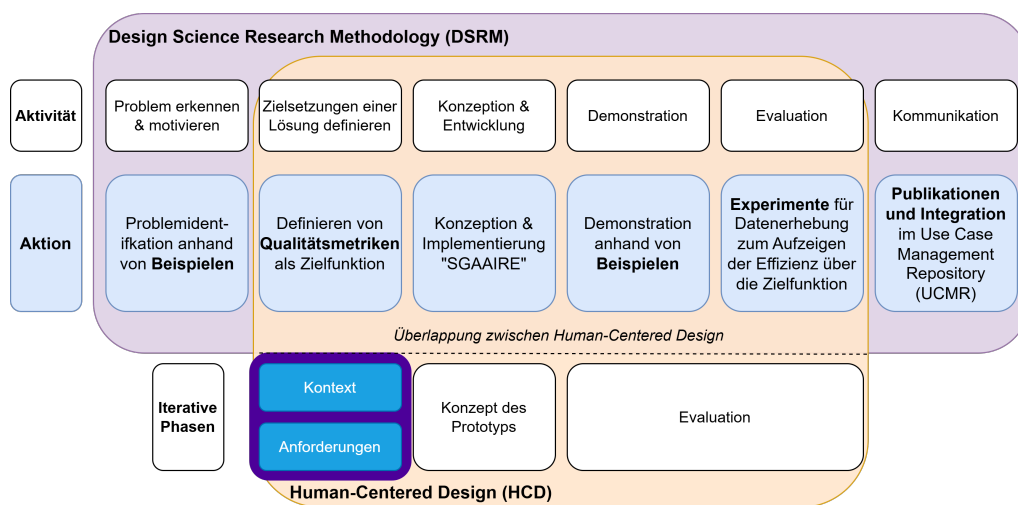
Abbildung 4.9: Retrieval Augmented Generation (RAG)

wie Wikipedia, ResearchGate sowie Domänen-spezifische Wissensbasen umfassen. Die abgerufenen Dokumente werden gemeinsam mit der Nutzereingabe in einen überarbeiteten Prompt integriert und dem Generator  $\mathcal{G}$  übergeben. Dieser nutzt die zusätzlichen Informationen zur Generierung einer Antwort mit Bezug auf externe Informationen. [GXG<sup>+</sup>24]



## 5 Anforderungsanalyse von SGAAIRE

Dieses Kapitel behandelt die Kontextanalyse sowie die Anforderungserhebung für die Entwicklung des KI-gestützten Systems *SGAAIRE*. Die Kontextanalyse beleuchtet insbesondere die Herausforderungen, die bei der Arbeit mit der IEC 62559 Use Case-Methodik, dem SGAM Framework sowie den eingesetzten Tools auftreten, insbesondere im Bereich der Erstellung. Abbildung 5.1 stellt die Einordnung der Ergebnisse (hervorgehoben in Lila) im Gesamtverfahren dar. Die Erkenntnisse aus dieser Analyse bilden die Grundlage für die folgenden Kapitel 7 und 8, in denen *SGAAIRE* eingeordnet wird. Ein nutzerzentrierter Ansatz für die Anforderungserhebung soll sicherstellen, dass die Bedürfnisse und Herausforderungen der Nutzenden berücksichtigt werden.



Quelle: Eigene Darstellung, angelehnt an [PTRC07]

Abbildung 5.1: Auswirkungen der Ergebnisse aus Kapitel 4 auf das Forschungsvorhaben

Sowohl für die Kontextanalyse als auch für die Anforderungserhebung wurde eine gemeinsame Umfrage durchgeführt, die in Abschnitt 5.1 näher erläutert wird. Auf Grundlage der Ergebnisse der Umfrage werden anschließend die nicht-funktionalen Anforderungen sowie die funktionalen Anforderungen abgeleitet. Diese dienen als direkte Anforderungen an das zu entwickelnde intelligente System *SGAAIRE* und werden in den Abschnitten 5.2 und 5.3 hergeleitet und beschrieben.

## 5.1 Anforderungs- und Kontexterhebung: Teilstandardisierter Fragebogen

Für die Konzeption von unterstützenden KI-Systemen in der IEC 62559 Use Case-Methodik und dem SGAM Framework ist es hilfreich, sowohl die Arbeitsweise als auch die damit verbundenen Herausforderungen zu analysieren, um ein fundiertes Verständnis des Kontextes zu gewinnen. Diese Analyse wird im Rahmen eines teilstandardisierten Fragebogens durchgeführt, der darauf abzielt, ein umfassendes Bild der relevanten Aspekte zu erhalten. Darüber hinaus dient der selbige Fragebogen der Erhebung funktionaler Anforderungen. Basierend auf bestehender Literatur werden funktionale Anforderungen abgeleitet, die anschließend von Anwendenden der IEC 62559 Use Case-Methodik und des SGAM Frameworks hinsichtlich ihrer Auswirkungen auf die Qualitätsverbesserung bewertet werden. Es wird angenommen, dass funktionale Anforderungen, die keine statistische Signifikanz erreichen, im Umsetzungsprozess keine wesentliche Verbesserung der Qualität bewirken. Aus diesem Grund werden sie im Rahmen der Nutzwertanalyse nicht weiter berücksichtigt. Anforderungen, die statistisch signifikant sind, werden hingegen als Entscheidungsalternativen in die Nutzwertanalyse integriert.

### 5.1.1 Aufbau und Entwicklung der Umfrage

Die im Rahmen dieser Studie durchgeführte Umfrage wurde als Online-Fragebogen konzipiert und mithilfe des Tools Limesurvey implementiert. Eine Instanz des Tools wurde durch *OFFIS e.V.* bereitgestellt. Die Wahl eines Online-Fragebogens erfolgte aufgrund der einfachen Zugänglichkeit für potenzielle Teilnehmende, der Möglichkeit, verschiedene Fragetypen zu nutzen, und der spezifischen Bereitstellung von Fragestellungen [JC07]. Dadurch konnte eine objektive Meinungsbildung einer heterogenen Gruppe von Teilnehmenden erreicht werden. Um die Zuverlässigkeit der Ergebnisse zu gewährleisten, war eine Voraussetzung für die Teilnahme, dass die Befragten bereits Erfahrung in der Erstellung von IEC 62559 Use Case-Beschreibungen und/oder SGAM Modellen hatten, was im Fragebogen entsprechend berücksichtigt wurde [KNSS24, KNN<sup>+</sup>24].

Ziel der Umfrage war es, die Erfahrungen einer vielfältigen Kohorte von IEC 62559 Use Case- und SGAM Modell-Erstellern mit unterschiedlichen Erfahrungsstufen zu erfassen und funktionale Anforderungen zu erheben. Die Teilnehmenden wurden dazu aufgefordert, die Herausforderungen, denen sie sowohl bei der Arbeit mit den entsprechenden Standards als auch bei der Anwendung der Toolchain begegneten, zu benennen. Der Aufbau des Fragebogens<sup>1</sup> umfasste folgende thematische Abschnitte:

- **Abschnitt A + B (Allgemeine Profilinformationen zur Erstellung von IEC 62559 Use Cases und SGAM Modellen):** Diese Abschnitte dienen der Erfassung der Erfahrungswerte der Teilnehmenden, ihrer Selbsteinschätzung hinsichtlich ihrer Expertise sowie der Dauer ihrer Beschäftigung mit den genannten Methodiken.

<sup>1</sup> Der vollständige Fragebogen ist im Online-Anhang im Ordner Online-Umfrage einsehbar.

- **Abschnitt C + E (Erstellung von IEC 62559 Use Cases und SGAM Modellen):** Diese Abschnitte befassen sich mit der Erhebung spezifischer Herausforderungen und der Bewertung der eingesetzten Werkzeuge im Rahmen der Methodiken.
- **Abschnitt D + F (Anforderungen an ein (KI-)System zur Unterstützung der Erstellung von IEC 62559 Use Cases und SGAM Modellen):** In diesen Frageblöcken werden Anforderungen an ein unterstützendes (KI-)System für die Erstellung von IEC 62559 Use Cases und SGAM Modellen erhoben. Die Teilnehmenden werden aufgefordert, die Auswirkungen funktionaler Anforderungen auf die Qualität einzuschätzen sowie gegebenenfalls zusätzliche Anforderungen zu formulieren.
- **Abschnitt G (Abschluss der Umfrage und Erfassung von Metadaten):** Dieser Abschnitt umfasst die Erhebung demografischer Daten der Teilnehmenden, die abschließende Danksagung sowie die finalen Angaben zur Umfrage.

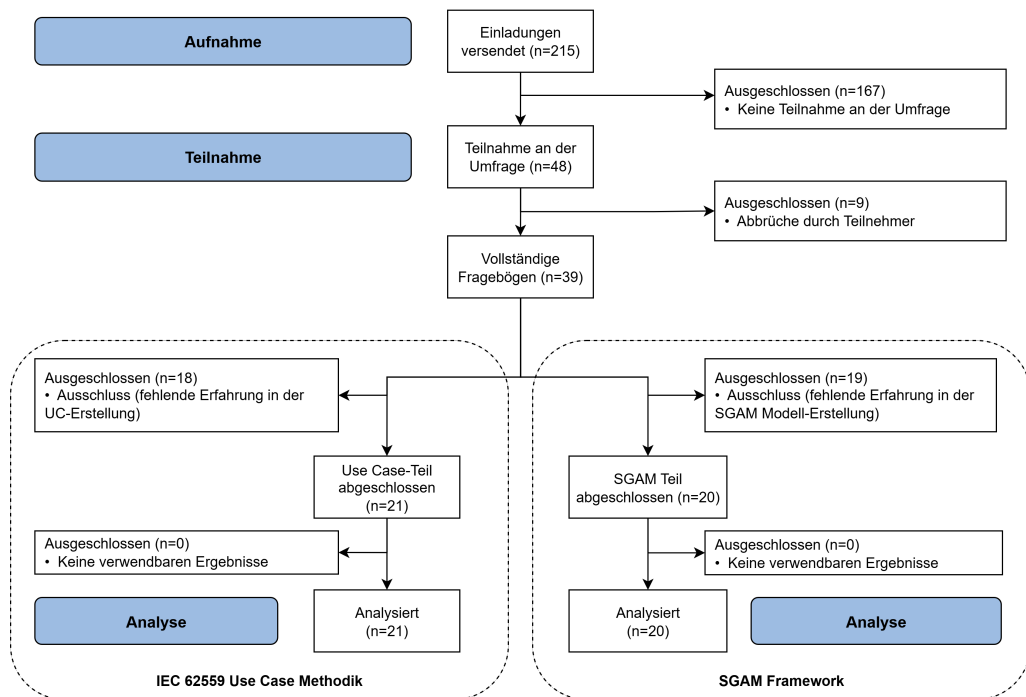
Die Abschnitte C bis G konnten nur von Teilnehmenden ausgefüllt werden, die in Abschnitt A und B angegeben hatten, bereits IEC 62559 Use Case-Beschreibungen und/oder SGAM Modelle erstellt zu haben. Basierend auf den Rückmeldungen zweier Pilotteilnehmender wurde die durchschnittliche Bearbeitungszeit der Umfrage auf etwa 28 Minuten geschätzt. Zur Wahrung der Anonymität wurde in Limesurvey die IP-Adressnachverfolgung deaktiviert und keine Zeitstempel gespeichert. Daher stehen keine personenbezogenen Daten zur Verfügung. [KNSS24, KNN<sup>+</sup>24]

### 5.1.2 Durchführung der Umfrage

Die Umfrage wurde zwischen Juni und Juli 2023 durchgeführt und war gezielt an Fachpersonen adressiert, die in den Projekten *int:net* und *Redispatch 3.0* sowie im Energiebereich von *OFFIS e.V.* involviert waren. Die Verteilung erfolgte über spezifische E-Mail-Verteiler und es wurde eine Teilnahme in deutscher und englischer Sprache ermöglicht. Es wurde bewusst auf Anreize wie Gutscheine verzichtet, um die Freiwilligkeit und Objektivität der Ergebnisse sicherzustellen. Die Daten der Online-Umfrage liegen im Online-Anhang vor<sup>2</sup>. [KNSS24, KNN<sup>+</sup>24]

Die Anzahl der Umfrageteilnahmen und die Filterung sind in Abbildung 5.2 als Flussdiagramm dargestellt.  $n_{invite} = 215$  Personen wurden eingeladen, basierend auf Schätzungen aus den E-Mail-Listen von *OFFIS e.V.* (74), dem *int:net*-Projekt (59) und dem *Redispatch 3.0*-Projekt (82). Davon nahmen  $n_{part} = 48$  an der Umfrage teil. Aufgrund des Fokus der Umfrage auf Erfahrung in der Erstellung von IEC 62559 Use Case-Beschreibungen und/oder SGAM Modellen qualifizierten sich  $n_{UC} = 21$  für den IEC 62559-UC-Teil und  $n_{SGAM} = 20$  für den SGAM-Teil. Alle der qualifizierten Teilnehmenden brachten verwendbare Resultate ein und wurden in der Analyse berücksichtigt. [KNSS24, KNN<sup>+</sup>24]

<sup>2</sup> vgl. Ordner „Online-Umfrage“ im Online-Anhang.



Quelle: Übersetzt aus [KNN<sup>+</sup>24]

Abbildung 5.2: Flussdiagramm der Durchführung der Studienteilnehmenden

### 5.1.3 Datenaufbereitung- und Analyse

Die vorliegende Analyse kombiniert qualitative Inhaltsanalyse mit statistischen Methoden, um ein umfassendes Bild der erhobenen Daten zu gewinnen. Dabei wurde insbesondere auf die systematische Erfassung und Kategorisierung der offenen Antworten Wert gelegt, während geschlossene Fragen mittels quantitativer Methoden ausgewertet wurden. In der Stichprobe wurden sowohl Fach- als auch Methodenexpertise aus unterschiedlichen Wissensständen und Hintergründen miteinbezogen (vgl. Abschnitt 5.1.3.1). Die qualitativen Daten umfassen den Kontext der Use Case-Methodik sowie des SGAM Frameworks (vgl. Abschnitt 5.1.3.2) und die von den Umfrageteilnehmenden erhobenen (funktionalen) Anforderungen. Die quantitativen Daten beziehen sich insbesondere auf die Resonanzwerte der funktionalen Anforderungen, die aus der Literatur abgeleitet wurden (vgl. Abschnitt 5.1.3.3).

#### 5.1.3.1 Beschreibung der Stichproben

An der Studie nahmen 48 Personen teil, von denen 39 die Umfrage vollständig ausgefüllt haben (vgl. Abbildung 5.2). Die Umfrage wurde sowohl auf Deutsch als auch auf Englisch angeboten. Die *Abschnitte A und B (Allgemeine Profilvereinerungen zur Er-*

stellung von IEC 62559 Use Cases und SGAM Modellen) erfassen die Erfahrungswerte der Teilnehmenden im Umgang mit den entsprechenden Methodiken. Diese sind in Tabelle 5.1 aufgeführt. Um an den inhaltlichen Sektionen teilnehmen zu können, mussten die Teilnehmenden bereits Erfahrung in der Erstellung von IEC 62559 Use Cases sowie SGAM Modellen nachweisen. Daher war der Zugriff auf diese Bereiche nur nach entsprechender Angabe der relevanten Informationen möglich. Für den Use Case-Teil qualifizierten sich 21 Personen, für den SGAM-Teil 20 Personen. Eine große Schnittmenge von 20 Teilnehmenden hatte Zugriff auf beide Bereiche. Am Ende der Umfrage konnten sie im *Abschnitt G (Abschluss der Umfrage und Erfassung von Metadaten)* ihre demografischen Daten angeben, die in Tabelle 5.2 dargestellt sind.

Die Ergebnisse aus Tabelle 5.1 zeigen, dass die Mehrheit der Teilnehmenden bereits Erfahrung mit IEC 62559 Use Cases und SGAM Modellen hat, jedoch insbesondere SGAM eine höhere Einstiegshürde zu haben scheint. Während sich viele als fortgeschritten einstufen, ist der Anteil an Experten vergleichsweise gering, was darauf hindeutet, dass eine tiefgehende Spezialisierung in diesen Methodiken noch nicht weit verbreitet ist. Zudem zeigt sich, dass die Nutzung meist moderat ist und nur wenige Personen diese intensiv anwenden, was möglicherweise darauf hindeutet, dass ihre Anwendung ein Nebenteil ihrer Tätigkeit ist und sie den Arbeitsalltag stärker auf andere Tätigkeiten fokussieren. Demografisch (vgl. Tabelle 5.2) betrachtet stammen die Teilnehmenden überwiegend aus technischen Disziplinen wie Elektrotechnik und Informatik und verfügen meist über einen Master- oder Promotionsabschluss.

### 5.1.3.2 Ergebnisse aus der Kontexterhebung

Die Ergebnisse der Diskussionen wurden im Rahmen der Publikation von [KNN<sup>+</sup>24] veröffentlicht. Die *qualitative Inhaltsanalyse* stellt eine systematische Methode zur Beschreibung und Analyse von erhobenen Daten dar und ist für die Interpretation von erhobenen qualitativen Daten geeignet [Sch12]. Hierbei stellt gemäß der modernen toolbasierten Unterstützung der Auswertung die Vorgehensweise der „Qualitativen Textanalyse“ nach Udo Kuckartz ein etabliertes Verfahren [KM14] dar. Ein besonderes Augenmerk liegt hierbei auf der Kategorisierung von Informationsblöcken [Sch14].

Zunächst wurden die Daten gesichtet und in strukturierte (z. B. demografische Angaben) sowie unstrukturierte (offene Antworten) Informationen unterteilt. Während geschlossene Fragen deduktiv kategorisiert wurden, erfolgte die Auswertung der offenen Antworten auf Basis einer induktiven Codierung. Dabei wurden wiederkehrende Muster identifiziert und in übergeordnete Kategorien überführt. Die qualitative Inhaltsanalyse folgte dabei den methodischen Prinzipien von Kuckartz [KM14]. Die Textsegmente wurden in kleinere Einheiten zerlegt und mit Codes versehen, wobei neue Codes induktiv ergänzt wurden, falls bestehende Kategorien nicht ausreichten. Diese schrittweise Verfeinerung ermöglichte eine detaillierte Analyse der thematischen Schwerpunkte innerhalb der Antworten. Abbildung 5.3 illustriert diesen Prozess. [KNN<sup>+</sup>24]

Nach Abschluss der Codierung erfolgte eine statistische Auswertung der Häufigkeiten. Für die statistische Analyse und Visualisierung wurden Python-Bibliotheken wie Pandas

Tabelle 5.1: Übersicht über die ungefilterten Metadaten zur Methodennutzung

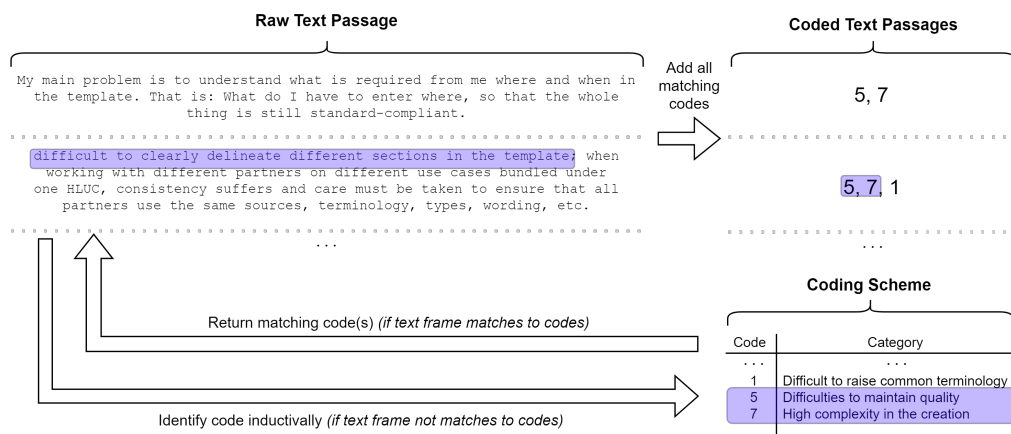
Quelle: [KNN<sup>+</sup>24]

Merkmal	UC: n (%)	SGAM: n (%)
<b>Erfahrung</b>		
Ja, auch IEC 62559 Use Cases / SGAM Modelle erstellt	21 (53,85 %)	20 (54,05 %)
Ja, aber noch keine IEC 62559 UC's / SGAM's erstellt	3 (7,69 %)	9 (24,32 %)
Nein, aber allgemein mit Use Cases vertraut	10 (25,64 %)	-
Nein	5 (12,82 %)	8 (21,62 %)
<b>Erfahrungslevel</b>		
Experte	2 (8,33 %)	3 (10,34 %)
Fortgeschritten	16 (66,67 %)	9 (31,03 %)
Einsteiger	6 (25 %)	17 (58,62 %)
<b>Arbeitsintensität</b>		
Hoch	3 (12,50 %)	2 (6,90 %)
Moderat	14 (58,33 %)	17 (58,62 %)
Gering	7 (29,17 %)	10 (34,48 %)
<b>Rollen</b>		
Methodenorientiert	21 (67,74 %)	18 (51,43 %)
Domänenorientiert	7 (22,58 %)	14 (40 %)
Teil von Standardisierungsgremien	3 (9,68 %)	3 (8,57 %)
<b>Erstellte Typen von IEC 62559 Use Cases</b>		
High-Level Use Case	13 (38,24 %)	-
Low-Level Use Case	5 (14,71 %)	-
Business-Level Use Case	6 (17,65 %)	-
System-Use Case	10 (29,41 %)	-

Tabelle 5.2: Übersicht über die demografischen Informationen

Quelle: [KNN<sup>+</sup>24]

Merkmal	Teilnehmende	Merkmal	Teilnehmende
<b>Alter</b>		<b>Bildungsabschluss</b>	
21 - 29 Jahre	8 (42,11 %)	Promotion	4 (20 %)
30 - 39 Jahre	6 (31,58 %)	Master	14 (70 %)
40 - 49 Jahre	3 (15,79 %)	Diplom	1 (5 %)
50 - 59 Jahre	2 (10,53 %)	Bachelor	1 (5 %)
<b>Bevorzugte Sprache</b>		<b>Studienrichtung</b>	
Deutsch	25 (64 %)	Elektrotechnik	6 (31,59 %)
Englisch	14 (36 %)	Informatik	6 (31,59 %)
		Wirtschaftswissenschaften	3 (15,79 %)
		Wirtschaftsinformatik	2 (10,53 %)
		Physik	1 (5,26 %)
		Chemie	1 (5,26 %)



Quelle: [KNN<sup>+</sup>24]

Abbildung 5.3: Exemplarische Codierung für "Herausforderungen bei der Erstellung von IEC 62559 Use Cases"

und Numpy für die Datenverarbeitung sowie Matplotlib und Seaborn für die grafische Darstellung genutzt [SJ19]. [KNN<sup>+</sup>24]

### Herausforderung mit der IEC 62559 Use Case-Methodik

Die *hohe Komplexität der Erstellung* (29,17 %) ist die größte identifizierte Herausforderung. Diese Komplexität ergibt sich aus der linearen Natur des Entwicklungsprozesses, der oft ein wiederholtes Überarbeiten erfordert. Ein fortgeschrittener Nutzender betonte, dass die mangelnde Klarheit in technischen Details und die Wahl des richtigen Abstraktionsniveaus zusätzliche Hindernisse darstellen. Die Erstellung von Diagrammen wurde als besonders herausfordernd beschrieben, wobei ein Mangel an geeigneten Werkzeugen festgestellt wurde. [KNN<sup>+</sup>24]

Die *Schwierigkeit, die Qualität aufrechtzuerhalten* (20,83 %) kann zu Missverständnissen im System-of-Interest führen, was zu Fehlern und höheren Kosten in späteren Projektphasen führt. Ein Methodenexperte wies darauf hin, dass die unklare Abgrenzung verschiedener Abschnitte in dem Template die Konsistenz beeinträchtigt. Die Beteiligung mehrerer Partner erhöht die Komplexität, da unterschiedliche Terminologien sowie Assets verwendet werden. [KNN<sup>+</sup>24]

Die *Schwierigkeit, eine einheitliche Terminologie zu etablieren* (16,67 %) wurde identifiziert. Diese Herausforderung ergibt sich aus der Vielfalt der Hintergründe und Perspektiven der beteiligten Experten. Ein Umfrageteilnehmer betonte, dass die Zusammenarbeit mit verschiedenen Partnern die Konsistenz gefährdet, wenn nicht sichergestellt wird, dass alle Partner dieselben Quellen und Terminologien verwenden. [KNN<sup>+</sup>24]

Weitere Herausforderungen umfassen die *mangelnde Verfügbarkeit von Informatio-*

nen (12,5 %), die Heterogenität des Vorwissens (8,33 %), die zeitaufwändige Erstellung (8,33 %) und die notwendige Verbesserung der Aktualisierungsprozesse (4,17 %). [KNN<sup>+</sup>24]

### Toolchain und Herausforderungen der IEC 62559 Use Case-Methodik

Die Analyse der verwendeten Werkzeuge im Kontext der IEC 62559 Use Case Methodik zeigt, dass *Textverarbeitungsprogramme* (94,4 %) und *Use Case Management Repositories* (72,2 %) am häufigsten genutzt werden. Während Textverarbeitungswerkzeuge (via dem IEC 62559-2 Word-Template) aufgrund ihrer Zugänglichkeit weit verbreitet sind, wird das UCMR insbesondere für die Speicherung und Verwaltung von Use Cases verwendet. [KNN<sup>+</sup>24]

Es wurde untersucht, warum das UCMR für die direkte Erstellung von IEC 62559 Use Cases nicht immer bevorzugt wird. Das *UCMR wurde von 50 % der Nutzenden als „noch unreif“* beschrieben, wobei technische Mängel wie fehlende Versionskontrolle, Instabilität und unzureichende Unterstützung bei der Dateneingabe kritisiert wurden. *30 % der Teilnehmenden* gaben an, keinen Zugang zum UCMR zu haben. Zudem empfanden *20 % der Nutzenden das Word-Template als benutzerfreundlicher*, besonders für die Zusammenarbeit mit Partnern. [KNN<sup>+</sup>24]

Weitere Herausforderungen betreffen die *mangelnde Benutzerfreundlichkeit des Word-Templates* (13,04 %). Das Template ist verständlich, aber oft fehlen klare Anleitungen und Beispiele für das Ausfüllen bestimmter Felder. Ein weiteres Problem ist der *hohe Aufwand beim Austausch von Word-Dokumenten* (8,70 %), da dezentral gespeicherte Dokumente die Konsistenzprüfung erschweren. Ein zentrales Problem ist auch die *mangelnde Akzeptanz des UCMR durch Projektpartner* (13,04 %). Ein Experte kommentierte, dass dedizierte Werkzeuge oft auf Widerstand stoßen, da Partner den zusätzlichen Aufwand scheuen. Dies kann die Zusammenarbeit und die Qualität der Use Cases beeinträchtigen. Zudem wurde die *mangelnde Einhaltung des Standards* (13,04 %) als Herausforderung genannt, da die Flexibilität der Word-Vorlage zu Inkonsistenzen führen kann. [KNN<sup>+</sup>24]

### Herausforderung mit dem SGAM Framework

Die Erstellung von SGAM Modellen ist mit einer Reihe von Herausforderungen verbunden, die sich aus der Komplexität des Modellierungsprozesses ergeben. Die Umfrageergebnisse zeigen, dass die *präzise Klassifizierung von Elementen in SGAM-Domänen und -Zonen* (30,77 %) die größte Herausforderung darstellt. Es fehlt häufig die notwendige technische Terminologie oder spezifische Informationen, um eine eindeutige Klassifizierung vorzunehmen. [KNN<sup>+</sup>24]

Ein weiteres zentrales Problem ist die *mangelnde Verfügbarkeit von Referenzinformationen* (23,08 %). Ein Umfrageteilnehmer mit regelmäßiger SGAM-Erfahrung betonte, dass die Erstellung eines konformen SGAM Modells schwierig ist, wenn externe Use Cases als Referenz dienen, jedoch dort wichtige Informationen fehlen. Dies führt

zu Lücken im Modell und erschwert die korrekte Darstellung der Systemarchitektur. Die *Schwierigkeit, eine einheitliche Terminologie zu etablieren* (15,38 %) wurde ebenfalls identifiziert. Ein Experte hob hervor, dass die Verwendung einheitlicher technischer Begriffe und Akteure innerhalb des SGAM Frameworks entscheidend ist, insbesondere in Projekten mit mehreren Partnern. Die mangelnde gemeinsame Sprache oder das fehlende Verständnis für den Informationsaustausch zwischen den Partnern kann zu Verzögerungen führen und erfordert Nacharbeiten. [KNN<sup>+</sup>24]

Weitere Herausforderungen umfassen die *Schwierigkeiten bei der korrekten Darstellung* (7,69 %), bei der Nutzende berichten, dass die Anzeigefelder in den Werkzeugen oft zu klein sind und manuell angepasst werden müssen. Die *fehlende Anleitung bei der Erstellung* (7,69 %) wurde ebenfalls als Problem genannt, da Nutzende oft Unterstützung oder Absprachen benötigen, um Entscheidungen während des Modellierungsprozesses besser zu begründen. Darüber hinaus wurde die *mangelnde Verständlichkeit des SGAM Frameworks* (7,69 %) kritisiert, da die Einarbeitung in das Framework oft zeitaufwendig ist. Schließlich wurde der *Mangel an Fachwissen* (7,69 %) als Herausforderung identifiziert, wobei ein Nutzender kommentierte: „Ein Narr mit einem Werkzeug bleibt ein Narr“, was darauf hinweist, dass viele Probleme auf mangelndes Know-how zurückzuführen sind und nicht auf die SGAM-Methodik selbst. [KNN<sup>+</sup>24]

### Toolchain und Herausforderungen der SGAM Framework

Zur Erstellung von SGAM Modellen kommen verschiedene Werkzeuge zum Einsatz. Microsoft PowerPoint ist mit 89,5 % das am häufigsten genutzte Tool zur Erstellung von SGAM Modellen, da es mit den bestehenden Templates eine Flexibilität und einfache Verfügbarkeit bietet. Das UCMR wird von 26,3 % der Nutzenden eingesetzt, um auf die zugehörigen IEC 62559 Use Cases zurückzugreifen. Zudem greifen 26,3 % der Nutzenden auf Visualisierungstools wie den SGAM Visualizer zurück. Generative KI-Tools wie ChatGPT wurden mit einer Nutzungsrate von 5,53 % hingegen nur selten verwendet (Zeitpunkt der Umfrage war Mitte 2023). Weitere unterstützende Werkzeuge sind die SGAM Toolbox, Online-Whiteboardtools sowie Referenzmaterialien wie bestehende SGAM Modelle und Fachliteratur. Erwähnt wurde ebenfalls das DISCERN Microsoft Visio SGAM-Template. [KNN<sup>+</sup>24]

Allerdings bringt die Verwendung dieser Werkzeuge auch spezifische Herausforderungen mit sich. Die *Darstellung komplexer SGAM Modelle* (28,57 %) ist für einige Modellierende problematisch, insbesondere wenn mehrere Layer überlagert werden müssen, was die Lesbarkeit erschwert. Zudem erfordert die Erstellung korrekter SGAM Modelle *ein hohes Maß an Fachwissen* (21,43 %), da die Zuordnung von Objekten zu den richtigen Layern, Domänen und Zonen korrekt erfolgen sollte. Ein weiteres Problem ist das *Fehlen von Informationen in den Use Case-Beschreibungen* (21,43 %), was die Modellierung erschwert. Herausforderungen bestehen auch in der Zusammenarbeit: *Fehlende Versionierungsmöglichkeiten* (7,14 %) und *eingeschränkte parallele Bearbeitungsmöglichkeiten* (14,29 %) führen zu Schwierigkeiten bei der Nachverfolgbarkeit von Änderungen und erschweren Abstimmungen. Zusätzlich kann *die Freiheit der PPTX-*

*SGAM-Vorlagen* (7.14 %) dazu führen, dass nicht syntaktisch korrekte Modelle entstehen, ähnlich wie es beim Word-Template des IEC 62559 Use Case-Templates der Fall ist. [KNN<sup>+</sup>24]

### 5.1.3.3 Ergebnisse aus der Anforderungserhebung

Die Ergebnisse dieses Abschnitts wurden im Rahmen der Publikation von [KNSS24] veröffentlicht. Ein Ziel der Umfrage war es, funktionale Anforderungen, die aus den Anforderungen an IEC 62559 Use Cases sowie SGAM Modellen abgeleitet wurden, von den Modellierenden bewerten zu lassen und zugleich weitere funktionale Anforderungen zu erheben. Die vorgeschlagenen funktionalen Anforderungen basieren auf den in Tabelle 6.5 (veröffentlicht in [KSS23]) definierten Qualitätsdimensionen, die Korrektheit, Konsistenz, Relevanz, Verständlichkeit, Eindeutigkeit und Erreichbarkeit umfassen. Die funktionalen Anforderungen wurden so formuliert, dass sie direkte Auswirkungen auf die genannten Qualitätsmerkmale haben. Darüber hinaus dienen die Qualitätsmerkmale als Grundlage für die Bewertung des implementierten Systems. In der Arbeit zur Herleitung der Qualitätsmerkmale wurden Anforderungen an IEC 62559 Use Cases (vgl. Tabelle 6.2), SGAM Modelle (vgl. Tabelle 6.3) sowie allgemeine Anforderungen (vgl. Tabelle 6.4) identifiziert, die in die Auswahl der funktionalen Anforderungen einfließen. Die Selektion der abgeleiteten funktionalen Anforderungen ist in Tabelle 5.3 aufgeführt. In dieser Tabelle enthalten sind zudem die Referenzen zu den Quellen (in dem Kontext die IEC 62559 Use Cases, SGAM Modell sowie allgemeinen Anforderungen).

Diese funktionalen Anforderungen wurden im Rahmen der Umfrage 5.3 von den Umfrageteilnehmenden vorgelegt, welche eine Bewertung der möglichen Auswirkung auf der Modellqualität tätigten. Hierfür wurde eine 5-stufige Likert-Scale verwendet, welche die Werte von 1 (Sehr irrelevant) bis 5 (Sehr relevant) berücksichtigt. Zudem konnten die Teilnehmenden individuelle Anmerkungen zu den Anforderungen bereitstellen. Die verwendete Likert-Scale folgt folglich einer ordinalen Struktur. Für jede zu bewertende funktionale Anforderung wurde ein Mittelwert bestimmt, welcher die *Resonanz* angibt. Trotz der ordinalen Natur der Likert-Skala wurde der Mittelwert berechnet, um eine zusammenfassende quantitative Einschätzung der Resonanz einer funktionalen Anforderung approximiert zu ermöglichen. Diese Entscheidung wurde getroffen, da der Mittelwert in diesem Kontext als nützliches Instrument zur Ermittlung von zentralen Tendenzen dient. Ergänzend wurden sowohl der Median als auch der Mittelwert im Boxplot (vgl. Abbildung 5.4) visualisiert, um die Verteilung und Streuung der Bewertungen darzustellen und die Variabilität sowie potenzielle Ausreißer der Reaktionen auf die funktionale Anforderung zu verdeutlichen.

Die erhobenen Daten galt es im Hinblick auf die Filterung der funktionalen Anforderungen zu interpretieren. Diese Filterung dient nicht der Selektion für das SGAAIRE-System, sondern verfolgt den Zweck, zunächst jene Anforderungen zu eliminieren, die von den befragten Modellierenden als weniger relevant für die qualitativen Auswirkungen eingeschätzt wurden. Hierzu wurden die relevanten funktionalen Anforderungen über Hypothesentesting gefiltert. Es wurde die alternative Hypothese  $H_1^{RQ_i}$  so formu-

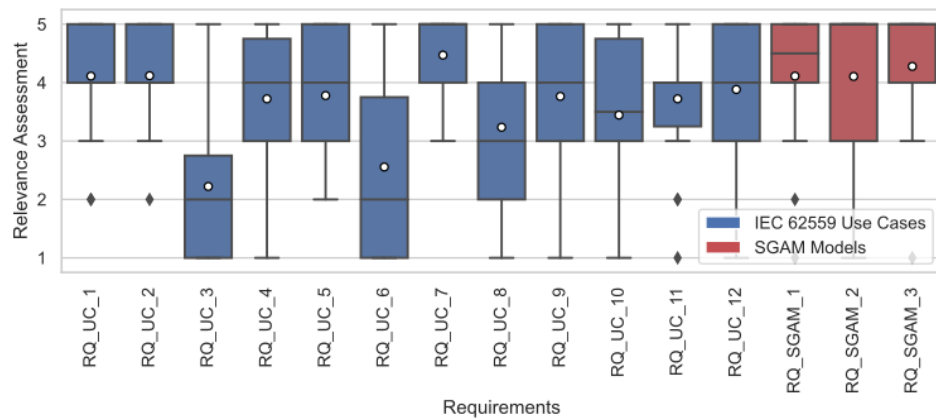
Tabelle 5.3: Übersicht aus den funktionalen Anforderungen, hergeleitet aus den Modellanforderungen (vgl. Abschnitt 6.2)

Quelle: Angelehnt an [KNSS24]

Anf.-ID.	Name	Resonanz	$H_1^{RQ^2}$
UC.1	Das System soll vorschlagen, Elemente (z. B. Akteure, Information Exchanges, Requirements, KPI, Keywords usw.) hinzuzufügen, zu entfernen oder zu ändern. <i>(Hergeleitet aus M_UC.6,7,8, M_G.3,6)</i>	4.11	Akzeptiert
UC.2	Das System soll in freien Textfeldern abweichende und nicht standardisierte Terme feststellen. <i>(Hergeleitet aus M_UC.12, M_G.1,2)</i>	4.11	Akzeptiert
UC.3	Das System soll den Titel von Use Cases validieren. <i>(Hergeleitet aus M_UC.2, M_G.6)</i>	2.22	Abgelehnt
UC.4	Das System soll Use Cases in SGAM-Zonen und Domänen einordnen. <i>(Hergeleitet aus M_UC.3, M_G.3,4,6)</i>	3.72	Akzeptiert
UC.5	Das System soll überprüfen, dass alle definierten Akteure in der vollständigen Beschreibung aufgeführt sind. <i>(Hergeleitet aus M_UC.15, M_G.3)</i>	3.77	Akzeptiert
UC.6	Das System soll den Level of Detail vom Use Case (z. B. High Level, Low Level, ...) vorschlagen. <i>(Hergeleitet aus M_UC.1, M_G.4,6)</i>	2.55	Abgelehnt
UC.7	Das System soll auf Inkonsistenzen bei Akteuren, Information Exchanges und Requirements hinweisen, die nicht dem Use Cases-Typ entsprechen. <i>(Hergeleitet aus M_UC.10, M_G.1,2,6)</i>	4.47	Akzeptiert
UC.8	Das System soll überprüfen, ob die Diagramme von dem Use Case Interaktionen enthalten. <i>(Hergeleitet aus M_UC.16, M_G.3,5)</i>	3.24	Abgelehnt
UC.9	Das System soll für Akteure die Akteurstypen durch Akteursnamen- und Beschreibung vorschlagen. <i>(Hergeleitet aus M_UC.3, M_G.3,4,6)</i>	3.76	Akzeptiert
UC.10	Das System soll Akteure in SGAM-Zonen und Domänen klassifizieren und diese im Further Information-Feld aufführen. <i>(Hergeleitet aus M_UC.9, M_G.1)</i>	3.44	Abgelehnt
UC.11	Das System soll Scenario Services in der Step-by-Step Analyse vorschlagen. <i>(Hergeleitet aus M_UC.10, M_G.3,6)</i>	3.88	Akzeptiert
UC.12	Das System soll mögliche Akteurguppen zur Vereinfachung vorschlagen. <i>(Hergeleitet aus M_UC.13, M_G.3,6)</i>	3.72	Akzeptiert
SGAM.1	Das System soll Arbeitsvorschläge für SGAM Modelle anhand von Use Cases Templates generieren. <i>(Hergeleitet aus allen SGAM und allgemeinen Anforderungen)</i>	4.11	Akzeptiert
SGAM.2	Das System soll SGAM Modelle anhand eines Use Cases prüfen auf Vollständigkeit, Redundanzen, Eindeutigkeiten, Korrektheit oder Verständlichkeit. <i>(Hergeleitet aus allen SGAM und allgemeinen Anforderungen)</i>	4.10	Akzeptiert
SGAM.3	Das System soll Wahrscheinlichkeiten für Inkonsistenzen innerhalb von SGAM Modellen erkennen und aufzeigen. <i>(Hergeleitet aus allen SGAM und allgemeinen Anforderungen)</i>	4.27	Akzeptiert

liert, dass eine vorgeschlagene funktionale Anforderung  $RQ_i$  einen positiven Einfluss auf die Qualitätsdimensionen gemäß der Benutzerbasis hat. Im konkreten wurden Hypothesentests unter Verwendung des einseitigen Einstichproben t-Tests (mit  $\mu = 3$  und einem Signifikanzniveau von  $\alpha = 0.05$ ) verwendet. Wenn das Ergebnis des statistischen Tests einen p-Wert ergibt, der kleiner oder gleich  $\alpha$  ist, wird die alternative Hypothese ( $H_1^{RQ_i}$ ) akzeptiert. Ansonsten, wenn der p-Wert  $\alpha$  überschreitet, wird die Nullhypothese ( $H_0^{RQ_i}$ ) akzeptiert und somit  $H_1^{RQ_i}$  abgelehnt. Die Ergebnisse der Hypothesentests sind ebenfalls in Tabelle 5.3 enthalten. [KNSS24]

Die im Rahmen der Umfrage identifizierten funktionalen Anforderungen ( $n_{UC} = 3$  und  $n_{SGAM} = 3$ , vgl. Tabelle 5.4) sowie die daraus abgeleiteten Perspektiven und Kommentare bieten einen Einblick in den Wünsche der Modellierenden.



Quelle: [KNSS24]

Abbildung 5.4: Boxplot der Resonanzwerte der erhobenen funktionalen Anforderungen für IEC 62559 Use Cases (n=18) und SGAM Modelle (n=19)

Für IEC 62559 Use Cases wurde von einem Teilnehmenden die Bedeutung einer Vereinheitlichung von Akteuren, Begriffen und Abhängigkeiten über mehrere Use Cases hinweg betont (UC\_X1). Diese Anforderung lässt sich mit den bestehenden funktionalen Anforderungen RQ\_UC\_1, RQ\_UC\_2 und RQ\_UC\_7 in Verbindung bringen. Das UCMR als Werkzeug bietet hierfür bereits in der Theorie eine Grundlage, da es Akteure, Anforderungen und Begriffe als Assets betrachtet, die über mehrere Use Cases hinweg genutzt werden können [GDU17]. In der Praxis zeigt sich jedoch, dass die Konsistenz dieser Assets nicht immer gewährleistet ist [GS15], was auf ein Verbesserungspotenzial hinweist. Darüber hinaus wurde der Wunsch geäußert, dass das System in der Lage sein soll, Synergien zwischen anderen Use Cases sowie Projekten zu identifizieren (UC\_X2), um die Wiederverwendung von Use Cases zu fördern und Redundanzen zu vermeiden. Ein weiterer Aspekt ist die Generierung von Sequenzdiagrammen aus der Step-by-Step-Analyse von IEC 62559 Use Cases (UC\_X3). Diese Anforderung unterstreicht die Notwendigkeit, die textuellen Beschreibungen von Use Cases in formalisierte, visuelle Darstellungen zu überführen, um die Nachvollziehbarkeit und Analysefähigkeit zu erhöhen. [KNSS24]

Auf der SGAM-Seite wurde der Bedarf an Klassifizierungsvorschlägen für Komponenten in Zonen und Domänen sowie an Protokollvorschlägen für die Kommunikationsschicht artikuliert (SGAM\_X1). Darüber hinaus wurde die Notwendigkeit hervorgehoben, die erforderlichen Systeme, Akteure und Abhängigkeiten aus der Beschreibung der IEC 62559 Use Cases zu identifizieren (SGAM\_X2), um diese manuell in das Modell zu integrieren. Beide Aspekte repräsentieren signifikante Teilmengen der funktionalen Anforderung RQ\_SGAM\_1. Zudem wurde der Vorschlag unterbreitet, dass das KI-System in der Lage sein sollte, Beziehungen zwischen den Interoperabilitätsschichten zu verstehen, zu überprüfen oder abzuleiten, um Protokolle oder weiterführende Anforderungen

für die zu implementierenden Systeme zu formulieren. Diese Fähigkeit spiegelt sich in sämtlichen erhobenen SGAM-bezogenen funktionalen Anforderungen wider. Ebenfalls wurde von einem Teilnehmenden die funktionale Anforderung formuliert, die Ausrichtung der Sequenzdiagramme der IEC 62559 Use Cases mit denen des zugehörigen SGAM Modells zu überprüfen, um sicherzustellen, dass die korrekten Systeme miteinander kommunizieren. Diese Anforderung stellt einen Teilbereich der funktionalen Anforderung RQ\_SGAM.2 dar. Alle von den Teilnehmenden erhobenen funktionalen SGAM Anforderungen werden als Teilmenge der aus der Tabelle 5.3 funktionalen Anforderungen betrachtet. [KNSS24]

Tabelle 5.4: Übersicht der von den Umfrageteilnehmenden vorgeschlagenen funktionalen Anforderungen

Quelle: Angelehnt an [KNSS24]

Anf.-ID.	Name
UC_X1	Das System soll eine Vereinheitlichung von Akteuren, Begriffen und Abhängigkeiten über mehrere IEC 62559 Use Cases hinweg sicherstellen. ( <i>Teilmenge von UC_1,2,7</i> )
UC_X2	Das System soll Synergien zwischen verschiedenen Projekten bei einer ähnlichen IEC 62559 Use Case-Beschreibung identifizieren.
UC_X3	Das System soll Sequenzdiagramme anhand der Step-by-Step Analyse von IEC 62559 Use Cases generieren.
SGAM_X1	Das System soll Klassifizierungsvorschläge für Komponenten in Zonen und Domänen sowie Protokollvorschläge für die Kommunikationsschicht bereitstellen. ( <i>Teilmenge von SGAM_2-3</i> )
SGAM_X2	Das System soll die notwendigen Systeme, Akteure und Abhängigkeiten aus der Beschreibung der IEC 62559-Anwendungsfälle identifizieren. ( <i>Teilmenge von SGAM_1-3</i> )
SGAM_X3	Das System soll anhand der Sequenzbeschreibungen von einem IEC 62559 Use Case das SGAM Modell herleiten. ( <i>Teilmenge von SGAM_1</i> )

Alle in Tabelle 5.3 aufgeführten funktionalen Anforderungen, für die die Alternativhypothese  $H_1^{RQ^1}$  akzeptiert wurde, sowie die funktionalen Anforderungen, die unter Berücksichtigung der über die Literatur erhobenen Anforderungen ermittelt wurden, werden im Rahmen der Nutzwertanalyse als Alternativen integriert. Hierbei wurden auch die in Tabelle 5.4 erhobenen Anforderungen berücksichtigt, die keine Teilmenge von bestehenden Anforderungen darstellen. Diese Integration wird im Abschnitt 5.3 näher erläutert sowie die Selektion der funktionalen Anforderungen für SGAAIRE.

## 5.2 Nicht-funktionale Anforderungen

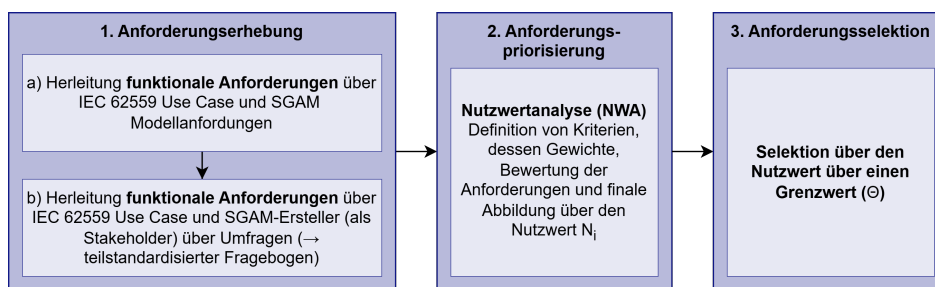
Die Entwicklung von SGAAIRE basiert auf einer Reihe nicht-funktionaler Anforderungen (Nicht-funktionale Anforderungen (NFA)), die sich aus den übergeordneten Forschungszielen und der Kontexterhebung ableiten. Diese Anforderungen sind entscheidend für die Qualität und Akzeptanz der entwickelten Lösung, da sie grundlegende Eigenschaften wie Integrationsfähigkeit, Wartbarkeit, Standardkonformität und Benutzerfreundlichkeit sicherstellen.

Ein zentrales Ziel von SGAAIRE ist die Verbesserung der Qualität von IEC 62559 Use Case-Beschreibungen und SGAM Modellen durch den Einsatz von KI. Dabei soll das System in bestehende Werkzeuge und Arbeitsabläufe integriert werden können, ohne deren Funktionalität einzuschränken oder bestehende Prozesse zu beeinträchtigen. Gleichzeitig muss SGAAIRE als eigenständige Softwarelösung wartbar und langfristig nutzbar bleiben. Die Berücksichtigung standardisierter Datenformate gewährleistet die Interoperabilität mit bestehenden Systemen, während der Fokus auf eine benutzerzentrierte Entscheidungsfreiheit sicherstellt, dass menschliche Fachexperten jederzeit die Kontrolle über die von SGAAIRE generierten Vorschläge behalten. Die folgenden nicht-funktionalen Anforderungen definieren diese grundlegenden Eigenschaften im Detail:

- **[REQ\_NFA\_1] Qualitätsverbesserung durch KI:** Der zentrale Fokus der Untersuchung im Rahmen des Forschungsvorhabens liegt auf der Steigerung der Qualität von IEC 62559 Use Case-Beschreibungen und SGAM Modellen mit Hilfe von KI. Dieses Ziel ist mit den vorliegenden NFAs verknüpft. Der Einsatz von SGAAIRE soll die Theorie stützen, dass KI-Systeme zur Verbesserung der Qualität beitragen können.
- **[REQ\_NFA\_2] Lose-gekoppelte Integration in UCMR:** Das in Abschnitt 3.2.3 aufgeführte UC(M)R fungiert einerseits als gemäß IEC 62559-1-Standard [IEC19] definiertes Instrument zur IEC 62559 Use Case-Verwaltung und andererseits als Teil mehrerer Forschungs- und Innovationsprojekte, wodurch es breite Anerkennung bei der Erstellung von Anwendungsfallbeschreibungen findet, was in der Kontexterhebung im Abschnitt 5.1.3.2 verdeutlicht ist. Das OFFIS entwickelt an einer eigenen UCMR-Instanz für Forschungsprojekte wie *Redispatch 3.0* und *Intelligente Digitalisierung der Energieversorgung zur Optimierung des Netzbetriebs und zur Erhöhung der Akzeptanz (IDiNA)*. Die lose-gekoppelte Integration in einem UCMR zielt darauf ab, einerseits die Akzeptanz von SGAAIRE zu fördern und andererseits die direkte Verwendung des UCMR zur kollaborativen Erstellung zu verbessern.
- **[REQ\_NFA\_3] Standalone Software System:** Mit dem Ziel der Wartbarkeit soll SGAAIRE als autarkes System konstruiert werden, das von anderen Systemen und Prozessen über eine externe Schnittstelle genutzt werden kann.
- **[REQ\_NFA\_4] Standardisierte Serialisierung:** Die Verwendung des IEC 62559-3 XML-Serialisierungsstandards für IEC 62559 Use Case-Beschreibungen [GDU17] und dem DISCERN SGAM Datenformat [GDU17] stellen anerkannte Formate dar.
- **[REQ\_NFA\_5] Benutzerzentrierte Entscheidungsfreiheit:** Das System agiert, indem es dem Nutzer Vorschläge präsentiert, die er akzeptieren oder ablehnen kann. Dadurch soll dem Benutzer die uneingeschränkte Entscheidungsbefugnis über seine IEC 62559 Use Case- und SGAM Modelle gewährt werden.

### 5.3 Funktionale Anforderungen

Die aus der Anforderungserhebung (vgl. Abschnitt 5.1) entstandenen funktionalen Anforderungen gilt es im Rahmen des Dissertationsvorhabens zu priorisieren und anschließend auf eine zu realisierende Menge für die Entwicklung des SGAAIRE-Systems zu selektieren. Hierfür gilt es Verfahren aus dem Forschungsfeld der Entscheidungsfindung anzuwenden. Das Rahmenwerk für die Anforderungspriorisierung bildet die Nutzwertanalyse (NWA). Diese sieht eine Definition von relevanten Kriterien und die Bestimmung dessen Gewichtung vor, um anschließend die Nutzwerte der Anforderungen zu berechnen. Basierend auf den Nutzwerten werden die Anforderungen über die Nutzung eines Schwellwerts selektiert. Dieser Prozess ist in der Abbildung 5.5 visualisiert.



Quelle: Eigene Darstellung

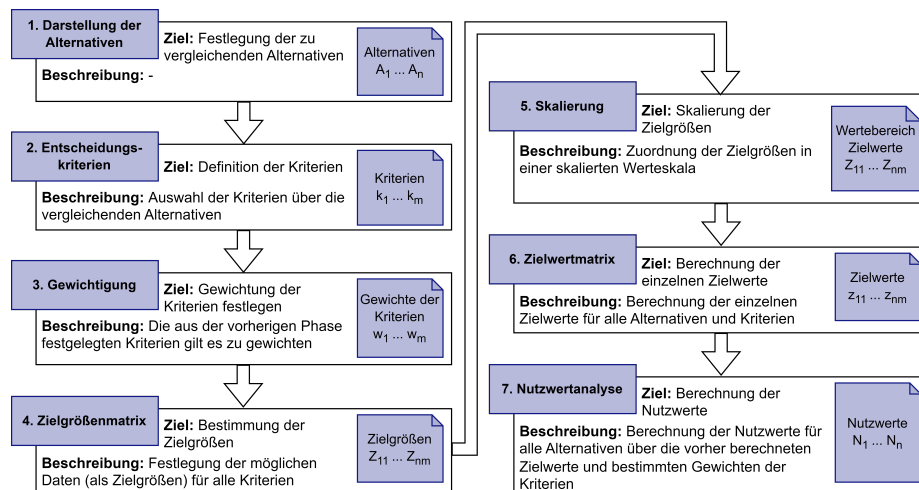
Abbildung 5.5: Prozess zur Herleitung der funktionalen Anforderungen

In den folgenden Unterabschnitten werden sowohl die Schritte der systematischen Anforderungspriorisierung über die NWA aufgezeigt als auch die Anforderungsselektion über einen Schwellwert zur Bestimmung der umzusetzenden funktionalen Anforderungen. Die durchgeführte NWA sowie deren einzelne Teilergebnisse sind im Online-Anhang<sup>3</sup> enthalten.

#### 5.3.1 Anforderungspriorisierung: Nutzwertanalyse (NWA)

Eine Nutzwertanalyse stellt ein methodologisches Konzept für die Lösung von komplexen, mehrdimensionalen Entscheidungsproblemen dar, welches sowohl in der Theorie als auch in der Praxis Anwendung findet [Zan14]. Ein Entscheidungsproblem stellt ein nicht-triviales Problem dar, wenn eine Auswahl aus zwei oder mehreren zu realisierenden Alternativen  $A_i$  mit  $i = 1, \dots, n$  und  $n \geq 2$  notwendig ist [Wit79]. In der Nutzwertanalyse wird das Selektionsproblem gelöst, indem die Nutzwerte  $N$  der Alternativen betrachtet werden. Für jede Alternative  $A_i$  wird der Nutzwert  $N_i$  über die Erfüllung der  $m$  Kriterien im Rahmen einer Präferenzfunktion  $N_i = PR(k_{i1}, \dots, k_{im})$  bestimmt [Wit79].

<sup>3</sup> vgl. Ordner „Nutzwertanalyse“ im Online-Anhang.



Quelle: Eigene Darstellung nach [Reu14]

Abbildung 5.6: Verwendete Schritte der Nutzwertanalyse

Bei der Durchführung einer NWA existiert bislang keine einheitliche Darstellung, jedoch lassen sich wesentliche Phasen identifizieren [Reu14] [Zan14] [Bec78]. Abbildung 5.6 führt sieben Phasen für die Durchführung der vorliegenden Nutzwertanalyse auf. Jede Phase ist mit einem Ziel und mit einem bestimmten Ergebnis als Artefakt verbunden.

In dieser Dissertation wird die Nutzwertanalyse als ein wesentlicher Bestandteil zur Anforderungspriorisierung von funktionalen Anforderungen für die Konzeption und Implementierung des intelligenten Systems betrachtet. Die funktionalen Anforderungen wurden sowohl aus den berücksichtigten (Modell-)Anforderungen von IEC 62559 Use Cases und SGAM Modellen abgeleitet als auch durch die Befragung von Fach- und Methodenexperten (Definition vgl. Kapitel 3) mittels eines *teilstandardisierten Fragebogens* erhoben, die bereits Erfahrung mit der Erstellung von IEC 62559 Use Cases und SGAM Modellen haben. Die identifizierten funktionalen Anforderungen weisen unterschiedliche Charakteristika auf, wie beispielsweise ihre Auswirkungen auf die Qualität, zusätzlichen Mehrwert bei der Systemnutzung, mögliche Umsetzungshürden, Implementierungsdauer und die Relevanz von Künstlicher Intelligenz. Zur Auswahl und Priorisierung der wichtigsten funktionalen Anforderungen wird die systematische NWA als fundierte Methode eingesetzt.

Im weiteren Verlauf des Abschnitts werden für die in Abbildung 5.6 aufgeführten Phasen sowohl notwendige Entscheidungen und dessen Begründungen als auch die Ergebnisse betrachtet.

### 5.3.1.1 Schritt 1: Darstellung der Alternativen

Innerhalb der Nutzwertanalyse werden alle aus Abschnitt 5.1.3.3 gelisteten erhobenen funktionalen Anforderungen miteinbezogen. Diese haben ihren Ursprung durch die Ana-

lyse der Modellanforderungen für IEC 62559 Use Cases und SGAM Modelle und der Ergebnisse des ausgeführten *teilstandardisierten Fragebogen* aus Abschnitt 5.1.

### 5.3.1.2 Schritt 2: Entscheidungskriterien festlegen

Die Auswahl der Entscheidungskriterien ist ein subjektiver Faktor der NWA und gilt nach [Kü14] als kreativer Prozess. Kriterien werden selektiert, welche den Hintergrund von a) der Zielsetzung und b) dem Entscheidungsproblem betreffen [Kü14, DIN20].

Der Themenbereich der *Anforderungspriorisierung* (engl. *Requirement Prioritization*) befasst sich unter anderem mit der Auswahl von relevanten Kriterien [ASIM14] [BA05] [RD15]. In der systematischen Literaturrecherche von [RD15] zur Identifikation von Anforderungskriterien für die Priorisierung konnten aus 83 analysierten Publikationen 760 Kriterien identifiziert werden. Diese wurden in sechs Primärkategorien und in 32 Sekundärkategorien unterteilt. Bei der Selektion der Kriterien werden infolgedessen die folgenden Kategorien betrachtet [RD15]:

- Nutzen
- Kosten
- Risiken
- Sanktionen oder deren Vermeidung
- Geschäftskontext
- Technischer Kontext und Anforderungscharakteristiken

Die Auswahl der Entscheidungskriterien soll auf die Entwicklung eines F&E Prototypen gemäß der Problemstellung und der einhergehenden Forschungsziele in Form eines KI-basierten Systems einzielen. Für die Auswahl der Entscheidungskriterien gilt es wesentliche Annahmen zu treffen, welche den zugrundeliegenden Kategorien zugeordnet werden können:

#### Nutzen

Die zu selektierenden funktionalen Anforderungen sollen zum einen einen positiven Effekt auf die Qualität für IEC 62559 Use Cases und SGAM Modelle gemäß des Forschungsziels erzielen. Zum anderen sieht das Rahmenwerk der Arbeit vor, die Bedürfnisse der Nutzenden zu kennen und zu verstehen, um Anforderungen im Nutzungskontext besser umsetzen zu können [Mag01]. Beide Aspekte gehen auf die Leistung bzw. die Relevanz einer Anforderung ein. Abgeleitet aus den Annahmen entstehen die folgende Kriterien:

**[k1] Auswirkung auf die Qualität** Die Auswirkungen auf die Qualität stellt eine subjektive quantitative Einschätzung der Umfrageteilnehmer aus dem *teilstandardisierten Fragebogen* dar. Dieses Kriterium soll die Auffassungen der Erstellenden von SGAM Modellen bzw. IEC 62559 Use Cases berücksichtigen und dient als ein Indikator für das Forschungsziel der *Forschungsfrage 2* für die Fehlerminimierung.

[k2] **Weitere Mehrwerte/Folgen** Die Umfrageteilnehmer aus dem *teilstandardisierten Fragebogen* konnten bei jeder vorgeschlagenen bzw. den eigenerstellten Anforderungen weitere Mehrwerte/Folgen und Kommentare angeben. Über das Kriterium gilt es das Meinungsbild von Erstellenden von SGAM Modellen bzw. IEC 62559 Use Cases interpretiert zu berücksichtigen.

### Kosten

Kosten bei der Anforderungspriorisierung werden in der Literatur zwischen generischen Typen und Zeitaufwände in der Entwicklung bzw. Wartung unterschieden [RD15]. Für die Arbeit entstanden keine zusätzlichen Kosten für Literatur, Standards oder Daten; der wesentliche Aufwand lag im Faktor „Zeit“, also in der Umsetzungsdauer:

[k3] **Umsetzungsdauer** Die Konzeption und Umsetzung einer Anforderung kann unterschiedliche Dauern aufweisen. Dieses Kriterium soll den eingeschätzten zeitlichen Aufwand darstellen.

### Risiken

In der Literatur wurden verschiedene Arten von Risiken analysiert, darunter technik-, stakeholder-, projekt- und anforderungsstatusbezogene Risiken [RD15]. Begründet durch die existierende Vielzahl an unterschiedlichsten Kriterien werden die wesentlichsten Risiken als Kriterien für die NWA selektiert. Hierbei stellen technische Hürden ein großes Risiko dar, da beispielsweise die vorliegende Datenbasis möglicherweise nicht für die Umsetzung einer Anforderung ausreichen kann. Abgeleitet aus der Annahme entsteht das folgende Kriterium:

[k4] **Realisierbarkeit** Technische Hürden (bspw. fehlende Daten) können den Erfolg der Umsetzung schädigen. Dieses Kriterium soll die eingeschätzte Wahrscheinlichkeit von technischen Hürden und deren Auswirkungen berücksichtigen.

### Sanktionen oder deren Vermeidung

Durch das Realisieren eines Forschungsvorhabens sind keine Kriterien aus der Sanktionen-Kategorie erforderlich.

### Geschäftskontext → Forschungskontext

Der Term „Geschäftskontext“ ist in dem Forschungsvorhaben nicht zielführend, da nicht der geschäftliche [RD15], sondern der wissenschaftliche Kontext im Vordergrund steht. Es werden Entscheidungskriterien benötigt, welche die erhobenen funktionalen Anforderungen im wissenschaftlichen Kontext verifizieren.

Eine Rahmenbedingung im Forschungsvorhaben stellt der Einsatz von KI-Algorithmen dar. Der Einsatz von KI-Algorithmen muss gegenüber konventionellen Techniken bei

den funktionalen Anforderungen über einen deutlichen Mehrwert begründet sein. Abgeleitet aus der Annahme entsteht das folgende Kriterium:

**[k5] Mehrwert von KI-Algorithmen** Aus der gegebenen Problemstellung und den Forschungsfragen heraus sollen insbesondere KI-Algorithmen für die Realisierung der Anforderungen umgesetzt werden. Das Verwenden von KI-Algorithmen (insb. zur Erzeugung von Mehrwerten) soll als Kriterium eingesetzt werden, um die Sinnhaftigkeit des KI-Einsatzes für eine konkrete Anforderung zu berücksichtigen.

#### Technischer Kontext und Anforderungscharakteristiken

Das umzusetzende KI-System soll für das Erreichen der Nutzenden innerhalb eines bestehenden Ökosystems, nämlich in dem UCMR integriert werden. Durch diese aus der Problemstellung gegebene nicht-funktionale Anforderung leitet sich ab, dass das umzusetzende System syntaktisch, semantisch und pragmatisch interoperabel mit dem UCMR gemäß den zugrundeliegenden Anforderungen eines UCMR [GDU17] sein muss. Abgeleitet aus der Annahme entsteht das folgende Kriterium:

**[k6] Interoperabilität mit dem UCMR** Das umzusetzende KI-System wird prototypisch an ein UCMR angebunden. Hieraus resultiert ein Bedarf an syntaktischer, semantischer und pragmatischer Interoperabilität. Funktionale Anforderungen sollen sowohl technisch als auch zweckmäßig daher zum UCMR interoperabel sein.

#### 5.3.1.3 Schritt 3: Gewichtung festlegen

Für die Gewichtung der Kriterien wird der *Analytic Hierarchy Process (AHP)* nach [Saa87] (auch bekannt als Saaty-Methode) genutzt, um die Subjektivität in der Gewichtung zu verringern und die Entscheidungen rational begründen zu können. AHP findet ebenfalls als das gängigste Werkzeug aus dem Themenbereich der Anforderungspriorisierung Anwendung [RD15]. Der Prozess sieht eine Gegenüberstellung der Kriterien vor, die anhand einer Skala miteinander verglichen werden. Die dabei entstehenden Werte (bzw. deren Kehrwerte) werden in einer Matrix dargestellt [Saa87]. Für die Gewichtung wird die Skala aus Tabelle A.39 zur Hand gezogen.

Die Bewertungsmatrix aus Tabelle A.40 verwendet die aus Tabelle A.39 aufgeführte Intensitätsskala für den paarweisen Vergleich der identifizierten Kriterien mit der Zieleigenschaft „Wichtigkeit“. Es wurden für die jeweiligen Kriterien die folgenden Annahmen für die Bewertung getroffen:

**k1: Auswirkung auf die Qualität** Das zentrale Ziel des vorliegenden Forschungsvorhabens besteht darin, eine Verbesserung der Qualität zu erreichen. Für die Beantwortung der Forschungsfrage ist es von Bedeutung, das Feedback der Umfrageteilnehmer (→ k2) angemessen zu berücksichtigen.

- k2: Weitere Mehrwerte/Folgen** Zusätzliche Mehrwerte und Konsequenzen, die sich aus der Umsetzung einer funktionalen Anforderung ergeben, sollen sich ebenfalls auf die Entscheidungsfindung auswirken, jedoch in Bezug auf den Aspekt  $k1$  eine sekundäre Rolle spielen. Die Umsetzungsdauer ( $k3$ ) hat im direkten Vergleich zu  $k1$  eine etwas größere Relevanz, dennoch nicht gleichwertig. Die anderen Kriterien werden in gleicher Weise wie bei  $k1$  berücksichtigt.
- k3: Zeitlicher Aufwand** Im direkten Vergleich zu den anderen Kriterien nimmt der zeitliche Aufwand eine eher untergeordnete Rolle ein.
- k4: Realisierbarkeit** Das Kriterium der Realisierbarkeit im Sinne technischer Hürden nimmt eine etwas höhere Bedeutung ein als der zeitliche Aufwand. Keine der zu betrachtenden Alternativen scheint zum aktuellen Zeitpunkt technisch unlösbar zu sein, weshalb dies die Begründung für die vergleichsweise geringere Gewichtung darstellt.
- k5: Mehrwert von KI-Algorithmen** Der Forschungskontext der Arbeit befasst sich mit KI-Algorithmen und deren Mehrwerten. Eine selektive Auswahl funktionaler Anforderungen, die besser mit konventionellen Techniken realisiert werden können, könnte das Ergebnis der Evaluation verfälschen und die Beantwortung der Forschungsfragen beeinträchtigen. Daher wurde diesem Aspekt im Vergleich zu den anderen Kriterien eine erheblich hohe Wichtigkeit zugeschrieben. Nur im Vergleich zur Integration im UCMR ( $k6$ ) wurde diesem Aspekt eine größere Bedeutung beigemessen, da das UCMR im methodischen Kontext als führendes Tool eine hohe Akzeptanz aufweist (vgl. Abschnitt 5.1.3.2) und die Mehrwerte in das Tool integriert werden sollen.
- k6: Interoperabilität mit dem UCMR** Im Mittelpunkt der Toolchain im methodischen Kontext stellt das UCMR ein weit verbreitetes Tool im energiebezogenen Anforderungsmanagement von Forschungs- und Innovationsprojekten dar [KNN<sup>+</sup>24]. Dadurch zeigt sich eine hohe Akzeptanz bei der Integration möglicher funktionaler Anforderungen, da solche Tools bereits in den Arbeitsprozessen integriert sind. Das am OFFIS Forschungsinstitut entwickelte UCMR wird ebenfalls in Projekten eingesetzt, insbesondere in der Verwaltung von Use Cases. Aufgrund dessen ist die Integrierbarkeit innerhalb des Tools von Relevanz und führt zu einer deutlich höheren Bedeutung im Vergleich zu den anderen Kriterien, da dies einen längerfristigen, praxisnahen Einsatz des Systems ermöglichen kann.

Basierend auf der aufgestellten Bewertungsmatrix werden sowohl die Kriteriengewichte als auch der *Consistency Ratio* ( $CR$ ) berechnet. Der  $CR$  stellt einen im AHP verankerten Index dar, welcher Widersprüche in der Selektion der Intensitätswerte numerisch aufzeigt [Saa87]. Eine Anpassung der Bewertungsmatrix ist nach [Saa87] erforderlich, sofern  $CR > 0.10$  ist. Grundlage beider Aspekte stellt die Berechnung des Eigenvektors  $EV$  dar. Die Elemente im Eigenvektor werden hierbei mit den Gewichten gleichgesetzt.

Der Eigenvektor wird über das Berechnen einer Grenzwertmatrix bestimmt. Hierbei wird die Bewertungsmatrix quadriert, die Reihensumme für alle Aktivitäten gebildet

und diese anschließend normalisiert. Dieser Prozess wird solange wiederholt bis die normalisierte Reihensumme stabil bleibt [MPR17]. Diese normalisierte Reihensumme entspricht dem Eigenvektor und folglich den Gewichten der Kriterien [MPR17]. Gemäß der Ausführung wurde in der vierten Iteration eine Stabilität des Eigenvektors erreicht (vgl. Anhang A.4.2). Somit konnten die Gewichte aus Tabelle 5.5 bestimmt werden.

Tabelle 5.5: Gewichte und Rangordnung der Kriterien

Kriterien	Gewicht $\omega_i$	Rang
[k1] Auswirkungen auf die Qualität	0,1251	3
[k2] Weitere Mehrwerte/Folgen	0,0886	4
[k3] Zeitlicher Aufwand	0,0432	6
[k4] Realisierbarkeit	0,0504	5
[k5] Mehrwert von KI-Algorithmen	0,4303	1
[k6] Interoperabilität mit dem UCMR	0,2624	2

Der CR wird über die Division des Consistence Index (CI) durch den stochastisch erzeugten Random Consistency Index (RI)-Index gemäß der Werteliste (vgl. Tabelle 5.6) nach [Saa87] berechnet. Sei  $\lambda_i$  der Eigenwert sowie  $n$  die Dimension, so können CR und CI über die folgende Formel [Saa87] berechnet werden:

$$\text{CR} = \frac{\text{CI}}{\text{RI}} \quad \text{CI} = \frac{\lambda_{max} - n}{n - 1} \quad \lambda_{max} = \frac{1}{n} \sum \lambda_i \quad (5.1)$$

Tabelle 5.7 zeigt die bestimmten Werte für CR, CI, den max. Eigenwert  $\lambda_{max}$  sowie RI auf. Da die Bewertungsmatrix 6 Dimensionen beinhaltet, wird  $RI = 1.24$  gemäß der Tabelle von 5.6 ausgewählt. Der berechnete CR-Index liegt mit 0,0815 unter dem Schwellwert von [Saa87] von 0,10, weswegen keine Anpassung der Intensitätswerte in der Bewertungsmatrix erforderlich sind.

Tabelle 5.6: Stochastisch von [Saa87] erzeugte RI-Werte mit n=500 zufällig generierte Bewertungsmatrizen

Dimension:	1	2	3	4	5	6	7	8	9	10
RI:	0	0	0,58	0,90	1,12	1,24	1,32	1,41	1,45	1,49

Tabelle 5.7: Bestimmte Werte für CR, CI,  $\lambda_{max}$  und RI auf Basis der vierten Iteration

Variable:	CR	CI	$\lambda_{max}$	RI
Wert:	0,0815	0,1011	6,505	1,24

Ein Problem entsteht durch die fehlende Bewertung der Qualität bei den funktionellen Anforderungen, die von den Umfrageteilnehmern aufgeführt wurden. Um dieses Problem anzugehen, werden für diese eine Einschätzung der  $k1$  Werte getätigt.

Eine theoretische Kausalität ist für eine komplette Objektivität keineswegs vorhanden, aber die Entscheidungsfindung über AHP unterstützt bei der Transparenz und bei der Konsistenz. [BECV08] führte kritisch auf, dass das mathematische Fundament in der Eigenwert-Methode zwar rund ist, aber dennoch die Anwendung nicht alle Intentionen berücksichtigen kann und dies im Sinne von AHP zu Regelbrüchen führt.

#### 5.3.1.4 Schritt 4: Erstellung einer Zielgrößenmatrix

In der Zielgrößenmatrix werden gemäß [Zan14] die Zielgrößen der Alternativen für die einzelnen Kriterien aufgeführt. Dafür werden die Messungen der Kriterien sowie deren Wertebereiche bestimmt. Die möglichen Wertebereiche sowie die Messungsarten der Kriterien sind in Tabelle A.42 aufgeführt.

#### 5.3.1.5 Schritt 5: Bestimmung einer Skalierung

Im Prozess der Zielwertberechnung ist es zunächst erforderlich, alle Zielgrößen numerisch aufzulisten. Darüber hinaus ist eine über alle Kriterien hinweg einheitliche Skalierung dieser Größen von Bedeutung, um eine Verfälschung der Gewichtung zu vermeiden. Zu diesem Zweck wurden die potenziellen Werte gemäß der Reihenfolge in Tabelle A.42 numerisch definiert und anschließend innerhalb eines Intervalls von  $[0, 1] \in \mathbb{R}$  proportional skaliert. Die skalierten Werte finden sich in Tabelle A.43 wieder.

#### 5.3.1.6 Schritt 6: Aufstellung der Zielwertmatrix

Unter Berücksichtigung der Ausprägungen der Alternativen ist es möglich, die Zielwerte anhand der Beschreibungen und skalierten Werte der Zielgrößen (vgl. Tabelle A.42 und A.43) zu ermitteln. Die Zielwertmatrix stellt eine Gegenüberstellung der Zielwerte sämtlicher zu analysierender Alternativen dar, die im Hinblick auf die Berechnung der Nutzwerte von Bedeutung sind. In der Tabelle 5.8 werden die Ausprägungswerte für alle funktionalen Anforderungen aufgeführt. Die durchgestrichenen Anforderungen dienen der besseren Übersicht und stellen die funktionalen Anforderungen dar, die aufgrund des Hypothesentests ausgeschlossen wurden (vgl. Abschnitt 5.1.3.3). Diese ergaben sich durch folgende Herangehensweise:

- **k1:** Mittelwert-Ableitung aus der Umfrage (vgl. Abschnitt 5.1). Für die von den Umfrageteilnehmenden erhobenen funktionalen Anforderungen wurde hier eine Einschätzung vorgenommen.
- **k2:** Rückmeldung der Umfrageteilnehmer zu den Kommentaren (vgl. Abschnitt 5.1). Aufgrund der begrenzten Anzahl von Kommentaren waren die Vorkommnisse gering. Bei fehlenden Kommentaren wurde die Kategorie „Neutral“ zugewiesen.
- **k3:** Erste Abschätzung der zur Umsetzung benötigten Zeitspanne.
- **k4:** Komplexitätsbewertung erfolgte durch die Berücksichtigung der Möglichkeit,

Tabelle 5.8: Ausprägungswerte der Alternativen

Anf.-ID. REQ- <i>j</i>	[ <i>k1</i> ]	[ <i>k2</i> ]	[ <i>k3</i> ]	[ <i>k4</i> ]	[ <i>k5</i> ]	[ <i>k6</i> ]
UC_1	4,11	Eher positiv	Hoch	Hoch	Hoch	Ja
UC_2	4,12	Neutral	Mittel	Gering	Mittel	Ja
<del>UC_3</del>	<del>2,22</del>	<del>Neutral</del>	<del>Gering</del>	<del>Gering</del>	<del>Niedrig</del>	Ja
UC_4	3,72	Eher positiv	Mittel	Hoch	Hoch	Ja
UC_5	3,78	Neutral	Gering	Gering	Niedrig	Ja
<del>UC_6</del>	<del>2,56</del>	<del>Neutral</del>	<del>Mittel</del>	<del>Mittel</del>	<del>Niedrig</del>	Ja
UC_7	4,47	Neutral	Hoch	Mittel	Mittel	Ja
<del>UC_8</del>	<del>3,23</del>	<del>Neutral</del>	<del>Mittel</del>	<del>Mittel</del>	<del>Niedrig</del>	Ja
UC_9	3,76	Eher positiv	Mittel	Hoch	Hoch	Ja
<del>UC_10</del>	<del>3,44</del>	<del>Neutral</del>	<del>Mittel</del>	<del>Mittel</del>	<del>Mittel</del>	Ja
UC_11	3,72	Neutral	Mittel	Gering	Mittel	Ja
UC_12	3,88	Neutral	Mittel	Mittel	Niedrig	Ja
SGAM_1	4,11	Neutral	Hoch	Mittel	Hoch	Ja, aber [...]
SGAM_2	4,11	Neutral	Hoch	Mittel	Hoch	Ja, aber [...]
SGAM_3	4,28	Neutral	Hoch	Mittel	Hoch	Ja, aber [...]
UC_X2	4,00	Neutral	Hoch	Hoch	Mittel	Ja
UC_X3	4,00	Neutral	Niedrig	Mittel	Niedrig	Ja

dass die Anforderung möglicherweise nicht über die Daten- und Wissensgrundlage realisiert werden kann.

- **k5:** Auf Grundlage der Beschreibung und aktueller Techniken wurde der Mehrwert von KI-Algorithmen, insbesondere im Vergleich zu deterministischen Verfahren, eingeschätzt.
- **k6:** Gestützt auf das vom OFFIS entwickelte UCMR wurde eine Einschätzung zur möglichen Integration der Anforderungen vorgenommen.

### 5.3.1.7 Schritt 7: Berechnung der Nutzwerte

Im abschließenden Schritt des Verfahrens erfolgt die Berechnung der (Teil-)Nutzwerte auf Grundlage der Zielwerte. Ein Teilnutzwert  $n_{ij}$  für das Kriterium  $k_i$  der Alternative  $A_j$  ergibt sich durch die Multiplikation der Gewichtung  $\omega_i$  mit dem Zielwert  $z_{ij}$  [Reu14]:

$$n_{ij} = \omega_i \cdot z_{ij} \quad (5.2)$$

Der gesamte Nutzwert  $N_j$  für die Alternative  $A_j$  wird als Summe der Teilnutzwerte über alle Kriterien gebildet [Reu14]:

$$N_j = \sum_{i=1}^K n_{ij} \quad (5.3)$$

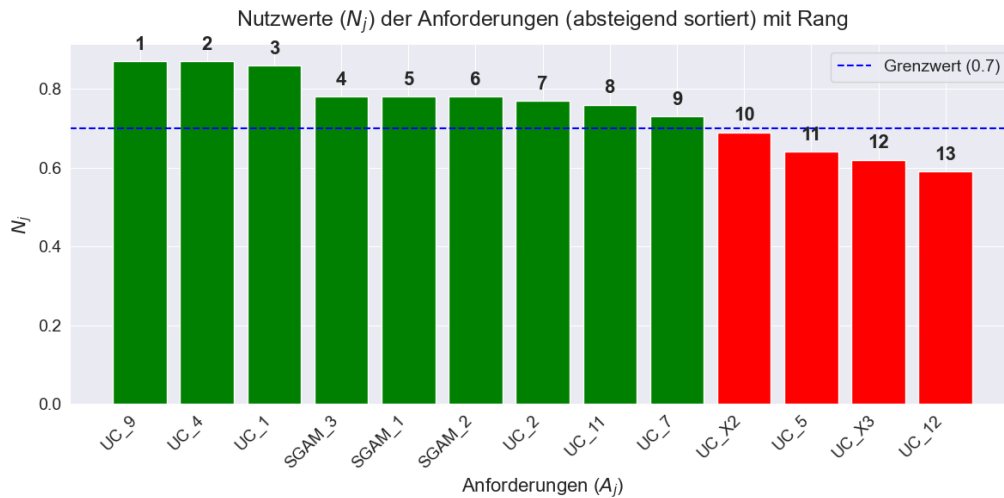
Die in Tabelle 5.9 aufgeführte Matrix veranschaulicht sowohl die erzielten Teilnutzwerte für jede Alternative als auch die endgültig berechneten Gesamtnutzwerte sowie ihre entsprechenden Rangordnungen. Die durchgestrichenen Anforderungen stellen aus Gründen der Übersichtlichkeit die funktionalen Anforderungen dar, die durch den Hypothesentest ausgeschlossen wurden (vgl. Abschnitt 5.1.3.3).

Tabelle 5.9: Aufführung der erzielten (Teil-)Nutzwerte der gegenübergestellten Alternativen

$A_j$	$n_{1j}$	$n_{2j}$	$n_{3j}$	$n_{4j}$	$n_{5j}$	$n_{6j}$	$N_j$	Rang
UC_1	0,10	0,07	0,00	0,00	0,43	0,26	<b>0,86</b>	3
UC_2	0,10	0,04	0,02	0,05	0,29	0,26	<b>0,77</b>	7
<del>UC_3</del>	<del>0,04</del>	<del>0,04</del>	<del>0,04</del>	<del>0,05</del>	<del>0,14</del>	<del>0,26</del>	<del><b>0,60</b></del>	14
UC_4	0,09	0,07	0,02	0,00	0,43	0,26	<b>0,87</b>	2
UC_5	0,09	0,04	0,04	0,05	0,14	0,26	<b>0,64</b>	12
<del>UC_6</del>	<del>0,05</del>	<del>0,04</del>	<del>0,02</del>	<del>0,03</del>	<del>0,14</del>	<del>0,26</del>	<del><b>0,56</b></del>	17
UC_7	0,11	0,04	0,00	0,03	0,29	0,26	<b>0,73</b>	9
<del>UC_8</del>	<del>0,07</del>	<del>0,04</del>	<del>0,02</del>	<del>0,03</del>	<del>0,14</del>	<del>0,26</del>	<del><b>0,58</b></del>	16
UC_9	0,09	0,07	0,02	0,00	0,43	0,26	<b>0,87</b>	1
<del>UC_10</del>	<del>0,08</del>	<del>0,04</del>	<del>0,02</del>	<del>0,03</del>	<del>0,29</del>	<del>0,26</del>	<del><b>0,73</b></del>	10
UC_11	0,09	0,04	0,02	0,05	0,29	0,26	<b>0,76</b>	8
UC_12	0,09	0,04	0,02	0,03	0,14	0,26	<b>0,59</b>	15
SGAM_1	0,10	0,04	0,00	0,03	0,43	0,17	<b>0,78</b>	5
SGAM_2	0,10	0,04	0,00	0,03	0,43	0,17	<b>0,78</b>	6
SGAM_3	0,10	0,04	0,00	0,03	0,43	0,17	<b>0,78</b>	4
UC_X2	0,10	0,04	0,00	0,00	0,29	0,26	<b>0,69</b>	11
UC_X3	0,10	0,04	0,05	0,03	0,14	0,26	<b>0,62</b>	13

### 5.3.2 Anforderungsselektion

Der letzte Schritt der Anforderungsanalyse besteht in der Selektion der Anforderungen. Hierbei wurden die in Tabelle 5.9 aufgeführten funktionalen Anforderungen über einen definierten Grenzwert  $\Theta = 0.7$  selektiert, welche anhand der Nutzwertanalyse bereits priorisiert wurden. Der Grenzwert für die Anforderungsselektion wurde gewählt, um eine ausgewogene Balance zwischen Qualität und Umfang der selektierten Anforderungen zu gewährleisten. Ein Wert oberhalb dieses Schwellenwerts stellt sicher, dass nur Anforderungen mit einer hinreichend hohen Nutzwertbewertung  $N_j$  berücksichtigt werden, was die Relevanz und Effektivität der Auswahl verbessert. Gleichzeitig bleibt die Selektion flexibel genug, um eine ausreichende Anzahl an Anforderungen für die Konzeption



Quelle: Eigene Darstellung

Abbildung 5.7: Selektion der funktionalen Anforderungen für SGAAIRE über den Grenzwert  $\Theta = 0.7$

des Systems zu erhalten. Abbildung 5.7 zeigt absteigend die Ränge der funktionalen Anforderungen auf, deren Nutzwerte sowie die Selektion der funktionalen Anforderungen über den Grenzwert  $\Theta$ .

Tabelle 5.10 zeigt die für die Konzeption des KI-Systems SGAAIRE selektierten funktionalen Anforderungen. Bei der Konzeption werden neben den funktionalen auch die nicht-funktionalen Anforderungen berücksichtigt (vgl. Abschnitt 5.2), ebenso wie die Wünsche der Umfrageteilnehmenden, die in Abschnitt 5.1.3.3 erörtert wurden. Diese Aspekte fließen im Kapitel 7 in die Gestaltung des KI-Systems ein. Zudem werden in der Konzeptionsphase potenzielle Herausforderungen identifiziert und Lösungen aus dem KI-Bereich angestrebt, um so die Integration zu gewährleisten.

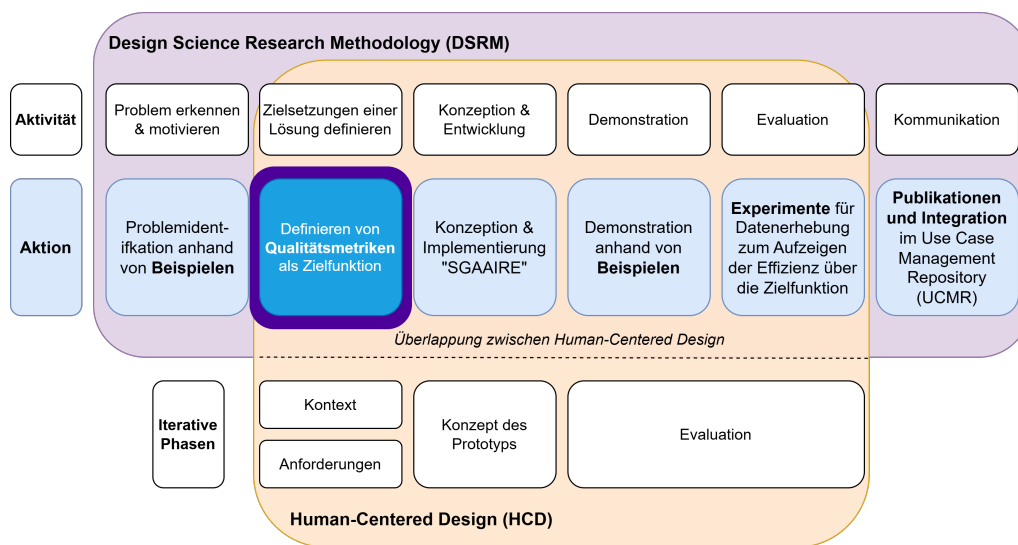
Tabelle 5.10: Übersicht der selektierten funktionalen Anforderungen

Anf.-ID.	Name
UC.1	Das System soll vorschlagen, Elemente (z. B. Akteure, Information Exchanges, Requirements, KPI, Keywords usw.) hinzuzufügen, zu entfernen oder zu ändern.
UC.2	Das System soll in freien Textfeldern abweichende und nicht standardisierte Terme feststellen.
UC.4	Das System soll Use Cases in SGAM-Zonen und Domänen einordnen.
UC.7	Das System soll auf Inkonsistenzen bei Akteuren, Information Exchanges und Requirements hinweisen, die nicht dem Use Cases-Typ entsprechen.
UC.9	Das System soll für Akteure die Akteurstypen durch Akteursnamen- und Beschreibungen vorschlagen.
UC.11	Das System soll Scenario Services in der Step-by-Step Analyse vorschlagen.
SGAM.1	Das System soll Arbeitsvorschläge für SGAM Modelle anhand von Use Cases generieren.
SGAM.2	Das System soll SGAM Modelle anhand eines Use Cases überprüfen auf Vollständigkeit, Redundanzen, Eindeutigkeiten, Korrektheit oder Verständlichkeit.
SGAM.3	Das System soll Wahrscheinlichkeiten für Inkonsistenzen innerhalb von SGAM Modellen erkennen und aufzeigen.



## 6 Konzeptionalisierung der Bewertungsmetriken

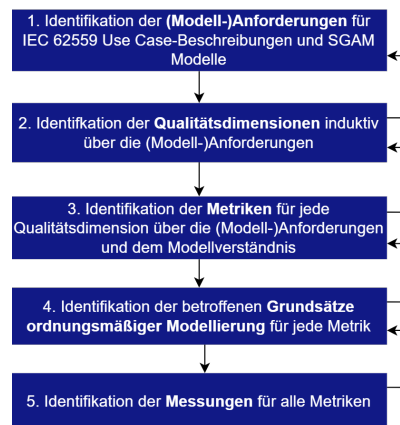
Im Rahmen dieses Kapitels wird das Ziel verfolgt, die Evaluierung der Qualität von IEC 62559 Use Case-Beschreibungen sowie SGAM Modellen zu unterstützen. Dieses Ziel erfordert die Entwicklung von Bewertungsmetriken, die auf Qualitätsmerkmalen basieren. Im Gesamtverfahren dienen diese Qualitätsmerkmale als Zielfunktion zur Bewertung der Effektivität und Effizienz von SGAAIRE, wie in Abbildung 6.1 veranschaulicht.



Quelle: Eigene Darstellung, angelehnt an [PTRC07]

Abbildung 6.1: Auswirkungen der Ergebnisse aus Kapitel 6 auf das Forschungsvorhaben

Diese Metriken sollen gemäß dem im Abschnitt 2.3 vorgestellten Vorgehensmodell (visualisiert in Abbildung 6.2) als ein zentrales Artefakt der Forschungsfrage F2.1 konzipiert werden. Zur Konzeption dieser Metriken werden zunächst die Begriffe „Qualität“ und „Fehler“ definiert und in den Kontext dieser Forschungsarbeit adaptiert. Im Anschluss erfolgt die Identifikation der Modellanforderungen durch eine Literaturrecherche. Anforderungen an IEC 62559 Use Case-Beschreibungen sowie SGAM Modelle werden erfasst und anschließend in Bezug auf das im Information Systems-Forschungsfeld vorzufindende Modellverständnis und die GoMs analysiert. Auf dieser Grundlage werden die primären Qualitätseigenschaften ausgewählt. Darüber hinaus ist es notwendig, Metriken zu entwickeln, die es ermöglichen, diese Qualitätseigenschaften objektiv zu messen. Diese aufgeführte Herangehensweise wurde in ihrer Gesamtheit einschließlich der Ergebnisse in der separaten Veröffentlichung [KSS23] konzipiert, umgesetzt und realisiert.



Quelle: Angepasst aus [KSS23]

Abbildung 6.2: Visualisierung des Prozesses zur Herleitung der Bewertungsmetriken

## 6.1 Adaptierung des Qualitäts- und Fehlerbegriffes

Der in **Forschungsfrage F2** eingeführte Begriff der *Qualitätsverbesserung* bildet die Grundlage zur Bewertung der erzielten Ergebnisse sowohl der konventionellen Methoden als auch der unterstützenden KI-Systeme. In diesem Abschnitt wird die Definition dieses Begriffs sowie die Methoden zur Messung seiner Ausprägung erörtert. Hierbei ist es zunächst von Bedeutung, die Definition von Qualität gemäß DIN EN ISO 9000:2015 (Qualitätsmanagementsysteme - Grundlagen und Begriffe) zu betrachten:

### Definition 9 (Qualität)

Grad, in dem ein Satz inhärenter Merkmale eines Objekts **Anforderungen erfüllt** [IN15]

Im Kontext der IEC 62559 Use Case-Beschreibungen sowie der SGAM Modelle ist es zunächst erforderlich, die inhärenten Merkmale festzulegen, die als *Qualitätseigenschaften* betrachtet werden sollen. Zusätzlich müssen Anforderungen definiert werden, um den Grad der Erfüllung dieser ausgewählten Qualitätseigenschaften messen zu können. Zur Unterstützung bei der Beantwortung der Forschungsfrage werden Metriken verwendet, die Reproduzierbarkeit und Transparenz gewährleisten sollen. Bestehende Qualitätsmetriken, wie beispielsweise das 3QM (eingesetzt im Kontext der Prozessmodellierung) [OBS12], verwenden absolute und relative Häufigkeiten von Fehlern. Ein Fehler wird gemäß DIN EN ISO 9000:2015 (Qualitätsmanagementsysteme - Grundlagen und Begriffe) wie folgt definiert:

### Definition 10 (Fehler)

Ein Fehler (bzw. die Nichtkonformität) ist nach DIN EN ISO 9000:2015 definiert als **Nichterfüllung einer Anforderung**. [IN15]

In diesem Zusammenhang wird deutlich, dass die Nichterfüllung einer Anforderung als Fehler betrachtet werden kann. Dies bedeutet, dass gemäß Definition 9 (Qualität) ein Fehler vorliegt, wenn eine Anforderung, die an ein inhärentes Merkmal gestellt wird, nicht erfüllt wird. Daraus ergibt sich die Implikation, dass die Verbesserung der Qualität mit der *Reduzierung von Fehlern* gleichgesetzt werden kann, sofern diese Fehler einer Qualitätseigenschaft zugeordnet sind. [KSS23]

Um die Qualitätsmerkmale und -metriken abzuleiten, müssen zunächst Modellanforderungen aus den IEC 62559 Use Case-Beschreibungen und den SGAM Modellen hergeleitet werden, wie im folgenden Abschnitt näher erläutert wird. Diese Ableitungen wurden im Rahmen der Veröffentlichung „Quality Properties of IEC 62559 Use Cases and SGAM Models“ auf der DACH+ Energy Informatics 2023 Konferenz erstellt. [KSS23]

### **6.2 Anforderungen an IEC 62559-2 Use Case-Beschreibungen und SGAM Modelle**

Im Rahmen der Ableitung von Qualitätsmerkmalen für IEC 62559 Use Case-Beschreibungen sowie SGAM Modelle erfolgte zunächst die Erfassung ihrer jeweiligen Anforderungen. Dieser Prozess wurde mittels einer Literaturrecherche durchgeführt, welche dazu diente, die relevanten Anforderungen zu identifizieren. Bei der Durchführung der Literaturrecherche wurden verschiedene Quellen herangezogen, darunter Standards und Referenzdokumente, methodisch orientierte Veröffentlichungen sowie anwendungsbezogene Publikationen [KSS23]. Diese sind in Tabelle 6.1 dargestellt und ihrer Kategorie zugewiesen:

- „S“ steht für die Literatur, die spezifisch für SGAM relevant ist.
- „U“ steht für die Literatur, die spezifisch für IEC 62559 Use Case-Beschreibungen relevant ist.
- „U/S“ steht für die Literatur, die sowohl für SGAM als auch für IEC 62559 Use Case-Beschreibungen von Bedeutung ist.

Die Herangehensweise zur Identifikation der Anforderungen aus der Literatur erfolgte auf induktive Weise. Dabei wurden die in den ausgewählten Quellen gefundenen Anforderungen systematisch erfasst und anschließend entsprechend ihrer Relevanz in die Kategorien der Anforderungen für IEC 62559 Use Case-Beschreibungen, SGAM Modelle und generische Anforderungen unterteilt. Die Ergebnisse der Anforderungserhebung über die Literatur sind in den Tabellen 6.2 (für die IEC 62559 Use Case-bezogenen Anforderungen), 6.3 (für die SGAM-bezogenen Anforderungen) und 6.4 (für die generischen Anforderungen) dargestellt. [KSS23]

Tabelle 6.1: Betrachtete Literatur zur Anforderungserhebung von IEC 62559 Use Case-Beschreibungen sowie SGAM Modellen

Quelle: Angelehnt an [KSS23]

Quelle	Kat.	Titel
[GDU17]	U/S	The Use Case and Smart Grid Architecture Model Approach
[IEC15]	U	IEC 62559 Use case methodology - Part 2: Definition of the templates
[CEN12]	U	CEN-CENELEC-ETSI: Smart Grid Reference Architectur
[Coc01]	U	Writing effective use cases
[CEN14]	U/S	SGAM User Manual - Applying, testing and refining the SGAM
[Bal16]	S	The Reference Designation System A common naming convention for systems and their elements
[GS15]	U	Towards Identifying an Approach for Consistency Checks to Smart Grid Descriptions
[NET <sup>+</sup> 14]	U/S	Towards Consistent Smart Grid Architecture Tool Support: From Use Cases to Visualization
[SUG <sup>+</sup> 14]	S	SGAM-based Methodology to analyse Smart Grid solutions in DISCERN research project
[URN <sup>+</sup> 19]	U/S	Applying the SGAM for Designing and Validating System-of-Systems in the Power and Energy Domain: A European Perspective
[NUEL16]	S	A standards-based approach for domain specific modelling of smart grid system architectures
[TGD13]	U	IEC/PAS 62559-Based Use Case Management for Smart Grids
[USS <sup>+</sup> 21]	U	Addressing quality properties in use case descriptions: adding assurance to a use case process
[GU15]	U	Supporting the Development of Smart Cities using a Use Case Methodology
[UG15]	S	Extending the SGAM for electric vehicles
[EU12]	S	Europäisches Architekturmodell für Smart Grids-Methodik und Anwendung der Ergebnisse der Arbeitsgruppe Referenzarchitektur des EU Normungsmandats M/490

### 6.3 Herleitung der (primären) Qualitätsdimensionen

Im Rahmen dieses Schrittes werden die primären Qualitätsdimensionen auf Grundlage der im vorherigen Abschnitt ermittelten (Modell-)Anforderungen festgelegt. Hierbei wird zunächst der aktuelle Stand bzgl. der Datenqualitätsmerkmale in Abschnitt 6.3.1 dargestellt. Für die Herleitung ist es zudem notwendig, das Modellverständnis der Information Systems (IS)-Forschung sowie die Grundsätze ordnungsmäßiger Modellierung in Abschnitt 6.3.2 zu betrachten, da IEC 62559 Use Case-Beschreibungen und SGAM Modelle als Ergebnisse eines Modellierungsprozesses im Requirement Engineering verstanden werden können [USD<sup>+</sup>13] und eine adäquate Modellierung Regeln unterliegt. Abschnitt 6.3.3 zeigt die Qualitätsdimensionen- sowie dessen Metriken auf.

#### 6.3.1 Bestehende Datenqualitätsmerkmale

Zunächst gilt es den Fokus auf Datenqualitätsmerkmale und deren Anwendung zu legen, um einen umfassenden Überblick zu gewinnen. Die ISO/IEC 25012 definiert allgemeine Datenqualitätsmerkmale im Bereich der Softwaretechnik. Diese werden unterteilt in inhärente Datenqualitätsmerkmale, systemabhängige Datenqualitätsmerkmale und die Kombination aus beiden [ISO08]. ISO/IEC 25012 definiert inhärente Datenqualität als den Grad, in dem Datenqualitätsmerkmale das intrinsische Potenzial haben, implizite Datenanforderungen zu erfüllen. Gemäß dieser Norm werden die folgenden Eigenschaften als integral für inhärente Datenqualität betrachtet: *Genauigkeit, Vollständigkeit, Konsistenz, Glaubwürdigkeit* und *Aktualität*. [KSS23]

Tabelle 6.2: Anforderungen für IEC 62559 Use Case-Beschreibungen (M.UC.ID)

Quelle: Basierend auf [KSS23]

ID	Name	[EUI2]	[UG15]	[GU15]	[USS+21]	[TGD13]	[NUEL16]	[URN+19]	[SUG+14]	[NET+14]	[GS15]	[Ba116]	[CEN14]	[Co01]	[CEN12]	[IEC15]	[GDU17]
1	Die Struktur der Anwendungsfälle sollte der Norm IEC 62559-2 entsprechen.	X			X	X											
2	Der Name des Anwendungsfalls sollte aus einem Verb und der Beschreibung bestehen		X	X													
3	Der Anwendungsfall sollte die korrekten Domänen- und Zoneninformationen gemäß dem Referenzmodell enthalten.		X														
4	Der Anwendungsfall sollte eine Beschreibung des Anwendungsbereichs enthalten, die die Systemgrenze definiert		X	X													
5	Der Anwendungsfall sollte eine Beschreibung im narrativen Teil zur Spezifikation der Funktionen enthalten		X	X		X											
6	Der Anwendungsfall sollte entsprechend der Art des Anwendungsfalles konsistente Akteure enthalten.		X				X	X									
7	Die Use Case-Akteure sollten gemäß den (Geschäfts-)Rollen Verantwortlichkeiten sein, die darauf abzielen, Folgendes zu unterstützen		X														
8	Die Akteure des Business Use Case sollten Rollen nach dem Conceptual Model oder dem Harmonisierten Rollenmodell verwenden		X														
9	Der Anwendungsfall sollte die Gruppierung von Akteuren (oder Rollen) zur Vereinfachung verwenden, wenn Gruppierungen vorhanden und notwendig sind (z. B. DSO).		X														
10	Die Use Case-Akteure und -Rollen haben eine konsistente Beziehung hinsichtlich der Verfeinerung (Spezialisierung)		X	X													
11	Der Anwendungsfall ist von konsistenter Natur (z. B. Business Use Cases, System Use Cases, ...)		X	X		X											
12	Der Anwendungsfall sollte gültige Referenzen enthalten (z. B. Normen, Gesetze, Vorschriften, Grid Codes, ...)		X			X											
13	Die Use Case Schritt-für-Schritt-Analyse sollte deutlich machen, welche Informationen in jedem Schritt weitergegeben werden		X			X											
14	Die schrittweise Analyse der Anwendungsfälle sollte die anwendbaren Anforderungen spezifizieren		X														
15	Die Akteure des Anwendungsfalls sollten in der vollständigen Beschreibung erwähnt werden	X		X													
16	Das Anwendungsfalldiagramm sollte Interaktionen zeigen		X														
17	Die Kurzbeschreibung sollte nur einen kleinen Überblick unter 11 Zeilen geben	X															
18	Die vollständige Beschreibung sollte die Fragen beantworten, wie, wo, wann, warum und unter welchen Bedingungen aus der Sicht des Benutzers etwas geschieht	X		X													

In dem Forschungsbeitrag von [SSPA<sup>+</sup>12] wurden 41 verschiedene sowie überschneidende Datenqualitätsdimensionen identifiziert, darunter Konsistenz, Genauigkeit und Vollständigkeit. Einige der Dimensionen werden in Frameworks wie dem 3QM-Framework verwendet, um die Qualität von Prozessmodellen quantitativ zu bewerten, wobei Metriken über die syntaktische/semantische Korrektheit, Relevanz, Vollständigkeit, Flexibilität, Eindeutigkeit und Verständlichkeit berücksichtigt werden [OBS12]. Zur Messung werden zwei Verhältnisse verwendet, die sich auf spezifische Vorkommnisse beziehen: Der absolute Wert  $A$  bezieht sich auf die Anzahl auftretender Mängel im Verhältnis zur entsprechenden Eigenschaft (z. B. Vollständigkeit), und der relative Wert  $R = 1 - \frac{A}{N}$

Tabelle 6.3: Anforderungen für SGAM Modelanforderungen (M\_SGAM\_ID)

Quelle: Basierend auf [KSS23]

ID	Name	[EU12]	[UG15]	[GU15]	[USS+21]	[TGD13]	[NUEL16]	[URN+19]	[SUG+14]	[NET+14]	[GS15]	[Bal16]	[CEN14]	[Coc01]	[CEN12]	[IEC15]	[GDU17]
1	Die Struktur des SGAM Modells sollte der SGAM-Referenzarchitektur entsprechen.							X							X		
2	Die Komponentenschicht sollte Komponenten enthalten (z. B. logische oder physische Akteure, Komponenten)														X		
3	Die Kommunikationsschicht sollte Verbindungen mit einer Beschreibung der Art und Weise der Kommunikation zwischen den kommunizierenden Akteuren enthalten														X		
4	Die Informationsschicht sollte Verbindungen mit einer Beschreibung des Informationsaustauschs zwischen den kommunizierenden Akteuren enthalten														X		
5	Die Funktionsschicht sollte Informationen über die Funktionalität enthalten					X									X		
6	Die Geschäftsschicht sollte einen Zweck des Anwendungsfalls enthalten														X		
7	Das Element sollte in die richtige Zone eingeordnet werden														X		
8	Das Element sollte in den richtigen Bereich eingeordnet werden														X		
9	Die Komponenten auf der „Component Layer“ sollten auf den richtigen Abstraktionsebenen angeordnet sein										X	X					
10	Die Mechanismen des Informationsaustauschs auf der „Communication Layer“ sollten auf einer identischen Abstraktionsebene angeordnet werden										X	X					
11	Der Informationsaustausch auf der „Information Layer“ sollte auf einer identischen Abstraktionsebene erfolgen.										X	X					
12	Die Funktionen auf „Functional Layer“ sollten auf einer identischen Abstraktionsebene angeordnet sein										X	X					
13	Die Zwecke auf „Business Layer“ sollten auf einer identischen Abstraktionsebene angeordnet werden										X	X					
14	Die Akteure aus dem „Component Layer“ sollten auch in den anderen Layern an der gleichen Stelle vorkommen											X					

Tabelle 6.4: Anforderungen für generische Modelanforderungen (M\_G\_ID)

Quelle: Basierend auf [KSS23]

ID	Name	[EU12]	[UG15]	[GU15]	[USS+21]	[TGD13]	[NUEL16]	[URN+19]	[SUG+14]	[NET+14]	[GS15]	[Bal16]	[CEN14]	[Coc01]	[CEN12]	[IEC15]	[GDU17]	
1	Die Beschreibung oder das Modell sollte verständlich sein														X	X	X	X
2	Die Beschreibung oder das Modell sollte eindeutig sein								X			X			X	X	X	X
3	Die Beschreibung oder das Modell sollte vollständig sein								X			X			X	X	X	X
4	Die Beschreibung oder das Modell sollte relevant sein											X			X	X	X	X
5	Die Beschreibung oder das Modell sollte syntaktisch korrekt sein														X	X	X	X
6	Die Beschreibung oder das Modell sollte syntaktisch korrekt sein														X	X	X	X

bezieht sich auf die Größe der Menge der absoluten Vorkommnisse [OBS12].

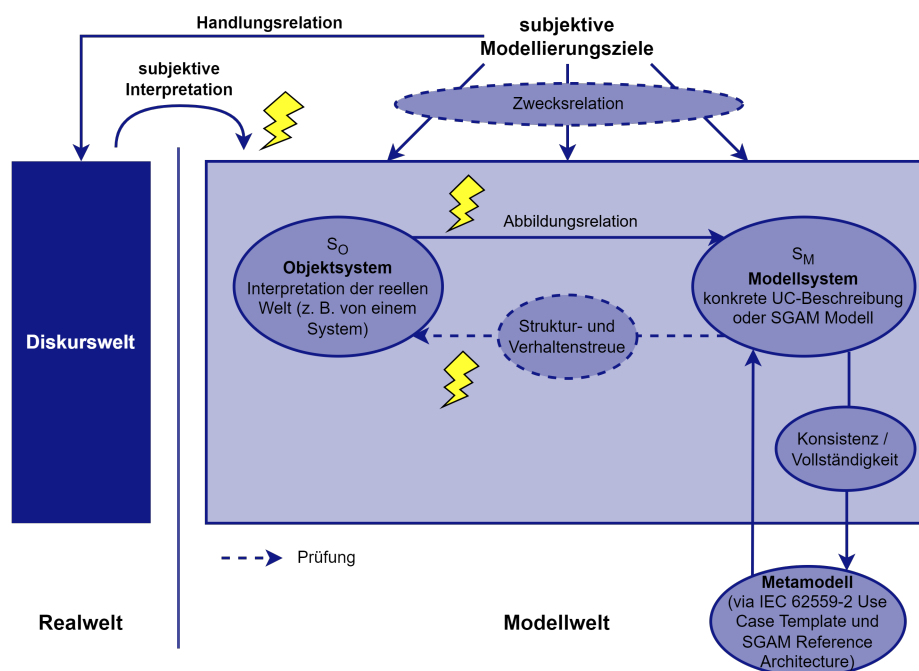
### 6.3.2 Herausforderungen bei Anwendung des Modellverständnisses

IEC 62559 Use Case-Beschreibungen sowie SGAM Modelle entsprechen Ergebnissen eines Modellierungsprozesses im Smart Grid Requirement Engineering [USD<sup>+</sup>13], welches wiederum als Teil der IS-Forschung betrachtet wird [USD<sup>+</sup>13]. Im IS-Feld haben sich die Grundsätze ordnungsmäßiger Modellierung (GoM) ([BRS95]) etabliert, welche Richtlinien für die Modellierung (insb. im Fokus auf Prozessmodelle) darstellen. Hierbei wird das Modellverständnis ([Ros13]) zur Hand gezogen, welches ebenfalls im IS entworfen wurde. Die GoM umfassen gemäß [BRS95] die folgenden:

- **Grundsatz der Richtigkeit:** Diese Richtlinie behandelt zwei Unteraspekte der Korrektheit: die syntaktische und semantische Korrektheit. Ein Modell gilt als syntaktisch korrekt, wenn alle Regeln einer Modellierungssprache (das Metamodell) eingehalten werden. Ein Modell gilt als semantisch korrekt, wenn die strukturelle und funktionale Korrektheit des Modells hinsichtlich des über die Abbildungsrelation repräsentierten Objektsystems korrekt ist.
- **Grundsatz der Relevanz:** Dieses Prinzip befasst sich mit der Modellierung der relevanten Aspekte für das jeweilige Modellierungsziel. Bei der Abstraktion in einem Modell müssen nicht alle Aspekte ständig dargestellt werden. [BRS95]
- **Grundsatz der Wirtschaftlichkeit:** Das Modellierungsziel sollte mit minimalem Aufwand erreicht werden. Das Prinzip besagt, dass ein Modell verfeinert werden sollte, bis die Kosten der Verfeinerung den resultierenden Nutzen der Verfeinerung ausgleichen. [BRS95]
- **Grundsatz der Klarheit:** Diese Richtlinie stellt sicher, dass ein Modell leicht interpretierbar ist, sodass es leicht lesbar, klar und verständlich ist. Das Prinzip der Klarheit zielt auch darauf ab, sicherzustellen, dass die Hierarchie, das Layout-Design und die Filterung für den Adressaten angemessen sind. [BRS95]
- **Grundsatz der Vergleichbarkeit:** Zwei Aspekte werden hier beleuchtet: Erstens müssen die Prozesse der realen Welt und der in der Modellierungssprache dokumentierten imaginären Welt im Modell vollständig identisch sein. Der zweite Aspekt zielt auf den Vergleich von Modellen ab, die in verschiedenen Modellierungssprachen erstellt wurden. Diese sollten gegenseitig übersetzbar sein. [BRS95]
- **Grundsatz des systematischen Aufbaus:** Dieses Prinzip betrifft die Anforderung an die interne Konsistenz zwischen Struktur- und Verhaltensmodellen. [BRS95]

In [KSS23] wurden die Herausforderungen des Modellverständnisses für die Erstellung von IEC 62559 Use Case-Beschreibungen sowie SGAM Modellen über die Beschreibung des System-of-Interest aufgeführt. Kernprobleme stellen die subjektive Interpretation der Diskurswelt in ein Objektsystem, die adäquate Transformation der Abbildungsrelation sowie falsche Ergebnisse bei den Überprüfungen der Struktur- und Verhaltenstreue dar. Diese in der Abbildung 6.3 markierten Kernprobleme können auf folgende Probleme zurückzuführen sein: [KSS23]

- Syntaktische, semantische und pragmatische Barrieren bei der Kommunikation mit den Stakeholdern [PKML22].
- Unterschiedliches Verständnis zum System-of-Interest sowie über die IEC 62559 Use Case-Methodik/SGAM Framework
- Unterschiedliche Schwerpunkte bzw. Hintergründe
- Fehler im Hinblick auf den erforderlichen Abstraktionsgrad
- Allgemeine Verständnis- und Kommunikationsprobleme



Quelle: Übersetzt aus [KSS23]

Abbildung 6.3: Vorliegendes Modellverständnis und Herausforderungen im Kontext

### 6.3.3 Qualitätsdimensionen- und Metriken

Auf Grundlage der Modellanforderungen wurden in [KSS23] die Qualitätsdimensionen sowie High-Level-Metriken und Ausprägungsmerkmale abgeleitet. Darüber hinaus wurden die Grundsätze ordnungsmäßiger Modellierung herangezogen, um die Merkmale den entsprechenden Modellierungsrichtlinien zuzuweisen. Die Qualitätsdimensionen, ihre Ausprägungsmerkmale sowie die verweisenden Anforderungen und GoM sind in Tabelle 6.5 aufgeführt.

Die Metriken dienen dazu, die Messbarkeit der Qualitätsdimensionen sicherzustellen. Da jedes Ausprägungsmerkmal anhand der Häufigkeit von Verletzungen quantifiziert werden kann, werden sowohl der absolute Wert als auch der relative Wert in Bezug auf die Anzahl der Vorkommnisse herangezogen.

## 6.4 Bewertungsmetriken für Qualitätsdimensionen

Im vorherigen Abschnitt 6.3 wurde gezeigt, dass Qualitätsdimensionen und deren Ausprägungsmerkmale mithilfe der vorhandenen Literatur identifiziert werden konnten. Dabei wurden auch das Modellverständnis im Bereich der Informationssystemforschung (gemäß [Ros13]) und die GoM herangezogen.

Die in Abschnitt 6.3.3 aufgeführten Ausprägungsmerkmale sowie deren Metriken beziehen sich auf den im Abschnitt 6.1 adaptierten Fehlerbegriff (im Sinne von Qualitätsdefizite) und somit gilt es diese Ausprägung innerhalb der IEC 62559 Use Case-Beschreibungen sowie SGAM Modelle zu minimieren. Hierbei können Qualitätsdefizite in beliebiger Anzahl auftauchen. Jedoch gilt es in der Bewertungsmetrik die Intensität des Qualitätsdefizits zu berücksichtigen. Während sich ein kleiner Flüchtigkeitsfehler in der Syntax weniger problematisch auswirken kann, könnte eine Fehlklassifikation eines Aktors (z. B. in einer anderen Domäne oder Zone) das Modell verfälschen und folglich treffen die Systemintegratoren falsche Annahmen. Basierend auf diesen Gegebenheiten, wurde die in Definition 11 aufgeführte Bewertungsfunktion basierend auf die in Abschnitt 6.3.3 aufgeführten Terminologien hergeleitet.

### **Definition 11** (Fehlerfunktion als Bewertungsmetrik)

Für ein Modellierungsartefakt (im Sinne einer IEC 62559 Use Case-Beschreibung oder eines SGAM Modells  $x$ ) kann die Fehlerfunktion  $E(x) \in \mathbb{R}$  angewandt werden, wenn eine zum Modellierungsergebnis zugehörige Menge an Qualitätsausprägungsmerkmalen  $Q(x)$  vorliegt, welche die Qualitätsdefizite aufzeigt. Für ein Qualitätsausprägungsmerkmal  $q \in Q(x)$  sind folgende Funktionen vorgesehen:

- i. **Quantität (die Wiederholung)** des Qualitätsausprägungsmerkmals innerhalb des Modells  $n(q, x) \in \mathbb{N}$
- ii. **Qualitätintensität (Gewicht)** eines Qualitätsausprägungsmerkmals  $s(q, x) \in \mathbb{R}$

Die Gesamtheit aller Qualitätsdefizite unter Berücksichtigung ihrer Quantitäten und abgebildeten Intensitäten bildet die Grundlage für die folgende Fehlerfunktion, welche als Bewertungsmetrik betrachtet wird:

$$E(x) = \sum_q^{Q(x)} n(q, x) * s(q, x) \quad (6.1)$$

Bei der Ableitung von Bewertungsmetriken ergibt sich ein Problem der subjektiven Interpretation, wie es in unserem Modellverständnis (gemäß [Ros13]) dargestellt ist (vgl. Abbildung 6.3). In diesem Modellverständnis wird angenommen, dass die globale Wahrheit in der Realwelt existiert und von Modellierern unter Verwendung ihrer individuellen Expertise für ihre subjektiven Modellierungsziele in die Modellwelt überführt wird. Dies kann zu Verzerrungen und Fehlern führen, da die globale Wahrheit nicht deterministisch bestimmbar ist. Es resultiert auch in der Tatsache, dass die Bewertung der Ausprägungsmerkmale, wie in Tabelle 6.5 aufgeführt, von jeder bewertenden Instanz (sei es eine Person oder ein Computersystem) individuell sein kann. [KSS23]

Um dennoch Tendenzen in Bezug auf die Erfüllungsgrade der Qualitätsdimensionen zu erhalten, besteht die Möglichkeit, die Bewertung mithilfe statistischer Methoden und unter Verwendung mehrerer Stichproben zu approximieren [GTP15]. Dies bedeutet, dass die Qualität einer bestimmten IEC 62559 Use Case-Beschreibung oder eines SGAM Modells anhand der Merkmale von einer Gruppe von  $n$ -bewertenden Instanzen beurteilt wird, um eine klarere Tendenz auf Basis dieser  $n$  Bewertungen zu erhalten. Ein Optimierungsproblem, das diese Problematik angemessen berücksichtigt, kann unter Verwendung der in Definition 12 dargestellten Zieldefinition formuliert werden. [KSS23]

**Definition 12** (Optimierungsziel über die Fehlerfunktion)

Das Hauptziel der Optimierung besteht in der Minimierung der Bewertungsmetrik gemäß Definition 11. Dieses Ziel bringt jedoch die Herausforderung mit sich, dass aufgrund der subjektiven Interpretation des System-of-Interest eine Unsicherheit bei der Qualitätsbewertung durch die bewertende Instanz vorhanden ist. Dies führt dazu, dass die Fehlerfunktion  $E(x)$  über  $n$  Stichproben mit  $\tilde{E}(x)$  approximiert wird, um dieses Ungenauigkeitsproblem zu berücksichtigen [KSS23]. Das Optimierungsproblem kann mathematisch durch die folgende Funktion (basierend auf [KSS23]) definiert werden:

$$\arg \min_x \tilde{E}(x) \approx \arg \min_x \sum_q^{Q(x)} n(q, x) * s(q, x) \quad (6.2)$$

Tabelle 6.5: Abgeleitete Qualitätsdimensionen und deren Metriken für IEC 62559 Use Case-Beschreibungen und SGAM Modelle nach dem Anforderungsmapping

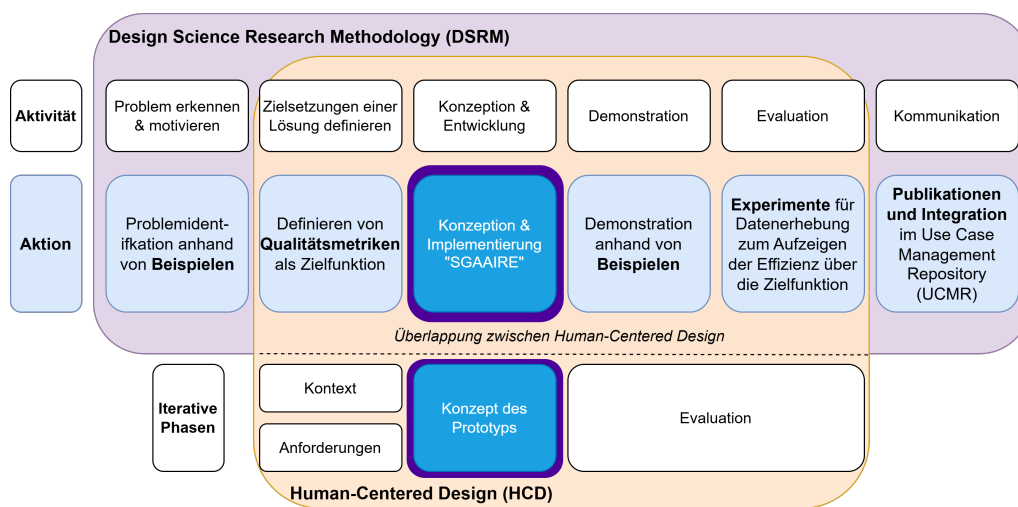
Quelle: Basierend auf [KSS23]

Qualitätsdimensionen und Ausprägungsmerkmal	Beschreibung (referenziert auf IEC 62559 Use Case-Beschreibungen und SGAM Modelle)	Bezogene Anforderungen			GoM
		M_UC	M_SGAMM_G		
<b>Korrektheit</b>					
Inkorrekte Klassifikation	Eine Klassifizierung in der Vorlage wurde fehlerhaft vorgenommen	1, 3	1, 7, 8	6	Korr.
Falsche Entitäten	Eine Entität hat in ihrem Inhalt Fehler	1, 11	1, 2, 9-13	5, 6	Korr.
Falsche Abstraktionsebene gemäß dem Objektsystem	Entitäten im Template/Modell haben gemäß des Objektsystems unterschiedliche Abstraktionsstufen	1, 10, 11	1, 9-13	6	Sys.A.
Nichtkonformitäten gemäß des Standards	Fehler in Bezug auf die Einhaltung zum Standard sind im Modell enthalten	1	1	1	Korr.
<b>Konsistenz</b>					
Inkonsistenz innerhalb des Modells	Inkonsistenzen innerhalb der Beschreibung/des Modells sind zu vermeiden	1	1	2	Vergl.
Inkonsistenz zwischen Use Cases und SGAM Modellen	Inkonsistenzen zwischen einer Use Case-Beschreibung und dem zugehörigen SGAM Modell sind gegeben	1	1	3, 4	Vergl.
<b>Vollständigkeit</b>					
Unvollständigkeiten gemäß des Standards	Standards sehen Felder, Ebenen oder Entitäten vor, die in der Beschreibung bzw. Modell fehlen	1, 2, 4-6, 12-18	1 - 10, 14	3, 4	Rel.
Fehlende Entitäten gemäß dem Objektsystem	Das Objektsystem sieht Entitäten vor, die im Modellsystem fehlen	1	1	3, 4	Rel.
<b>Relevanz</b>					
Irrelevante Entitäten gemäß dem Objektsystem	Unnötige Entitäten, die zur Veranschaulichung nicht notwendig sind (z. B. keine zusätzlichen Mehrwerte), sind Teil des Modells/der Beschreibung	1	1	4	Rel., Wirt.
<b>Nachvollziehbarkeit</b>					
Abweichende Bezeichnungen	In der Beschreibung/dem Modell wurden unzulässige/abweichende Begriffe verwendet	1, 12, 18	1, 3, 4	1	Klar.
Nicht genormte Bezeichnungen und Abkürzungen	In der Vorlage/dem Modell werden nicht genormte Bezeichnungen und/oder Abkürzungen verwendet	1, 18	1, 3, 4	1	Klar.
<b>Eindeutigkeit</b>					
Redundanz gemäß dem Objektsystem	Informationen mit der gleichen Bedeutung sind mehrfach in der Beschreibung/dem Modell vorhanden	1	1	2	Sys.A.
<b>Erreichbarkeit</b>					
Unleserliche Darstellungen oder Texte	Texte oder Darstellungen haben eine unklare Darstellung	1, 2, 4, 5, 13, 15-18	1, 14	1	Klar.



## 7 Konzeptionalisierung von SGAAIRE

In diesem Kapitel wird die Konzeption von SGAAIRE als Softwaresystem behandelt. Unter Berücksichtigung der nicht-funktionalen Anforderungen aus Abschnitt 5.2, der funktionalen Anforderungen aus Abschnitt 5.3 sowie Betrachtung des Kontexts aus Abschnitt 5.1.3.2 wird ein Prototyp entwickelt, der gezielt in den Prozess der Erstellung von IEC 62559 Use Cases und SGAM Modellen integriert werden soll. Im Gesamtverfahren stellt SGAAIRE das zentrale Artefakt dieser Dissertation dar, dessen Konzeption und Entwicklung im Prozess in Abbildung 6.1 dargestellt sind.



Quelle: Eigene Darstellung, angelehnt an [PTRC07]

Abbildung 7.1: Auswirkungen der Ergebnisse aus Kapitel 7 auf das Forschungsvorhaben

Die Konzeptualisierung von SGAAIRE erfordert eine Berücksichtigung der Rahmenbedingungen bei der Entwicklung von IEC 62559 Use Cases und SGAM Modellen. Aus diesem Grund wird im Abschnitt 7.1 der zugrunde liegende Entwicklungsprozess ausgewählt, analysiert und systematisch eingeordnet. Darüber hinaus werden die identifizierten Herausforderungen im Rahmen der Kontextanalyse (vgl. Abschnitt 5.1.3.2) näher betrachtet und darauf aufbauend zentrale Konzeptentscheidungen abgeleitet. Nach diesen Einordnungen werden in den Abschnitten 7.2 und 7.3 Übersichten über SGAAIRE in Form von statischen und dynamischen Systemkonzepten aufgezeigt. Die Konzeption der essenziellen Komponenten des Systems werden im Rahmen des Abschnitts 7.4 beleuchtet.

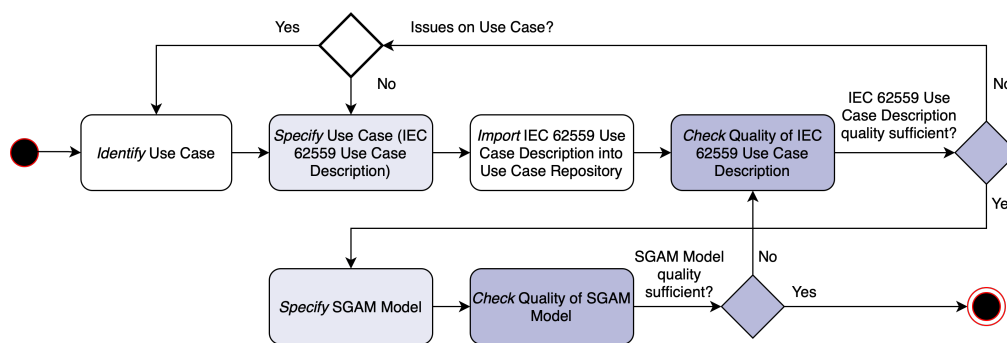
## 7.1 Zugrundeliegender Prozess

Bereits in Abschnitt 5.1.3.2 wurde aufgezeigt, dass bei der Anwendung der IEC 62559 Use Case-Methodik sowie des SGAM Frameworks unterschiedliche Prozesse und eine ausgeprägte Heterogenität vorherrschen. Für die Modellierenden ergeben sich daraus vielfältige Herausforderungen, die die Erstellung adäquater IEC 62559 Use Cases und SGAM Modelle zu einer anspruchsvollen und keineswegs trivialen Aufgabe machen. Da sich diese Dissertation gleichermaßen auf die IEC 62559 Use Case-Methodik sowie das SGAM Framework fokussiert und dabei deren wechselseitige Beziehungen berücksichtigt, wird ein *Use Case-orientierter Prozess* herangezogen, der insbesondere im Rahmen der Kontextanalyse [KNN<sup>+</sup>24] als verbreiteter Ansatz gilt.

Weitere Ansätze sind vor allem im Bereich des Model-Based Systems Engineering (MBSE) verortet und agieren weitgehend losgelöst von der Erstellung von SGAM Modellen auf Basis von IEC 62559 Use Cases. Auf Seiten der IEC 62559 Use Case Methodik kommt insbesondere das Werkzeug *MODSARUS*<sup>®</sup>, eine Erweiterung von *Sparx Enterprise Architect*, zum Einsatz [EDF]. Für die SGAM Modellierung wiederum wird häufig die *SGAM Toolbox* innerhalb von *Enterprise Architect* verwendet [VNE<sup>+</sup>25]. In beiden Fällen erfolgt die Modellierung zentral über die UML, wobei einzelne Modellelemente und deren semantische Relationen formal abgebildet werden. Während die SGAM Toolbox eine strukturierte Modellierung entlang der fünf SGAM Ebenen erlaubt, ermöglicht *MODSARUS*<sup>®</sup> die systematische Erfassung der Anforderungen der einzelnen Akteuren. Zudem können in *MODSARUS*<sup>®</sup> IEC 62559 Use Cases aus dem Tool generiert werden.

Der hier zugrundeliegende *Use Case-orientierte Prozess* (visualisiert in Abbildung 7.2) beginnt mit der Identifikation eines konkreten Anwendungsfalls, beispielsweise dem Laden eines EVs unter Berücksichtigung dynamischer Stromtarife. Dieser Anwendungsfall wird anschließend im Rahmen der IEC 62559 Use Case-Methodik spezifiziert. Gemäß dem IEC 62559-1 Standard [IEC19] wird der IEC 62559 Use Case innerhalb eines UCMR bereitgestellt. Im Anschluss erfolgt eine Überprüfung der inhaltlichen Qualität. Ist diese nicht ausreichend, wird der Prozess der Erstellung, bzw. gegebenenfalls der Identifikation, iterativ wiederholt. Bei ausreichender Qualität wird in einem nächsten Schritt das zugehörige SGAM Modell spezifiziert. Auch dessen Qualität wird überprüft. Sollte diese nicht zufriedenstellend sein, erfolgt eine kritische Rückkopplung auf Basis der IEC 62559 Use Case-Beschreibung. Bei hinreichender Qualität dieser Beschreibung wird die Erstellung des SGAM Modells erneut durchlaufen. Dieser iterative und qualitätsgesicherte Prozess stellt sicher, dass das SGAM Modell und die IEC 62559 Use Case-Beschreibung inhaltlich konsistent zueinander sind. [KNN<sup>+</sup>24, KLS25]

Ein KI-System wie SGAAIRE kann insbesondere in den Phasen der Spezifikation und Qualitätssicherung in den beschriebenen Prozess integriert werden. In Abbildung 7.2 sind die jeweiligen Prozessphasen farblich hervorgehoben, wobei die Farben die folgenden Bedeutungen repräsentieren:



Quelle: [KLS25]

Abbildung 7.2: Zugrundeliegender Prozess der Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen

- **Hintergrund mit hellerem Blau:** Zeigt Aktivitäten auf, die durch die Nutzung von SGAAIRE beeinflusst werden. Dies entsprechen Bereitstellungen von Hinweisen sowie ggf. Vorschläge (z. B. Generieren von SGAM Arbeitsmodelle) und zielen darauf ab, die Erstellung weitestgehend zu beeinflussen.
- **Hintergrund mit dunklerem Blau:** Zeigt Aktivitäten oder Entscheidungen auf, die durch SGAAIRE in Rahmen von gezielter Qualitätsbewertung beeinflusst werden. Es werden Hinweise sowie ggf. Vorschläge zur Qualitätsverbesserung bereitgestellt welche auf indirekte Verbesserungen abzielen.
- **Weißer Hintergrund:** Zeigt Aktivitäten auf, die durch die Nutzung von SGAAIRE nicht beeinflusst werden.

Konzeptionell betrachtet beeinflusst SGAAIRE wesentlich die iterative Ausgestaltung der Spezifikations- und Qualitätssicherungsphasen. Das System wirkt dabei als aktiver kognitiver Impulsgeber innerhalb des Modellierungsprozesses. Indem es auf Inkonsistenzen, Unschärfen oder semantische Lücken in den IEC 62559 Use Cases sowie den zugehörigen SGAM Modellen hinweist, initiiert SGAAIRE gezielte Reflexionsprozesse bei den Modellierenden. Im Idealfall führt dies zu einem bewussten Überdenken zuvor getroffener Modellierungsentscheidungen, sei es in der strukturellen Beschreibung von Funktionen, in der Zuordnung von Akteuren, oder das adäquate Positionieren von Elementen innerhalb des SGAM. Die Modellierenden werden nicht ersetzt, sondern in ihrer Analyse- und Urteilsfähigkeit gestärkt. Langfristig kann ein solches System zur Etablierung eines qualitätsorientierten Modellierungsverständnisses beitragen und die Diskussionen verbessern.

Tabelle 7.1: Einschätzung der identifizierten Herausforderungen auf die funktionalen Anforderungen

ID	Anforderung	Hohe Komplexität	Qualität aufrechtzuerhalten	Einheitliche Terminologie	Präzise SGAM Klassifikationen	Informationsverfügbarkeit	Korrekte Darstellung	Fehlende Anleitung	Fachwissen nötig
UC 1	Elemente vorschlagen, ändern, entfernen	+	+	+		+	+	+	+
UC 2	Nicht-standardisierte Terme erkennen	+	+	++	+				
UC 4	Einordnung in SGAM-Zonen und -Domänen	+	+		++		+	+	+
UC 7	Inkonsistenzen bzgl. Use Case-Typ erkennen	+	+				+	+	+
UC 9	Vorschlag Akteurstypen durch Beschreibung	+	+	+			+	+	+
UC 11	Vorschlag von Scenario Services	+	+	+			+	+	+
SGAM 1	SGAM Modellvorschläge aus Use Cases	+	+	+	++	+	+	+	+
SGAM 2	Prüfung von SGAM Modellen	+	+	+	++	+	+	+	+
SGAM 3	Erkennung von Inkonsistenzen	+	+	+	++		+	+	+

### 7.1.1 Betrachtung der Herausforderungen

Im Rahmen der Kontextanalyse konnten die in Abschnitt 5.1.3.2 dargestellten Herausforderungen identifiziert werden, die der Entwicklung von IEC 62559 Use Cases und SGAM Modellen zugrunde liegen. Einzelne dieser Herausforderungen und Komplexitäten können durch den gezielten Einsatz von SGAAIRE womöglich abgeschwächt werden. Eine Einordnung der wesentlichsten identifizierten Herausforderungen und deren potenzieller Mitigation durch SGAAIRE ist in Tabelle 7.1 dargestellt. In der Tabelle sind die identifizierten funktionalen Anforderungen den typischen Herausforderungen bei der Modellierung gegenübergestellt; dabei kennzeichnet „++“ eine starke und „+“ eine teilweise adressierende Wirkung der jeweiligen Anforderung auf die jeweilige Herausforderung. Hierbei kann folgendes identifiziert werden:

- Die **hohe Komplexität in der Erstellung** liegt in der Natur des Anforderungsmanagements [PR15], da sowohl Fach-, Domänen- als auch Methodenwissen erforderlich ist (vgl. Kapitel 3). Das KI-basierte System SGAAIRE kann diese Komplexität zwar nicht vollständig auflösen, sie jedoch durch gezielte Empfehlungen und Vorschläge teilweise eingrenzen, ein Vorgehen, das auch im konventionellen Anforderungsmanagement Anwendung findet [RDCC25]. Grundsätzlich kann jede der betrachteten Anforderungen zur Reduktion dieser Komplexität beitragen.
- Auch die **Aufrechterhaltung der Qualität** stellt eine zentrale Herausforderung dar. Die wiederholte Bearbeitung von IEC 62559 Use Cases und SGAM Modellen in verschiedenen Projektphasen erfordert eine systematische Planung und kann nicht durch SGAAIRE kompensiert werden. Dennoch kann das System in der Anforderungsanalysephase wertvolle Unterstützung anbieten. Basierend auf der Einschätzung durch

Modellierende besteht das Potenzial, dass alle betrachteten Anforderungen die Modellqualität verbessern können (vgl. Nutzwertanalyse in Abschnitt 5.3.1).

- Ein typisches Problem im dokumentbasierten Anforderungsmanagement ist das Fehlen einer **einheitlichen Terminologie** [PR15]. SGAAIRE kann durch seine Vorschläge die Diskussion zwischen Stakeholdern anregen. Insbesondere die Fähigkeit, nicht standardisierte Terme zu erkennen, ermöglicht es, gezielte Ergänzungen für das Glossar der IEC 62559 Use Case-Beschreibung vorzuschlagen.
- Ein weiterer häufig genannter Aspekt ist die **präzise Klassifikation von Elementen bzw. Use Cases in SGAM-Domänen und -Zonen**. Hierbei bietet SGAAIRE insbesondere im SGAM-Kontext direkte Unterstützung durch automatisierte Einordnungen und Vorschläge.
- Die Herausforderung der **mangelnden Informationsverfügbarkeit** ergibt sich daraus, dass bei der Erhebung von Anforderungen nicht alle notwendigen Informationen bereits vorliegen. Zwar kann SGAAIRE fehlende Informationen nicht ersetzen, es kann jedoch relevante Diskussionen mit Stakeholdern initiieren und somit zur Informationserschließung beitragen, insbesondere bei den aufgeführten, anregungsorientierten Anforderungen.
- Die **korrekte Darstellung**, insbesondere im Rahmen der SGAM Modellierung, wurde ebenfalls als Problem benannt. SGAAIRE kann hier Unterstützung bieten, etwa durch Hinweise zur normkonformen Anwendung des IEC 62559-2 Templates. Da SGAM-Elemente zusätzlich grafisch darzustellen sind, liegt die endgültige visuelle Ausgestaltung jedoch bei externen Werkzeugen wie dem SGAM Visualizer [GDU17] oder der SGAM Toolbox [VNE<sup>+</sup>25]. SGAAIRE kann jedoch standardkonforme serialisierte Dateien bereitstellen, die als Eingabe für diese Tools dienen.
- Ein ähnliches Bild zeigt sich bei der **fehlenden Anleitung zur Modellerstellung**. SGAAIRE kann durch lokale und globale Vorschläge nicht nur Transparenz schaffen, etwa im Sinne von XAI, sondern auch die regelkonforme Anwendung der relevanten Standards erleichtern.
- Abschließend wurde die Notwendigkeit eines hohen Maßes an **Fachwissen** betont. Zwar kann ein KI-basiertes System wie SGAAIRE kein Fachwissen ersetzen, jedoch Impulse setzen und fundierte Diskussionen fördern. Dabei ist zu beachten, dass die generierten Vorschläge stets passiv und nicht als normative Wahrheit zu verstehen sind.

Basierend auf den analysierten Herausforderungen und dem Beitrag der funktionalen Anforderungen lassen sich Handlungsempfehlungen für die Konzeption von SGAAIRE ableiten. Zunächst sollte das System modular aufgebaut sein, um unterschiedliche Unterstützungsarten, etwa Vorschläge zur Terminologie, Modellkonsistenz oder SGAM-Klassifikation, gezielt und kontextsensitiv bereitstellen und diese Komponenten wiederverwenden zu können. Die Bereitstellung von nachvollziehbaren Vorschlägen mit

erklärenden Hinweisen ist dabei essenziell, um Vertrauen bei den Modellierenden zu schaffen und den XAI-Ansatz zu stärken. Ziel ist nicht die Automatisierung, sondern die qualitätsgesicherte, interaktive Assistenz bei der Modellerstellung zu sein.

### 7.1.2 Einordnung in der Projektlandschaft

Unabhängig vom konkreten Erstellungsprozess wirkt sich die Anwendung der IEC 62559 Use Case Methodik sowie des SGAM Frameworks auf den Gesamtverlauf eines Projekts aus. Eine beispielhafte Einordnung lässt sich anhand des in der Wirtschaftsinformatik sowie in der Praxis etablierten V-Modells vornehmen [AABW19], wie in Abbildung 7.3 dargestellt. Insbesondere in den frühen Phasen mit geringem Detaillierungsgrad (etwa bei der Definition von Systemanforderungen und dem Systementwurf) entfalten diese Methodiken eine zentrale Wirkung, da sie maßgeblich zur Erhebung adäquater Anforderungen beitragen. Werden Anforderungen jedoch fehlerhaft identifiziert oder Akteure, Standards bzw. Technologien unzutreffend berücksichtigt, kann dies in späteren Projektphasen, vor allem im Systementwurf und bei der Ableitung der Systemarchitektur, zu weitreichenden Fehlentscheidungen führen, mit potenziellen Kaskadeneffekten [HSD<sup>+</sup>04]. Auch in den testbezogenen Phasen, die gemäß dem V-Modell typischerweise gegen Ende eines Projekts stattfinden, können sich diese Fehler fortsetzen. Wenn in der Abnahmephase fehlerhafte oder unvollständige IEC 62559 Use Cases sowie SGAM Modelle als Referenz dienen, kann es zu False Negatives in den Tests kommen: Die Tests selbst werden zwar erfolgreich abgeschlossen, übersehen jedoch reale funktionale oder strukturelle Defizite im System die dann in der Laufzeit auftreten. Gekoppelt an jene Phasen, in denen IEC 62559 Use Cases und SGAM Modelle wirksam werden, hat somit auch SGAAIRE einen direkten Einfluss auf die Qualität der Ergebnisse.

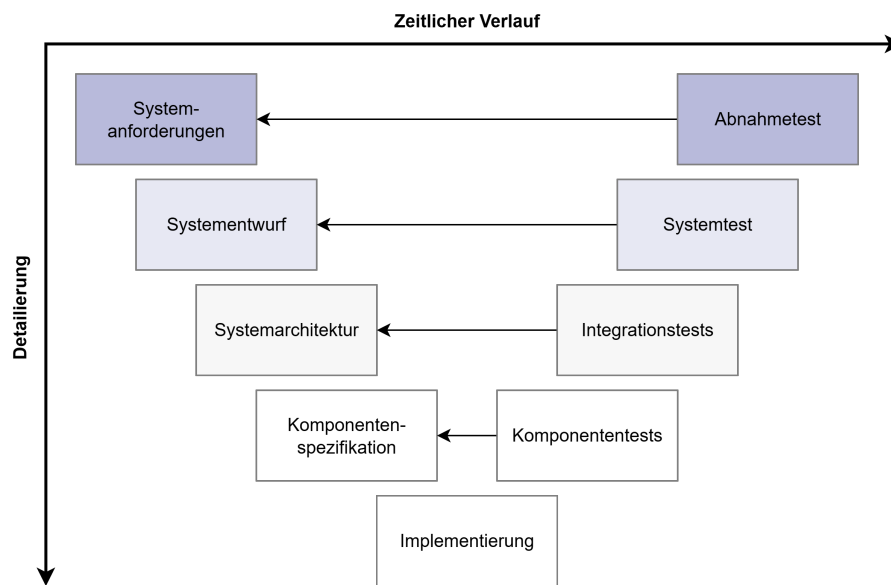
## 7.2 Statisches Systemkonzept

Dieser Abschnitt beleuchtet das statische Systemkonzept von SGAAIRE über die Verwendung von strukturellen Modelle. Das statische Systemkonzept bedient sich an die in Software Engineering etablierte Definition 13 von [Som12].

**Definition 13** (Strukturelle Modelle)

**Strukturelle Modelle** beschreiben die *statische Systemstruktur* in Form von Objektklassen und ihrer Beziehungen. Wichtige Beziehungen, die auf dieser Stufe dokumentiert werden können, sind Verallgemeinerungsbeziehungen (Vererbung), Verwendet-/Wird-verwendet-von- sowie Kompositionsbeziehungen. [Som12]

Unter Berücksichtigung der in Abschnitt 7.1 beschriebenen Abgrenzung sowie des aktuellen Forschungsstands zur Anwendung der IEC 62559 Use Case-Methodik und des SGAM Frameworks im Anforderungsmanagement (vgl. Kapitel 3) wird in Abschnitt 7.2.1 ein Systemkontextmodell entwickelt, das die Integration von SGAAIRE in den Verwendungs- und Erstellungsprozess veranschaulicht. Die Definition 14 des Begriffs



Quelle: Basierend auf [AABW19]

Abbildung 7.3: Schematische Darstellung des V-Modells

„Systemkontextmodell“ gemäß [Som12] wird als Referenz herangezogen. Auf Grundlage des erhobenen Systemkontexts sowie der funktionalen Anforderungen ist die Systemarchitektur für SGAAIRE im Abschnitt 7.2.2 aufgeführt, zur Verdeutlichung der einzelnen Systemkomponenten.

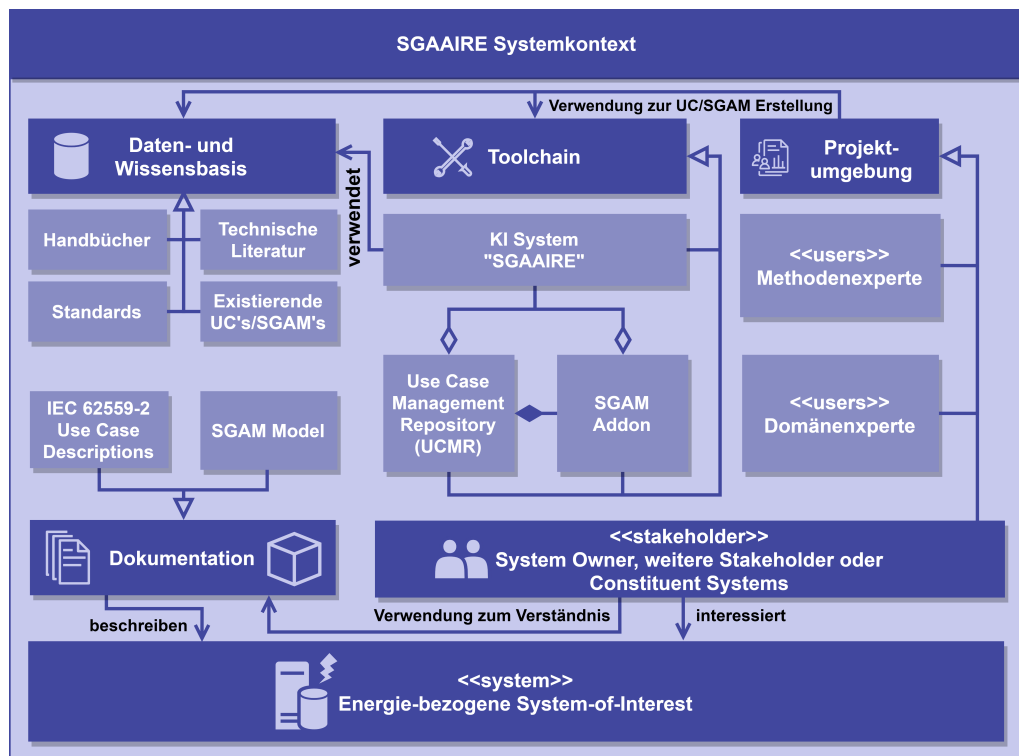
**Definition 14** (Systemkontextmodell)

Ein **Systemkontextmodell** ist ein strukturelles Modell, das die anderen Systeme in dieser Umgebung veranschaulicht. [Som12]

### 7.2.1 Systemkontext

Das in Abbildung 7.4 dargestellte Systemkontextmodell integriert sich in den Prozess der Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen. Diese Integration erfolgt auf Grundlage der in Abschnitt 5.1.3.2 erörterten Umfrageergebnisse, welche im Abschnitt 7.1 eingegrenzt wurden. Hierbei arbeiten Methoden- und Domänenexperten gemeinsam an der Erstellung von Anwendungsfallbeschreibungen und SGAM Modellen unter Verwendung der Toolchain sowie der vorgesehenen Daten- und Wissensbasis, um das System-of-Interest zu verdeutlichen. [KNSS24]

Das Systemkontextmodell illustriert die Integration von SGAAIRE als eine Erweiterung der Toolchain zur Unterstützung bei der Erstellung von IEC 62559-2 Use Case-Beschreibungen und SGAM Modellen. In diesem Zusammenhang wird SGAAIRE als



Quelle: [KNSS24]

Abbildung 7.4: Systemkontextmodell von SGAAIRE

eine Erweiterung des UCMR konzeptualisiert. Die weitverbreitete Nutzung des UCMR in Forschungs- und Entwicklungsprojekten zur Verwaltung von Anwendungsfällen, wie etwa im DISCERN, Redispatch 3.0 oder IDiNA, lässt die Annahme zu, dass das UCMR in der Modellierungspraxis als praktikabel und für Modellierer leicht zugänglich betrachtet werden kann (vgl. die nicht-funktionale Anforderung REQ\_NFA\_2 in Abschnitt 5.2). [KNSS24]

Für die SGAM bezogenen funktionalen Anforderungen ist die Integration eines SGAM Addons UCMR erforderlich. Dies ergibt sich aus der Tatsache, dass der Standard IEC 62559-1:2019 [IEC19] keine Unterstützung für die Speicherung und Verarbeitung von SGAM abseits der Domänen- und Zonenklassifikation vorsieht [KNSS24]. Basierend auf den funktionalen Anforderungen im Bezug auf SGAM wird das UCMR mittels eines SGAM Addons um die folgenden Funktionen erweitert:

- **Verwaltung von SGAM Modellen:** Auf der UCMR-Oberfläche können die Create Read Update Delete (CRUD)-Operationen für SGAM Modelle über das DISCERN SGAM Datenformat vollzogen werden. Diese Funktion ist vorgesehen, um die Anforderung REQ\_SGAM\_1 und REQ\_SGAM\_3 zu realisieren.

- **Zuweisen von SGAM Modellen zu IEC 62559 Use Cases:** SGAM Modelle können IEC 62559 Use Cases (und umgekehrt) zugewiesen werden. Diese Funktion ist vorgesehen, um die Anforderung *REQ\_SGAM.2* zu realisieren.

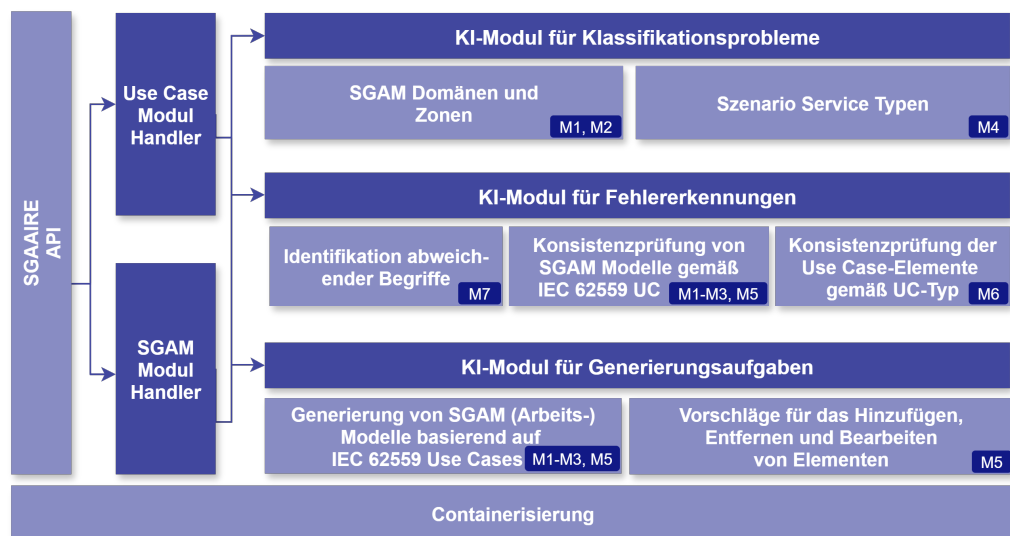
Um KI-Systeme wie SGAAIRE zu realisieren, ist eine geeignete Daten- oder Wissensbasis von entscheidender Bedeutung [KNSS24]. Die in diesem Forschungsprojekt betrachteten KI-Bereiche erfordern alle Daten und/oder Wissen, um KI-Modelle zu erstellen [KNSS24]. Das KI-System kann hierbei auf die von Domänen- und Methodenexperten verwendete verfügbare Daten- und Wissensbasis zurückgreifen [KNSS24]. In einer vorherigen Arbeit [KNSS24] wurden die folgenden Arten von Daten- und Wissensbasen unterschieden:

- **Existierende IEC 62559 Use Case-Beschreibung und SGAM Modelle:** Die Anwendung von KI erfordert die Gewinnung von Regeln und statistischen Erkenntnissen aus Daten aus dem Anwendungsbereich. Diese Daten können Use Case- und SGAM Modelle darstellen und weisen häufig Heterogenität auf sowie möglicherweise Ungenauigkeiten [PKML22]. Um das KI-Lernen zu erleichtern, müssen diese Daten in einem zentralen Datensatz aggregiert werden, unter Berücksichtigung relevanter rechtlicher Aspekte, und sie müssen sorgfältig aufbereitet werden. Der resultierende Datensatz kann zudem für Systemtests und Demonstrationszwecke eingesetzt werden. Solche Daten sind in Projektergebnisdokumenten sowie in öffentlichen UCMR-Instanzen verfügbar, wie zum Beispiel im BRIDGE-Repository [DAZ<sup>+</sup>23b]. [KNSS24]
- **Standards, Handbücher und technische Literatur:** Die Integration von Standards, Handbüchern und technischer Literatur aus dem Energiebereich ist ein entscheidender Ansatz, um das Kontextwissen zu nutzen und ein grundlegendes Rahmenwerk für die Implementierung des KI-Systems bereitzustellen. Durch bewährte Praktiken, Roadmaps und Übersichtsarbeiten, wie sie in [URMG<sup>+</sup>11] und [USD<sup>+</sup>13] beschrieben sind, lassen sich relevante Standards, Handbücher und technische Literatur im Energiekontext ermitteln. [KNSS24]
- **Erarbeitete Regeln aus weiteren Arbeiten** Im Energiekontext wurden mehrere Arbeiten durchgeführt, um Akteure zu identifizieren, Klassifizierungen vorzunehmen und einen allgemeinen Rahmen zu etablieren. Diese Arbeiten können als Grundlage zur Ableitung von Regeln für die Entwicklung von KI-Modellen berücksichtigt werden. Solche Arbeiten sind z. B. das ENTSO-E Harmonised Role Model [ENT22], der Standard IEC 62357-1 [IEC16] oder die Betrachtung des amerikanischen Cybersecurity NISTIR 7628 Standard in das SGAM Framework [URS14]. [KNSS24]

## 7.2.2 Systemarchitektur

Abbildung 7.5 zeigt die SGAAIRE Systemarchitektur, die die funktionalen Anforderungen integriert, welche empirisch erhoben und in Abschnitt 5.3 detailliert dargestellt sind. Diese Anforderungen sind problemspezifischen Aufgaben zugeordnet und logisch innerhalb von KI-Modulen organisiert. Im Rahmen dieser Arbeit umfasst ein KI-Modul sowohl KI- bzw. ML-Modelle als auch prozeduralen Code. Diese Designentscheidung wird insbesondere bei der Erstellung von SGAM Modellen aus bestehenden IEC 62559 Use Cases deutlich. Während bestimmte Elemente dieser Beschreibungen deterministisch extrahiert werden können (bspw. Akteure), ist für die vollständige Generierung eines SGAM Modells ein KI-gestützter Ansatz erforderlich, insbesondere aufgrund von fehlenden Kommunikationsdetails, fehlende Informationen in der IEC 62559 Use Case-Beschreibung sowie der Notwendigkeit zur Domänen- und Zonenklassifikation die nicht im Template vorgesehen ist. [KLS25]

Das System ist dementsprechend so konzipiert, dass es sowohl die Arbeit mit der IEC 62559 Use Case-Methodik als auch das SGAM Framework unterstützt. Die Architektur verfügt über eine offen zugängliche HTTP-API, welche die Nutzbarkeit des Systems verbessert [KLS25]. Mit Fokus auf Portabilität, Wartbarkeit und schnelle Bereitstellung integriert die Architektur zudem Containerisierungskonzepte [KLS25]. Im folgenden werden die Komponenten der Systemarchitektur detaillierter beleuchtet:



Quelle: Übersetzt aus [KLS25]

Abbildung 7.5: Systemarchitektur von SGAAIRE

- **SGAAIRE API:** Die SGAAIRE API dient als externe Schnittstelle, welche zur Verwendung von SGAAIRE in Tools, wie dem UCMR, integriert werden kann. Die über die API entgegengenommen Anfragen werden an die Handler weitergeleitet.

- **Use Case- und SGAM Modul Handler:** Die Handler sind verantwortlich für die Interpretation von Anfragen, Aufrufen von KI-Modulen und die Aufbereitung von Antwortdaten. Bei der Interpretation wird zunächst die syntaktische Richtigkeit der im Payload enthaltenen IEC 62559-3 XML-serialisierten Use Cases oder DISCERN-Datenformat XML-serialisierten SGAM Modelle überprüft. Aus dieser Überprüfung können (Meta-)Daten extrahiert werden, wie beispielsweise der Use Case-Typ oder der Grad der Vollständigkeit, die für die Nutzung der KI-Module erforderlich ist. Da eine Anfrage mehrere Aufrufe von KI-Modulen enthalten kann, werden die Ergebnisse dieser Aufrufe in einer Antwort auf die Anfrage zusammengefasst und an die SGAAIRE API zurückgegeben.
- **KI-Module:** Die Geschäftslogik des Systems ist in den KI-Modulen verankert, die den funktionalen Anforderungen aus Abschnitt 5.3 zugeordnet sind. Während der Konzeption und Entwicklung der individuellen Module werden die Konzepte der im Kapitel 4 beschriebenen künstlichen Intelligenz angewendet. Eine vertiefte Darstellung der Konzeption der KI-Modelle wird in dem Abschnitt 7.4 behandelt.
- **Containerisierung:** Für die Bereitstellung des Systems auf beliebige Infrastrukturen wird eine Software Containerisierung zur Hand gezogen.

### 7.3 Dynamisches Systemkonzept

Dieser Abschnitt beleuchtet das dynamische Systemkonzept von SGAAIRE über die Verwendung von UML-Diagrammen. Das dynamische Systemkonzept bedient sich der im Software Engineering etablierten Definition 15 von [Som12].

**Definition 15** (Dynamische Modelle)

**Dynamische Modelle** beschreiben die *dynamische Systemstruktur* und zeigen das Zusammenspiel zwischen den Systemobjekten. Interaktionen, die dokumentiert werden können, sind zum Beispiel die Abfolge von Dienstanfragen durch Objekte sowie Zustandsänderungen, die von diesen Objektinteraktionen angestoßen werden. [Som12]

Diese Definition umfasst die Modellierung von dynamischen Aspekten wie Interaktionen und Zustandsänderungen. Durch die spezifizierten funktionalen Anforderungen (vgl. Abschnitt 5.3) sowie die nicht-funktionalen Anforderungen (vgl. Abschnitt 5.2) für SGAAIRE können sowohl Interaktionen mit externen Systemen als auch mit den Nutzern identifiziert werden. Um diese Interaktionen zu modellieren, bietet sich die Verwendung sowohl von Anwendungsfällen als auch Sequenzmodellen an, die mithilfe von UML-Anwendungsfalldiagrammen und -Sequenzdiagrammen erstellt werden können [Som12]. Beide Modellarten stellen Interaktionsmodelle gemäß Definition 16 dar. Da keine Zustandsänderungen aufgrund der vortrainierten KI-Modelle während der Laufzeit des Systems auftreten, ist die Modellierung von Zustandsänderungen nicht erforderlich.

**Definition 16** (Interaktionsmodell)

Ein **Interaktionsmodell** ist ein dynamisches Modell, das zeigt, wie das System während der Benutzung mit seiner Umwelt zusammenspielt. [Som12]

## 7.3.1 Anwendungsfälle

Das UML-Anwendungsfalldiagramm, welches in Abbildung 7.6 visualisiert ist, integriert die funktionalen Anforderungen, die im Abschnitt 5.3 ausgewählt wurden, als Basis für die Anwendungsfälle. Zudem werden die Eigenschaften der nicht-funktionalen Anforderungen aus Abschnitt 5.2 berücksichtigt. Dabei ist vorgesehen, dass das System ausschließlich KI-basierte Anwendungsfälle verwendet (*REQ\_NFA\_1*), als eigenständiges Softwaresystem realisiert wird (*REQ\_NFA\_3*), und die Nutzerunterstützung durch das Aufzeigen von Vorschlägen gewährleistet (*REQ\_NFA\_5*).

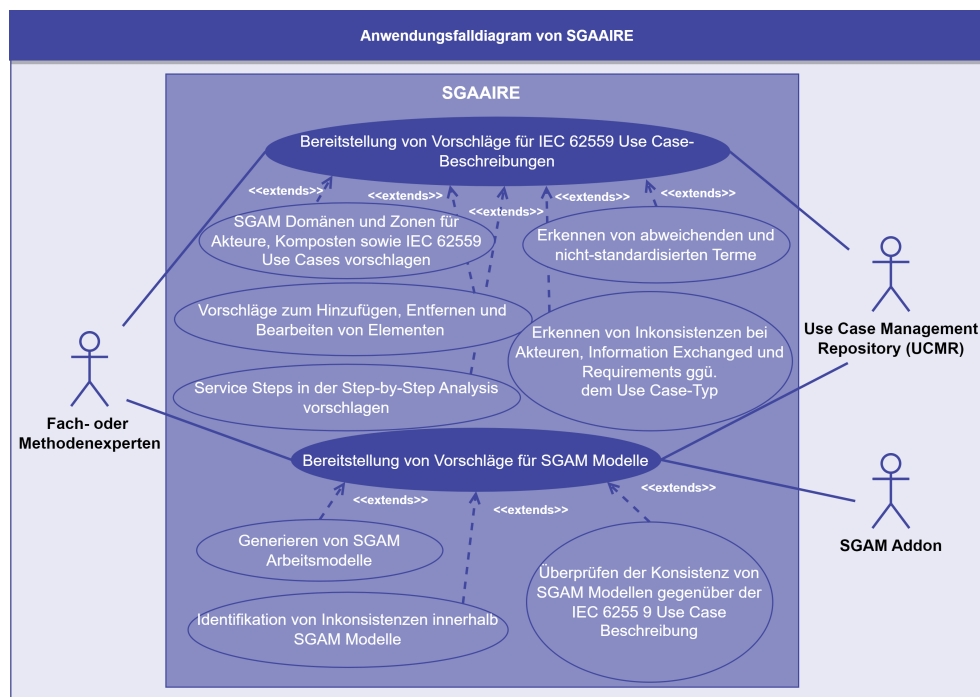


Abbildung 7.6: UML-Anwendungsfalldiagramm für die Interaktionen mit SGAAIRE für die IEC 62559 Use Case-bezogenen funktionalen Anforderungen

Für sämtliche im Diagramm aufgeführten Anwendungsfälle wurde im Einklang mit den funktionalen Anforderungen eine detaillierte Beschreibung der Anforderungen erstellt. Dies erfolgte mittels des bereitgestellten Vorlagenformats von Cockburn [Coc01]. Darüber hinaus wurden potenzielle Ausgaben innerhalb klar definierter Rahmenbedingungen durch Verweise auf Abbildungen im Anhang veranschaulicht. Die entsprechenden Referenzen können in Tabelle 7.2 nachgelesen werden.

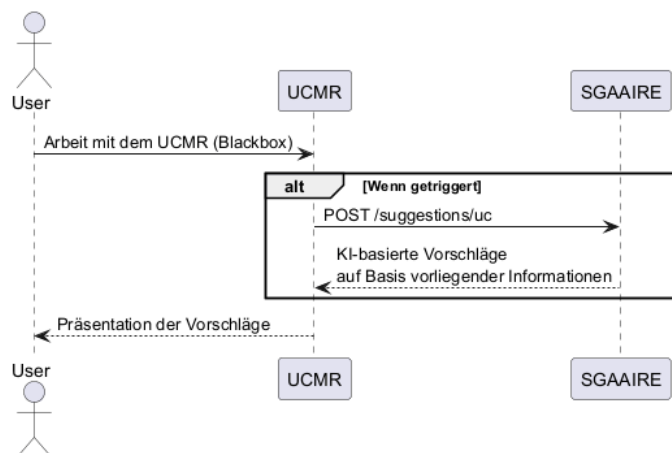
Tabelle 7.2: Referenzen der Anwendungsfallbeschreibungen sowie Abbildungen für den Use Case-Teil

Anwendungsfall	Funktionale Anforderung	Referenz zur Anwendungsfallbeschreibung und verdeutlichenden Abbildungen
1. Bereitstellung von Vorschläge für IEC 62559 Use Case-Beschreibungen	<i>Alle UC-bezogenen Anforderungen</i>	Tabelle A.46 (S. 259)
1.1 SGAM Domänen und Zonen für Akteure, Komponenten sowie IEC 62559 Use Cases vorschlagen	RQ_UC_4 und RQ_UC_9	Tabelle A.47 (S. 260)
1.2 Erkennen von abweichenden und nicht-standardisierten Terme	RQ_UC_2	Tabelle A.48 (S. 260)
1.3 Vorschläge von Hinzufügen, Entfernen und Bearbeiten von Elementen	RQ_UC_1	Tabelle A.49 (S. 261)
1.4 Service Steps in der Step-by-Step Analysis vorschlagen	RQ_UC_11	Tabelle A.50 (S. 261)
1.5 Erkennen von Inkonsistenzen bei Akteuren, Informationsaustausche und Anforderungen ggü. dem Use Case-Typ	RQ_UC_7	Tabelle A.51 (S. 262)
2. Bereitstellung von Vorschläge für SGAM Modelle	<i>Alle SGAM-bezogenen Anforderungen</i>	Tabelle A.52 (S. 263)
2.1 Generieren von SGAM Arbeitsmodelle	RQ_SGAM_1	Tabelle A.53 (S. 264)
2.2 Überprüfen der Konsistenz von SGAM Modellen gegenüber der IEC 62559 Use Case-Beschreibungen	RQ_SGAM_2	Tabelle A.54 (S. 264)
2.3 Identifikation von Inkonsistenzen innerhalb SGAM Modelle	RQ_SGAM_3	Tabelle A.55 (S. 265)

### 7.3.2 Interaktionen

Die Darstellung in Abbildungen 7.7 und 7.8 veranschaulichen die Interaktionen mit SGAAIRE im Kontext der an die IEC 62559 Use Case- sowie SGAM-bezogenen Interaktionen mit dem UCMR. Das Konzept sieht keine feste Integration für SGAAIRE innerhalb eines UCMR vor, sondern eine entkoppelte Interaktion über Schnittstellenkommunikation. Im Falle von SGAAIRE erfolgt die Kommunikation über eine webbasierte HTTP-API. Diese architektonische Wahl bewirkt den Vorzug, SGAAIRE nicht auf eine spezifische UCMR-Instanz zu beschränken, sondern erlaubt ihre anpassungsfähige Integration in diverse Werkzeuge und Prozesse gemäß Nutzerentscheidungen.

Für die Verwendung von SGAAIRE innerhalb eines UCMR bedarf es der Implementierung folgender essenzieller Aspekte für die IEC 62559 Use Case-bezogenen Anforderungen:

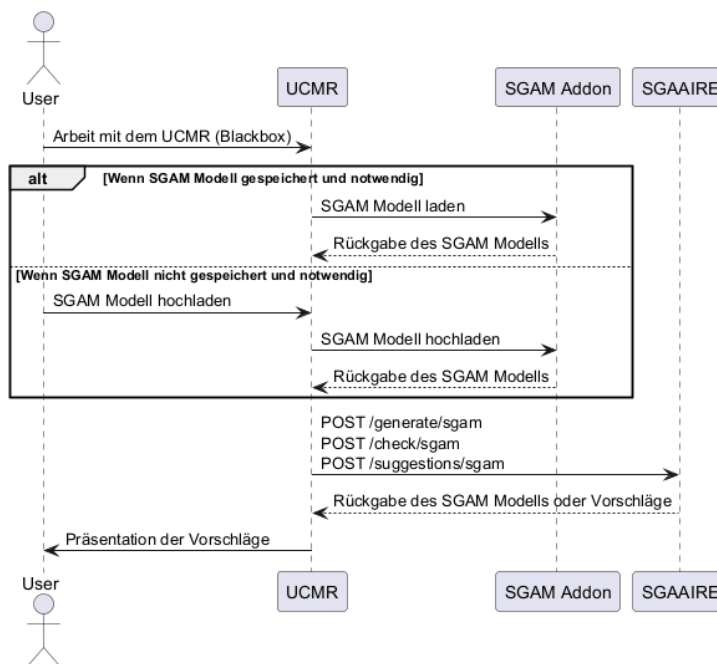


Quelle: Eigene Darstellung

Abbildung 7.7: Sequenzdiagramm für die Interaktionen mit SGAAIRE für die IEC 62559 Use Case-bezogenen funktionalen Anforderungen

- **Setzen eines Auslösers für den Aufruf von SGAAIRE:** Mittels eines Auslösers muss die Initiierung der Interaktion mit SGAAIRE ermöglicht werden. Hierfür bieten sich verschiedene Optionen an, die über das UCMR vorgesehene Erstellungs Oberfläche von IEC 62559 für Use Cases integriert werden können. Eine Möglichkeit besteht darin, eine *direkte Nutzeraufforderung* (z. B. über eine Schaltfläche) zu nutzen. Darüber hinaus können *indirekte Nutzeraufforderungen*, die sich aus den Regelungen des UCMR ergeben (z. B. basierend auf eingegebenen Informationen auf der Erstellungs Oberfläche) als Auslöser der Kommunikation dienen.
- **Integrierung der technischen SGAAIRE-Kommunikation:** Das UCMR muss eine technische Kommunikation mit SGAAIRE aufbauen. Zu diesem Zweck empfängt SGAAIRE Anfragen über die HTTP POST Schnittstelle, die im Payload die XML-Serialisierung von Use Cases gemäß IEC 62559-3 Standard enthält. Für das Interpretieren der HTTP Antworten ist ein spezifisches Datenmodell vorgesehen, das im Abschnitt 8.1 basierend auf den Ausgaben der implementierten Anforderungen konstruiert ist.
- **Präsentation der Vorschläge:** Das UCMR muss die von SGAAIRE erstellten Vorschläge dem Nutzer auf eine benutzerfreundliche Art und Weise präsentieren. Des Weiteren sollen die Optionen zur Annahme oder Ablehnung dieser Vorschläge bereitgestellt werden.

Die SGAM-bezogenen Anforderungen sehen im Vergleich zu den IEC 62559 Use Case-bezogenen Anforderungen eine separate Form der Integration vor. In diesem Kontext wird das Konzept des SGAM Addons eingeführt, wie es in Abschnitt 7.2.1 beschrieben wurde. Die Generierung von SGAM-Arbeitsmodellen gemäß der Anforderung



Quelle: Eigene Darstellung

Abbildung 7.8: Sequenzdiagramm für die Interaktionen mit SGAAIRE für die SGAM-bezogenen funktionalen Anforderungen

zung *REQ\_SGAM\_1*, die Überprüfung der Konsistenz von SGAM Modellen mit der zugehörigen IEC 62559 Use Case-Beschreibung gemäß *REQ\_SGAM\_2* sowie die Erkennung allgemeiner Inkonsistenzen gemäß *REQ\_SGAM\_3* werden direkt vom Nutzer aufgerufen. Um die SGAM-bezogenen Anforderungen von SGAAIRE in einem UCMR effizient nutzen zu können, sind folgende Funktionen zu integrieren:

- **Integrierung der SGAM Addon-Konzept:** Durch die Implementierung des SGAM Addon-Konzepts können die definierten SGAM-Aspekte innerhalb des UCMRs integriert werden.
- **Integrierung der technischen SGAAIRE-Kommunikation:** Das UCMR muss eine technische Kommunikation mit SGAAIRE aufbauen. Zu diesem Zweck empfängt SGAAIRE Anfragen über HTTP POST-Schnittstellen. Die Anfragen enthalten XML-Serialisierungen von Use Cases gemäß dem IEC 62559-3-Standard sowie von SGAM Modellen gemäß dem DISCERN SGAM-Datenformat. Für das Interpretieren der HTTP Antworten ist ein spezifisches Datenmodell vorgesehen, das im Abschnitt 8.1 basierend auf den Ausgaben der implementierten Anforderungen konstruiert ist.
- **Präsentation der Vorschläge:** Das UCMR muss die von SGAAIRE erstellten Vorschläge dem Nutzer auf eine benutzerfreundliche Art und Weise präsentieren.

## 7.4 Konzeptualisierung der Anwendungsfälle

Dieser Abschnitt beschreibt die Konzeption derjenigen KI-Module, die gemäß den funktionalen Anforderungen der IEC 62559 Use Cases sowie der SGAM Modelle in der Systemübersicht als Anwendungsfälle vorgesehen sind (vgl. Abschnitt 7.3.1). Die folgenden Unterabschnitte analysieren für jedes KI-Modul und dessen zugehörigen Anwendungsfällen, wie die funktionale Umsetzung auf konzeptioneller Ebene realisiert werden kann. Dabei wird insbesondere betrachtet, in welchen Fällen eine prozedurale Umsetzung ausreichend ist und wo der Einsatz KI-gestützter Methoden zur Komplexitätsbewältigung erforderlich oder vorteilhaft erscheint. Die konkrete Modellkonzeption und Auswahl geeigneter KI-Modelle erfolgt im Anschluss in Abschnitt 7.5.

### 7.4.1 KI-Modul für Klassifikationsprobleme

Das KI-Modell für Klassifikationsprobleme zielt darauf ab, Eingaben (Instanzen) automatisch einer *endlichen Menge von Klassen* zuzuordnen. Dabei generiert oder liefert das Modell Vorschläge, die jeweils einer dieser vorgegebenen Klassen entsprechen. Im Wesentlichen werden Vorschläge im Rahmen der SGAM Domänen und Zonen betrachtet sowie für Service-Typen in der Step-by-Step Analyse des IEC 62559 Use Case-Template.

#### 7.4.1.1 Vorschläge für SGAM Domänen und Zonen

Im Rahmen der Modellierung von Use Cases gemäß IEC 62559 sowie der SGAM Modellierung ist eine systematische Einordnung in die Domänen und Zonen des SGAM Frameworks erforderlich (vgl. Abschnitt 3.3.1). Auf der Ebene der IEC 62559 Methodik betrifft dies insbesondere die Klassifikation von Akteuren, Informationsobjekten, Standards sowie der Anwendungsfall selbst [IEC15]. Die Struktur des SGAM Frameworks, das eine dreidimensionale Modellierung vorsieht, erzwingt darüber hinaus eine explizite räumliche und funktionale Einordnung dieser Elemente in den beiden Achsen [CEN12]. An dieser Stelle setzt das KI-Modul an, indem es sowohl bei der initialen Einordnung als auch bei der Validierung bestehender SGAM Domänen und Zonen-Zuordnungen unterstützend wirkt. Ziel ist es, potenziell fehlerhafte Klassifikationen zu erkennen und so falsche Systemannahmen zu vermeiden, die sich aus einer inkorrekten Übertragung vom *System-of-Interest* in das *System-of-Systems* ergeben könnten.

Eine präzise Einordnung ermöglicht im Weiteren auch die Ableitung von normativen und regulatorischen Anforderungen anhand der jeweiligen SGAM-Kategorie. Arbeiten wie die Seamless Integration Architecture (SIA) [IEC16] oder die *IEC Standards Map* [IEC22] referenzieren auf das SGAM Framework und stellen Standards und Rahmenbedingungen einzelner Akteure dar.

Das vorgesehene Modul besteht aus zwei voneinander unabhängigen Textklassifikationsmodellen auf Basis von (L)LMs. Diese Modelle analysieren jeweils die in der IEC 62559 Use Case-Beschreibung oder im SGAM Modell enthaltenen Informationen, um

geeignete Klassifikationsvorschläge abzuleiten. Im Einzelnen handelt es sich um den *M1: SGAM Domänen-Klassifikator* (vgl. Abschnitt 7.5.1) sowie den *M2: SGAM Zonen-Klassifikator* (vgl. Abschnitt 7.5.2).

#### 7.4.1.2 Vorschläge für Services in Scenario Steps

Innerhalb des IEC 62559 Use Case-Templates werden im Feld *Service* Angaben zum Informationsfluss für die einzelnen Schritte eines Szenarios gemacht. Dieses Feld beschreibt, in welcher Form Informationen innerhalb einer Aktivität verwendet, verarbeitet oder übertragen werden, etwa ob Informationen angefragt, gepusht, periodisch übermittelt, gelöscht oder modifiziert werden (vgl. Abschnitt 3.2.1.1). An diesem Punkt setzt das KI-Modul an, indem es sowohl bei der initialen Einordnung als auch bei der Validierung bestehender Service-Klassifikationen unterstützend wirkt. Ziel ist es, potenzielle Fehlklassifikationen zu identifizieren, um dadurch fehlerhafte Systemannahmen zu vermeiden, die sich aus einer inkorrekten Überführung vom *System-of-Interest* in das *System-of-Systems* ergeben könnten.

Insbesondere für Modellierende die nicht den Hintergrund der Informationsmodellierung haben, die mit der IEC 62559 Methodik arbeiten, kann eine automatisierte Klassifikation eine Unterstützung darstellen. Häufig werden in der Praxis die feinen Unterschiede zwischen den verschiedenen Service-Typen nicht eindeutig erkannt oder aus einer fachlich abweichenden Perspektive interpretiert [KNN<sup>+</sup>24]. Ein KI-gestütztes Modul kann hier dazu beitragen, die Konsistenz und semantische Genauigkeit der Service-Angaben zu erhöhen.

Das vorgesehene Modul basiert auf einem Textklassifikationsmodell unter Verwendung von (L)LMs. Es analysiert die jeweils vorliegenden Beschreibungen eines Szenarioschritts bzw. einer Aktivität und leitet daraus passende Klassifikationsvorschläge für den Service-Typ ab. Das entsprechende Modell wird als *M4: Service Step-Klassifikator* vorgesehen (vgl. Abschnitt 7.5.4).

#### 7.4.2 KI-Modul für Fehlererkennungen

Das KI-Modul zur Fehlererkennung verfolgt das Ziel, für gegebene Eingaben (Instanzen) eine Liste von Empfehlungen variabler Länge zurückzugeben. Obwohl dem Modul im Kern ebenfalls Klassifikationsmodelle zugrunde liegen, werden diese als Teilaspekte eingebettet und durch fest kodierte, prozedurale Logiken ergänzt, um die gewünschte Funktionalität abzubilden. Im Wesentlichen umfasst das Modul drei zentrale Aufgabebereiche: Die Identifikation abweichender oder nicht-standardisierter Terme, die Überprüfung der Konsistenz von SGAM Modellen im Vergleich zu einem vorliegenden IEC 62559 Use Case sowie die Detektion von Inkonsistenzen in Bezug auf den jeweils deklarierten Use Case-Typ.

### 7.4.2.1 Erkennung von abweichenden und nicht-standardisierten Termen

Der Begriff „Term“ bezieht sich auf jegliche Wörter, weswegen für genauere Betrachtung des Anwendungsfalls es notwendig ist, den Begriff „Term“ auf das standardisierte Template einzugrenzen. Das IEC 62559-2 Use Case Template sieht in Abschnitt 7 („Common terms and definitions“) vor, dass die relevanten Terme für den Anwendungsfall in diesem Abschnitt aufgeführt werden sollen und diese automatisiert mithilfe der vorliegenden Informationen erstellt werden [IEC15]. Die aufgeführten Terme repräsentieren primär Akronyme, wie im Beispielteil des Standards veranschaulicht ist [IEC15]. Basierend auf dieser Eingrenzung sieht das Modul die folgenden Ziele vor:

1. **Akronym-Erkennung und Langform-Zuordnung:** Das Modul soll Akronyme erkennen und überprüfen, ob diesen eine ausgeschriebene Langform zugeordnet ist. Die Langform kann sowohl direkt im Fließtext als auch in den verschiedenen Elementen des Use Cases (z. B. im Akteursteil, Glossar, KPI-Bereich) definiert sein.
2. **Vorschläge bei fehlender Definition:** Das Modul soll dem Nutzer bei fehlender Langform-Definition mögliche Vorschläge unterbreiten. Diese Vorschläge können vom Anwendenden dann im Glossar aufgenommen werden.
3. **Erkennung von inkorrekten Definitionen:** Zusätzlich soll das Modul die Funktionalität bereitstellen, inkorrekte Definitionen zu erkennen, um den Anwender auf mögliche Fehler aufmerksam zu machen.

Die Bereiche der automatisierten Erkennung von Akronymen (Acronym Identification (AId)) sowie der Ermittlung ihrer korrekten kontextuellen Bedeutung (Acronym Disambiguation (AD)) werden in der Fachliteratur diskutiert [VDTN20] und erfahren speziell im biomedizinischen Sektor erhöhte Aufmerksamkeit [SH03]. In diesen Zusammenhängen können sowohl konventionelle Ansätze wie die Anwendung RegExs oder Algorithmen als auch Verfahren der Künstlichen Intelligenz (insbesondere Feature Extraction und Deep Learning) eingesetzt werden [SH03] [VDTN20]. Die Ergebnisse diverser Modelle verdeutlichen in der Untersuchung von [VDTN20], dass sowohl herkömmliche als auch KI-Ansätze vergleichbare Resultate bei der automatisierten Erkennung von Akronymen (AId) und der Klärung ihrer kontextuell richtigen Bedeutung (AD) erzielen. Dennoch wird aufgezeigt, dass durch den Einsatz von KI-Ansätzen die Metriken für Recall und F1-Score verbessert werden können [VDTN20]. Durch die gegebene nicht-funktionale Anforderung *REQ\_NFA\_1* gilt es im Konzept die Verfahren aus dem KI-Bereich zu verwenden.

Das zu entwerfende Modul sieht vor, die *Acronym Identification (Modul M7a, vgl. Abschnitt 7.5.7)* und *Acronym Disambiguation (Modul M7b, vgl. Abschnitt 7.5.8)* einzusetzen, um die im Abschnitt benannten Ziele zu erreichen. Dieses Konzept wird in Abbildung 7.9 dargestellt. Akronyme und deren Definitionen werden durch das AId-Model bestimmt, indem es jedes Freitextfeld der IEC 62559 Use Case-Beschreibung

durchläuft. Die gewonnenen Informationen werden im Acronym Dictionary gespeichert. Darüber hinaus werden benutzerdefinierte Elemente wie Akteure oder Informationsobjekte in das Acronym Dictionary aufgenommen. Das Acronym Dictionary soll bereits eine Vielzahl von Abkürzungen aus bestehenden Verzeichnissen enthalten. Das Acronym Disambiguation-Model hat die Aufgabe, fehlende sowie falsche Definitionen für einen bestimmten Term anhand des Kontextes sowohl des zugehörigen Textelements als auch der IEC 62559 Use Case-Beschreibungen zu identifizieren.

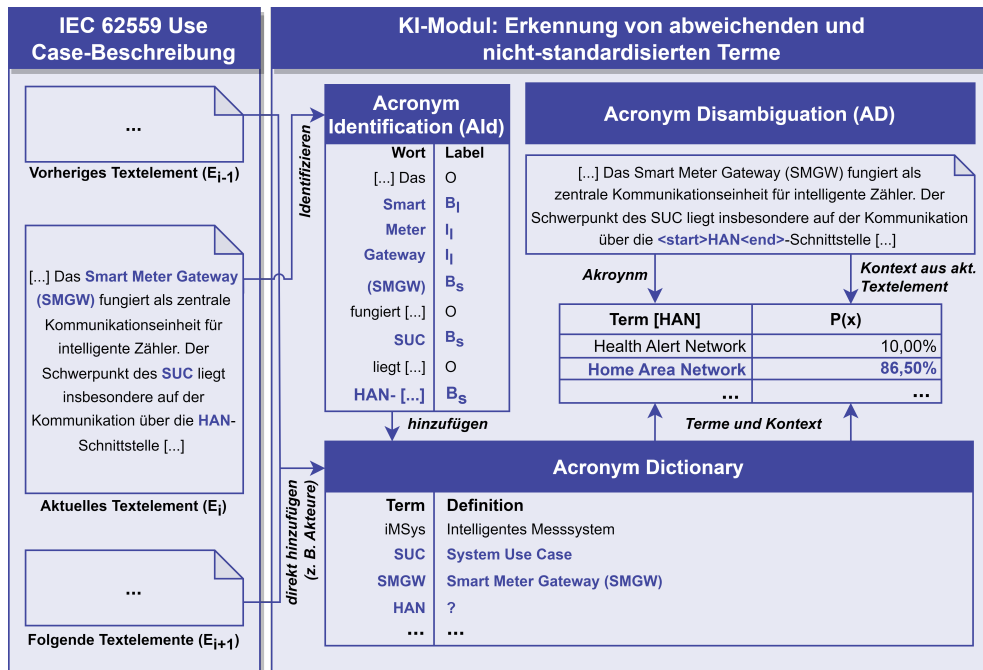


Abbildung 7.9: Visualisierung des Konzeptes für die Verwendung von Acronym Identification (AId) und Acronym Disambiguation (AD) für das KI-Modul zur Erkennung von abweichenden und nicht-standardisierten Termen

#### 7.4.2.2 Überprüfung der Konsistenz von SGAM Modellen gegenüber IEC 62559 Use Cases

Bei der Erstellung eines SGAM Modells sollte sichergestellt werden, dass dieses inhaltlich konsistent zur zugehörigen IEC 62559 Use Case-Beschreibung ist. Einzelne Aspekte der Use Case-Beschreibung, etwa Objectives, KPIs, Akteure, Informations- und Kommunikationsflüsse, sollen sich strukturell und semantisch im SGAM Modell widerspiegeln. Das Ziel des KI-Moduls besteht darin, gezielte Hinweise und Empfehlungen bereitzustellen, um die Übereinstimmung zwischen SGAM Modell und IEC 62559 Use Case-Beschreibung sicherzustellen oder zu verbessern. Darüber hinaus können die generierten Empfehlungen auch dazu beitragen, inkonsistente oder unvollständige Angaben innerhalb der IEC 62559 Use Case-Beschreibung selbst zu identifizieren und somit als

Rückkopplung in den Modellierungsprozess einfließen.

Der in diesem Konzept verfolgte Ansatz basiert auf der Generierung eines SGAM Modells aus der zugrundeliegenden IEC 62559 Use Case-Beschreibung. Dieses automatisch abgeleitete Modell dient als *Referenzmodell* im Rahmen des anschließenden Vergleichs mit einem manuell erstellten SGAM Modell (*Evaluationsmodell*). Ein konsistentes SGAM Modell sollte im Idealfall mit dem Referenzmodell inhaltlich übereinstimmen, abgesehen von rein sprachlichen Variationen (z. B. Synonyme, unterschiedliche Benennungen mit gleicher Bedeutung oder generelle textuelle Abweichungen). Die methodische Herleitung und technische Umsetzung der SGAM Modellgenerierung auf Basis eines IEC 62559 Use Cases wird ausführlich in Abschnitt 7.4.3.1 dargestellt.

Bei dem vorgeschlagenen Vergleichsansatz werden alle fünf Interoperabilitätsschichten separat betrachtet. Für jede Schicht wird geprüft, inwiefern die vorhandenen Elemente inhaltlich und strukturell mit denjenigen übereinstimmen, die aus dem Use Case erwartbar wären. Das Verfahren berücksichtigt dabei die folgenden Typen von Inkonsistenzen:

- **Strukturelle Inkonsistenzen:** Elemente, die im SGAM-Modell vorhanden sind, jedoch keine Entsprechung im Use Case besitzen, oder umgekehrt im SGAM Modell fehlen. Ein abstraktes Beispiel ist in Abbildung 7.10a dargestellt. Im Beispiel sind die Elemente *B*, *D* und *E* sowohl im IEC 62559 Use Case als auch im SGAM Modell vorhanden, während *A* ausschließlich im SGAM Modell und *C* ausschließlich im IEC 62559 Use Case enthalten ist.
- **Klassifikatorische Abweichungen:** Elemente, deren Domänen- oder Zonenklassifikation nicht mit den aus dem Use Case abgeleiteten Einordnungen übereinstimmt. Ein abstraktes Beispiel ist in Abbildung 7.10b dargestellt. Im Beispiel ist das Element *B* sowohl im IEC 62559 Use Case als auch im SGAM Modell enthalten. Es wurde jedoch im IEC 62559 Use Case (im UCMR) der Domäne *Customer* zugeordnet, während es im SGAM Modell der Domäne *DER* zugeordnet ist.
- **Verbindungstechnische Inkonsistenzen:** Verbindungen zwischen Elementen, die entweder fehlen, überzählig sind oder semantisch nicht plausibel erscheinen. Ein abstrakteres Beispiel ist in Abbildung 7.10c dargestellt. Darin sind die Elemente *A* und *B* sowohl im SGAM Modell als auch gemäß der IEC 62559 Use Case-Beschreibung korrekt miteinander verbunden. Die Verbindung zwischen *B* und *C* fehlt jedoch im SGAM Modell, obwohl sie laut Use Case erforderlich wäre, während die Verbindung zwischen *A* und *C* im SGAM Modell vorhanden ist, jedoch aus Sicht der IEC 62559 Use Case-Beschreibung redundant ist.

Durch die Verwendung der Funktionalität zur Generierung von SGAM Modellen kommen die Modelle *M1: SGAM Domanen Klassifikator*, *M2: SGAM Zonen Klassifikator*, *M3: SGAM Interoperabilitatslayer Klassifikator* sowie *M5: SGAM Element Proposer* zum Einsatz. Um rein sprachlich-syntaktische Unterschiede (etwa unterschiedliche Benennungen bei gleicher Bedeutung) nicht als echte Inkonsistenzen zu werten, wird zusätzlich das Pre-trained Language Model RoBERTa [LOG<sup>+</sup> 19] eingesetzt. Auf dieser

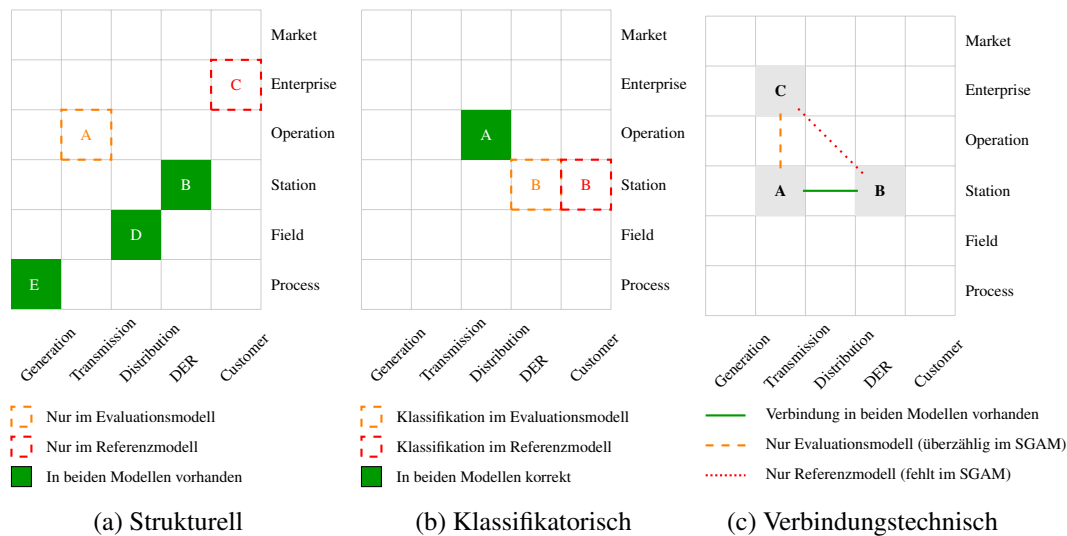


Abbildung 7.10: Beispiele von Inkonsistenzen zw. SGAM Modell und IEC 62559 UC

Basis wird der *BERTScore* [ZKW<sup>+</sup>20] zwischen den Elementbezeichnungssequenzen des Evaluations- und Referenzmodells berechnet. Eine Konstante als Schwellwert (z. B. > 0,9) dient dazu, semantisch gleichwertige, aber sprachlich variierende Bezeichnungen zu erkennen und entsprechend als konsistent zu klassifizieren.

### 7.4.2.3 Erkennung von Inkonsistenzen gegenüber dem Use Case-Typ

Die Unterscheidung bestimmter Use Case-Typen, etwa (generischere) High Level Use Cases oder (konkretere) Low Level Use Cases, ist in verschiedenen Projekten explizit gefordert (z. B. in TDX-Assist [LCR<sup>+</sup>19]). Vor diesem Hintergrund soll das KI-Modul überprüfen, ob die im IEC 62559 Use Case verwendeten Elemente mit dem im Feld „Nature of the use case“ (Abschnitt 1.7 des Templates) deklarierten Use Case-Typ konsistent sind. Ziel des Moduls ist es, potenzielle Inkonsistenzen zwischen der strukturellen Granularität der Modellinhalte und dem deklarierten Use Case-Typ zu erkennen. Den Nutzenden sollen daraufhin gezielte Hinweise gegeben werden, ob Anpassungen entweder auf Ebene einzelner Use Case-Elemente oder in der übergeordneten Anwendungsfallbeschreibung erforderlich sind.

Das dem SGAM Framework zugrundeliegende Konzept der Abstraktionsebenen wurde bereits in Abbildung 3.13 erläutert und lässt sich über bestehende Mappings auf IEC 62559 Use Cases gemäß Abbildung 3.14 anwenden. Das KI-Modul operationalisiert diese Mappings, indem es Business Use Cases, konzeptionelle Use Cases und High Level Use Cases einer *höherabstrakten Ebene* zuordnet, während System Use Cases, Technical Use Cases sowie Low Level Use Cases der *feingranulareren Ebene* zugeordnet werden. Ein High-Level Use Case sollte gemäß dem zugrunde liegenden Konzept überwiegend Elemente enthalten, die höheren Abstraktionsebenen zuzuordnen sind, analog dazu sollte ein Low-Level Use Case vorwiegend aus niedriggranularen Elementen bestehen. Bei

festgestellten Inkonsistenzen weist das KI-Modul die Nutzenden gezielt darauf hin, dass die gewählte Ausprägung einzelner Elemente sowie die Angabe des Use Case-Typs möglicherweise nicht konsistent bzw. nicht adäquat gewählt wurden.

Die Einordnung des Abstraktionsniveaus erfolgt konzeptuell über den Einsatz eines separaten KI-Modells, dem *M6: Abstraktionsgradregressor* (vgl. Abschnitt 7.5.6). Dabei handelt es sich um ein Regressionsmodell, das Use Case-Elemente auf einer kontinuierlichen Skala bewertet: Je höher der ausgegebene Wert, desto niedriger ist der Abstraktionsgrad des Elements, d.h. desto stärker ist es einem Low-Level-Kontext zuzuordnen.

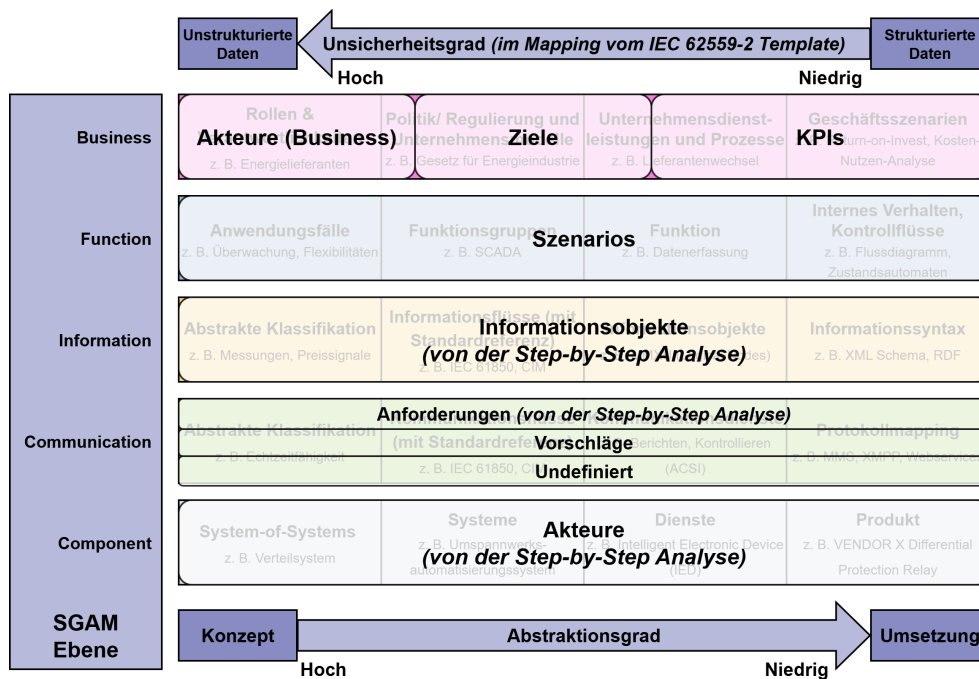
### 7.4.3 KI-Modul für Generierungsaufgaben

Das KI-Modul für Generierungsaufgaben verfolgt das Ziel, auf Basis gegebener Eingaben (Instanzen) ein Objekt oder eine Liste von Objekten zu erzeugen. Dabei kommen neben Klassifikationsmodellen auch weiterführende KI-Modelle zum Einsatz, die als Teilsysteme in das Gesamtsystem integriert und durch fest kodierte, prozedurale Logiken ergänzt werden, um die angestrebte Funktionalität zuverlässig umzusetzen. Im Kern beschäftigt sich dieses Modul mit der automatisierten Generierung von SGAM Modellen auf Grundlage von IEC 62559 Use Cases. Darüber hinaus sollen auf Basis bestehender Elemente eines IEC 62559 Use Cases gezielt Vorschläge zum Hinzufügen, Entfernen oder Bearbeiten einzelner Komponenten abgeleitet werden.

#### 7.4.3.1 Generierung von SGAM Modellen über IEC 62559 Use Cases

Die grundlegende Logik zur Erstellung von SGAM Modellen auf Basis von IEC 62559 Use Cases wurde bereits in den Arbeiten von [FSBU20] und [GDU17] behandelt. In diesen wurde eine High-Level Zuordnung einzelner Elemente des IEC 62559 Use Case-Templates zum SGAM Framework vorgenommen. Eine direkte 1:1-Zuordnung über einen konventionellen, regelbasierten Ansatz ist jedoch nicht ohne Weiteres möglich. Dies liegt einerseits daran, dass das Template sehr konkret ausgefüllt sein muss, um eine valide Ableitung zu ermöglichen, und andererseits daran, dass SGAM Zonen und Domänen nicht für Informations- und Kommunikationsflüsse explizit angegeben werden können. Letzteres wird jedoch in bestehenden UCR-Tools, wie dem OFFIS UCMR oder dem BRIDGE Smart Grid Repository, durch ergänzende Funktionen ermöglicht, was jedoch keinem Pflichtfeld entspricht und nicht von allen Use Case-Erstellern ausgefüllt wird [KNN<sup>+</sup>24]. Dieser Teil des KI-Moduls befasst sich im Wesentlichen mit dem Generieren von SGAM Modellen auf Basis von IEC 62559 Use Case-Beschreibungen durch Verwendung von sequenzieller sowie KI-basierter Logik.

Abbildung 3.15 zeigt einen möglichen Mapping-Ansatz zur Erstellung von SGAM Modellen auf Basis von IEC 62559 Use Case-Beschreibungen. Liegt ein Use Case in der maschinenlesbaren XML-Serialisierung gemäß IEC 62559-3 vor, können bestimmte Informationen systematisch zugeordnet werden. Gleichzeitig ergeben sich jedoch auch Herausforderungen, insbesondere aufgrund heterogener Strukturen und Erstellungsstile,



Quelle: Erweitert auf Basis von [CEN14]

Abbildung 7.11: Konzeptuelle Zuordnung der Sektionen des IEC 62559 Use Case-Templates im Rahmen des KI-basierten Mappings

und nicht zuletzt der Hintergründe der Erstellenden. Die nachfolgende Auflistung differenziert zwischen den einzelnen Interoperabilitätsschichten des SGAM Frameworks und bildet die Basis des in Abbildung 7.11 aufgeführten Mappings:

- **Business Layer:** Dieser Layer lässt sich in der Regel über Informationen zum Anwendungsbereich (*Scope*), zu Zielen (*Objectives*), KPIs, (Meta-)Akteuren, Referenzen sowie Anforderungen erschließen. Aufgrund der Vielzahl an Ausprägungen und der hohen Heterogenität in der Darstellung ist ein direktes Mapping jedoch nicht trivial. Die Extraktion relevanter Informationen erfordert den Einsatz textanalytischer Verfahren, um die Inhalte zu selektieren und ggf. zu kürzen. Ein rein regelbasierter, KI-freier Ansatz ist zum Zeitpunkt dieser Arbeit daher nur eingeschränkt praktikabel.

Der zugrundeliegende KI-Ansatz berücksichtigt insbesondere die Bezeichnungen der KPIs, der Business-Akteure sowie der Objectives. Mithilfe des *M3: SGAM Interoperabilitätsschicht Klassifikators* kann überprüft werden, ob diese Elemente dem Business Layer zuzuordnen sind. Sollte dabei kein entsprechender Bezug zur Business-Ebene erkannt werden, können ergänzend auf Basis der vorhandenen Informationen durch das Modell *M5: SGAM Element Proposer* geeignete Vorschläge für potenzielle Elemente im Business Layer generiert werden.

- **Function Layer:** Dieser Layer kann primär über die Beschreibung des Use Cases, beteiligte Akteure sowie über Szenarien erschlossen werden. Während die Szenariennamen potenziell direkt gemappt werden können, bietet der Einsatz eines KI-Modells Vorteile durch die Ergänzung von Vorschlägen zu weiteren Aspekten. Ein regelbasiertes Mapping ist hier bedingt möglich, sofern entsprechende Szenariennamen im Use Case enthalten sind und klar zuordenbar bleiben.

Das direkte Mapping wird durch einen unterstützenden KI-Ansatz ergänzt, indem zunächst überprüft wird, ob es sich bei dem vorliegenden Szenario um ein Element des Function Layers handelt - basierend auf dem *M3: SGAM Interoperabilitätslayer-Klassifikator*. Sollte dies nicht zutreffen, generiert das Modell *M5: SGAM Element Proposer* auf Grundlage der vorhandenen Verbindungen einen geeigneten Vorschlag für eine alternative Zuordnung.

- **Information Layer:** Die Informationsobjekte innerhalb der Step-by-Step-Analyse sind in der Regel klar strukturiert und können direkt ohne den Einsatz KI-gestützter Verfahren gemappt werden, vorausgesetzt, sie sind im Use Case vollständig vorhanden und korrekt serialisiert.
- **Communication Layer:** Die relevanten Angaben zu Kommunikationsflüssen finden sich sowohl in der Schritt-für-Schritt-Analyse als auch in den Absätzen der Referenzen bzw. Anforderungen. In der Praxis zeigt sich jedoch, dass diese Angaben häufig auf nicht-funktionale Anforderungen beschränkt sind und nur selten konkrete Kommunikationsstandards benennen. Dies erschwert die automatische Zuordnung erheblich. Ein direktes Mapping ohne KI-Unterstützung ist daher nur eingeschränkt realisierbar.

Wie bereits für den Business- und (unterstützend auch für den) Function Layer wird das KI-Modell *M3: SGAM Interoperabilitätslayer-Klassifikator* eingesetzt, um zu überprüfen, ob die vorliegenden sowie verlinkten Requirements oder Referenzen Elemente des Communication Layers darstellen. Sollte dies nicht der Fall sein, werden auf Grundlage der vorhandenen Informationen durch das Modell *M5: SGAM Element Proposer* geeignete Vorschläge zur Ergänzung generiert. Ergibt sich dabei jedoch keine einheitliche Klassifikation über die Mehrzahl der Kombinationen aus beteiligten Akteuren und zugehörigen Informationsobjekten, so werden im ungünstigsten Fall „Undefined / Proprietär“-Platzhalter eingefügt.

- **Component Layer:** Die Zuordnung zur physischen Systemebene kann über die in der Step-by-Step-Analyse verwendeten Akteure erfolgen. Allerdings besteht das strukturelle Problem, dass Akteure im IEC 62559 Use Case-Template nicht verpflichtend in SGAM-Domänen und -Zonen klassifiziert werden müssen. In der Praxis fehlen diese Angaben daher häufig oder sind inkonsistent. Ein regelbasiertes Mapping ist somit nur dann möglich, wenn SGAM-bezogene Angaben explizit vorhanden, korrekt und maschinenlesbar serialisiert sind.

Die in der Schritt-für-Schritt-Analyse referenzierten Akteure können, sofern keine entsprechende Zuordnung bereits im UCMR erfolgt ist, mithilfe der KI-Modelle *M1: SGAM Domänen-Klassifikator* und *M2: SGAM Zonen-Klassifikator* automatisiert den entsprechenden SGAM-Domänen und -Zonen zugeordnet werden.

Der vorgeschlagene hybride Ansatz zur Generierung von SGAM Modellen kombiniert regelbasierte Verfahren mit KI-basierten Modellen. Insbesondere kommen dabei die Klassifikatoren für SGAM-Domänen (*M1*, vgl. Abschnitt 7.5.1), Zonen (*M2*, vgl. Abschnitt 7.5.2) sowie Interoperabilitätsebenen (*M3*, vgl. Abschnitt 7.5.3) zum Einsatz. Diese Modelle dienen dazu, vorhandene Elemente eines Use Cases strukturiert in das SGAM Modell einzuordnen und insbesondere deren Zugehörigkeit zu Funktions- und Businesslayern zu überprüfen. Da IEC 62559 Use Cases in der Praxis häufig heterogen ausgefüllt werden, ist es darüber hinaus notwendig, auf Basis der verfügbaren Informationen auch Vorschläge für potenziell fehlende SGAM-Elemente abzuleiten. Zu diesem Zweck wird das Modell *M5: SGAM Element Proposer* (vgl. Abschnitt 7.5.5) eingeführt. Es handelt sich hierbei um ein KI-Modell auf Basis eines Causal Language Model-Ansatzes, das geeignete Ergänzungen generieren kann, sofern Elemente für das SGAM Modell fehlen oder unzureichend sind.

#### 7.4.3.2 Vorschläge zum Hinzufügen, Entfernen und Bearbeiten von Elementen

Im Rahmen der IEC 62559 Use Case Erstellung besteht die Herausforderung, alle relevanten Modellelemente (wie Akteure, Informations- und Kommunikationsflüsse, funktionale Anforderungen oder betriebliche Ziele) vollständig, konsistent und korrekt abzubilden. Da die Erfassung dieser Inhalte fehleranfällig ist, kann es zu Unvollständigkeitsen, Redundanzen oder fehlerhaften Beschreibungen kommen. Das Ziel des KI-Moduls besteht darin, gezielte Vorschläge zur Ergänzung, Korrektur oder Entfernung dieser Elemente bereitzustellen.

Ein KI-basierter Ansatz kann die Modellierung durch die gezielte Generierung von Vorschlägen zu Ergänzungen, Korrekturen oder Anpassungen relevanter IEC 62559 Use Case-Elemente unterstützen. Da die möglichen Vorschläge potenziell sehr fein granular sein können, ist es notwendig, den Suchraum einzugrenzen. Das SGAM Framework bietet hierfür eine geeignete Struktur, indem es die betrachteten Entitäten systematisch entlang der fünf SGAM-Interoperabilitätsebenen organisiert (vgl. Abbildung 3.14). Zur Ableitung der Vorschläge wird das KI-Modul zur SGAM Modellgenerierung herangezogen, insbesondere das Modell *M5: SGAM Element Proposer* (vgl. Abschnitt 7.5.5). Dieses analysiert bestehende Elementkombinationen innerhalb eines Use Cases und generiert daraus plausible Ergänzungen für solche Inhalte, die in der ursprünglichen IEC 62559-Beschreibung nicht explizit benannt wurden (z. B. fehlende Kommunikationsstandards, Informationsobjekte oder Akteure). Die automatisch erzeugten Elemente werden in das SGAM Modell eingebettet. Diese können dann anschließend als Referenz für Empfehlungen innerhalb der IEC 62559 Use Case-Beschreibung dienen, die den Modellierenden auf potenzielle Lücken oder Verbesserungsmöglichkeiten hinweisen.

Der vorgeschlagene Ansatz ermöglicht eine Ableitung relevanter Inhalte aus IEC 62559 Use Case-Beschreibungen für jede der fünf SGAM-Interoperabilitätsebenen. Dabei werden potenzielle Modelllücken oder -überschneidungen identifiziert und gezielt adressiert. Die Analyse erfolgt entlang folgender Kategorien:

- **Fehlende Elemente:** Über die SGAM Interoperabilitätsebenen, denen auf Grundlage der IEC 62559 Use Case-Beschreibung keine oder nur unzureichende Inhalte zugeordnet werden können (z. B. fehlende Akteure oder nicht beschriebene Informationsobjekte), generiert das System Vorschläge zur inhaltlichen Ergänzung über das KI-Modell *M5: SGAM Element Proposer*.

Ein Beispiel für ein fehlendes Element ist in Abbildung 7.12a dargestellt. Die Abbildung zeigt ein generiertes SGAM Modell für eine beliebige SGAM-Schicht. Während die Elemente *A* und *B* sowohl in der IEC 62559 Use Case-Beschreibung als auch im generierten SGAM Modell enthalten sind, wurde das Element *C* ausschließlich durch das SGAM Modell ergänzt. Dieses Element stellt somit einen Vorschlag dar, der auf potenzielle Lücken in der ursprünglichen Modellierung hinweist.

- **Redundante Elemente:** Elemente, die mehrfach erfasst oder nicht verwendet werden, etwa doppelt aufgeführte Akteure oder Kommunikationsobjekte ohne Bezug zu einem konkreten Szenarienschritt, werden als potenziell überflüssig identifiziert.

Ein Beispiel ist in Abbildung 7.12b dargestellt, in der ein SGAM Modell für eine exemplarische Interoperabilitätsschicht visualisiert wird. Die Elemente *A*, *B* und *C* sind sowohl im generierten SGAM Modell als auch in der zugrundeliegenden IEC 62559 Use Case-Beschreibung vorhanden. Allerdings wurde das Element *C* im Use Case doppelt aufgeführt, sodass das System eine Empfehlung zur Überprüfung und möglichen Bereinigung redundanter Einträge ausspricht.

- **Inkorrekte Elemente:** Falls Elemente innerhalb der IEC 62559 Use Case Beschreibung fälschlich klassifiziert oder einer nicht passenden SGAM-Schicht zugeordnet wurden (überprüft mit dem KI-Modell *M3: Interoperabilitätslayer Klassifikator*, vgl. Abschnitt 7.5.3) werden diese dann inkorrekt eingeordnet. Mittels des Modells *M5: SGAM Element Proposer* werden gezielte Alternativvorschläge generiert.

Ein exemplarischer Fall ist in Abbildung 7.12c dargestellt, in der ein SGAM Modell für eine beliebige Interoperabilitätsschicht visualisiert wird. Die Elemente *A*, *B* und *C* sind sowohl im generierten SGAM Modell als auch in der zugrundeliegenden IEC 62559 Use Case-Beschreibung enthalten. Allerdings wurde das Element *C* im Use Case inhaltlich oder strukturell fehlerhaft eingeordnet. Das generierte SGAM Modell zeigt eine korrigierte Zuordnung, auf deren Grundlage dem Modellierenden eine gezielte Empfehlung zur Anpassung oder Korrektur des ursprünglichen Eintrags bereitgestellt werden kann.

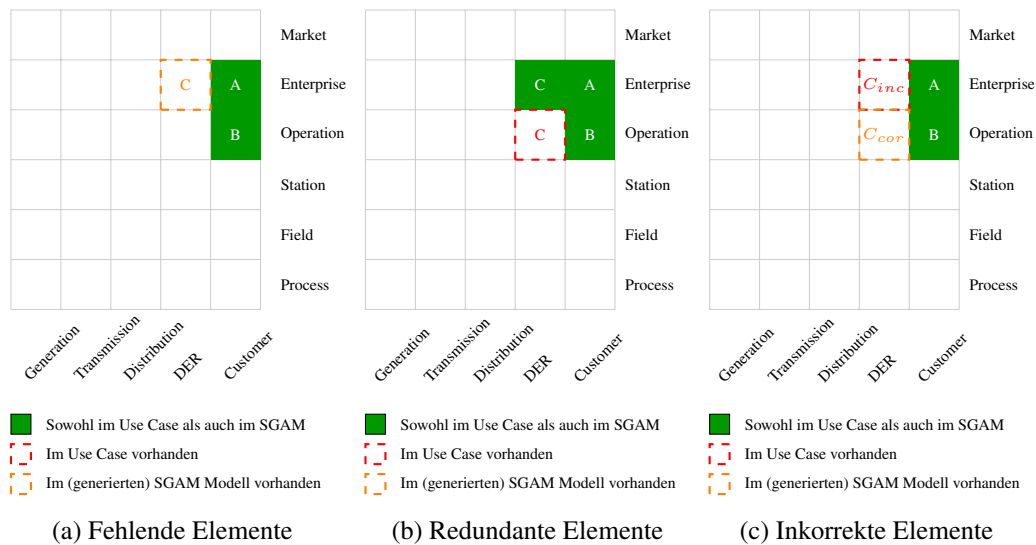


Abbildung 7.12: Beispiele zur Elementableitung für die IEC 62559 Use Case-Methodik mit Hilfe des SGAM Frameworks

## 7.5 Konzeptionalisierung der KI-Modelle

Der vorherige Abschnitt 7.4 zeigt auf, wie ein KI-basiertes System die identifizierten Anwendungsfälle adressiert. Hierbei wurde kritisch betrachtet, welche Funktionalitäten im Rahmen dessen über prozedurale Algorithmen realisiert werden können und welche mittels KI-Modelle realisiert werden müssen bzw. unterstützt werden. Dieser Abschnitt befasst sich mit der Konzeption der KI-Modelle zur Realisierung der KI-Module. Hierbei können die Anwendungsfälle gemappt werden auf die vorgesehenen KI-Modelle, was in Abbildung 7.13 tabellarisch dargestellt wurde.

Die Entwicklung von solchen KI-Modellen erfordert eine adäquate Daten- und Wissensbasis. Insbesondere in den betrachteten KI-Disziplinen (vgl. Kapitel 4) bildet diese Grundlage das Fundament für die angestrebte Funktionalität. Für die Umsetzung der konzipierten Anwendungsfälle sind die im Abschnitt 7.2.1 aufgeführten Datenquellen vorgesehen. Aufgrund der Heterogenität der betrachteten Datenquellen ist es erforderlich, einen möglichst allgemeinen gemeinsamen Nenner für die Modellierung zu identifizieren. Daraus ergibt sich, dass die Eingaben für die Modelle in ihrer Komplexität möglichst reduziert werden sollten, ohne dabei relevante semantische Informationen zu verlieren. In der Betrachtung der zugrundeliegenden Daten- und Wissensquellen konnte festgestellt werden, dass sich für die Mehrzahl der kontextuellen Modelle die beiden Merkmale *Name* und *Beschreibung* konsistent extrahieren lassen. Folglich ist jeder Akteur, Standard und Aktivität mit diesen Merkmalen technisch abgebildet. Die Realisierung des Datensatzes wird explizit in Abschnitt 8.2 behandelt. Dementsprechend wird im Folgenden ein Element formal als Tupel aus Name und Beschreibung aufgefasst:

$$\mathcal{X}_{\text{elem}} \subseteq \mathcal{N} \times \mathcal{D},$$

wobei  $\mathcal{N}$  die Menge aller möglichen Namen und  $\mathcal{D}$  die Menge aller zugehörigen Beschreibungen bezeichnet. Das Tupel  $(n, d) \in \mathcal{X}_{\text{elem}}$  bildet somit die Eingabestruktur für die vorgesehenen Sprachmodelle, sofern nicht explizit anderes erwähnt wird.

Im folgenden Abschnitt werden auf konzeptioneller Ebene die acht KI-Modelle aus Abbildung 7.13 vorgestellt und formal beschrieben, die als Basis zur Realisierung des Systems verwendet werden.

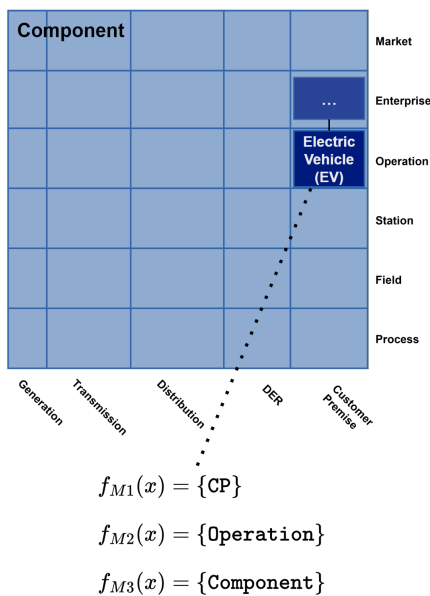
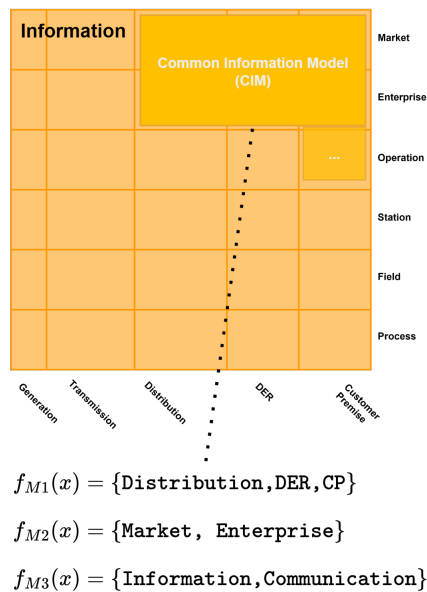
### 7.5.1 SGAM Domänen Klassifikator (M1)

Dieses Modell dient der Zuordnung eines gegebenen Elements zu den fünf SGAM Domänen. Die Funktion ist formal definiert als  $f_{M1} : \mathcal{X}_{\text{elem}} \rightarrow \{0, 1\}^5$ . Als Eingabe erhält das Modell ein Element  $x \in \mathcal{X}_{\text{elem}}$ , bestehend aus Name und Beschreibung; die Ausgabe ist ein binärer Vektor, der angibt, welche der Domänen *Generation*, *Transmission*, *Distribution*, *DER* und *Customer* zutreffen (Multi-Label Klassifikation).

Die Abbildungen 7.14 und 7.15 veranschaulichen exemplarisch die Anwendung des Modells  $M1$  auf den Akteur EV beziehungsweise den Standard CIM. Im Rahmen der Klassifikation wird neben dem Namen auch eine Beschreibung berücksichtigt, welche aus Gründen der Übersichtlichkeit in den Abbildungen nicht dargestellt ist. Diese Tupels führen zur Zuordnung des EV zur Customer Premise-Domäne, während der CIM-Standard aufgrund seiner inhaltlichen Beschreibung den Domänen Distribution, Distributed Energy Resources sowie Customer Premise zugewiesen wird.

	IEC 62559 Use Case-Methodik					SGAM Framework		
	1.1 - SGAM Domänen und Zonen vorschlagen	1.2 - Erkennen von abweichenden und nicht-standardisierten Terme	1.3 - Vorschläge von Hinzufügen, Entfernen und Bearbeiten von Elementen	1.4 - Service Steps in der Step-by-Step Analysis vorschlagen	1.5 - Erkennen von Inkonsistenzen [...] ggü. dem Use Case-Typ	2.1 - Generieren von SGAM Arbeitsmodelle	2.2 - Überprüfen der Konsistenz [...] gegenüber der UC-Beschreibungen	2.3 - Identifikation von Inkonsistenzen innerhalb SGAM Modelle
<b>M1:</b> SGAM Domänen Klassifikator	■					■	■	■
<b>M2:</b> SGAM Zonen Klassifikator	■					■	■	■
<b>M3:</b> SGAM Interoperabilitäts-layer Klassifikator			■			■	■	■
<b>M4:</b> Service Step Klassifikator				■				
<b>M5:</b> SGAM Element Proposer			■			■	■	■
<b>M6:</b> Abstraktionsgrad Regressor					■			■
<b>M7a:</b> Acronym Identifikation		■						
<b>M7b:</b> Acronym Disambiguation		■						

Abbildung 7.13: Modellübersicht von SGAAIRE

Abbildung 7.14: Einordnung des Akteurs *EV* im SGAMAbbildung 7.15: Einordnung des Standards *CIM* im SGAM

### 7.5.2 SGAM Zonen Klassifikator (M2)

Dieses Modell dient der Zuordnung eines gegebenen Elements zu den sechs SGAM Zonen. Die Abbildung erfolgt gemäß  $f_{M2} : \mathcal{X}_{\text{elem}} \rightarrow \{0, 1\}^6$ . Basierend auf einem Element mit Name und Beschreibung gibt das Modell einen binären Vektor zurück, der die Zugehörigkeit zu den Zonen *Process*, *Field*, *Station*, *Operation*, *Enterprise* und *Market* angibt (Multi-Label Klassifikation).

Die Abbildungen 7.14 und 7.15 illustrieren beispielhaft die Anwendung des Modells *M2* auf den Akteur *EV* sowie den Standard *CIM*. Die daraus resultierenden Zuordnungen weisen das *EV* der Zone *Operation* zu, während der *CIM*-Standard den Zonen *Market* und *Enterprise* zugeordnet wird.

### 7.5.3 SGAM Interoperabilitätslayer Klassifikator (M3)

Dieses Modell dient der Zuordnung eines gegebenen Elements zu den fünf SGAM Interoperabilitätsebenen. Die zugrundeliegende Funktion ist  $f_{M3} : \mathcal{X}_{\text{elem}} \rightarrow \{0, 1\}^5$ . Ausgehend von Name und Beschreibung eines Elements liefert das Modell einen binären Vektor, der Mehrfachzuweisungen zu den Layern *Business*, *Function*, *Information*, *Communication* und *Component* erlaubt (Multi-Label Klassifikation).

Die Abbildungen 7.14 und 7.15 illustrieren exemplarisch die Anwendung des Modells *M2* auf den Akteur *EV* sowie den Standard *CIM*. Die resultierenden Klassifikationen ordnen das *EV* der Komponentenebene zu, während der *CIM*-Standard sowohl der

Informations- als auch der Kommunikationsebene zugewiesen werden kann. Eine solche Mehrfachzuordnung ist auch im Rahmen von [IEC21b] vorgesehen, da CIM Inhalte adressiert und den beiden Ebenen zugeordnet werden können. Die inhaltliche Beschreibung kann dabei als Indikator für eine bevorzugte Einordnung in eine der beiden Ebenen dienen.

#### 7.5.4 Service Step Klassifikator (M4)

Dieses Modell klassifiziert eine Aktivität (bzw. Szenarioschritt) eines Szenarios innerhalb des IEC 62559 Use Case-Templates hinsichtlich seiner semantischen Rolle in der Kommunikation über das Feld „Service“. Es wird formal beschrieben durch  $f_{M4} : \mathcal{X}_{\text{step}} \rightarrow \mathcal{Y}_S$ . Die Eingabe besteht aus Name und Beschreibung einer Aktivität; die Ausgabe ist eine Klasse aus der Menge  $\mathcal{Y}_S = \{\text{CREATE, GET, CHANGE, ...}\}$  gemäß des Standards (Multi-Class Klassifikation). Ein Beispiel ist in Abbildung 7.16 erkennbar, in welchem ein Austausch von Anreiztabellen zwischen dem Energy Broker und dem Customer Energy Manager stattfindet. Hierbei wird die Aktivität als REPORT klassifiziert.

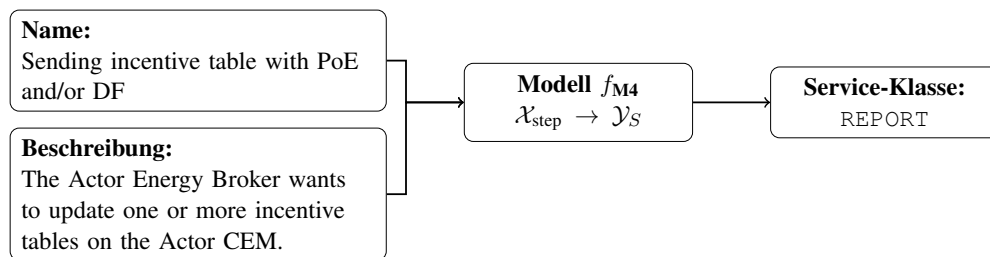


Abbildung 7.16: Beispielhafte Klassifikation einer Aktivität im Feld „Service“ im IEC 62559 Use Case-Template

#### 7.5.5 SGAM Element Proposer (M5)

Das im Folgenden vorgestellte Modell zielt darauf ab, semantisch geeignete SGAM-Elemente auf einer definierten Ziel-Interoperabilitätsebene vorherzusagen. Als Eingabe dienen bestehende Elemente anderer Interoperabilitätsebenen, welche als kontextuelle Information zur Ableitung genutzt werden. Ziel ist es, auf Basis semantischer Relationen innerhalb eines SGAM Modells eine konsistente und plausible Vervollständigung vorzunehmen. Eine schematische Darstellung möglicher Kombinationen von SGAM-Elementen ist in Abbildung 7.17 dargestellt.

Formell lässt sich das Problem wie folgt beschreiben: Gegeben sei eine Eingabesequenz  $\mathcal{I} = [(l_1, o_1), (l_2, o_2), \dots, (l_n, o_n)]$ , wobei jedes Tupel  $(l_i, o_i)$  eine Interoperabilitätsebene  $l_i \in \mathcal{L}$  und ein zugehöriges Objektelement  $o_i \in \mathcal{O}$  repräsentiert. Gesucht ist ein Objekt  $\hat{o}_{\text{target}} \in \mathcal{O}$ , das auf einer festgelegten Ziel-Interoperabilitätsebene  $l_{\text{target}} \in \mathcal{L}$  semantisch konsistent zur Eingabesequenz passt.

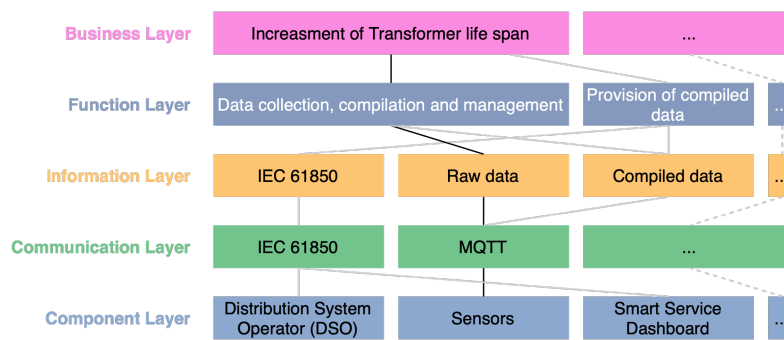


Abbildung 7.17: Mögliche Elementkombinationen innerhalb eines SGAM Modells

Das Modell basiert auf einem Causal Language Modelling-Ansatz, welcher die Eingabesequenz autoregressiv verarbeitet und die wahrscheinlichste Fortsetzung in Form des Zielobjekts  $\hat{o}_{target}$  erzeugt:  $\hat{o}_{target} = \arg \max_{o \in \mathcal{O}} P(o | \mathcal{I}, l_{target})$ .

Eine Illustration der Funktionsweise des Modells im Hinblick auf Ein- und Ausgabestrukturen ist in Abbildung 7.18 dargestellt. Dabei wird auf Basis semantisch miteinander verknüpfter Informationen aus der Funktions-, Informations- und Komponentenebene ein geeignetes Element für die Kommunikationsebene vorgeschlagen. Im gezeigten Beispiel wird deutlich, dass für die beteiligten Akteure *Energy Broker* und *Delivery Broker*, sowie für kontextbezogene Informationen zu *Incentive Tables* im Bereich von Energieproduktion und -verbrauch der Einsatz des Standards CIM als geeignet erscheint. Dies liegt unter anderem darin begründet, dass das CIM-Metamodell für den Marktbereich bereits relevante Aspekte berücksichtigt [IEC21b], jedoch Spezifika in der betrachteten Modellinstanz fehlen, die im Rahmen des vorliegenden Projekts adressiert werden können.

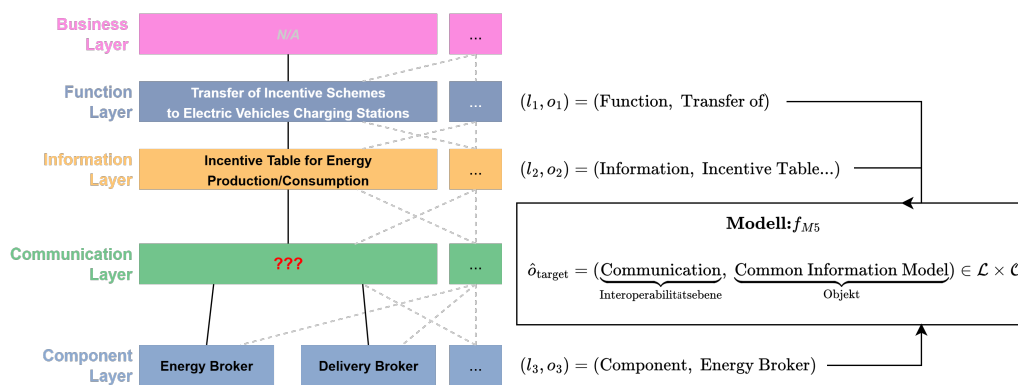


Abbildung 7.18: Beispielhafte Anwendung des Modells für Erhalten der Vorschläge für die Kommunikationsebene

### 7.5.6 Abstraktionsgrad Regressor (M6)

Dieses Modell schätzt den konzeptionellen Abstraktionsgrad eines Elements auf einer kontinuierlichen Skala. Hierbei orientiert sich das Modell gemäß den unterschiedlichen Abstraktionsgemäß für die SGAM Interoperabilitätsebenen gemäß Abbildung 3.13. Die Funktion ist gegeben durch  $f_{M6} : \mathcal{X}_{\text{elem}} \rightarrow \mathbb{R}$ . Basierend auf Name und Beschreibung eines Elements gibt das Modell eine reelle Zahl zurück, die wie folgt diskretisiert wird:

$$L_{\text{hoch}} = \{y \mid y < 0,5\}, \quad L_{\text{eher hoch}} = \{y \mid 0,5 \leq y < 1,5\}, \\ L_{\text{eher niedrig}} = \{y \mid 1,5 \leq y < 2,5\}, \quad L_{\text{niedrig}} = \{y \mid y \geq 2,5\}.$$

Abbildung 7.19 zeigt exemplarisch die Einordnung des Abstraktionsgrads für ein virtuelles Kraftwerk sowie ein Smart Meter Gateway. Je höher der Wert, desto konkreter ist das Element; je niedriger, desto abstrakter.

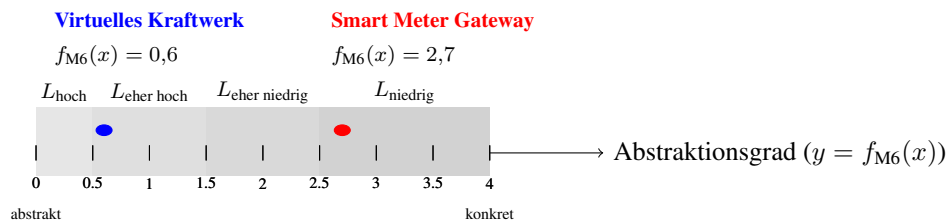


Abbildung 7.19: Einordnung zweier Systeme nach Abstraktionsgrad

### 7.5.7 Acronym Identification (M7a)

Dieses Modell identifiziert Akronyme und deren potenzielle Langformen innerhalb von Texten. Acronym Identification wird von [VDN<sup>+</sup>20] als „NER“-Problem beschrieben, welche einem Teilbereich der Token Klassifikation entspricht. Gegeben ist ein Satz mit den Wörtern  $S = [w_1, w_2, \dots, w_n]$ . Die Aufgabe besteht darin, für jedes Wort die zugehörigen Labels  $L = [l_1, l_2, \dots, l_n]$  zu bestimmen, wobei  $l_i \in \{B_s, I_s, B_l, I_l, O\}$ . Diese Labels entsprechen der BIO-Tagging-Notation, bei der  $B_s$  den Beginn und  $I_s$  den zugehörigen Inhalt des Akronyms repräsentieren, während  $B_l$  und  $I_l$  den Beginn und zugehörigen Inhalt der Langform darstellen. Das Label  $O$  wird für andere Wörter verwendet. Für die Entwicklung von Acronym Identification-Modellen werden Datensätze, wie SciAI, AESM oder LSAEF eingesetzt. [VDN<sup>+</sup>20]

Die Eingabe des zu konzipierenden Modells ist eine Sequenz von Tokens  $(w_1, \dots, w_T)$ , z. B. aus Fließtextfeldern oder strukturierten Elementen (z. B. Akteure, KPIs). Das Modell gibt für jedes Token ein Label  $l_i \in \{B_s, I_s, B_l, I_l, O\}$  aus, wobei Akronyme und Langformen gemäß BIO-Notation erkannt werden. Abbildung 7.9 veranschaulicht anhand des Satzteils „Das Smart Meter Gatewayx (SMGW) fungiert [...] SUC liegt [...] HAN. [...]“, welche Labels den einzelnen Tokens zugewiesen werden. Formal:  $f_{M7a} : \mathcal{X}_{\text{text}} \rightarrow \mathcal{L}^T$ , wobei  $\mathcal{L} = \{B_s, I_s, B_l, I_l, O\}$ .

### 7.5.8 Acronym Disambiguation (M7b)

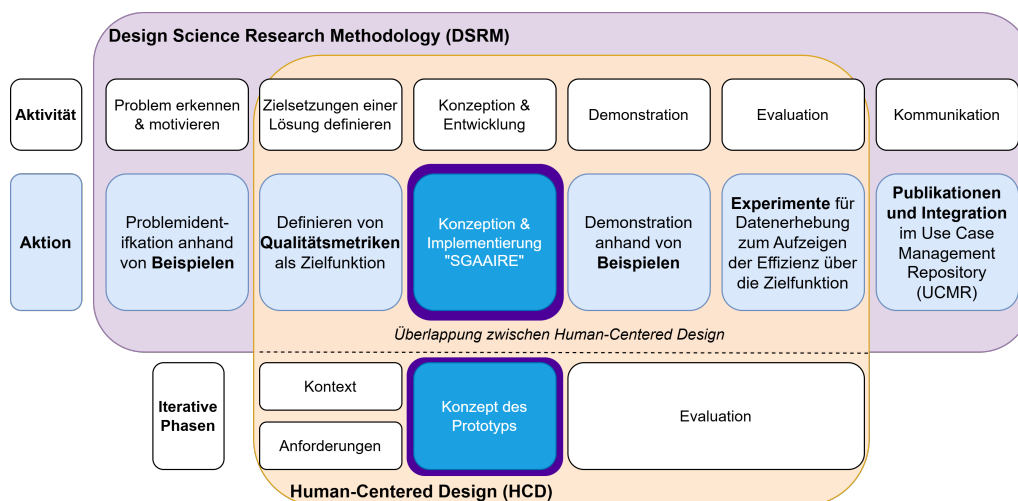
Acronym Disambiguation wird von [VDN<sup>+</sup>20] als ein textuelles Klassifikationsproblem beschrieben, das darauf abzielt, die spezifische Bedeutung eines gegebenen Akronyms innerhalb eines Satzes zu identifizieren. Angenommen sei ein Satz mit den Wörtern  $S = [w_1, w_2, \dots, w_n]$ , und es existiere ein Index  $t$ , der auf ein Akronym  $w_t$  im Satz verweist. Für dieses Akronym kommen eine Reihe möglicher Langformen in Betracht, dargestellt als die Menge  $L = \{l_1, l_2, \dots, l_m\}$ , welche in Acronym Dictionaries aufgeführt sind. Das Ziel besteht darin, die korrekte Langform  $l_i$  für das Akronym  $w_t$  zu ermitteln. Für die Entwicklung von Acronym Disambiguation-Modellen werden Datensätze wie beispielsweise Science WISE, NOA oder SciAD genutzt. Der Datensatz SciAD setzt sich aus einer Vielzahl wissenschaftlicher Veröffentlichungen unterschiedlicher Domänen zusammen. [VDN<sup>+</sup>20]

Das zu konzipierende Modell nutzt den umgebenden Kontext des Akronyms sowie weitere relevante IEC 62559 Use Case-Bestandteile (z. B. Glossar oder Akteursliste). Abbildung 7.9 zeigt beispielhaft welche potentielle Optionen für den Term HAN im Satz „*Das Smart Meter Gatewayx (SMGW) fungiert [...] SUC liegt [...] HAN. [...]*“ mit Berücksichtigung des Kontextes des Satzes zur Verfügung stehen. Formal:  $f_{M7b} : (\mathcal{X}_{\text{text}}, a) \rightarrow \hat{l} \in \mathcal{L}_a$ , wobei  $a \in \mathcal{A}$  ein Akronym und  $\mathcal{L}_a$  die Menge möglicher Langformen ist.



## 8 Implementierung von SGAAIRE

Dieses Kapitel beschreibt die konkrete Umsetzung des KI-Systems SGAAIRE, das als zentrales Artefakt dieser Arbeit entwickelt wurde. Die Konzeption wurden bereits in Kapitel 7 dargelegt. Im Gesamtverfahren stellt SGAAIRE das zentrale Artefakt dieser Dissertation dar, dessen Umsetzung in Abbildung 6.1 veranschaulicht ist.



Quelle: Eigene Darstellung, angelehnt an [PTRC07]

Abbildung 8.1: Auswirkungen der Ergebnisse aus Kapitel 8 auf das Forschungsvorhaben

Im folgenden Abschnitt 8.1 wird zunächst die Implementierung des Gesamtsystems erläutert. Dabei werden die ausgewählten Techniken, eingesetzte Bibliotheken sowie die Integration des Systems in ein UCMR detailliert dargestellt. Da SGAAIRE maßgeblich auf den Einsatz von KI-Modellen basiert, ist eine qualitative sowie quantitative Datenbasis für das Training dieser Modelle unerlässlich. Die systematische Erhebung und Überführung der dafür benötigten Daten wird im Abschnitt 8.2 beschrieben. Diese Arbeit verwendet bestehende, vortrainierte State-of-the-Art-Sprachmodelle als Grundlage für die domänenspezifische Feinabstimmung auf die in SGAAIRE vorgesehenen Aufgaben. Die Auswahl geeigneter Modelle, die Durchführung des Fine-Tunings sowie die Evaluation der trainierten Modelle im Hinblick auf ihre Performanz in spezifischen KI-Aufgaben werden im folgenden Abschnitt 8.3 detailliert beschrieben. In diesem Abschnitt wird außerdem anhand der trainierten KI-Modelle die Integration innerhalb der Anwendungsfälle von SGAAIRE beschrieben.

## 8.1 Systemumsetzung

Dieser Abschnitt beschreibt die Implementierung von SGAAIRE. Darin werden die getroffenen Architektur- und Technologieentscheidungen erläutert, einschließlich der verwendeten Programmiersprachen, des Aufbaus der Laufzeitumgebung sowie der integrierten Verfahren. Abschließend wird die Integration von SGAAIRE in das UCMR dargestellt sowie die implementierte API zur Nutzung des Systems.

### 8.1.1 Programmiersprache

Durch die im Systemkonzept vorgesehene Anbindung über eine HTTP-API ist die ursprüngliche Abhängigkeit von Java, der im UCMR verwendeten Programmiersprache, aufgehoben. Dies ermöglicht eine flexiblere Auswahl der Programmiersprache insbesondere hinsichtlich der KI-Komponenten. Für die Entwicklung von SGAAIRE wird *Python* als Programmiersprache verwendet.

In der aktuellen Forschung und Praxis gilt Python als De-facto-Standard für die Entwicklung von Anwendungen der Künstlichen Intelligenz [Sta25]. Gründe hierfür sind die niedrige Einstiegshürde, die hohe Lesbarkeit sowie die umfassende Verfügbarkeit von spezialisierten Bibliotheken wie TensorFlow [Dev22], PyTorch [IPK21] oder Transformers von Hugging Face [Jai22]. Eine weitere größere Alternative stellt C/C++ dar. C/C++ hingegen wird bevorzugt in produktiven KI-Systemen eingesetzt, in denen Performanz, Ressourcenmanagement und Latenz eine zentrale Rolle spielen, etwa bei eingebetteten Systemen, KI in Computerspielen oder in der Robotik [GBC16]. Für SGAAIRE ist dies jedoch nicht erforderlich, da im Rahmen des prototypischen Aufbaus keine nicht-funktionalen Anforderungen an Echtzeitfähigkeit gestellt wurden.

### 8.1.2 Bibliotheken und Tools

Die Programmierung von SGAAIRE wurde weitestgehend in *JetBrains PyCharm* [Jet25] als Python Integrated Development Environment (IDE) sowie in Teilen in *Visual Studio Code* [Mic25] entwickelt. Für die Versionierung des Sourcecodes wurde ein *internes Repository im OFFIS GitLab*<sup>1</sup> verwendet, welche auch Continuous Integration/Delivery Funktionalitäten bereitstellt. Für die Entwicklung der KI-Modelle wurden Jupyter Notebooks [RPG17] verwendet, die eine interaktive, dokumentationsnahe (durch die Integration von Markdown) und reproduzierbare Arbeitsumgebung bieten. Durch die Blockstruktur der Notebooks können einzelne Codezellen (bspw. zur Evaluierung eines trainierten Modells) gezielt und unabhängig erneut ausgeführt werden, was eine effizientere Programmierung unterstützt.

Für die Umsetzung der KI-Modelle wurde auf die Bibliothekssuite von Hugging Face zurückgegriffen, insbesondere auf die Bibliotheken *Transformers*, *Datasets*, *Accelerate* und *Evaluate* [Jai22]. Hugging Face stellt zum Zeitpunkt der Dissertation die größte

<sup>1</sup> <https://gitlab.offis.de/official/fbe-p.ucmr/SGAAIRE>

offene Community zur Entwicklung, Nutzung und Bereitstellung von den State-of-the-Art Open Source Sprachmodellen (vgl. Abschnitt 4.3) über den Hugging Face Hub dar. Die Bibliothek Transformers dient der allgemeinen Arbeit mit (Large) Language Models, insbesondere als Assistenz für das Training sowie die Inferenz vortrainierter Modelle. Datasets ermöglicht die strukturierte und effiziente Aufbereitung von Datensätzen zur Weiterverwendung im Hugging-Face-Ökosystem. Die Bibliothek Accelerate wird in dieser Arbeit eingesetzt, um die Ausführung und das Training von Deep-Learning-Modellen abstrahiert auf unterschiedlichen Hardwarekonfigurationen effizient umzusetzen. Die Bibliothek Evaluate wurde zur Bewertung von Sprachmodellen verwendet. Sie unterstützt eine Vielzahl standardisierter Metriken, die über den Hugging Face Hub bereitgestellt werden, darunter unter anderem F1-Score, BERTScore und METEOR.

Hugging Face integriert eine umfassende Unterstützung für die führenden Backends im Bereich des KI-Trainings, insbesondere für PyTorch, TensorFlow und JAX [Jai22]. Für die Umsetzung der in dieser Arbeit verwendeten KI-Modelle wurde als Backend *PyTorch* gewählt, da dieses Framework auch für das Vortraining der meisten eingesetzten Sprachmodelle verwendet wurde und somit eine hohe Kompatibilität erwartet wird. Eine Konvertierung von PyTorch-Modellen in TensorFlow (und umgekehrt) ist möglich, jedoch mit zusätzlichem Implementierungsaufwand und potenziellen Kompatibilitätsproblemen verbunden, insbesondere im Hinblick auf Modellarchitektur, Tokenizers und Framework-spezifischen Operationen [Jai22].

Die Bibliothek *scikit-learn* [Kra16] wurde ebenfalls zur Berechnung klassischer Evaluationsmetriken eingesetzt, insbesondere zur Erstellung von Konfusionsmatrizen sowie zur Generierung von Klassifikationsberichten. Für die allgemeine Arbeit mit Datensätzen sowie deren Vorverarbeitung und Analyse kamen die Bibliotheken *pandas* [McK11] und *NumPy* [HMW<sup>+</sup>20] zum Einsatz.

Für die Bereitstellung der HTTP-Schnittstellen wurde das Python-basierte Framework *Flask* [Gri18] verwendet. Es ermöglichte eine schlanke, modulare und gut dokumentierte Umsetzung der serverseitigen Logik, die insbesondere für prototypische Webanwendungen geeignet ist. Zur Containerisierung und für die plattformunabhängige Bereitstellung der Anwendung kam *Docker* [Mer14] sowie *docker compose* zum Einsatz. Durch die Verwendung von Docker sowie *docker compose* konnten konsistente Umgebungen geschaffen und die Reproduzierbarkeit der Systembereitstellung sowohl lokal als auch auf Laufzeitumgebung (vgl. Abschnitt 8.1.3) erreicht werden.

Weitere unterstützende Bibliotheken wurden im Rahmen der Entwicklung eingesetzt. Eine vollständige Übersicht ist der Datei *requirements.txt* sowie jeweils dem ersten Codeblock der verwendeten Jupyter Notebooks zu entnehmen.

### 8.1.3 Laufzeitumgebung

Das System wurde auf einer konventionellen Virtuellen Maschine (VM) mit 20 GB Arbeitsspeicher und 4 vCores einer Intel Xeon Gold 6248R CPU (3 GHz) betrieben. Eine dedizierte Grafikeinheit ist nicht integriert. Diese Laufzeitumgebung kam insbesonde-

re im Rahmen der Demonstration sowie der Evaluierung zum Einsatz (vgl. Kapitel 9). Ein Vorteil der Nutzung dieser VM liegt in den geringen Betriebskosten und der einfachen Zugänglichkeit. Ein Nachteil besteht jedoch in den vergleichsweise längeren Inferenzzeiten, da moderne KI-Modelle insbesondere von CUDA-Kernen auf Nvidia-Grafikkarten profitieren [MCL<sup>+</sup>18]. Durch die Integration von Pufferungs- und Caching-Verfahren (vgl. Abschnitt 8.1.4) konnten die erhöhten Inferenzzeiten im Kontext der Evaluation auf ein akzeptables Maß reduziert werden. Für eine produktive Laufzeitumgebung wäre der Einsatz einer VM mit GPU-Unterstützung im Hinblick auf die Inferenzlatenz mit CUDA-fähiger Grafikkarte empfehlenswert.

### 8.1.4 Pufferung- und Cachingverfahren für Inferenzen

Aufgrund der längeren Inferenzzeiten, die durch den Einsatz von CPUs anstelle von GPUs entstehen, war im Rahmen dieser Arbeit die Entwicklung eines geeigneten Workarounds erforderlich. Der Einsatz von Pufferungs- und Cachingtechniken soll sicherstellen, dass Ergebnisse deterministischer Modellen (vgl. Definition 17) für einen längeren Zeitraum im System zur Verfügung stehen, um so unnötige Inferenzschritte zu vermeiden und folglich die Prozessorlast intelligent zu regeln.

**Definition 17** (Deterministisches Modell)

Ein Modell wird als *deterministisch* bezeichnet, wenn seine Inferenzfunktion  $f_M(x)$  für eine gegebene Eingabe  $x \in \mathcal{X}$  stets denselben Ausgang  $y \in \mathcal{Y}$  liefert. Das bedeutet, dass  $f_M(x) = y$  für alle Ausführungen invariant ist, sofern die Ausführung unter identischen Rahmenbedingungen erfolgt. [GBC16]

Im Rahmen dieser Arbeit wurde eine einfache cache-basierte Inferenzlogik implementiert (vgl. Algorithmus 1). Dabei werden Eingabe/Ausgabe-Tupel  $(x, y)$  in einem Key-Value-Speicher abgelegt. Der Algorithmus prüft zunächst, ob für eine gegebene Eingabe  $x$  bereits ein Ergebnis im Speicher vorhanden ist. Ist dies nicht der Fall, wird die Ausgabe  $y$  mittels eines KI-Modells berechnet und das Tupel  $(x, y)$  dem Speicher hinzugefügt.

---

**Algorithmus 1:** Cache-basierte Inferenz mit Zwischenspeicherung

---

**Input:** Modelleingabe  $x$   
**Output:** Modellantwort  $y$   
**if** `is_cached(x)` **is false** **then**  
    |  $y \leftarrow \text{model}(x)$   
    | `cache_store(x, y)`  
**else**  
    |  $y \leftarrow \text{cache_load}(x)$   
**return**  $y$

---

Eine Anfrage an die SGAAIRE API kann zu mehreren Modellaufrufen führen, insbesondere wenn innerhalb der Anfrage oder über verschiedene Anfragen hinweg identische Eingaben auftreten. Dies ist beispielsweise der Fall, wenn mehrere Modellierende

an einem IEC 62559 Use Case oder einem SGAM Modell arbeiten und ähnliche Strukturelemente abfragen. Durch die Implementierung einer cache-basierten Inferenzlogik können wiederholte Modellberechnungen vermieden und die Anzahl tatsächlicher Inferenzaufrufe reduziert werden.

### 8.1.5 SGAAIRE API

Die Systembereitstellung orientiert sich an der nicht-funktionalen Anforderung *REQ\_NFA\_3: Standalone Software System* (vgl. Abschnitt 5.2) und erfolgt in Form einer eigenständigen Softwarelösung. Für die Umsetzung einer leichtgewichtigen, plattformunabhängigen API wird die in Python etablierte Bibliothek *Flask* zur Entwicklung von HTTP Schnittstellen eingesetzt [Gri18]. Zur Verbesserung der Transparenz, Wartbarkeit und Interoperabilität erfolgt die Schnittstellenbeschreibung zusätzlich mit *Swagger(UI)* und *OpenAPI* [PRT22].

Eine über SwaggerUI visualisierte Variante der SGAAIRE-API ist in Abbildung 8.2 dargestellt. Sie verdeutlicht die Trennung zwischen den Operationen, die unmittelbar die KI-Modelle adressieren, und den im Systemkonzept vorgesehenen Handlern für SGAM Modelle sowie IEC 62559-Use Cases. Die in diesem Kontext bereitgestellten Endpunkte lassen sich in die drei funktionalen Gruppen einteilen:

- **KI Modelle:** Diese Gruppe umfasst Endpunkte zur direkten Nutzung der in SGAAIRE integrierten KI-Modelle. Dazu gehören Klassifikations- und Regressionsaufgaben, die automatische Generierung von Vorschlägen sowie das Management der implementierten Caching-Mechanismen.

Über `/model/classify` können Elemente (z. B. Domänen, Zonen oder Service Steps) anhand von Namen und optionalen Beschreibungen einem Klassifikations- bzw. Regressionsproblem zugeordnet werden. Diese Funktionalität entspricht den in der API integrierten KI-Modellen *M1*, *M2*, *M3*, *M4*, *M6* und *M7b*. Der Endpunkt `/model/classify/token` ermöglicht darüber hinaus eine tokenbasierte Analyse von Textblöcken, um undefinierte oder fehlerhafte Begriffe zu identifizieren (KI-Modell *M7a*). Mit `/model/proposal` können schließlich Vorschläge für SGAM Modellelemente generiert werden, die sich an bestehenden Relationen orientieren (KI-Modell *M5*).

Ergänzend stellen die Caching-APIs (`/model/cache`, `/model/cache/load`, `/model/cache/load-all`, `/model/cache/store`) eine Verwaltung der Zwischenspeicherung der Modellergebnisse sicher, wodurch Rechenaufwand reduziert und Reproduzierbarkeit verbessert werden sollen.

- **SGAM-bezogene KI-Modelle:** Diese Endpunkte dienen der Qualitätssicherung sowie der automatischen Generierung von SGAM-Arbeitsmodellen und adressieren die SGAM-bezogenen Anwendungsfälle. Über `/sgam/generate` können SGAM-Arbeitsmodelle direkt aus IEC 62559 Use Case-Beschreibungen erzeugt werden (*Anwendungsfall 2.1*). `/sgam/check` ermöglicht eine Prüfung dieser Modelle unter

SGAM Models		Suggestions related to SGAM Model creation and checking	^
POST	/sgam/check		∨
POST	/sgam/consistency		∨
POST	/sgam/generate		∨
IEC 62559-2 Use Case Descriptions		Suggestions related to IEC 62559-2 use case descriptions creation and checking	^
POST	/uc/suggestions		∨
AI Model		AI-related API endpoints that directly calls the models of interest	^
GET	/model/cache	Clear the cache by removing all entries	∨
POST	/model/cache/load	Load data from the cache	∨
POST	/model/cache/load-all		∨
POST	/model/cache/store	Store data in the cache	∨
GET	/model/classify		∨
GET	/model/classify/token		∨
GET	/model/proposal		∨

Quelle: Eigene Darstellung

Abbildung 8.2: Übersicht der SGAAIRE API

Einbeziehung der zugrunde liegenden Use Case-Beschreibungen (*Anwendungsfall 2.2*). Über `/sgam/consistency` können isolierte Konsistenzanalysen auf Basis des SGAM Modells selbst realisiert werden (*Anwendungsfall 2.3*).

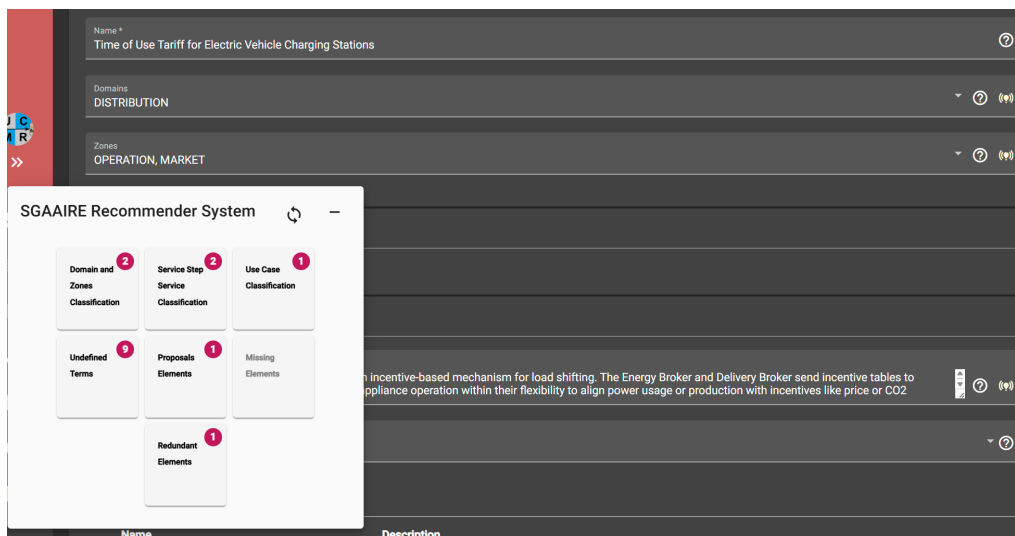
- **IEC 62559-2 Use Case-bezogene KI-Modelle:** Diese Kategorie ist auf die Qualitätssicherung der Anwendungsfallbeschreibungen ausgerichtet und adressiert die UC-bezogenen Anwendungsfälle. Der Endpunkt `/uc/suggestions` erzeugt Vorschläge für die in den *Anwendungsfällen 1.x* vorgesehenen Aufgaben.

Die Containerisierung von SGAAIRE erfolgt mittels Docker und `docker-compose`. In der `docker-compose`-Konfiguration werden die Ports 8080 und 5000 für die Kommunikation zwischen API und Proxy-Diensten exponiert, während über definierte Umgebungsvariablen (u. a. `VIRTUAL_HOST` und `VIRTUAL_PATH`) die Anbindung an den Reverse Proxy realisiert wird. Zur Sicherung der Persistenz werden die trainierten

Modelle und Cache-Dateien über Volumes eingebunden, sodass der Systemzustand auch nach Neustarts konsistent erhalten bleibt. Die Erzeugung des lauffähigen Containers basiert auf einem generischen `Dockerfile`, das eine Python-3.10-Umgebung bereitstellt und somit eine reproduzierbare Grundlage für die Ausführung des Systems schafft.

### 8.1.6 Integration im UCMR

Im Hinblick auf eine praxisorientierte Anwendbarkeit von SGAAIRE wurde die nicht-funktionale Anforderung *REQ\_NFA\_2: Lose-gekoppelte Integration in UCMR* definiert (vgl. Abschnitt 5.2). Zu diesem Zweck erfolgte die prototypische Einbettung von SGAAIRE in das Frontend des OFFIS UCMR. Das UCMR selbst ist in Java unter Verwendung des Spring Frameworks als Backend implementiert, während das Frontend in Angular realisiert ist. Die Containerisierung vom UCMR wurde ebenfalls mittels Docker bzw. docker-compose umgesetzt, weshalb SGAAIRE als zusätzlicher Dienst innerhalb des Ökosystems integriert werden kann. Die Interoperabilität zwischen dem OFFIS UCMR und SGAAIRE wird durch die Nutzung der IEC 62559-3 Use Case XML-Serialisierung sowie, auf Seiten des SGAM Frameworks, des DISCERN XML-Formats erreicht mit Verwendung einer HTTP-API von SGAAIRE, welche über Flask bereitgestellt wird. Das Frontend von SGAAIRE wurde an den folgenden Punkten im OFFIS UCMR integriert:



Quelle: Eigene Darstellung

Abbildung 8.3: Screenshot zur Illustration des IEC 62559 Use Case-Hub

- **Beim Editieren von IEC 62559-Anwendungsfällen** wird den Modellierenden ein Vorschlagshub (vgl. Abbildung 8.3) angezeigt, der auf der jeweils vorliegenden Anwendungsfallbeschreibung basiert. Dieser Hub umfasst verschiedene Kategorien, dar-

unter die Domänen- und Zonenklassifikation, die Service-Step-Klassifikation, die Use Case-Klassifikation, undefinierte Terme sowie Vorschläge zu fehlenden, redundanten oder zu ergänzenden Elementen. Die Vorschläge werden gemäß der Logik der KI-Module generiert (vgl. Abschnitt 7.4) und in strukturierter Form bereitgestellt. Zusätzlich werden innerhalb der Editorseite Glühbirnen-Icons verwendet, um an den entsprechenden Eingabefeldern visuell darauf hinzuweisen, dass ein konkreter Vorschlag vorliegt. Einzelne Empfehlungen, beispielsweise zur Service-Step-Klassifikation (vgl. Abbildung 8.4), werden mit erläuternden Hinweisen sowie einem Wahrscheinlichkeitswert transparent dargestellt.

Step No.	Triggering event	Producer	Receiver	Name
1a	The Actor Energy Broker wants to update one or more incentive tables on the Actor CEM.	Energy Broker	Customer Energy Manager	Sending incentive table with (P) PoE and/or DF
1b	The Actor Energy Broker wants to update one or more incentive tables on the Actor CEM.	Energy Broker	Customer Energy Manager	Sending incentive table with (P) DF
2	-	-	Charging Station	Sending pricing information to the charging station

The AI support system "SGAAIRE" has a recommendation for the Activity '1a' (Part of Scenario Transfer of Incentive Schemes to Custom Energy Manager and Charging Stations) based on the composition of the name and the description of the activity.

The system recommends the following service step(s):

- REPORT (Score 3.3776)

Dismiss    Apply

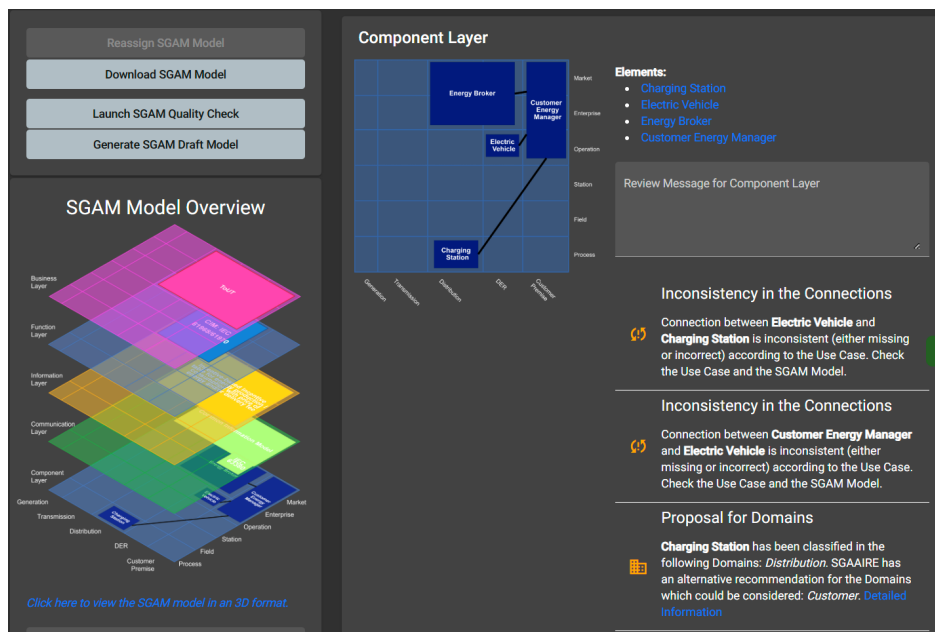
Quelle: Eigene Darstellung

Abbildung 8.4: Screenshot zur Illustration eines Service Step-Vorschlag für ein IEC 62559 Use Case

Für die prototypische Implementierung wurde insbesondere das Angular Materials [Kot20] Komponentenmodul `MatSnackBar` eingesetzt. Beim Aufruf der Editorseite wird der Vorschlagshub auf Grundlage der aktuell vorliegenden IEC 62559-Use Case-Beschreibung (über die XML-Serialisierung) initial geladen. Darüber hinaus kann der Ladevorgang durch das Aktualisieren-Icon im Hub zur Laufzeit erneut ausgelöst werden.

- **In der SGAM-Ansicht des Anwendungsfalls** wurden zusätzliche Funktionalitäten in Form von Buttons zur Erstellung eines SGAM-Qualitätsberichts sowie zur Generierung eines SGAM-Arbeitsmodells integriert.

Der Qualitätsbericht berücksichtigt insbesondere die funktionalen Anforderungen auf Seiten des SGAM und prüft diese im Hinblick auf Konsistenz und potenzielle Fehler. Dabei werden die entsprechenden Informationen aus der verknüpften IEC 62559 Use Case-Beschreibung einbezogen. Die in Abbildung 8.5 dargestellte Implementierung stellt eine prototypische Realisierung innerhalb des OFFIS UCMR dar, die vor allem für das in Kapitel 9 beschriebene Evaluationsszenario verwendet wird. Die Ansicht sieht eine layer-separierte Struktur vor, in der für jede SGAM-Ebene spezifische Empfehlungen bereitgestellt werden, die jeweils erläutert und begründet sind.



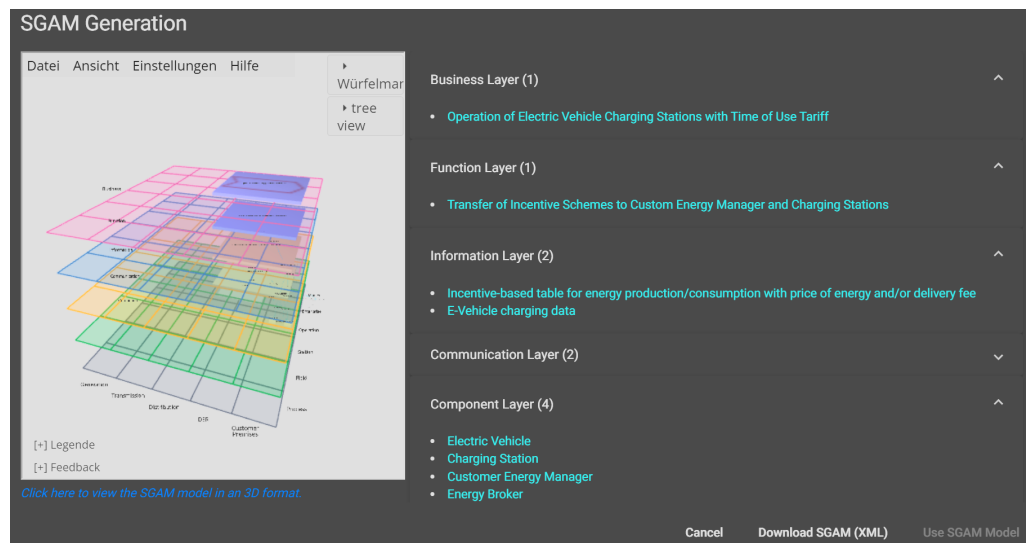
Quelle: Eigene Darstellung

Abbildung 8.5: Screenshot zur Illustration des SGAM Qualitätschecks

Über den Button zur Generierung eines SGAM Arbeitsmodells öffnet sich ein `MatDialog` (vgl. Abbildung 8.6), in dem eine vorläufige Ansicht mit einer Kurzübersicht des Modells dargestellt wird, wofür der SGAM Visualizer [GDU17] das Modell in 3D visualisiert. Zudem zeigt der Dialog die potenziellen Objekte. Das generierte SGAM Arbeitsmodell kann zum aktuellen Stand der Arbeit als XML-Datei exportiert werden, wobei die DISCERN-XML-Serialisierung verwendet wird. Gleichzeitig dient es als Referenzmodell für weitere Analysen und Vergleiche innerhalb des SGAM-Kontexts.

## 8.2 Datenerhebung- und Überführung

Mit der Qualität und Quantität des Datensatzes steht und fällt die Leistungsfähigkeit von KI-Modellen [KJH<sup>+</sup>19, KA18]. Neben der Auswahl passender Datensätze für die einzelnen Modelltypen ist insbesondere bei nicht-großen LMs eine ausreichende Datenmenge entscheidend, um eine adäquate Performanz zu erzielen. Gleichzeitig spielt die Datenqualität auch für LMs und LLMs eine zentrale Rolle, da insbesondere die Vorteile des Few-Shot Learnings nur bei qualitativ hochwertigen Trainingsdaten voll zur Geltung kommen [KJH<sup>+</sup>19, BMR<sup>+</sup>20]. Der Prozess zur Datenerhebung für diese Dissertation lässt sich in drei wesentliche Schritte zusammenfassen, die sich auch in der Literatur widerspiegeln [HKP12]. Diese Schritte bilden den Kern des Prozesses zum Datensatz und werden in den folgenden Unterabschnitten im Detail ausgeführt:



Quelle: Eigene Darstellung

Abbildung 8.6: Screenshot zur Generierung von SGAM Modellen

- **Datenerhebung:** Sammlung der relevanten Informationen aus den zugrunde liegenden Quellen, beispielsweise aus IEC 62559 Use Cases, SGAM Modellen sowie Literatur wie Standards, Referenzmodelle, Übersichtsarbeiten etc. Hierbei wurden sowohl strukturierte als auch unstrukturierte Daten berücksichtigt.
- **Datenüberführung:** Transformation der erhobenen Daten in ein einheitliches, maschinenlesbares Format, das als Input für die Trainingsprozesse geeignet ist. Dies beinhaltet die Überführung in eine Input/Output-Form, die Normalisierung von Begrifflichkeiten, die Beseitigung von Inkonsistenzen sowie die Anreicherung der Daten durch Annotationen.
- **Datenbereitstellung:** Ablage der überführten Daten in einem konsistenten Datensatz, der für weitere Analysen und die spätere Anwendung im entwickelten Bewertungsverfahren genutzt werden kann. Die Bereitstellung umfasst dabei sowohl die Speicherung in strukturierter Form als auch die Dokumentation der enthaltenen Inhalte.

### 8.2.1 Datenerhebung

Die Datenerhebung bildet die Grundlage für das Fine-Tuning der Modelle. Dabei wurden sowohl bestehende IEC 62559 Use Cases als auch SGAM Modelle gesammelt. Ergänzend wurden aus weiterer Literatur zusätzliche Daten und Ableitungsregeln integriert, um den Datensatz durch höherwertige Datensätze anzureichern.

### 8.2.1.1 IEC 62559 Use Cases und SGAM Modelle

Verschiedene IEC 62559 Use Cases und SGAM Modelle wurden aus Quellen wie dem BRIDGE Use Case Repository [DAZ<sup>+</sup>23a], dem EPRI Use Case Repository [Ele25], internen Use Cases und SGAM Modellen sowie aus Projektergebnissen wie aus TDX-Assist [LCR<sup>+</sup>19] oder CoordiNet [GMIFM<sup>+</sup>23] berücksichtigt und extrahiert [KLS25].

Für die Datenerhebung war maßgeblich, dass Use Cases dem IEC 62559-2 Use Case-Template [IEC15] folgen und SGAM Modelle das SGAM Framework [CEN12] konsistent abbilden. Im EPRI Use Case Repository finden sich vermehrt Use Cases, die nicht strikt dem IEC 62559-2 Template entsprechen bzw. eine abweichende Strukturierung aufweisen. Auf SGAM-Seite wurden ausschließlich Modelle berücksichtigt, die die fünf SGAM-Ebenen (Business, Function, Information, Communication, Component) sowie die Domänen und Zonen konsistent referenzieren und nachvollziehbar verknüpfen. Die folgenden Tabellen 8.1 und 8.2 zeigen eine Übersicht der selektierten Datenquellen.

Tabelle 8.1: Übersicht der extrahierten IEC 62559 Use Cases

Projekt / Quelle	Anzahl Use Cases
Accept (D2.1) [LVHD23]	14
BRIDGE Accept [DAZ <sup>+</sup> 23a]	4
CoordiNet [GMIFM <sup>+</sup> 23]	3
Crossbow [DAZ <sup>+</sup> 23a]	1
E-Balance Plus [DAZ <sup>+</sup> 23a]	5
ELAND [DAZ <sup>+</sup> 23a]	3
GIFT [DAZ <sup>+</sup> 23a]	1
MAESHA [DAZ <sup>+</sup> 23a]	1
OneNet [DAZ <sup>+</sup> 23a]	15
PARITY [DAZ <sup>+</sup> 23a]	2
PUC [DAZ <sup>+</sup> 23a]	1
TIGON [DAZ <sup>+</sup> 23a]	1
Generic BRIDGE UCs [DAZ <sup>+</sup> 23a]	2
XFLEX [DAZ <sup>+</sup> 23a]	2
EDF [DAZ <sup>+</sup> 23a]	2
Elering [DAZ <sup>+</sup> 23a]	19
GIFT (Projekt) [DAZ <sup>+</sup> 23a]	4
InteGrid [DAZ <sup>+</sup> 23a]	4
OFFIS [DAZ <sup>+</sup> 23a]	2
TDX-Assist [LCR <sup>+</sup> 19]	11
<b>Summe</b>	<b>97</b>

Tabelle 8.2: Übersicht der extrahierten SGAM Modelle

Projekt / Quelle	Anzahl SGAM Modelle
CoordiNet [GMIFM <sup>+</sup> 23]	13
Flexgrid [DAZ <sup>+</sup> 23a]	11
GIFT [DAZ <sup>+</sup> 23a]	4
MAESHA [DAZ <sup>+</sup> 23a]	6
TDX-Assist [LCR <sup>+</sup> 19]	13
TwinErgy [DAZ <sup>+</sup> 23a]	29
Crossbow [DAZ <sup>+</sup> 23a]	1
Interne SGAM Modelle	2
<b>Summe</b>	<b>79</b>

### 8.2.1.2 Weitere Datenquellen

Neben den erfassten IEC 62559 Use Cases und SGAM Modellen wurden für die Analyse auch ergänzende Referenzdokumente und Hilfestellungen aus dem Domänenbereich berücksichtigt. Diese enthalten beispielsweise Klassifikationen oder Zusatzinformationen, die zur Herleitung der Datensätze berücksichtigt werden. Insbesondere für die klassifikationsorientierten Modelle (Interoperabilitätsebenen, Domänen, Zonen, etc.) können Regeln abgeleitet werden. Die folgende Übersicht fasst die eingebundenen zusätzlichen Dokumente zusammen:

- **European Network of Transmission System Operators for Electricity (ENTSO-E) Harmonized Role Model** (*entso-e\_role\_model.xml*): Definiert die relevanten Rollen und Akteure im europäischen Strommarkt [ENT22].
- **Smart Grid Standardization Map** (*smart\_grid\_standard\_map.xml*): Übersicht über einschlägige Smart-Grid-Standards und deren Einordnung in die SGAM-Domänen und -Ebenen [IEC22].
- **IEC 62357-1** (*iec62357-1.xml*): Teil der Referenzarchitektur für (die Selektion) energiebezogene Kommunikationsstandards [IEC16].
- **NIST Logical Reference Model** (*nist\_logical\_reference\_model.xml*): Logisches Referenzmodell für die US-amerikanische Smart Grid-Architektur [NEU16].
- **SGAM User Manual** (*sgam\_user\_manual\_(example)\_a/b/c.xml*): Handreichungen und Beispiele zur praktischen Anwendung des SGAM Frameworks aus dem offiziellen Handbuch [CEN14].
- **Ambient Assisted Living (AAL)** (*aal.xml*): In SGAM Domänen und Zonen klassifizierte Beispielsakteure- und Komponenten aus dem Umfeld AAL, zur Abbildung von Anwendungsfällen außerhalb des klassischen Energiesektors. Extrahiert aus internen Arbeiten.

## 8.2.2 Datenüberführung

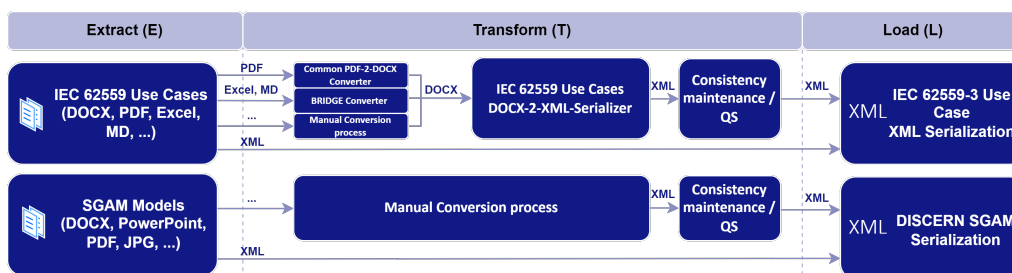
Die vorliegenden Daten sind so aufzubereiten, dass sie für das Training der jeweiligen KI-Modelle nutzbar werden. Die Datenerhebung erfolgt primär auf Basis bestehender IEC 62559 Use Cases sowie SGAM Modelle (vgl. Abschnitt 8.2.1.1), wird jedoch durch weitere Quellen, insbesondere relevante Literatur und anerkannte Standards, ergänzt (vgl. Abschnitt 8.2.1.2). Da die IEC 62559 Use Cases und die SGAM Modelle in heterogenen Datenformaten und -strukturen vorliegen (u. a. DOCX-, PDF- und Markdown-Dateien, in einzelnen Fällen auch XML-Dateien sowie Bilddateien für SGAM Modelle), ist eine Überführung in ein standardisiertes Format erforderlich, was die einheitlichere Weiterverarbeitung der Daten sowie eine erhöhte Reproduzierbarkeit der Ergebnisse ermöglicht.

### 8.2.2.1 Überführung in XML-Repräsentationen von Use Cases und SGAM Modellen

Für die Überführung der heterogen vorliegenden IEC 62559 Use Cases sowie SGAM Modelle wird eine dedizierte Datenpipeline eingesetzt, deren Aufbau sich insbesondere an etablierten Vorgehensweisen aus dem Bereich der Datenanalyse orientiert [HKP12]. Abbildung 8.7 veranschaulicht die einzelnen Phasen des Prozesses sowie die dabei verwendeten Werkzeuge, die in den Gesamtworkflow integriert sind. Im Rahmen dieser Arbeit wurden die folgenden Schritte durchlaufen:

- *Extract (E)*: Unterschiedliche IEC 62559 Use Cases und SGAM Modelle wurden aus verschiedenen Quellen extrahiert (vgl. Abschnitt 8.2.1).
- *Transform (T)*: Im Transformationsschritt werden die heterogenen Ausgangsdaten in ein einheitliches, standardisiertes Format überführt. Sofern die IEC 62559 Use Cases in PDF-Dateien vorlagen, erfolgte zunächst eine Konvertierung in das DOCX-Format unter Verwendung gängiger PDF-zu-DOCX-Tools, um eine Weiterverarbeitung zu ermöglichen. Das BRIDGE Repository [DAZ<sup>+</sup>23a] nutzt für die Verwaltung ein Markdown-Format; im offiziellen Repository steht jedoch ein BRIDGE-Konverter zur Verfügung, der die Dateien direkt in eine IEC 62559-3 Use Case XML-Serialisierung überführt [DAZ<sup>+</sup>23a]. Für weitere Dateiformate sowie bei auftretenden strukturellen Inkonsistenzen waren manuelle Nachbearbeitungsschritte erforderlich, bspw. die Korrektur fehlerhafter Templatestrukturen oder die Rekonstruktion in das standardisierte DOCX-Template. Die Transformation der SGAM Modelle gestaltet sich anders, da diese in einer Vielzahl heterogener PowerPoint-Vorlagen, Bilddateien sowie Abbildungen in PDF-Form vorliegen. Dies macht vorwiegend umfangreiche manuelle Konvertierungsschritte erforderlich. [KLS25]

Der im Rahmen dieser Arbeit entwickelte *IEC 62559 DOCX-2-XML Serializer* stellt ein eigenständiges Softwareartefakt als Nebenprodukt dar, das die halb-automatisierte Transformation von strukturell heterogenen IEC 62559-2 Use Cases in eine standardkonforme XML-Repräsentation gemäß IEC 62559-3 unterstützt. Der Serializer ist als Spring-Shell-Komponente realisiert und kann über eine Kommandozeilenschnittstelle genutzt werden. Er unterstützt die Eingabe von Use Case-Dateien in unterschiedlichen Formaten (z. B. DOCX, PDF, Markdown), erkennt diese anhand der Dateiendung und wählt darauf basierend den geeigneten Parser aus. Die Verarbeitung umfasst dabei mehrere Schritte: Das Einlesen und Parsen des Quelldokuments, die Normalisierung der Inhalte sowie Reparatur inkonsistenter Referenzen, und die Serialisierung in ein valides XML-Dokument mittels Jakarta XML Binding (JAXB), das dem Schema der IEC 62559-3 entspricht. Da im Prozess aufgrund der Heterogenität der Datenstrukturen Fehler auftreten können, werden den Anwendenden Vali-



Quelle: Eigene Darstellung

Abbildung 8.7: Datenpipeline für die Überführung in einheitliche XML-serialisierte IEC 62559 Use Cases sowie SGAM Modelle

```

<?xml version="1.0" encoding="utf-8"?>
<UseCaseRepository>
  <ActorLibrary>
    <Actor>
      <mRID>3b6d4cd1-7cd7-473e-af25-2cd3f990d415</mRID>
      <name>Electric Vehicle</name>
      <description>...</description>
      <type>Internal</type>
    </Actor>
    ...
  </ActorLibrary>
  <UseCaseLibrary>
    <UseCase>
      <name>Time of Use Tariff for Electric Vehicle Charging Stations</name>
      <ActorGrouping>
        <name>Application-related Actors</name>
        <description>Interest in charging/producing energy... </description>
        <FurtherActorInformation>
          <Actor>
            <mRID>3b6d4cd1-7cd7-473e-af25-2cd3f990d415</mRID>
            ...
          </Actor>
          ...
        </FurtherActorInformation>
      </ActorGrouping>
      ...
    </UseCase>
  </UseCaseLibrary>
  ...
</UseCaseRepository>

```

Listing 1: Ausschnitt aus der IEC 62559-3 Use Case XML-Serialisierung für den Anwendungsfall „Verwendung von dynamischen Stromtarifen für EV-Ladestationen“

dierungshinweise und unterstützende Meldungen bereitgestellt. Diese umfassen beispielsweise den Hinweis auf fehlerhafte Akteursangaben oder ungültige Referenzen. Für jede Ausführung des Tools ist ein Konsistenzabgleich mit der jeweiligen Quelldatei erforderlich. Dabei müssen erkannte Fehler manuell korrigiert werden, bis die erzeugte Zielfdatei semantisch konsistent zur ursprünglichen Quelldatei vorliegt. Semantische und pragmatische Fehler, die bereits in den Quelldateien enthalten sind, werden im Rahmen der Transformation jedoch bewusst nicht korrigiert, sondern in die resultierenden XML-Dateien übernommen.

- *Load (L)*: Nach Abschluss der Transformation liegen die resultierenden XML-Dateien in standardkonformer Form vor und werden zentral in einem GitLab-Repository<sup>2</sup> bereitgestellt [KLS25].

Beispielhafte Ausschnitte der IEC 62559-3 Use Case XML-Serialisierung sowie der DISCERN XML-Serialisierung sind in den Listings 1 und 2 dargestellt. Diese beziehen sich auf den Anwendungsfall „Verwendung von dynamischen Stromtarifen für EV-Ladestationen“.

### 8.2.2.2 Überführung in KI-Modelle

Für das Training der einzelnen KI-Modelle müssen die zugrunde liegenden Daten in Übereinstimmung mit dem Lösungs- und Dimensionsraum ( $L, D$ ) für alle Datenquel-

<sup>2</sup> [https://gitlab.offis.de/official/fbe-p\\_ucmr/sgaaire.dataset](https://gitlab.offis.de/official/fbe-p_ucmr/sgaaire.dataset)

```

<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns2:SGAM xmlns:ns2="http://www.discern.eu/XSDSchemas/DISCOVERNSgam">
  <ComponentLayer>
    <Component>
      <id>C-1</id>
      <name>EV Charging Station</name>
      <description>...</description>
      <Cell>
        <domain>Distribution</domain>
        <zone>Process</zone>
      </Cell>
    </Component>
    ...
    <Component>
      <id>C-4</id>
      <name>Customer Energy Manager</name>
      <description>...</description>
      <Cell>
        <domain>CustomerPremises</domain>
        <zone>Operation</zone>
      </Cell>
    </Component>
    <CommunicationLink>
      <SourceComponent><id>C-4</id></SourceComponent>
      <TargetComponent><id>C-1</id></TargetComponent>
    </CommunicationLink>
    ...
  </ComponentLayer>
  ...
</ns2:SGAM>

```

Listing 2: Ausschnitt aus der DISCERN XML-Serialisierung für den Anwendungsfall „Verwendung von dynamischen Stromtarifen für EV-Ladestationen“

len einheitlich vorliegen. Der Lösungsraum  $L$  beschreibt dabei die möglichen Ausgaben eines Modells, beispielsweise in Form von Klassifikationslabels oder generierten Tokens. Der Dimensionsraum  $D$  hingegen umfasst die Eingaben, die in einer Sequenzrepräsentation über Tokens kodiert und dem Modell zugeführt werden, um die jeweiligen Ausgaben zu erzeugen. Damit Eingabe- und Ausgabebetupel gebildet werden können, muss der Dimensionsraum über alle berücksichtigten Datenarten hinweg konsistent vorhanden sein. Tabelle 8.3 zeigt für die Modelle M1 - M6 den jeweiligen Lösungs- und Dimensionsraum sowie die Felder aus den zugrunde liegenden Standards und definierten Mappings, die als Datenbasis verwendet werden.

Die Modelle  $M7a$  (*Acronym Identification*) und  $M7b$  (*Acronym Disambiguation*) basieren nicht auf dem in dieser Arbeit erstellten Datensatz. Für  $M7a$  wird der Datensatz von [VDTN20] mit  $N = 17,473$  Sequenzen verwendet (Train: 14,006  $\approx$  80 %, Val: 1,717  $\approx$  10 %, Test: 1,750  $\approx$  10 %). Für  $M7b$  wird der MadDog Datensatz von [PBVDCN21] mit  $N = 46 \times 10^6$  Einträgen herangezogen, der analog in Anteile von ca. 80 % Training, 10 % Validierung und 10 % Test aufgeteilt wird.

### 8.2.3 Datenbereitstellung

Die aufbereiteten Daten werden in Form von CSV-Dateien bereitgestellt, welche im Rahmen von SGAAIRE und des Modelltrainings mithilfe von `pandas` in Python eingelesen werden können. Dabei erfolgt eine standardisierte Aufteilung in 80 % Trainings- und 20 % Testdaten, die jeweils in separaten Dateien abgelegt sind. Tabelle 8.4 gibt ei-

Modell	Lösungs- und Dimensionsraum ( $L, D$ )	Datenbasis
M1: SGAM Domänen Klassifikator	$L$ : Generation, Transmission, Distribution, DER, Customer $D$ : Name, Description	<ul style="list-style-type: none"> <li>• <b>IEC 62559-3 XML Use Case-Serialisierung:</b> (CustomInformation), (Actor)</li> <li>• <b>DISCERN SGAM XML-Format:</b> IdentifiedObject</li> <li>• <b>Regeln:</b> Alle weitere Datenquellen (vgl. Abschnitt 8.2.1.2)</li> </ul>
M2: SGAM Zonen Klassifikator	$L$ : Process, Field, Station, Operation, Enterprise, Market $D$ : Name, Description	vgl. M1: SGAM Domänen Klassifikator
M3: SGAM Interoperabilitätslayer Klassifikator	$L$ : Component, Communication, Information, Function, Business $D$ : Name, Description	<ul style="list-style-type: none"> <li>• <b>IEC 62559-3 XML Use Case-Serialisierung:</b> Mapping gemäß [FSBU20]</li> <li>• <b>DISCERN SGAM XML-Format:</b> IdentifiedObject</li> <li>• <b>Regeln:</b> Standardsangaben als Communication- oder Information-Label gemäß [SKKK24], Akteure als Component-Label</li> </ul>
M4: Service Step Klassifikator	$L$ : CREATE, GET, CHANGE, DELETE, CANCEL, CLOSE, EXECUTE, REPORT, TIMER, REPEAT $D$ : Name, Description	<ul style="list-style-type: none"> <li>• <b>IEC 62559-3 XML Use Case-Serialisierung:</b> Activity</li> </ul>
M5: SGAM Element Proposer	$L$ : beliebiges Objekt $D$ : Component, Communication, Information, Function, Business	<ul style="list-style-type: none"> <li>• <b>DISCERN SGAM XML-Format:</b> IdentifiedObject (Permutationen aus allen verbundenen Elemente innerhalb eines SGAM Modells)</li> </ul>
M6: Abstraktionsgrad Regressor	$L$ : Reelle Werte $x \in \mathbb{R}$ (Level-Cluster) $D$ : Name, Description	<ul style="list-style-type: none"> <li>• <b>IEC 62559-3 XML Use Case-Serialisierung:</b> Mapping der Anwendungshilfe [FSBU20]</li> <li>• <b>DISCERN SGAM XML-Format:</b> Mapping auf Basis der verschiedenen Abstraktionslevel im SGAM [CEN14]</li> </ul>

Tabelle 8.3: Datenquellen für die KI-Modelle (M1-M6)

ne Übersicht über die Anzahl der für das Training verwendeten Daten, einschließlich der jeweiligen Aufteilung sowie der zusätzlich künstlich angereicherten Daten (vgl. Abschnitt 8.2.4). Die aufgelisteten Dateien sind im `data`-Verzeichnis des Git-Repositories von SGAAIRE<sup>3</sup> jeweils für die KI-Modelle in `training.csv` bzw. `text.csv` abgelegt. Die bereitgestellten Dateien bilden die Grundlage für das Training der KI-Modelle, da sie die für den Dimensions- und Lösungsraum ( $L, D$ ) relevanten Werte enthalten. Diese liegen in Form von Tupeln  $(d, l)$  mit  $d \in D$  (Eingabedaten) und  $l \in L$  (Ausgabedaten) vor und können somit unmittelbar für die Modellanpassung genutzt werden. Ein exemplarischer Ausschnitt der Trainingsdaten ist in Listing 3 dargestellt. Dieser verdeut-

<sup>3</sup> [https://gitlab.offis.de/official/fbe-p\\_ucmr/SGAAIRE](https://gitlab.offis.de/official/fbe-p_ucmr/SGAAIRE)

```

name,description,Generation,Transmission,Distribution,DER,Customer
Transformer,...,0.0,1.0,1.0,0.0,0.0
CIM: IEC 61968/61970,0.0,1.0,1.0,0.0,0.0
Smart Meter,...,0.0,0.0,0.0,0.0,1.0
Plug-In Electric Vehicle,...,0.0,0.0,0.0,0.0,1.0
...

```

Listing 3: Ausschnitt aus der Trainings CSV-Datei für das KI-Modell M1: SGAM Domänen-Klassifikator

licht, dass sowohl die Eingabedaten (`name`, `description`) als auch die korrespondierenden Ausgabedaten (z. B. `Generation`, `Transmission`, `Distribution`, `DER`, `Customer`) explizit enthalten sind. Sämtliche Schritte der Datenerhebung und -überführung wurden in Form einer Pipeline konzipiert und durchgeführt. Die jeweiligen Zwischenergebnisse wurden in getrennten Verzeichnissen abgelegt, um sowohl die Nachvollziehbarkeit als auch die Reproduzierbarkeit des Gesamtprozesses sicherzustellen.

Tabelle 8.4: Übersicht der Datensätze und deren Aufteilung für die KI-Modelle M1–M6

Quelle: Angelehnt an [KLS25]

Modell	Rohdaten	Data Augmentation	Datensatz (Training 80 % / Test 20 %)
<b>M1: SGAM Domänen Klassifikator</b>	664	605	1,269 (1,048 / 221)
<b>M2: SGAM Zonen Klassifikator</b>	664	411	1,075 (859 / 216)
<b>M3: SGAM IOP-Layer Klassifikator</b>	2,912	0	2,912 (2,329 / 583)
<b>M4: Service Step Klassifikator</b>	691	245	936 (748 / 188)
<b>M5: SGAM Element Proposer</b>	76,657	0	76,657 (60,527 / 16,130)
<b>M6: Abstraktionsgrad Regressor</b>	2,433	0	2,433 (1,945 / 488)

#### 8.2.4 Künstliche Datenanreicherung mittels Data Augmentation

Insbesondere für die Modelle M1 (SGAM Domänen Klassifikator), M2 (SGAM Zonen Klassifikator) und M4 (Service Step Klassifikator) steht im Vergleich zu den übrigen Modellen eine geringere Menge an Rohdaten für das Training zur Verfügung. Es wurden Techniken der Data Augmentation eingesetzt, um die Trainingsdaten zu erweitern, eine größere Varianz in den Daten zu erzeugen und insbesondere den nicht-großen Sprachmodellen eine erweiterte Datenbasis bereitzustellen. Im Folgenden werden die eingesetzten Verfahren beschrieben:

- **Künstliche Generierung zusätzlicher Datensätze mittels ChatGPT-4:**

Zur Erzeugung der künstlichen Datenpunkte für den SGAM Domänen- sowie Zonen-Klassifikator und dem Service Step-Klassifikators wurde OpenAI ChatGPT mit GPT-

4 [OAA<sup>+</sup>24] eingesetzt. Hierfür wurde ein Prompt-basiertes Vorgehen gewählt, bei dem GPT-4 angewiesen wurde, eine vordefinierte CSV-Struktur in englischer Sprache zu befüllen.

Der Prompt (exemplarisch in Listing 4 dargestellt) spezifizierte dabei sowohl die erforderlichen Eingabemerkmale (*name, description*) als auch die gewünschten Klassifikationsausgaben (bspw. für den SGAM-Domänenklassifikator: *Generation, Transmission, Distribution, DER, Customer*). Um die semantische Konsistenz sicherzustellen, wurden den Prompts domänenspezifische Definitionen beigegeben. So wurde etwa für die Domäne (Bulk) Generation die Definition aus dem SGAM User Manual [CEN14] verwendet. Das Modell wurde angewiesen, pro Klasse 50 Datenpunkte in einer CSV-Datei zu generieren. Dieser Prozess wurde für sämtliche Klassen und für alle drei Klassifikatoren wiederholt. Da GPT-4 nicht in allen Fällen exakt 50 Einträge erzeugte, entstanden leichte Abweichungen in der Datensatzgröße, die im Kontext der synthetischen Datengenerierung jedoch akzeptabel waren.

```

Generate 50 data points with the following CSV-structure components that
→ are classified in the "(Bulk) Generation" Domain according the Smart
→ Grid Architecture Model (SGAM) Framework, where description should be
→ a two sentence description as a definition of the name:

name, description, Generation, Transmission, Distribution, DER, Customer

Definition of Generation: "Representing generation of electrical energy in
→ bulk quantities typically connected to the transmission system, such
→ as by fossil, nuclear and hydro power plants, off-shore wind farms,
→ large scale solar power plant (i.e. PV, CSP)."
```

Listing 4: Prompt zur Generierung synthetischer Datenpunkte für den SGAM Domänen Klassifikator für die Generation-Domäne

- **Synthetische Vervollständigung fehlender Eingabefelder mittels FLAN-T5:**

Der erhobene Datensatz besteht im Wesentlichen aus Projektergebnissen, bei denen nicht in allen Fällen sämtliche Felder adäquat ausgefüllt wurden. Da die relevanten Eingabemerkmale für das Finetuning der Modelle insb. den Namen sowie die Beschreibung umfassen, war es erforderlich, dass beide Merkmale gleichzeitig vorliegen. Dies war jedoch nicht bei allen IEC 62559-Use Cases oder SGAM Modellen gegeben.

Die Vollständigkeit des Datensatzes hinsichtlich der Beschreibungen wurde dadurch gewährleistet, dass fehlende Einträge automatisiert mittels FLAN-T5 aus den entsprechenden Namen abgeleitet wurden. Der Ablauf dieser Transformation ist in Algorithmus 2 spezifiziert. Auf diese Weise konnte sichergestellt werden, dass jeder Datensatz für das nachfolgende Finetuning über eine konsistente Eingabestruktur verfügte.

- **Oversampling:** Für die Klassifikationsmodelle wurde Oversampling mittels Weighted Sampling angewandt, um eine ausgewogene Klassenverteilung sicherzustellen

[KLS25]. Der eigentliche Trainingsdatensatz wird durch das Oversampling nicht modifiziert und ist folglich nicht in Tabelle 8.4 berücksichtigt. Vielmehr wird dieser Schritt im Rahmen der Vorverarbeitung zum Modelltraining realisiert und wirkt sich ausschließlich auf die Zusammenstellung der Trainingsinstanzen während des Modelltrainings aus.

### 8.3 Feinabstimmung, Selektion und Integration von KI-Modellen

In diesem Abschnitt wird die Umsetzung und Auswahl der eingesetzten KI-Modelle für SGAAIRE sowie dessen Integration auf Basis der zuvor dargestellten Konzeption (vgl. Abschnitt 7.5) beschrieben. Zunächst erfolgt die Auswahl geeigneter, vortrainierter Sprachmodelle als Grundlage. Diese werden anschließend auf den erhobenen Datensatz (vgl. Abschnitt 8.2) für die spezifischen Aufgaben feinjustiert. Die resultierenden Modelle werden verglichen und im Rahmen von SGAAIRE selektiert.

#### 8.3.1 Selektion der vortrainierten Sprachmodelle

Die grundlegende Beschreibung der Sprachmodelle wurde bereits im Abschnitt 4.3 dargelegt. Tabelle 8.5 bietet einen Überblick über die in dieser Arbeit für das Fine-Tuning herangezogenen Sprachmodelle (LMs) sowie Tabelle 8.6 für die verwendeten großen Sprachmodelle (LLMs). Berücksichtigt werden dabei sowohl aktuelle Modelle in ihren Basis- als auch, sofern verfügbar, in ihren Large-Varianten. Ergänzend wird mit EnergyBERT ein speziell für den Energiesektor entwickeltes Sprachmodell auf Basis von BERT einbezogen, das auf einem Korpus von rund 1,2 Millionen Volltextpublikationen aus den Bereichen Energie- und Materialwissenschaften trainiert wurde ([TYJB24]). [KLS25]

#### 8.3.2 Fine-Tuning der vortrainierten Sprachmodelle

Das Training und die Nutzung der KI-Modelle erfolgte unter Verwendung der Transformer-Bibliothek von Hugging Face [Jai22]. Die Hyperparameter orientieren sich an etablierten

---

#### Algorithmus 2: Generierung fehlender Beschreibungen mittels FLAN-T5

---

**Input:** Datensatz  $D = \{(name_i, description_i)\}_{i=1}^n$

**Output:** Datensatz  $D' = \{(name_i, description'_i)\}_{i=1}^n$  mit  $\forall i : description'_i \neq \varepsilon$

$D' \leftarrow \emptyset$

**foreach**  $(name_i, description_i) \in D$  **do**

**if**  $description_i = \varepsilon$  **then**

$description'_i \leftarrow \text{FLAN\_T5}(' \text{Explain:}' + name_i)$

**else**

$description'_i \leftarrow description_i$

$D' \leftarrow D' \cup \{(name_i, description'_i)\}$

**return**  $D'$

---

Tabelle 8.5: Überblick der verwendeten Sprachmodelle (LMs)

Quelle: [KLS25]

	BERT	DistilBERT	RoBERTa	DeBERTa	GPT-2	EnergyBERT
<b>Autor</b>	[DCLT19]	[SDCW20]	[LOG <sup>+</sup> 19]	[HLGC21, HG23]	[RWC <sup>+</sup> 19]	[TYJB24]
<b>Veröffentlichung</b>	2018	2019	2019	2020 (v3 2023)	2019	2022
<b>Max. Länge</b>	512 Token	512 Token	512 Token	512 Token	1024 Token	512 Token
<b>Modell-Name(n)</b>	• bert-base-uncased • bert-large-uncased	• distilbert-base-uncased	• roberta-base • roberta-large	• deberta-v3-base • deberta-v3-large	• gpt2 • gpt2-large	• EnergyBERT
<b>Typ</b>	Pre-Trained, Encoder-only	Pre-Trained, Encoder-only	Pre-Trained, Encoder-only	Pre-Trained, Encoder-only	Pre-Trained, Decoder-only	Kontinuierlich PT, Encoder-only
<b>Größe</b>	~450 MB, ~1,4 GB	~300 MB	~600 MB, ~1,5 GB	~750 MB, ~1,7 GB	~550 MB, ~3,5 GB	438 MB
<b>Parameter</b>	110 Mio., 335 Mio.	67 Mio.	125 Mio., 355 Mio.	86 Mio., 304 Mio.	137 Mio., 812 Mio.	110 Mio.

Tabelle 8.6: Überblick der verwendeten großen Sprachmodelle (LLMs)

Quelle: [KLS25]

	GPT-4	LLaMA	FLAN-T5
<b>Autor</b>	[OAA <sup>+</sup> 24]	[TLI <sup>+</sup> 23]	[CHL <sup>+</sup> 22]
<b>Veröffentlichung</b>	2024	2023	2022
<b>Max. Länge</b>	128k Token	128k Token	512 Token
<b>Modell-Name(n)</b>	• gpt-4o • gpt-4o-mini	• LLaMA 3.1 8B	• t5-base • t5-large
<b>Typ</b>	Pre-Trained, Decoder-only	Pre-Trained, Decoder-only	Pre-Trained, Encoder/Decoder
<b>Größe</b>	n/a	~16 GB	~900 MB, ~3,13 GB
<b>Parameter</b>	12 Mrd., 8 Mrd.	8 Mrd.	248 Mio., 783 Mio.

Konventionen im Bereich des Fine-Tunings vortrainierter Sprachmodelle und sind in Tabelle 8.7 aufgeführt. Das Modelltraining wurde GPU-basiert auf einem Desktop-System mit einer Nvidia RTX 3090 (24 GB GPU-Speicher) durchgeführt. [KLS25]

Tabelle 8.7: Auswahl der Hyperparameter beim Fine-Tuning

Quelle: [KLS25]

	Models M1 - M4, M6	Model M5
<b>Batch Size</b>	8	16
<b>Epochen</b>	20	20
<b>Lernrate</b>	$2 \cdot 10^{-5}$	$2 \cdot 10^{-5}$
<b>Sequenzlänge</b>	Gemäß Modell; Truncation	256 Tokens, Truncation & Padding

Sowohl im Training als auch in der Inferenz müssen die Eingaben in eine tokenisierte Sequenzform transformiert werden, die den Modellen als Input dient. Dabei ergeben sich zwei grundsätzliche Ausprägungen:

- **Modelle M1 - M4 und M6:** Bei diesen Modellen werden die Name- und Beschreibungstapel  $((n, d) \in \mathcal{X}_{\text{elem}})$  innerhalb der Eingabesequenz berücksichtigt, wodurch die Einführung weiterer spezieller Tokens eine eindeutige Strukturierung der Sequenz gewährleistet wird. Die Repräsentation orientiert sich an der in der Sprachmodel-

lierung etablierten Notation mit [CLS] (Klassifikationstoken) und [SEP] (Separatortoken) und wird durch die anwendungsfallspezifischen Tokens  $\langle start-name \rangle$ ,  $\langle end-name \rangle$ ,  $\langle start-description \rangle$  und  $\langle end-description \rangle$  ergänzt. Auf diese Weise entsteht eine formale Sequenzstruktur für Training und Inferenz, wobei die Platzhalter Name und Description durch kontextspezifische Inhalte substituiert werden (z. B. Name = „Electric Vehicle“, Description = entsprechende Beschreibung): [KLS25]

[CLS] $\langle start-name \rangle$ Name $\langle end-name \rangle$  $\langle start-description \rangle$ Description $\langle end-description \rangle$ [SEP]

- **Modell M5:** Für dieses Modell ist hinsichtlich der Struktur und vorgesehenen Eingaben (vgl. Abschnitt 7.5.5) eine andere Eingabesequenz notwendig. Hierbei muss sichergestellt werden, dass auf alle existierenden Layerinformationen ein entsprechender Output erzeugt wird, über  $\mathcal{I} = [(l_0, o_0), \dots, (l_n, o_n)]$ ,  $l_i \in L$ ,  $o_i \in O$ . Gesucht ist ein Zielobjekt  $o_p \in O$ , das auf einer festgelegten Ziel-Interoperabilitätsebene  $l_p \in L$  vorgeschlagen werden soll. Um die Ebenenzugehörigkeit explizit zu kennzeichnen, wird jedes Sequenzelement mit zusätzlichen anwendungsspezifischen Token für alle Interoperabilitätslayer versehen (z. B. [Component], [Information]). Die Eingabesequenz endet mit der gesuchten Interoperabilitätsebene  $l_p$ , wobei das zugehörige Objektelement  $o_p$  vom Modell bestimmt werden muss und beim Training vorliegt. Die Sequenzrepräsentation lässt sich wie folgt darstellen: [KLS25]

[CLS]  $[l_0 \in L] o_0 \in O \dots [l_n \in L] o_n \in O \rightarrow [l_p \in L] o_p \in O$  [SEP]

Die Trainingsstrategie umfasste eine epochenweise Evaluation sowie eine gleichzeitige Speicherung der Modelle, wobei die Anzahl der zu behaltenden Checkpoints auf zwei limitiert wurde. Es wurde nach jeder Epoche das Modell übernommen, das den höchsten F1-Wert (bei den Klassifikationsmodellen), Accuracy (bei dem Regressionsmodell) sowie bei dem CLM-Modell die geringste Loss (über CrossEntry) erreichte.

Als Aktivierungsfunktion für die Multi-Label Klassifikation wurde anstelle der voreingestellten Aktivierungsfunktion im Klassifikationskopf der Architektur eine Sigmoid-Funktion eingesetzt, sodass jedem Datenpunkt mehrere Klassenlabels zugeordnet werden konnten. Da die Sigmoid-Funktion einen kontinuierlichen Wertebereich  $\sigma(x) \in (0, 1)$  liefert, soll eine binäre Entscheidung darüber getroffen werden, ob ein Label einem Objekt entspricht oder nicht. Zu diesem Zweck wird pro Epoche ein adaptives Schwellenwertkriterium  $\Theta$  bestimmt. Werte oberhalb von  $\Theta$  werden dabei auf 1, Werte unterhalb auf 0 diskretisiert. Die Bestimmung von  $\Theta$  erfolgt so, dass primär die Accuracy maximiert und (bei Gleichstand) sekundär der F1-Score optimiert wird.

Der Einsatz von CLM ist im Rahmen des Modells *M5 (SGAM Element Proposer)* zur Generierung kohärenter Textsequenzen vorgesehen. CLM findet insb. in Decoderbasierten Sprachmodellen Anwendung. Alternativ kann Textgenerierung über CLM jedoch auch in Encoder-only Architekturen durch den Einsatz geeigneter Wrapper realisiert werden, indem MLM sequentiell angeordnet wird [Jai22]. Für FLAN-T5, das als Encoder-Decoder-Modell aufgebaut ist, ist ein direkter Einsatz von CLM nicht vorgesehen. Es wird auf Seq2Seq als Alternative zurückgegriffen, welches bereits in Abschnitt 4.4.4 erläutert wurde.

Eine Besonderheit stellt die Verwendung der GPT-4o sowie GPT-4o-Mini Modelle dar. Diese liegen im Gegensatz zu den anderen Modellen nicht als Open-Source- bzw. Weight-Modelle vor, sondern werden im Rahmen der OpenAI API nachtrainiert [OAA<sup>+</sup>24]. Für jedes Modell wurden dabei jeweils drei Epochen mit variabler Batch-Size und einem Lernrate Multiplikator von 1.8 durchlaufen. Ebenfalls über die Verwendung der ChatGPT API wurde jeder Datenpunkt klassifiziert und im Rahmen der Vorverarbeitung in die jeweiligen CSV-Dateien überführt.

### 8.3.3 Bewertungskriterien für die Selektion der KI-Modelle

Um die Leistung der trainierten KI-Modelle beurteilen zu können, ist die Auswahl geeigneter Bewertungsmetriken erforderlich. Diese wurden bereits in Abschnitt 4.4 im Rahmen der Betrachtung der KI-bezogenen Anwendungsbereiche vorgestellt. Die Auswahl der Metriken erfolgte auf Grundlage der jeweiligen Aufgabenstellung:

- **Text- und Tokenklassifikationsmodelle:** Für die Klassifikationsmodelle *M1 (SGAM Domänen)*, *M2 (SGAM Zonen)*, *M3 (Interoperabilitätsebenen)*, *M4 (Service Step)*, *M7a (Acronym Identifikation)* sowie *M7b (Acronym Disambiguation)* wurden Accuracy, Precision, Recall sowie der F1-Score, als harmonisches Mittel von Precision und Recall, betrachtet. Der F1-Score wurde als primäre Metrik für die Auswahl des Modells im SGAAIRE-System herangezogen, um ein ausgewogenes Verhältnis zwischen Precision und Recall sicherzustellen [KLS25].
- **Textregressionsmodell:** Für das Regressionsmodell *M6 (LoA Regressor)* wurden der MSE, der MAE sowie das Bestimmtheitsmaß  $R^2$  als Bewertungsmetriken herangezogen. MSE wurde als primäres Kriterium für die Modellauswahl im SGAAIRE-System gewählt, da dieser größere Fehler stärker gewichtet und damit eine strengere Bewertung ermöglicht. [KLS25]
- **Textgenerierungsmodell:** Für das Textgenerierungsmodell *M5 (SGAM Proposer)* stellt die Bewertung sowohl der syntaktischen als auch der semantischen Übereinstimmung mit den Referenztexten eine besondere Herausforderung dar. Da die generierten Ausgaben in der Regel nicht in Form kohärenter, durchgängiger Sätze vorliegen, ist ein direkter Vergleich erschwert. Darüber hinaus müssen Aspekte wie Kohärenz, Relevanz und sprachlicher Stil berücksichtigt werden. Als Annäherung an eine Bewertung werden für die syntaktische Analyse die Metriken METEOR (eine Erweiterung der BLEU-Metrik) sowie ROUGE (n-Gram-Überlappung) herangezogen, während für die semantische Analyse BERTScore (unter Verwendung der Kosinus-Ähnlichkeit zur kontextuellen Ausrichtung) eingesetzt wird. Der BERTScore wird auf Basis des Standardmodells roberta-large berechnet, welches eine hohe Korrelation mit menschlichen Bewertungen sowohl auf Satz- als auch auf Systemebene aufweist [ZKW<sup>+</sup>20]. Für die Selektion wird primär der BERTScore als Metrik zur semantischen Validierung herangezogen. Sekundär findet ROUGE Anwendung, tertiär METEOR, jeweils mit Betrachtung der F1-Werte. [KLS25]

### 8.3.4 SGAM Domänen Klassifikator (M1)

Auf Grundlage der in Tabelle 8.8 dargestellten Ergebnisse wird deutlich, dass die in der Dissertation betrachteten Encoder-basierten Modelle bessere Resultate erzielen als die Decoder-basierten Sprachmodelle. Insbesondere die Modelle *roberta-large* und *deberta-v3-large* erreichen mit  $F1_{roberta-large} = 0.83$  bzw.  $F1_{deberta-v3-large} = 0.82$  die besten Werte im Vergleich. Zwischen beiden Modellen lassen sich nur geringe Unterschiede feststellen, wobei *roberta-large* in allen betrachteten Metriken leicht überlegen und zudem kleiner ist und daher als leistungsfähigstes Modell für diese Aufgabe betrachtet werden kann. Demgegenüber zeigen die kompakten Varianten wie *distilbert-base-uncased* oder *roberta-base* etwas schwächere Ergebnisse (mit  $F1_{distilbert-base-uncased} = 0.79$  bzw.  $F1_{roberta-base} = 0.80$ ), liegen jedoch immer noch in einem akzeptablen Leistungsbereich. Die domänenspezifische Variante *EnergyBERT* liefert mit  $F1_{EnergyBERT} = 0.79$  eine vergleichbare Leistung wie die allgemeinen Baselines, ohne dabei einen erkennbaren Vorteil im energietechnischen Kontext zu zeigen.

Deutlich schwächer schneiden die generativen Decoder-Modelle (*gpt2(-large)*, *gpt4o* sowie *Llama-3.1-8B*) ab, die sowohl bei Accuracy als auch bei F1-Werten unterhalb der besten Encoder-Modelle liegen. Vor diesem Hintergrund wird *roberta-large* als Referenzmodell für das SGAAIRE-System ausgewählt, da es die beste Balance zwischen Genauigkeit und Robustheit bietet. Die Ergebnisse verdeutlichen zudem, dass eine weitere Feinabstimmung größerer generativer Modelle keinen zusätzlichen Nutzen erwarten lässt, da die Leistungsdifferenzen bereits klar zu Gunsten der Encoder-basierten Architekturen ausfallen.

Tabelle 8.8: Modellergebnisse für M1: SGAM Domänen Klassifikation

Quelle: Basierend auf [KLS25]

Model	M1: SGAM Domänen Klassifikation			
	Accuracy	Precision	Recall	F1
<i>bert-base-uncased</i>	0.74	0.81	0.79	0.80
<i>bert-large-uncased</i>	0.74	0.80	0.80	0.80
<i>distilbert-base-uncased</i>	0.72	0.81	0.79	0.80
<i>roberta-base</i>	0.72	0.81	0.76	0.79
<b><i>roberta-large</i></b>	<b>0.77</b>	<b>0.85</b>	<b>0.81</b>	<b>0.83</b>
<i>deberta-v3-base</i>	0.76	0.81	0.77	0.79
<i>deberta-v3-large</i>	0.77	0.85	0.79	0.82
<i>EnergyBERT</i>	0.73	0.80	0.79	0.79
<i>gpt2</i>	0.69	0.80	0.75	0.77
<i>gpt2-large</i>	0.73	0.77	0.82	0.79
<i>gpt4o-mini</i>	0.71	0.78	0.78	0.78
<i>gpt4o</i>	0.65	0.80	0.68	0.73
<i>Llama-3.1-8B</i>	0.70	0.84	0.67	0.74

Das Modell wird im Rahmen von SGAAIRE an verschiedenen Stellen eingesetzt. Unter anderem in:

- **Anwendungsfall 1.1 (SGAM-Domänen und -Zonen für IEC 62559 Use Cases bzw. Akteure vorschlagen):** Der Pseudocode in Algorithmus 3 beschreibt, wie die SGAM Domänen und Zonen eines IEC 62559 Use Cases ermittelt werden. Im Anwendungsfall 1.1 (SGAM-Domänen und -Zonen für IEC 62559 Use Cases bzw. Akteure vorschlagen) wird das KI-Modell als Teilmenge des Generierungsmodells eingesetzt, um Vorhersagen über die relevanten Domänen und Zonen zu treffen. Für die Akteursklassifikation hingegen wird explizit das M1-Modell verwendet, wie in Algorithmus 4 dargestellt. Im Anschluss wird die Konvexität der Domänen und Zonen überprüft und bei Bedarf in der Listenaufführung angepasst. Dies ist erforderlich, da ein KI-Modell beispielsweise die Domänen Distribution und Customer Premises vorschlagen kann, ohne DER zu berücksichtigen, welches jedoch zur Wahrung der Domänenkonvexität enthalten sein muss. Ein Beispiel für Domänenkonvexität ist in Abbildung 8.8, für Zonenkonvexität in Abbildung 8.9 dargestellt. Liegt die Domänen- bzw. Zonenrepräsentation in konvexer Form vor, wird anschließend geprüft, ob eine Informationsangabe fehlt oder ob ein Missmatch zwischen Use Case und SGAM Modell besteht. Falls dies der Fall ist, erzeugt das Verfahren automatisch einen Vorschlag, der neben dem Typ (Domänen- bzw. Zonenvorschlag) auch eine Objektreferenz (den betreffenden Use Case oder Akteur), die zugehörigen Domänen und Zonen sowie eine Bewertung in Form einer Punktzahl enthält.

---

### Algorithmus 3: Vorschläge für SGAM Domänen und -Zonen für ein Use Case

---

```

Input: Use Case UC
Output: List of Suggestions suggestions
Initialize empty list suggestions;
Generate SGAM draft model from UC;
Extract domains, zones from sgam;
Ensure convexity of domains, zones;
Compare domains, zones with UC classification;
if mismatch in domains  $\vee$  mismatch in zones then
    Create suggestion with type, object.ref and suggested.domains resp. .zones;
    Append suggestion to suggestions;
return suggestions

```

---



---

### Algorithmus 4: Vorschläge für SGAM Domänen und -Zonen für ein Akteur

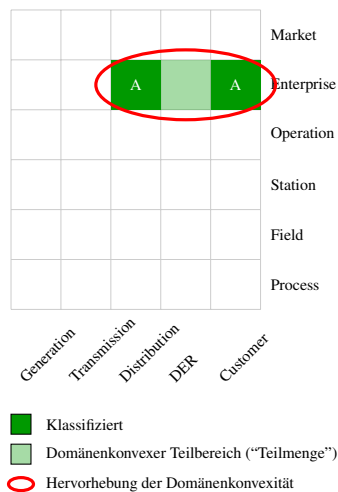
---

```

Input: Actor actor
Output: List of Suggestions suggestions
Initialize empty list suggestions;
Run model inference on (actor.name, actor.description);
Ensure convexity of domains, zones;
Compare domains, zones with actor classification;
if (mismatch in domains)  $\vee$  (mismatch in zones) then
    Create suggestion with type, object.ref, suggested.domains / suggested.zones;
    Append suggestion to suggestions;
return suggestions

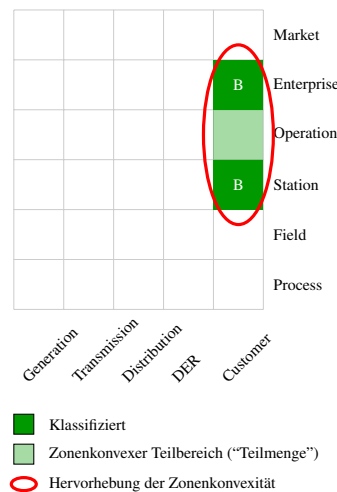
```

---



Quelle: Eigene Darstellung

Abbildung 8.8: Domänenkonvexität



Quelle: Eigene Darstellung

Abbildung 8.9: Zonenkonvexität

- **Alle SGAM-bezogenen Anwendungsfälle:** Die SGAM-Domänen- und Zonenklassifikation wird eingesetzt, um bei der Generierung von SGAM-Arbeitsmodellen auf Grundlage von IEC 62559 Use Cases Akteuren mit fehlenden Angaben zu Domänen / Zonen eine Zuweisung zu ermöglichen. Auf dieser Basis können die übergeordneten Ebenen systematisch durch die Analyse der Relationen zwischen den Akteuren abgeleitet und eingeordnet werden. Die zugrunde liegende Logik ist im vorherigen Abschnitt beschrieben und im dort dargestellten Pseudocode des Algorithmus 4 formalisiert.

Für die Konsistenzsicherung sowie Fehlerüberprüfung werden die SGAM Domänen-, Zonen- sowie Interoperabilitätsebenenmodelle eingesetzt, um jedes Element hinsichtlich der adäquaten Verwendung von Domänen, Zonen sowie Ebenenzuweisungen zu überprüfen. Algorithmus 5 zeigt den Pseudocode zur Überprüfung der Elemente innerhalb eines SGAM Modells. Hierbei wird für jedes Element der SGAM-Domänen-, Zonen- sowie Interoperabilitätsebenenklassifikator angewandt. Dadurch, dass bei den KI-Modellen Schwankungen und Abweichungen (z. B. False Positives oder False Negatives) auftreten können, liegt der Fokus insbesondere auf deutlichen Abweichungen. Geringe oder unsichere Tendenzen mit niedrigen Scores werden hingegen weniger stark berücksichtigt. Zu diesem Zweck werden Schwellenwerte definiert, die auf Grundlage empirischer Tests und Kalibrierung auf einen Wert von 2 gesetzt wurden. Werte unterhalb dieses Schwellenwertes werden nicht berücksichtigt und als „don't care“ eingeordnet. Sofern die verbleibenden Domänen, Zonen bzw. Interoperabilitätsebenen nicht mit den definierten Klassifikationen übereinstimmen, werden entsprechende Vorschläge generiert.

---

**Algorithmus 5:** Überprüfung von SGAM Modellen hinsichtlich Domänen, Zonen und Interoperabilitätsebenen
 

---

```

Input: SGAM Modell sgam
Output: List of Suggestions suggestions
Initialize empty list suggestions;
Extract all SGAM elements elements from sgam;
foreach elem ∈ elements do
  Read defined layer, domains, zones from elem;
  Run classification for layer on (elem.name, elem.description) → layer_scores;
  Run classification for domains on (elem.name, elem.description) → domain_scores;
  Run classification for zones on (elem.name, elem.description) → zone_scores;
  Filter domain_scores by thresholds (keep ≥ TH.DOMAIN, drop fuzzy < TH.DOMAIN) →
    domains_proposed;
  Filter zone_scores by thresholds (keep ≥ TH.ZONE, drop fuzzy < TH.ZONE) → zones_proposed;
  Filter layer_scores by threshold TH.LAYER → layers_proposed;
  Ensure convexity of domains and domains_proposed;
  Ensure convexity of zones and zones_proposed;
  if defined layer ∉ layers_proposed or domains ≠ domains_proposed or zones ≠
    zones_proposed then
    Create suggestion with type, object_ref=elem.name,
      suggested_layer=layers_proposed or suggested_domains=domains_proposed
      or suggested_zones=zones_proposed;
    Append suggestion to suggestions;
  return suggestions

```

---

### 8.3.5 SGAM Zonen Klassifikator (M2)

Die Ergebnisse in Tabelle 8.9 zeigen ein ähnliches Bild wie bereits bei der SGAM Domänenklassifikation (vgl. Abschnitt 8.3.4). Encoder-basierte Modelle erzielten durchweg bessere Performanz als die generativen Decoder-Modelle. Auffällig ist, dass die Werte insgesamt enger beieinanderliegen und die Unterschiede zwischen den Modellvarianten geringer ausfallen als bei M1. Das Modell *distilbert-base-uncased* erreicht mit  $F1_{distilbert-base-uncased} = 0.81$  das beste Ergebnis. Modelle wie *roberta-large* mit  $F1_{roberta-large} = 0.80$  oder EnergyBERT mit  $F1_{EnergyBERT} = 0.80$  liefern zwar vergleichbare Ergebnisse, bringen jedoch keinen signifikanten Mehrwert gegenüber der kompakteren Architektur.

Die generativen Decoder-Modelle (*gpt2(-large)*, *gpt4o* sowie *Llama-3.1-8B*) schneiden erneut schwächer ab und bestätigen die bereits bei M1 gemachte Beobachtung, dass Encoder-Architekturen für Klassifikationsaufgaben im SGAM-Kontext überlegen sind. Vor diesem Hintergrund wird für die Zonenklassifikation *distilbert-base-uncased* als Modell in SGAAIRE ausgewählt, da es eine gute Performanz bei gleichzeitig reduzierten Rechenressourcen ermöglicht.

Die Integration des Modells in das Gesamtsystem wurde bereits im Rahmen der SGAM-Domänenklassifikation (vgl. Abschnitt 8.3.4) beschrieben und kann analog übertragen werden.

Tabelle 8.9: Modellergebnisse für M2: SGAM Zonen Klassifikation

Quelle: Basierend auf [KLS25]

Model	M2: SGAM Zonen Klassifikation			
	Accuracy	Precision	Recall	F1
bert-base-uncased	0.73	0.79	0.82	0.80
bert-large-uncased	0.73	0.81	0.77	0.79
distilbert-base-uncased	0.73	0.84	0.78	0.81
roberta-base	0.74	0.82	0.78	0.80
roberta-large	0.74	0.83	0.78	0.80
deberta-v3-base	0.73	0.81	0.78	0.80
deberta-v3-large	0.74	0.80	0.79	0.80
EnergyBERT	0.74	0.83	0.78	0.80
gpt2	0.69	0.83	0.79	0.81
gpt2-large	0.70	0.79	0.77	0.78
gpt4o-mini	0.72	0.82	0.77	0.80
gpt4o	0.66	0.83	0.65	0.73
Llama-3.1-8B	0.63	0.80	0.62	0.70

### 8.3.6 SGAM Interoperabilitätslayer Klassifikator (M3)

Die Ergebnisse in Tabelle 8.10 verdeutlichen, dass die SGAM IOP Klassifikation insgesamt höhere Werte erzielt als die Domänen- (M1) und Zonenklassifikation (M2) (vgl. Abschnitt 8.3.4 sowie 8.3.5). Auffällig ist, dass neben den Encoder-basierten Modellen auch einige der nicht-großen Decoder-basierten Sprachmodelle vergleichbare Ergebnisse erreichen. Ein deutlicher Ausreißer zeigt sich jedoch beim gpt4o-mini, das mit einem  $F1_{gpt4o-mini} = 0.39$  eine deutlich geringere Performanz aufweist. Vor diesem Hintergrund wurde auf ein Training des größeren gpt4o-Modells verzichtet, da keine Verbesserung zu erwarten gewesen wäre. Ebenso zeigt das Llama-3.1-8B Modell mit einem  $F1_{Llama3.1-8B} = 0.76$  nur eingeschränkte Ergebnisse und liegt damit deutlich unterhalb der besten Modelle. Die insgesamt höchsten Werte sowohl bei F1-Score als auch bei Accuracy werden durch *roberta-large* erzielt. Daher wurde dieses Modell im Rahmen von SGAAIRE als Referenzmodell für die Klassifikation der SGAM Interoperabilitätslayer ausgewählt.

Im Rahmen von SGAAIRE findet das selektierte Modell an den folgenden Stellen Anwendung:

- **Anwendungsfall 1.3 (Vorschläge von Hinzufügen, Entfernen und Bearbeiten von Elementen) und folglich Anwendungsfall 2.1 (Generieren von SGAM Arbeitsmodellen):** Die Logik dieses Anwendungsfalls umfasst die Generierung von SGAM Modellen gemäß dem in Abschnitt 7.4.3.2 beschriebenen Konzept, das implizit auch die Verwendung des M3-Modells vorsieht. Ziel ist es sicherzustellen, dass die im Anwendungsfall spezifizierten Elemente konsistent den entsprechenden SGAM-Ebenen zugeordnet werden können. Dadurch wird gewährleistet, dass die in Abbildung 7.11 dargestellten Mappings eingehalten werden. Potenziell fehlerhafte Modellierungen (bspw. die fälschliche Verwendung von Kommunikationsprotokollen als Akteursan-

Tabelle 8.10: Modellergebnisse für M3: SGAM IOP Klassifikation

Quelle: Basierend auf [KLS25]

Model	M3: SGAM IOP Klassifikation			
	Accuracy	Precision	Recall	F1
bert-base-uncased	0.86	0.90	0.89	0.90
bert-large-uncased	0.87	0.90	0.90	0.90
distilbert-base-uncased	0.84	0.88	0.90	0.89
roberta-base	0.86	0.89	0.91	0.90
<b>roberta-large</b>	<b>0.89</b>	<b>0.91</b>	<b>0.92</b>	<b>0.91</b>
deberta-v3-base	0.87	0.91	0.91	0.91
deberta-v3-large	0.87	0.89	0.90	0.90
EnergyBERT	0.85	0.89	0.89	0.89
gpt2	0.85	0.91	0.89	0.90
gpt2-large	0.87	0.90	0.92	0.91
gpt4o-mini	0.25	0.47	0.34	0.39
gpt4o	-	-	-	-
Llama-3.1-8B	0.61	0.89	0.66	0.76

gaben) sollen im Rahmen dieses Prozesses identifiziert und den Anwenderinnen und Anwendern mit dem M5-Modell (SGAM Element Proposer) als Vorschläge zur Reevaluierung bereitgestellt werden.

- **Anwendungsfall 2.2 (Überprüfung der Konsistenz [...] gegenüber der IEC 62559 Use Case-Beschreibung) und Anwendungsfall 2.3 (Identifikation von Inkonsistenzen innerhalb von SGAM Modellen):** Die Umsetzung dieser beiden Anwendungsfälle basiert auf der in Abschnitt 8.3.4 beschriebenen Logik.

### 8.3.7 Service Step Klassifikator (M4)

Auf Grundlage der in Tabelle 8.11 dargestellten Ergebnisse lässt sich feststellen, dass sich alle feinabgestimmten Modelle in einem angemessenen Leistungsbereich bewegen und die Unterschiede lediglich in Nuancen erkennbar sind. Besonders hervorzuheben ist, dass die Encoder-basierten Modelle roberta-large und deberta-v3-base identische Werte über alle betrachteten Metriken hinweg erzielen. Vor diesem Hintergrund wurde *roberta-large* als Referenzmodell für das SGAAIRE-System ausgewählt. Zur Vermeidung zusätzlicher ökonomischer Kosten sowie zur Schonung von Rechenressourcen wurde zudem entschieden, ausschließlich das Modell GPT4o-mini zu trainieren und auf eine Feinabstimmung des größeren GPT4o-Modells zu verzichten, da bereits über die vorherigen trainierten Modelle eine Tendenz erkennbar ist.

Das Modell wird im **Anwendungsfall 1.4 (Vorschlag von Service Steps in der Step-by-Step-Analyse)** eingesetzt. Ziel ist es, für fehlende Services geeignete Vorschläge zu generieren sowie zu prüfen, ob die im jeweiligen Use Case spezifizierten Services adäquat eingesetzt werden. Falls dies nicht der Fall ist, werden entsprechende Empfehlungen bereitgestellt. Grundlage hierfür ist die Inferenz über die Modellklassifikation,

Tabelle 8.11: Modellergebnisse für M4: Service Step Klassifikation

Quelle: Basierend auf [KLS25]

Model	M4: Service Step Klassifikation			
	Accuracy	Precision	Recall	F1
bert-base-uncased	0.86	0.86	0.86	0.86
bert-large-uncased	0.84	0.84	0.84	0.84
distilbert-base-uncased	0.85	0.85	0.85	0.85
roberta-base	0.85	0.85	0.85	0.85
<b>roberta-large</b>	<b>0.87</b>	<b>0.87</b>	<b>0.87</b>	<b>0.87</b>
deberta-v3-base	0.87	0.87	0.87	0.87
deberta-v3-large	0.86	0.86	0.86	0.86
EnergyBERT	0.86	0.86	0.86	0.86
gpt2-large	0.86	0.86	0.86	0.86
gpt2	0.84	0.84	0.84	0.84
gpt4o-mini	0.83	0.83	0.83	0.83
gpt4o	-	-	-	-
Llama-3.1-8B	0.82	0.82	0.82	0.82

wobei lediglich Ergebnisse oberhalb eines definierten Schwellwertes  $\Theta$  berücksichtigt werden. Im Anschluss werden für fehlende oder inkorrekt zugeordnete Service-Step-Typen Vorschläge erzeugt. Jeder Vorschlag enthält dabei die Referenz auf die entsprechende Aktivität, die Referenz auf das übergeordnete Szenario sowie die Klassifikationsergebnisse einschließlich ihrer jeweiligen Punktwerte. Diese Logik ist im Algorithmus 6 dargestellt.

---

**Algorithmus 6:** Bereitstellung der Service-Vorschläge für Scenario Steps
 

---

**Input:** IEC 62559 Use Case UC**Output:** List of Suggestions suggestions

Initialize empty list suggestions;

Load scenarios from UC;

**foreach** activity  $\in$  scenarios **do**

Run model inference on (activity, description);

    Get labels above threshold  $\Theta$ ;    **if** service  $\notin$  selected labels **then**        Collect all classifications with score  $\geq \Theta$ ;        **if** classifications  $\neq \emptyset$  **then**

Create suggestion with type, object\_ref, parent\_ref, classifications;

Append suggestion to suggestions;

**return** suggestions;

### 8.3.8 SGAM Element Proposer (M5)

Die in Tabelle 8.12 dargestellten Ergebnisse verdeutlichen die Vorteile insbesondere von Decoder- und Encoder-Decoder-basierten Sprachmodellen für die betrachtete Textgenerierungsaufgabe. Für deberta-v3-base bzw. -large und distilbert liegen keine Ergebnisse vor, da diese in HuggingFace keine Wrapper für den Einsatz im CLM-Setting bereitstellen. Deutlich erkennbar ist die Leistungstärke großer Sprachmodelle wie FLAN-

T5 und GPT4o(-mini), die insbesondere bei den Metriken METEOR und ROUGE1 überlegen sind und somit syntaktisch bessere Ausgaben erzeugen als ggü. den Encoder-basierten Modellen. Zudem weisen diese Modelle konsistent einen BERTScore (bei F1) von  $F1 = 0.97$  auf, was eine hohe semantische Qualität der generierten Texte widerspiegelt. Unter Berücksichtigung der F1-Werte von ROUGE1, BERTScore und METEOR erzielte *FLAN-T5* die besten Gesamtergebnisse und wurde daher für die Integration in das System ausgewählt. Die Modelle GPT-4o und GPT-4o-mini erzielten nahezu identische Ergebnisse, sind jedoch nur über die OpenAI API verfügbar.

Tabelle 8.12: Modellergebnisse für M5: SGAM Element Proposer

Quelle: Basierend auf [KLS25]

Model	M5: SGAM Element Proposer						
	METEOR	ROUGE1 mid			BERTScore		
		F1	Precision	Recall	F1	Precision	Recall
bert-base-uncased	0.68	0.69	0.63	0.77	0.88	0.88	0.89
bert-large-uncased	0.61	0.68	0.61	0.78	0.88	0.88	0.88
roberta-base	0.80	0.79	0.86	0.77	0.95	0.95	0.94
roberta-large	0.69	0.76	0.77	0.77	0.94	0.94	0.94
EnergyBERT	0.58	0.71	0.69	0.77	0.89	0.88	0.88
gpt2	0.71	0.70	0.62	0.80	0.93	0.94	0.94
gpt2-large	0.76	0.78	0.71	0.87	0.95	0.95	0.96
flan-t5-small	0.90	0.83	0.82	0.85	0.97	0.97	0.97
<b>flan-t5-large</b>	<b>0.93</b>	<b>0.86</b>	<b>0.83</b>	<b>0.90</b>	<b>0.97</b>	<b>0.97</b>	<b>0.98</b>
gpt4o-mini	0.92	0.85	0.84	0.84	0.97	0.97	0.97
gpt4o	0.92	0.85	0.85	0.85	0.97	0.97	0.97
Llama-3.1-8B	0.71	0.80	0.70	0.62	0.94	0.94	0.93

Innerhalb von SGAAIRE kommt das selektierte Modell in den folgenden Bereichen zur Anwendung:

- **Anwendungsfall 2.1 (Generieren von SGAM Arbeitsmodellen):** Im Wesentlichen findet das Modell Einsatz in der Generierung von SGAM (Arbeits-)Modellen auf Basis von einer IEC 62559 Use Case-Beschreibung. Der in Algorithmus 7 dargestellte Pseudocode gibt einen Überblick über die Implementierung und realisiert das in Abbildung 7.11 vorgesehene Mapping, wobei das M5-Modell zum Einsatz kommt. Der Fokus liegt in Realisierung auf der Szenariobeschreibung eines ausgefüllten IEC 62559 Use Case. Das M5-Modell wird eingesetzt, um Kommunikationsstandardempfehlungen zu generieren, die in der Template lediglich indirekt vorgesehen sind, sowie um fehlende Elemente zu ergänzen. Dazu zählen fehlende Akteure (insbesondere auf der Komponentenebene und bei High-Level Use Cases die Businesssebene), Szenariobeschreibungen (für die Funktionsebene), Informationsobjekte (für die Informationsebene) sowie KPIs bzw. Use Case-Ziele (für die Geschäftsebene). Ergänzend dazu erfolgt die Klassifikation der Elemente über die anderen Modelle (insbesondere M1 und M2 für die SGAM-Domänen und -Zonen), während die Zuordnung zu den Interoperabilitätsebenen über M3 überprüft wird.

---

**Algorithmus 7: Generieren von SGAM Arbeitsmodellen basierend auf einer IEC 62559 Use Case-Beschreibung**


---

**Input:** IEC 62559 Use Case *UC*, Configurations *cfg*  
**Output:** SGAM Draft Model *SGAM*

Preprocess *cfg*;  
Determine abstraction level *nature*  $\in$  {Low, High, Mixed} by considering *UC.nature* **or** if not exists applying M6 inference;  
Set layer strategy based on *nature*;  
Load *actors*, *scenarios*, *kpis*, *objectives*, *information\_objects* from *UC*;  
**foreach** *actor*  $\in$  *actors* **do**  
    **if** *actor.domains* =  $\emptyset$  **or** *actor.zones* =  $\emptyset$  **then**  
        Run model inference *M1* and *M2* on (name, description);  
        Assign domains and zones to *actor*;  
Get all (*actor*, *information\_object*) pairs from scenario description;  
**foreach** *pair* (*actor*, *information\_object*)  $\in$  *all pairs* **do**  
    **if** *pair.requirements* =  $\emptyset$  **or** any requirement not mapped to communication layer by *M3* **then**  
        Run model inference *M5* on all relevant layers;  
        Assign inference result as communication approach to the *pair*  
Build *component\_layer* from (*actor*, *information\_object*) pairs **or** *M5* (based on *cfg* and *nature*);  
Build *information\_layer* from (*actor*, *information\_object*) pairs **or** *M5* (based on *cfg* and *nature*);  
Build *function\_layer* from *scenarios* **or** *M5* (based on *cfg* and *nature*);  
Build *business\_layer* from *kpis/objectives/actors* **or** *M5* (based on *cfg* and *nature*);  
Build *communication\_layer* from (*actor*, *information\_object*) pairs **or** *M5* (based on *cfg* and *nature*);  
Assemble *SGAM* from all constructed layers;  
**return** *SGAM*;

---

- **Anwendungsfall 1.3 (Vorschläge von Hinzufügen, Entfernen und Bearbeiten von Elementen):** Die Verwendung des M5-Modells erfolgt hierbei indirekt über die Generierung von SGAM Modellen (vgl. Anwendungsfall 2.1). Den Anwendenden von SGAAIRE werden dabei Empfehlungen für das Hinzufügen von Elementen bereitgestellt, die sich aus einem generierten SGAM Modell ergeben.
- **Anwendungsfall 2.2 (Überprüfung der Konsistenz [...] gegenüber der IEC 62559 Use Case-Beschreibung) sowie Anwendungsfall 2.3 (Identifikation von Inkonsistenzen innerhalb von SGAM Modellen):** Die Verwendung des M5-Modells in diesen Anwendungsfällen erfolgt ebenfalls indirekt durch die Generierung von SGAM Modellen (vgl. Anwendungsfall 2.1). Aus der Menge potenziell generierbarer Modellkombinationen wird jenes SGAM Modell selektiert, das die geringste Anzahl an Modifikationsvorschlägen erfordert. Die Konsistenzprüfung validiert, ob sämtliche Relationen zwischen den Komponenten gemäß der Szenarioanalyse der IEC 62559 Use Case-Beschreibung entsprechen. Darüber hinaus wird verifiziert, dass alle in der IEC 62559 Use Case-Beschreibung spezifizierten Entitäten (z. B. Akteure, Informationsobjekte und, sofern vorhanden Kommunikationsspezifikationen) vollständig integriert sowie Elemente im SGAM Modell ggü. des Use Cases redundant sind.

### 8.3.9 Abstraktionsgrad Regressor (M6)

Die in Tabelle 8.13 dargestellten Ergebnisse verdeutlichen, dass sich die Modelle insgesamt in einem vergleichbaren Leistungsbereich bewegen, wobei Encoder-basierte Sprachmodelle tendenziell die besseren Resultate erzielen. Im Hinblick auf das Bestimmtheitsmaß  $R^2$  erreicht roberta-large die besten Werte, während zugleich das mittlere quadratische Fehlermaß  $MSE_{roberta-large} = 0.19$  ebenfalls im oberen Leistungsbereich liegt. Einen Sonderfall stellt gpt4o-mini dar: Zwar weist dieses Modell mit  $MSE_{gpt4o-mini} = 0.03$  einen sehr geringen Fehlerwert auf, jedoch geht dies mit dem niedrigsten Bestimmtheitsmaß  $R^2_{gpt4o-mini} = 0.73$  einher, was auf eine eingeschränkte Erklärungs- und Prognosefähigkeit hinweist. Für SGAAIRE wurde daher, unter Priorisierung des Bestimmtheitsmaßes  $R^2$ , das Modell roberta-large ausgewählt. Zur Reduktion von Kosten und Rechenaufwand wurde lediglich das Modell GPT4o-mini anstelle von gpt4o trainiert, da sich die Leistungstendenzen bereits aus den zuvor trainierten Modellen ableiten ließen.

Tabelle 8.13: Modellergebnisse für M6: LoA Regression

Quelle: Basierend auf [KLS25]

Model	M6: LoA Regression		
	MSE	MAE	$R^2$
bert-base-uncased	0.19	0.21	0.82
bert-large-uncased	0.20	0.23	0.82
distilbert-base-uncased	0.20	0.23	0.82
roberta-base	0.24	0.26	0.78
<b>roberta-large</b>	<b>0.19</b>	<b>0.23</b>	<b>0.83</b>
deberta-v3-base	0.29	0.39	0.74
deberta-v3-large	0.30	0.39	0.72
EnergyBERT	0.20	0.25	0.82
gpt2-large	0.21	0.20	0.81
gpt2	0.23	0.30	0.79
gpt4o-mini	0.03	0.03	0.73
gpt4o	-	-	-
Llama-3.1-8B	0.27	0.22	0.75

Innerhalb von SGAAIRE wird das M6-Modell in den folgenden Anwendungsfällen wie folgt eingesetzt:

- **Anwendungsfall 1.5 (Erkennung von Inkonsistenzen [...] ggü. dem Use Case-Typ):**

Für diesen Anwendungsfall wird das Modell M6 herangezogen, um den Abstraktionsgrad der einzelnen Elemente zu bestimmen und daraus gruppenspezifische Mittelwerte für Akteure, Informationsobjekte und Anforderungen zu berechnen. Der resultierende globale Mittelwert bildet die Grundlage für die Bewertung des Abstraktionsgrads, wie in Algorithmus 8 dargestellt. Sofern in der Use Case-Beschreibung die Art (Nature) angegeben ist, werden die Grenzwerte gemäß Abschnitt 7.5.6 herangezogen; liegt diese Angabe nicht vor, erfolgt die Einschätzung durch das Modell M6 auf

Basis des Use Case-Namens sowie der Beschreibung. High-Level-Anwendungsfälle werden durch einen Wertebereich von  $(-\infty, 1.5]$  charakterisiert, während Low-Level-Anwendungsfälle Werte im Bereich  $(1.5, \infty)$  aufweisen. Liegt der globale Mittelwert außerhalb des für den jeweiligen Use Case-Typ definierten Intervalls, werden die abweichenden Gruppen identifiziert, aufbereitet und in Form eines konsolidierten Vorschlags mit den entsprechenden Werten zurückgegeben.

---

**Algorithmus 8:** Bestimmung der Abstraktionsgrade innerhalb des Anwendungsfalls

---

**Input:** IEC 62559 Use Case UC  
**Output:** List of Suggestions suggestions  
Initialize empty list suggestions;  
Load actors, information objects, requirements from UC;  
Set elements = union of all the mentioned UC elements;  
**foreach** element  $\in$  elements **do**  
    | Run model inference  $M6$  on (name, description)  $\rightarrow$  value;  
    | Accumulate overall mean and group means;  
Determine bounds (min, max) from UC.nature (if available), **otherwise** infer them using model  $M6$  based on UC.name and UC.description.  
**if** overall mean  $\notin$  bounds **then**  
    | Collect groups with means outside bounds;  
    | Create suggestion with type, object\_ref=UC.nature, parent\_ref, mean, group\_means;  
    | Append to suggestions;  
**return** suggestions;

---

- **Anwendungsfall 2.3 (Identifikation von Inkonsistenzen innerhalb von SGAM Modellen):**

Für den vorliegenden Anwendungsfall wird das Modell  $M6$ , analog zu Anwendungsfall 1.5, eingesetzt, wie in Algorithmus 9 dargestellt. Dabei werden die einzelnen Elemente des vorliegenden SGAM Modells durch das Modell  $M6$  eingeschätzt und den jeweiligen Interoperabilitätsebenen zugeordnet. Auf Basis der Mittelwerte erfolgt anschließend die Klassifikation des Abstraktionsgrads: High-Level für Werte  $\leq 0.5$ , Rather High-Level für  $(0.5, 1.0]$ , Mixed für  $(1.0, 2.0]$ , Rather Low-Level für  $(2.0, 2.5]$  sowie Low-Level für  $> 2.5$ . Ergänzend wird die Standardabweichung zur Bewertung der Konsistenz herangezogen und wie folgt klassifiziert:  $< 0.25$  = Very Low,  $\leq 0.5$  = Low,  $\leq 0.75$  = Moderate,  $\leq 1.0$  = High,  $> 1.0$  = Very High. Grundsätzlich sollte die Standardabweichung sowohl global als auch innerhalb der einzelnen Schichten möglichst gering ausfallen, während die Abstraktionsklasse dem jeweiligen Anwendungsfall entsprechen sollte. Die Bewertungsergebnisse werden den Modellierenden als Bericht bereitgestellt.

---

**Algorithmus 9: Bestimmung der Abstraktionsgrade im SGAM**


---

```

Input: SGAM Model SGAM
Output: Report report
Initialize empty lists reports, elements_values;
Load all elements from SGAM;
foreach element  $\in$  elements do
    Run model inference M6 on (name, description)  $\rightarrow$  value;
    Append value to elements_values[layer];
foreach layer  $\in$  elements_values do
    Compute mean, std of values;
    Determine LoA of layer based on the mean and std;
    Create report and append to reports;
Compute overall mean, std across all layers;
Determine LoA of SGAM based on the mean and std;
Create report and append to reports;
return report;

```

---

### 8.3.10 Acronym Identification und -Disambiguation (M7)

Die beiden Modelle Acronym Identification (M7a) und -Disambiguation (M7b) stellen im Vergleich zu den übrigen Modellen dieser Arbeit einen Sonderfall dar. Während in anderen Bereichen eine breite Auswahl an Modellarchitekturen untersucht wurde, beschränkt sich die Analyse hier auf eine gezielte Auswahl weniger Modelle. Diese Entscheidung beruht auf zwei zentralen Gründen: Erstens handelt es sich bei den Aufgaben um kontextunabhängige Modelle, die in der Literatur bereits umfassend untersucht und evaluiert wurden [VDN<sup>+</sup>20, VDTN20]. Zweitens zeigen vorhandene Studien sowie Voruntersuchungen in dieser Arbeit, dass selbst etablierte Baseline-Modelle bereits eine sehr hohe Leistung erzielen, sodass ein zusätzlicher Vergleich mit weiteren Modellvarianten nur geringen wissenschaftlichen Mehrwert versprochen hätte [VDN<sup>+</sup>20].

Für die Auswahl des Modells M7a (Acronym Identification) zeigt sich, dass *deberta-v3-base* mit  $F1_{deberta-v3-base} = 0.93$  das beste Ergebnis erzielt. Allerdings besteht die Herausforderung darin, dass der zugrundeliegende Einsatzbereich längere Textpassagen und komplexe Textelemente umfasst, wodurch das Modell lineare Abhängigkeiten über größere Eingabelängen hinweg verarbeiten muss. Dies führt zu einer deutlichen Skalierung der Inferenzzeit, die eine leistungsfähigere Laufzeitumgebung als die im Rahmen dieser Arbeit verfügbare Konfiguration (vgl. Abschnitt 8.1.3) erfordert. Bei größeren Modellen kann die Inferenzzeit daher stark ansteigen, was die Praxistauglichkeit für eine systematische Anwendung einschränkt. Vor diesem Hintergrund wurde für M7a die Auswahl auf *distilbert-base-uncased* mit  $F1_{distilbert-base-uncased} = 0.91$  getroffen, da dieses Modell ein günstigeres Verhältnis von Laufzeiteffizienz und Genauigkeit bietet und somit für den praktischen Einsatz im Rahmen der Modellumgebung am geeignetsten erscheint.

Für die Aufgabe des Modells M7b (Acronym Disambiguation) erreicht *bert-base-uncased* mit  $F1_{bert-base-uncased} = 0.97$  einen sehr hohen Wert. Da dieses Modell im vorgesehenen Anwendungsbereich nicht auf vollständige Texte angewendet wird, sondern ausschließlich auf Textelemente mit erkannten Akronymen, relativiert sich die Pro-

blematik längerer Inferenzzeiten. Die selektive Nutzung innerhalb von SGAAIRE macht bert-base-uncased durch seine eher geringere Größe nichtsdestotrotz daher zu einer geeigneten Wahl für M7b.

Tabelle 8.14: Modellergebnisse für M7a & b: Acronym Identification und -Disambiguation

Quelle: Basierend auf [KLS25]

Model	M7a: Acronym Identification				M7b: Acronym Disambiguation			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
bert-base-uncased	0.98	0.91	0.90	0.90	0.98	0.96	0.99	0.97
deberta-v3-base	0.99	0.93	0.93	0.93	–	–	–	–
distilbert-base-uncased	0.98	0.89	0.90	0.90	–	–	–	–
EnergyBERT	0.98	0.91	0.90	0.91	–	–	–	–

Beide Modelle M7a und M8b werden ausschließlich im Rahmen des **Anwendungsfalls 1.2 (Erkennen von abweichenden und nicht-standardisierten Termen)** verwendet. Algorithmus 10 beschreibt eine High-Level-Herangehensweise zur Realisierung des Anwendungsfalls unter Verwendung beider Modelle. Zunächst werden sämtliche Texte sowie die im Anwendungsfall enthaltenen Definitionen aus bestehenden Sektionen (z. B. Akteure, Custom Information, Common Terms and Definitions, Anforderungen, Informationsobjekte) extrahiert. Anschließend identifiziert das Modell *M7a* aus den Texten potenzielle Akronyme, Terme und zugehörige Definitionen. Bereits bekannte Definitionen werden dabei aus der Kandidatenliste entfernt. Für die verbleibenden unbekannteren Akronyme erfolgt eine Iteration, bei der unter Rückgriff auf das AD-Dictionary von MadDog [PBVDCN21] sowie mit Unterstützung des Modells *M7b* Wahrscheinlichkeiten für die möglichen Definitionen berechnet werden. Ergibt sich dabei eine eindeutige Übereinstimmung, wird die entsprechende Definition als konkreter Vorschlag übernommen und das Akronym nicht weiter als undefiniert betrachtet. Liegen hingegen mehrere plausible Optionen vor, werden diese in Form einer nach Wahrscheinlichkeit gerankten Vorschlagsliste aufbereitet und den Modellierenden als Empfehlung zur Verfügung gestellt.

---

#### Algorithmus 10: Identifikation und Aufbereitung undefinierter Begriffe

---

```

Input: IEC 62559 Use Case UC
Output: List of Suggestions term_suggestions
Initialize empty list term_suggestions;
Extract all texts from UC;
Extract all definitions from UC;
Identify potential acronyms and terms based on texts with model inference M7a; remove already
  defined acronyms (acc. definitions);
foreach acronym  $\in$  acronyms do
  Compare with all identified terms and lookup-dictionary entries;
  Estimate likelihoods and store candidate mappings with model inference M7b;
  if not exactly one definition found then
    Add to term_suggestions with multiple definitions and ranked suggestions;
return term_suggestions;

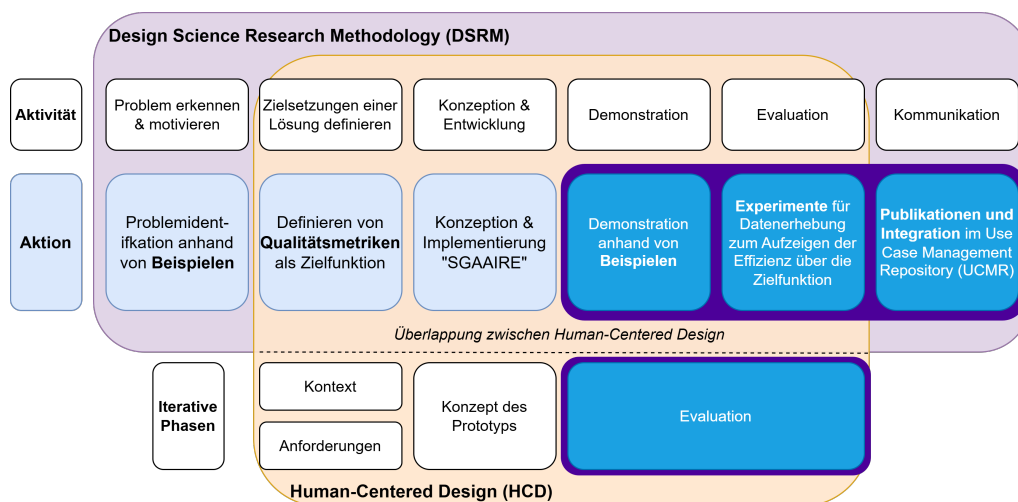
```

---



## 9 Evaluation und Diskussion

Dieses Kapitel widmet sich der Demonstration und Evaluation von SGAAIRE als Artefakt im Rahmen des Design Science Research Frameworks. Ziel ist es zu untersuchen, inwiefern SGAAIRE als repräsentatives Artefakt zu einer signifikanten Qualitätsverbesserung in der Anwendung der IEC 62559 Use Cases und SGAM Modelle beitragen kann. Zunächst wird im Abschnitt 9.1 die Wirksamkeit des Artefakts anhand zweier ausgewählter Demonstrationsbeispiele überprüft. Zur Evaluierung dieser Ergebnisse wird SGAAIRE im Rahmen einer Randomized Controlled Trial (RCT) (randomisierte kontrollierte Studie) durch Probanden in einem kontrollierten Experiment anhand eines Fallbeispiels getestet. Diese methodische Evaluation wird in Abschnitt 9.2 detailliert dargestellt und umfasst ein heterogenes Teilnehmerfeld aus den Bereichen Forschung, Praxis, Standardisierung und Studierende. Die aus der Dissertation gewonnenen Ergebnisse werden im Abschnitt 9.3 kritisch diskutiert und in den Kontext der vorliegenden Forschung eingeordnet. Abschließend fasst Abschnitt 9.4 die wissenschaftlichen und technischen Beiträge der Dissertation zusammen und stellt deren Relevanz für die (weitere) Forschung sowie praktische Anwendungen heraus. Im Gesamtverfahren repräsentiert Abschnitt 9.1 die Demonstration, Abschnitt 9.2 die Evaluation und Abschnitt 9.4 die Kommunikation gemäß dem Vorgehensmodell in Abbildung 9.1.



Quelle: Eigene Darstellung, angelehnt an [PTRC07]

Abbildung 9.1: Auswirkungen der Ergebnisse aus Kapitel 9 auf das Forschungsvorhaben

## 9.1 Demonstration

Um die Wirksamkeit von SGAAIRE zu testen, wurde das umgesetzte System auf zwei Fallbeispiele (bestehend aus einem mit Fehler versehenen IEC 62559 Use Case sowie SGAM Modell) angewandt. Ziel ist es, dass das System in der Lage ist, möglichst alle eingebauten Fehler korrekt zu identifizieren, ohne dabei Falschfehler zu identifizieren und idealerweise zusätzlich bislang unentdeckte Fehler aufzudecken. Die folgenden Abschnitte beschreiben die Anwendungsfälle sowie die gezielt integrierten Fehler im Detail und diskutieren die erhobenen Ergebnisse.

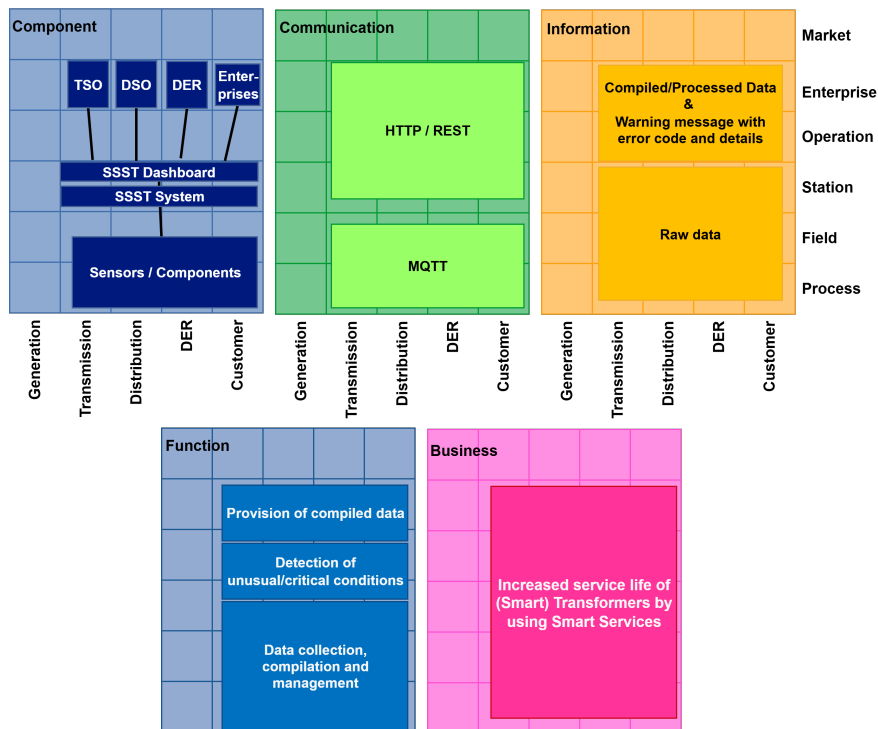
### 9.1.1 Fallbeispiel A: Optimierung der Lebensdauer von (smarten) Transformatoren mithilfe von Smart Services

Das erste Fallbeispiel befasst sich mit dem Anwendungsfall „Optimierung der Lebensdauer von (smarten) Transformatoren mithilfe von Smart Services“ und orientiert sich hierbei im Rahmen der Studie von [KAS23]. Der Anwendungsfall befasst sich mit der Entwicklung eines datengetriebenen Rahmenwerks zur Optimierung der Lebensdauer und Leistungsfähigkeit intelligenter Transformatoren im Rahmen des nationalen Forschungs- und Entwicklungsprojekts STIM. Hierbei werden die Konzepte des Condition Monitoring, der Predictive Maintenance sowie des Asset Managements als Smart Services implementiert, um eine fundierte Entscheidungsgrundlage für die Instandhaltung bereitzustellen. Das zentrale System, der Smart Service for Smart Transformers, basiert auf modernen Methoden der Künstlichen Intelligenz und des Maschinellen Lernens, darunter Künstliche Neuronale Netze. Diese Verfahren sollen eine kontinuierliche Erfassung, Analyse und Bewertung von Sensordaten ermöglichen, um Aspekte wie die Gesundheitsstatusbewertung, Fehlerdiagnose, Restlebensdauer- und Lebenszykluskostenschätzung sowie die Mustererkennung und Risikobewertung zu realisieren.

Für den vorliegenden Anwendungsfall sind sowohl eine Referenz- als auch eine fehlerhafte Version der IEC 62559 Use Case-Beschreibung sowie des zugehörigen SGAM Modells im Online-Anhang verfügbar<sup>1</sup>. Abbildung 9.2 stellt das Referenz-SGAM Modell für den Anwendungsfall dar, während Abbildung 9.3 das Evaluation-SGAM Modell mit absichtlich eingefügten Fehlern zeigt. Die Referenzversion des IEC 62559 Use Case ist sowohl in Abschnitt A.2 als auch im Online-Anhang<sup>1</sup> dokumentiert, während die fehlerhafte Variante ausschließlich im Online-Anhang<sup>1</sup> verfügbar ist. Im Rahmen des Demonstrationsbeispiels wurden die in Tabelle 9.1 (für den IEC 62559 Use Case) sowie Tabelle 9.2 (für das SGAM Modell) dokumentierten Fehler in die Evaluationsversion integriert. Dabei wurden die Fehler den identifizierten Qualitätsmerkmalen (vgl. Abschnitt 6.3.3) zugewiesen. Hierbei kennzeichnen „++“ die Qualitätsmerkmale, die adressiert werden, während „+“ für eine schwache Adressierung steht.

Die Fehlerauswahl basiert auf den Anforderungen der IEC 62559 Use Cases und SGAM Modelle ([IEC15, CEN14, GDU17, KSS23]) und umfasst typische sowie kom-

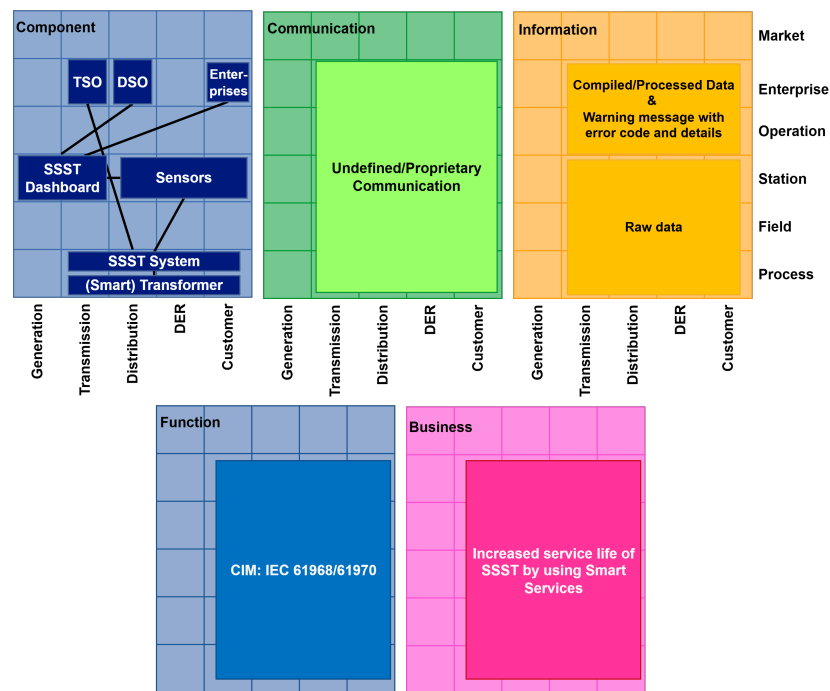
<sup>1</sup> vgl. Ordner „Demonstration/Fallbeispiel A“ im Online-Anhang (vgl. Abschnitt A.1)



Quelle: Eigene Darstellung

Abbildung 9.2: SGAM Modell (Referenz) für Fallbeispiel A

plexe Herausforderungen, die bei der Modellierung auftreten. Dabei handelt es sich bei dem fehlerbehafteten Modell um eines, das in der Praxis nur selten vorkommt oder unter begrenzten Erfahrungswerten entwickelt wurde. Solche Probleme sind in öffentlich verfügbaren Use Cases und SGAM Modellen erkennbar und führen je nach Akteur oder Anwendungskontext zu unterschiedlichen Interpretationen. Während High-Level-Use Cases andere Schwerpunkte setzen als Low-Level-Modelle, unterscheiden sich auch die Schwerpunkte von Fach- und Domänenexperten. Inkonsistenzen stellen ein zentrales Problem dar, insbesondere hinsichtlich der Übereinstimmung zwischen einem IEC 62559 Use Case und dem entsprechenden SGAM Modell, wie bereits in [GS15] hervorgehoben. Dies zeigt sich insbesondere in den Fehlern 13-14 und 17-18. Anforderungsmanagement ist per Definition interdisziplinär und involviert Stakeholder aus verschiedenen Fachbereichen [IEC15]. Die Klassifikation und Auswahl relevanter Elemente (sichtbar in den Fehlern 1-5, 9-12 und 16-18) wird sowohl durch energietechnisches Fachwissen als auch durch Kenntnisse der IKT beeinflusst. Fehlertypen treten in vielfältiger Form auf, wobei unbeabsichtigte Fehler über alle Kategorien hinweg zu beobachten sind. Besonders problematisch sind unklare oder undefinierte Akronyme, die die Lesbarkeit erschweren und die Interpretationszeit verlängern [VDN<sup>+</sup>20], was sich exemplarisch in Fehler 8 widerspiegelt. [KLS25]



Quelle: Eigene Darstellung

Abbildung 9.3: SGAM Modell (Evaluation) für Fallbeispiel A

SGAAIRE wurde angewendet mit dem Ziel, die eingebauten Fehler zu identifizieren und ggf. zu verbessern. Es gilt daher zu evaluieren, ob das System in der Lage ist, alle integrierten Fehler weitestgehend zu erkennen. Das Optimierungsziel von SGAAIRE besteht darin, alle integrierten Fehler korrekt zu erkennen (*True Positives*  $t_p$ ), Fehlerfreies nicht zu identifizieren (*True Negatives*  $t_n$ ) und falsche Positive (*False Positives*  $f_p$ ) sowie falsche Negative (*False Negatives*  $f_n$ ) zu minimieren. Idealerweise sollte SGAAIRE auch Fehler im Referenzmodell/-beschreibungen identifizieren, die nicht in der „perfekten“ Grundwahrheit berücksichtigt sind. [KLS25]

Mithilfe von SGAAIRE konnten alle integrierten Fehler bis auf Fehlernummer 11 identifiziert werden. Zusätzlich erkannte das System neun weitere Fehler, die nicht bewusst integriert wurden (vgl. Tabelle A.44). Gleichzeitig identifizierte SGAAIRE sechs vermeintliche Fehler, die sich jedoch als korrekt erwiesen. Auch diese sind in Tabelle A.44 aufgeführt und können über das Beispiel im Online-Anhang<sup>1</sup> sowie der Online-Instanz von SGAAIRE reproduziert werden. Betroffene Qualitätsmerkmale, die mit „–“ gekennzeichnet sind, weisen darauf hin, dass bei der Anwendung ein grober Fehler auftreten kann, während „–“ auf einen vergleichsweise geringfügigen Fehler hinweist. Zur Bewertung der Leistung der Anwendung wurde eine Konfusionsmatrix (vgl. Abbildung 9.4) erstellt, um die Effektivität des Systems darzustellen.

Die Bestimmung von True Negatives stellt innerhalb des IEC 62559-2 Use Case Template sowie dem SGAM Modell aufgrund der Komplexität in der Zählweise (durch die Granularität) eine Herausforderung dar. Dadurch können weder die Genauigkeit noch die Spezifität-Metriken berechnet werden. Da die Bestimmung von True Negatives schwierig ist, können weder die Genauigkeit noch die Spezifität berechnet werden. Allerdings können aus den bestehenden Daten heraus die Präzision, Sensitivität (Recall) und F1-Score berechnet werden: [KLS25]

$$\text{prec} = \frac{18}{18 + 6} = 0.75, \quad \text{rec} = \frac{18}{18 + 1} = 0.95, \quad \text{f1} = 2 \cdot \frac{0.75 \cdot 0.95}{0.75 + 0.95} = 0.84$$

Bei der Anwendung konnte SGAAIRE neun weitere (nicht-eingebaute) Fehler identifizieren, bei deren Berücksichtigung sich die Metriken positiv verändern. Folglich lassen sich bestimmen:

$$\text{prec} = \frac{18 + 9}{18 + 9 + 6} = 0.81, \quad \text{rec} = \frac{18 + 9}{18 + 9 + 1} = 0.96, \quad \text{f1} = 0.88$$

Die Ergebnisse zeigen, dass die Präzision von SGAAIRE in einem akzeptablen Bereich liegt, die Sensitivität hoch ist und der F1-Score, der beide Faktoren berücksichtigt, ebenfalls ein hohes Niveau aufweist. Das System kann jedoch in Bezug auf die Präzision

Tabelle 9.1: Fallbeispiel A: Beschreibung der eingebetteten Fehler im IEC 62559 Use Case

Quelle: Übersetzt und angepasst aus [KLS25]

Fehlerkategorie	Beschreibung	Betroffene Qualitätsmerkmale
<b>Fehler 1: Use Case – SGAM-Domänenklassifikation</b>	• Falsche Domänenklassifikation, da der Use Case (laut Geltungsbereich) TSO, DSO, DER und CP berücksichtigt.	Korrektheit (++) , Nachvollziehbarkeit (+)
<b>Fehler 2: Use Case – SGAM-Zonenklassifikation</b>	• Falsche Zonenklassifikation, da der Use Case systeminterne und systemexterne Akteure in den unteren fünf Zonen einbezieht.	Korrektheit (++) , Nachvollziehbarkeit (+)
<b>Fehler 3: Use Case-Typ</b>	• Abstraktionsebene als Business High-Level definiert, tatsächlich jedoch Technical High-Level.	Korrektheit (++) , Relevanz (++) , Konsistenz (++) , Nachvollziehbarkeit (+) , Zugänglichkeit (++)
<b>Fehler 4 - 7: Szenario Step Services</b>	• Fehlende Service-Klassifikationen (2x). • Semantische Abweichungen in Service-Klassifikationen (2x).	Vollständigkeit (++) , Korrektheit (+)
<b>Fehler 8: Fehlende Kommunikation</b>	• Fehlende Schnittstellenkommunikation (z. B. CIM, CGMES, IEC 61850).	Vollständigkeit (+) , Nachvollziehbarkeit (++)
<b>Fehler 9: Akteure</b>	• Unvollständige Akteursbeschreibung: Nur TSO erwähnt, DSO, DER und Unternehmen fehlen.	Vollständigkeit (++) , Eindeutigkeit (+)
<b>Fehler 10: Texte</b>	• Ungenau definierte Begriffe, mangelnde Klarheit (6x).	Nachvollziehbarkeit (++) , Zugänglichkeit (+) , Vollständigkeit (+)

Tabelle 9.2: Fallbeispiel A: Beschreibung der eingebetteten Fehler im SGAM Modell

Quelle: Übersetzt und angepasst aus [KLS25]

Fehlerkategorie	Beschreibung	Betroffene Qualitätsmerkmale	Qualitätsmerkmale
<b>Fehler 9: Falsche Domänenklassifikation</b>	• Falsche Domänenklassifikation für das SSST-Dashboard, da der Geltungsbereich keine Erzeugung vorsieht. • Folglich muss die Erzeugung auch in den anderen Layern angezeigt werden.	Korrektheit Zugänglichkeit Nachvollziehbarkeit	(++), (+), (+)
<b>Fehler 10 und 11: Falsche Zonenklassifikation</b>	• Sensoren gehören gemäß SGAM User Manual [CEN14] in die Prozesszone, wobei eine Platzierung in der Feldebene für zusätzliche Datenerfassung ebenfalls zulässig ist. • Das SSST-System, als „Prozess“ klassifiziert, ist zu nah am Energieprozess, da es aggregierte Daten verarbeitet und den Betrieb unterstützt.	Korrektheit Zugänglichkeit Nachvollziehbarkeit	(++), (+), (+)
<b>Fehler 12: Falsche Layer-Zuweisung</b>	• Die Nutzung von „CIM: IEC 61968/61970“ ist für der Funktionslayer ungeeignet, da es primär die Interoperabilität für die Kommunikations- und Informationslayer sicherstellt.	Korrektheit Zugänglichkeit Nachvollziehbarkeit	(++), (+), (+)
<b>Fehler 13 und 14: Inkonsistenzen zwischen Use Case und SGAM</b>	• Inkonsistenz zwischen SGAM und Use Case-Beschreibung: Die Komponente „Sensoren“ wird im Use Case als „Sensoren/Komponenten“ bezeichnet. • Eine weitere Inkonsistenz besteht in der SGAM-Verbindung zwischen TSO und SSST-Dashboard, die im Use Case nicht spezifiziert ist.	Konsistenz Nachvollziehbarkeit Korrektheit	(+), (+), (+)
<b>Fehler 16: Irrelevante Elemente</b>	• Irrelevante Elemente ohne direkten Interoperabilitätsbedarf, wie der (Smart) Transformer. • Da Sensoren, das System und das Dashboard einzeln im SGAM Modell betrachtet werden, ist der (Smart) Transformer überflüssig.	Relevanz Eindeutigkeit	(++), (++)
<b>Fehler 17 und 18: Fehlende Elemente</b>	• DER als Akteur fehlt, obwohl er laut Geltungsbereich vorgesehen ist. • Der Kommunikationslayer spezifiziert eine proprietäre oder undefinierte Kommunikation. Die Nutzung eines konformen Standards wäre für die Vollständigkeit und Interoperabilität geeigneter.	Vollständigkeit Konsistenz	(++), (++)

in zukünftigen Anwendungen weiter optimiert werden. Während SGAAIRE als Expertensystem genutzt werden kann, sollten Vorschläge und Empfehlungen mit Vorsicht betrachtet und verantwortungsvoll übernommen werden. Dies spiegelt auch die Erwartungen wider, die beim Einsatz generativer KI-Systeme wie ChatGPT bestehen (vgl. [AAZG<sup>+</sup>23]). [KLS25]

Das *Fallbeispiel A: Optimierung der Lebensdauer von (smarten) Transformatoren mithilfe von Smart Services* zeigt, dass das System in der Lage ist, integrierte Fehler effektiv zu identifizieren und somit die Qualität zu verbessern. Die genaue Ermittlung einer Effizienz sowie ein Nachweis der praxisbezogenen Tauglichkeit erfolgt durch den Einsatz von SGAAIRE im Rahmen eines RCT-Experiments mit Probanden zur Qualitätssicherung (vgl. Abschnitt 9.2).

		Actual defects (via experts)	
		No defects	Defects
Identified defects (SGAIRE)	No defects	True Negative N/A	False Negative 1
	Defects	False Positive 6	True Positive 18 (+9)

Quelle: Eigene Darstellung

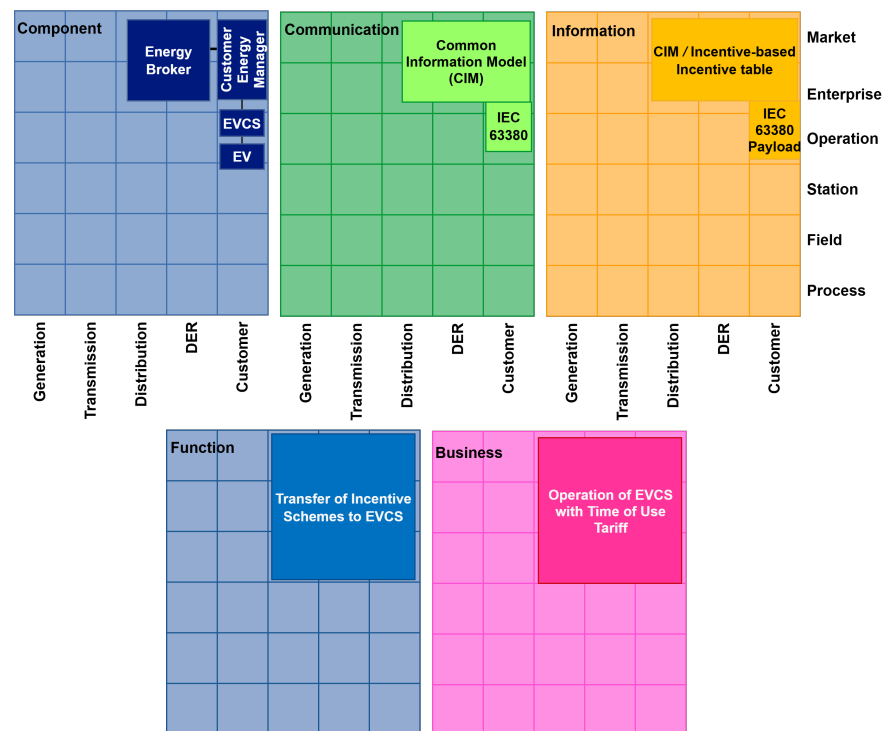
Abbildung 9.4: Fallbeispiel A: Konfusionsmatrix zur Fehleridentifikation durch die Anwendung von SGAIRE

### 9.1.2 Fallbeispiel B: Verwendung von dynamischen Stromtarifen für EV-Ladestationen

Das zweite Fallbeispiel befasst sich mit dem Anwendungsfall „Verwendung von dynamischen Stromtarifen für EV-Ladestationen“ und wurde aufgrund der einfachen Erreichbarkeit zudem innerhalb der Evaluation (vgl. Abschnitt 9.2) als Evaluationsbeispiel verwendet.

In diesem Anwendungsfall werden dynamische Stromtarife (Time of Use Tariffs) eingeführt mit dem Fokus auf das Laden von E-Fahrzeugen. Es wird ein anreizbasierter Mechanismus zur netzdienlichen Steuerung von Ladevorgängen für EVs vorgesehen, der auf die wachsenden Herausforderungen durch die zunehmende Elektrifizierung des Verkehrs und die Integration regenerativer Energien reagiert. Um lokale Überlastungen im Verteilnetz zu vermeiden, senden Energieversorgungsdienstleister oder Verteilnetzbetreiber zeitabhängige Anreiztabellen an den Customer Energy Manager (CEM) eines Gebäudes oder Ladeclusters. Dieser bewertet die übermittelten Anreize (wie z. B. Strompreis- oder CO<sub>2</sub>-Signale) und steuert die Ladevorgänge der angeschlossenen EVs innerhalb ihrer Flexibilitätsgrenzen. Mithilfe mehrerer Tarifstufen (Tiers), die jeweils ein Set an Anreizinformationen enthalten, entscheidet der CEM über Ladezeitpunkte und -dauern. Abhängig von regulatorischem Rahmen und Vertragslage können sowohl ESP als auch DSO Anreize bereitgestellt werden. Die beteiligten Akteure sind der Energy Broker (Optimierungseinheit des ESP), der Delivery Broker (DSO oder ESP) und der CEM. [KVN<sup>+</sup>26]

Für das vorliegende Fallbeispiel B sind analog zu Fallbeispiel A sowohl eine Referenz- als auch eine fehlerhafte (Evaluation) Version der IEC 62559 Use Case-Beschreibung



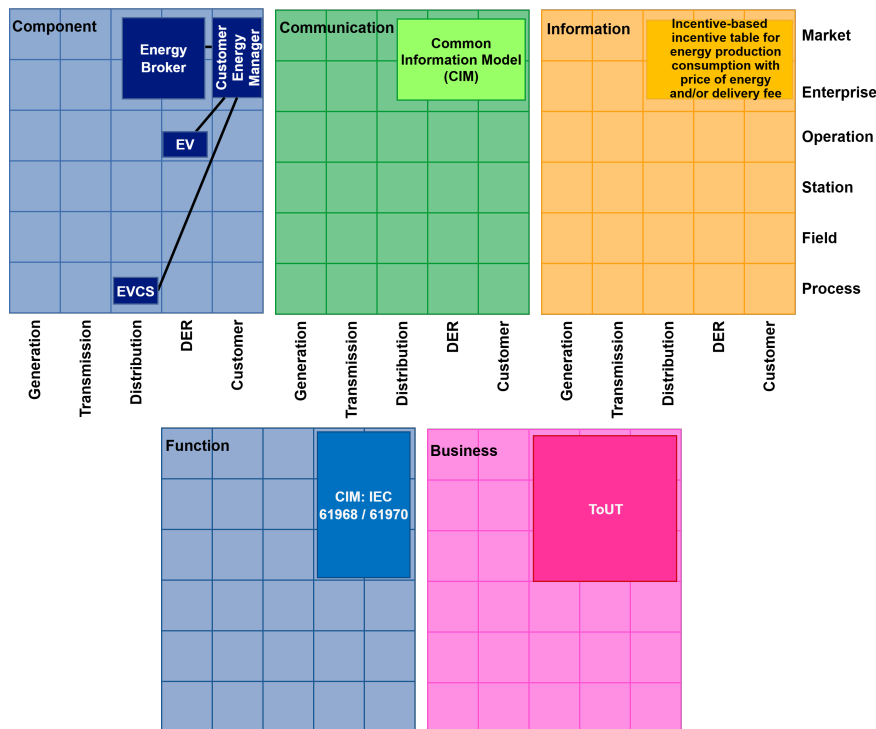
Quelle: Eigene Darstellung

Abbildung 9.5: SGAM Modell (Referenz) für Fallbeispiel B

sowie des zugehörigen SGAM Modells im Online-Anhang verfügbar<sup>2</sup>. Abbildung 9.5 zeigt das Referenzmodell, Abbildung 9.6 die fehlerhafte Evaluationsversion. Die Referenzversion des IEC 62559 Use Case wird sowohl in Abschnitt A.2 als auch im Online-Anhang<sup>1</sup> bereitgestellt, wohingegen die fehlerhafte Variante ausschließlich im Online-Anhang<sup>1</sup> enthalten ist. Die dokumentierten Fehler (vgl. Tabellen 9.3 und 9.4) wurden gezielt den Qualitätsmerkmalen aus Abschnitt 6.3.3 zugeordnet. „++“ kennzeichnet eine starke, „+“ eine schwächere Adressierung. Analog zu Fallbeispiel A basiert die Fehlerauswahl in Fallbeispiel B auf den Anforderungen der IEC 62559 Use Cases und SGAM Modelle ([IEC15, CEN14, GDU17, KSS23]) und umfasst typische sowie komplexe Modellierungsprobleme, welche bereits im Abschnitt 9.1.1 erläutert wurden und sich an diesen orientiert. In dem Abschnitt sind ebenfalls das Optimierungsziel beschrieben, die auch für das Demonstrationsbeispiel gelten. [KVN<sup>+</sup>26]

Mithilfe von SGAAIRE konnten ebenfalls bis auf Fehlernummer 6 alle eingebauten Fehler identifiziert werden. Des Weiteren konnte das System auch noch weitere nicht-definierte Terme identifizieren, weshalb die True Positives um einen halben Punkt steigen. SGAAIRE identifizierte zudem einen vermeintlichen Fehler, nämlich dass im IEC 62559 Use Case die Anforderung „Undefined/Proprietary Communication“ hinzugefügt

<sup>2</sup> vgl. Ordner „Demonstration/Fallbeispiel B“ im Online-Anhang (vgl. Abschnitt A.3)



Quelle: Eigene Darstellung

Abbildung 9.6: SGAM Modell (Evaluation) für Fallbeispiel B

werden sollte. Eine mögliche Erklärung hierfür liegt darin, dass das KI-System mit dem „Element Proposer“-Modul keinen passenden Standard gefunden hatte. Ein möglicher passender Standard wäre die Erstellung eines CIM-Profiles für dynamische Stromtarifdaten oder die Verwendung der EEBUS Spezifikation (vgl. Tabelle A.45). Diese Ergebnisse können über das Beispiel ebenfalls im Online-Anhang<sup>2</sup> sowie der Online-Instanz von SGAAIRE reproduziert werden. Zur Bewertung der Leistung der Anwendung wurde eine Konfusionsmatrix (vgl. Abbildung 9.7) erstellt, um die Effektivität des Systems darzustellen.

Die Ergebnisse des vorliegenden Demonstrationsbeispiels orientieren sich in weiten Teilen am Demonstrationsbeispiel A (vgl. Abschnitt 9.1.1). Sowohl die Präzision als auch die Sensitivität und der F1-Score liegen jeweils über 90 % und sind damit im gegebenen Kontext als sehr gut zu bewerten. Auch in diesem Fallbeispiel zeigt sich, dass die generierten Vorschläge und Empfehlungen größtenteils sinnvolle Resultate liefern. Dennoch sollten sie – analog zum Einsatz generativer KI-Systeme wie ChatGPT (vgl. [AAZG<sup>+</sup>23]) – mit der gebotenen Sorgfalt geprüft und verantwortungsbewusst übernommen werden.

Das Fallbeispiel B: Verwendung von dynamischen Stromtarifen für EV-Ladestationen zeigt ebenfalls, dass das System in der Lage ist, integrierte Fehler effektiv zu identi-

Tabelle 9.3: Fallbeispiel B: Beschreibung der eingebetteten Fehler im IEC 62559 Use Case

Quelle: Übersetzt aus [KVN<sup>+</sup>26]

Fehlerkategorie	Beschreibung	Betroffene Qualitätsmerkmale
<b>Fehler 1: Use Case – Domänenklassifikation</b>	Die Klassifikation ist inkorrekt, da die Akteure den Domänen DSO, DER und CP zugeordnet werden können. Dies ergibt sich aus der Akteursliste und der Scope-Beschreibung des Use Cases.	Korrektheit (++), Nachvollziehbarkeit (+)
<b>Fehler 2: Use Case – Zonenklassifikation</b>	Die Zonenklassifikation verstößt gegen das SGAM Konvexitätsprinzip, da ein Use Case mit Market und Operation auch Enterprise umfassen muss. Zudem ist der Energy Service Provider der Zone „Enterprise“ zuzuordnen.	Korrektheit (++), Nachvollziehbarkeit (+)
<b>Fehler 3: Use Case-Typ</b>	Die Abstraktionsebene ist fälschlicherweise als Technical Low-Level definiert, obwohl es sich um einen High-Level Business Use Case handelt.	Korrektheit (++), Relevanz (++), Konsistenz (++), Erreichbarkeit (++), Nachvollziehbarkeit (+)
<b>Fehler 4 - 5: Scenario Step Services</b>	Im Step 1a fehlt eine Klassifikation des Informationsflusses. Die Klassifikation für 1b sollte EXECUTE statt REMOVE sein.	Vollständigkeit (++), Korrektheit (++), Nachvollziehbarkeit (++)
<b>Fehler 6: Fehlende Standardangabe</b>	Es fehlen Angaben zur Schnittstellenkommunikation, wodurch unklar bleibt, welche Protokolle genutzt werden sollen.	Vollständigkeit (++), Erreichbarkeit (+)
<b>Fehler 7: Redundanter Akteur</b>	Der Akteur „Energy Service Provider“ ist redundant, da er nicht explizit im Use Case als Akteur vorgesehen ist. Dies kann zu Missverständnissen führen.	Relevanz (++), Unzweideutigkeit (+)
<b>Fehler 8: undefinierte Acronyme</b>	Acronyme wie ICT, ToUT, ESP, EV/s, DF, PoE und IEC sind nicht definiert.	Nachvollziehbarkeit (++), Eindeutigkeit (++), Erreichbarkeit (+), Vollständigkeit (+)

Tabelle 9.4: Fallbeispiel B: Beschreibung der eingebetteten Fehler im SGAM Modell

Quelle: Übersetzt aus [KVN<sup>+</sup>26]

Fehlerkategorie	Beschreibung	Betroffene Qualitätsmerkmale
<b>Fehler 9 und 10: Falsche Domänenklassifikation</b>	Die Klassifikation der <i>Charging Station</i> und des <i>Electric Vehicle</i> als Teil der <i>Distribution</i> -Domäne ist fehlerhaft. Beide Komponenten gehören z. B. zu <i>Customer Premises</i> .	Korrektheit (++), Erreichbarkeit (+), Nachvollziehbarkeit (+)
<b>Fehler 11: Falsche Zonenklassifikation</b>	Die <i>Charging Station</i> wurde fälschlicherweise der <i>Process-Zone</i> zugeordnet, obwohl sie als Aggregation mehrerer Komponenten in die <i>Operation-Zone</i> gehört.	Korrektheit (++), Erreichbarkeit (+), Nachvollziehbarkeit (+)
<b>Fehler 12: Unpassende Funktion</b>	Die Nutzung von <i>CIM: IEC 61968/61970</i> für die Function Layer ist unpassend, da diese Standards primär die Interoperabilität auf der Kommunikations- und Informationslayer sicherstellen, aber nicht (ausreichend) auf funktionale Interoperabilität.	Korrektheit (++), Erreichbarkeit (+), Nachvollziehbarkeit (+)
<b>Fehler 13: Function Layer Domänen- und Zonenklassifikation</b>	Die Elemente der Function Layer sollten die aus den unteren Layern angegebene Domänen und Zonen ( <i>Distribution, DER, Customer Premises</i> ) entsprechend reflektieren.	Korrektheit (++), Konsistenz (++)
<b>Fehler 14: Fehlende Verbindung</b>	Im SGAM Modell fehlt die direkte Verbindung zwischen <i>Electric Vehicle</i> und <i>Charging Station</i> , während eine fehlerhafte Verbindung zwischen <i>CEM</i> und <i>EV</i> existiert. Die Modellierung sollte mit der Use Case-Beschreibung konsistent sein.	Konsistenz (++), Nachvollziehbarkeit (++), Korrektheit (++)
<b>Fehler 15: Unklare Definition von Business-Elementen</b>	Der Term <i>ToUT</i> ist zu unspezifisch und erfordert eine genauere Beschreibung.	Nachvollziehbarkeit (++), Konsistenz (+)
<b>Fehler 16: Fehlendes Informationsobjekt</b>	Das im Szenario vorgesehene Objekt <i>E-Vehicle Charging Data</i> fehlt im Modell und sollte aus Vollständigkeitsgründen integriert werden.	Vollständigkeit (++), Konsistenz (++), Relevanz (++), Nachvollziehbarkeit (++), Korrektheit (+)

fizieren und somit zur Qualitätsverbesserung beizutragen. Die Demonstration anhand von zwei Beispielen ist ein Indikator für die Effektivität des Systems, stellt jedoch keinen Beleg dafür dar, dass IEC 62559 Use Case-Erstellende sowie SGAM Modellierende

		Actual defects (via experts)	
		No defects	Defects
Identified defects (SGAAIRE)	No defects	True Negative N/A	False Negative 1
	Defects	False Positive 1	True Positive 16 (+0,5)

Quelle: Eigene Darstellung

Abbildung 9.7: Fallbeispiel B: Konfusionsmatrix zur Fehleridentifikation durch die Anwendung von SGAAIRE

in jedem Fall eine Qualitätsverbesserung durch den Einsatz des Systems erzielen. Aus diesem Grund wurde das vorliegende Fallbeispiel im Rahmen der Evaluation (vgl. Abschnitt 9.2) als Experimentbeispiel herangezogen.

$$\text{prec} = \frac{16}{16 + 1} = 0.94, \quad \text{rec} = \frac{16}{16 + 1} = 0.94, \quad \text{f1} = 2 \cdot \frac{0.94 \cdot 0.94}{0.94 + 0.94} = 0.94$$

$$\text{prec} = \frac{16 + 0.5}{16 + 1} = 0.94, \quad \text{rec} = \frac{16 + 0.5}{16 + 0.5 + 1} = 0.94, \quad \text{f1} = 2 \cdot \frac{0.94 \cdot 0.94}{0.94 + 0.94} = 0.94$$

## 9.2 Evaluation

Über die Demonstrationsbeispiele konnte bereits gezeigt werden, dass SGAAIRE grundsätzlich zur Steigerung der Qualität innerhalb von IEC 62559 Use Cases und SGAM Modellen beitragen kann. Da die praktische Nutzung jedoch durch Modellierende mit unterschiedlichem Erfahrungsstand vom Anfänger- bis zum Expertenniveau sowohl im methodischen als auch im domänenspezifischen Bereich erfolgt, ist eine empirische Untersuchung des Nutzerkreises erforderlich.

Dieser Abschnitt behandelt die Evaluation sowie die damit einhergehende Datenerhebung, die zur Beantwortung der Forschungsfragen hinsichtlich des Beitrags SGAAIRE zur Qualitätsverbesserung intelligenter Systeme herangezogen werden. Ziel der Evaluation ist es, empirische Daten zu erheben, um belastbare Aussagen über die Auswirkungen des Systems treffen zu können. Der Schwerpunkt liegt dabei insbesondere auf der Qualitätssteigerung. Zusätzlich wird untersucht, ob der Einsatz von SGAAIRE

möglichen Verzerrungen bspw. bei der Identifikation nicht eingebauter Fehler hervorruft. Diese Effekte wurden im Rahmen eines Experiments erfasst und anschließend analysiert.

Zur Durchführung des Experiments wurde eine systematische Vorgehensweise gewählt, die in Abbildung 9.8 dargestellt ist. Sie definiert die Schritte zur Erhebung, Kodierung und Analyse der Daten. Im Einzelnen umfasst das Vorgehen folgende Schritte:

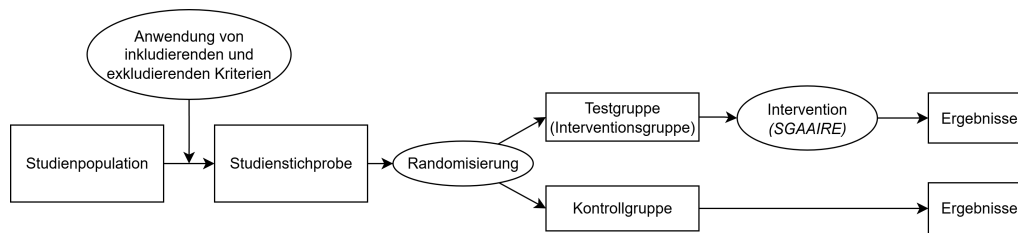


Quelle: Basierend auf [KVN<sup>+</sup>26]

Abbildung 9.8: Vorgehensweise zur Erhebung, Kodierung und Analyse der Evaluationsdaten

- **Datenerhebung durch ein Randomized Controlled Trial (RCT):** Um Daten aus der Empirie zu gewinnen, wurde ein online-basiertes RCT durchgeführt. Abbildung 9.9 zeigt den Aufbau der vorliegenden RCT-Studie. Dabei wurde eine Testgruppe (bzw. Interventionsgruppe), die mit dem KI-basierten Empfehlungssystem SGAAIRE arbeitete, mit einer Kontrollgruppe verglichen, die den konventionellen Modellierungsansatz ausschließlich auf Grundlage des bestehenden UCMR nutzte. Beide Vorgehensweisen orientieren sich im Abschnitt 7.1 beschriebenen Prozess. [KVN<sup>+</sup>26]

Da die Teilnehmenden online und asynchron am Experiment teilnehmen können, müssen sowohl die Gruppenzuweisung (Kontroll- oder Testgruppe) als auch der Validierungsprozess automatisiert erfolgen. Darüber hinaus soll die Benutzerfreundlichkeit des Systems (sowohl mit als auch ohne SGAAIRE) direkt durch die Teilnehmenden evaluiert werden. Hierfür sind zusätzlich zum Experiment folgende Komponenten integriert: [KVN<sup>+</sup>26]



Quelle: Konkretisiert aus [Hou15]

Abbildung 9.9: Aufbau der vorliegenden RCT

- **Eine Pre-Survey** zur Erfassung von Metadaten zu Expertise, Erfahrung und Rollen der Teilnehmenden in Bezug auf die IEC 62559 Use Case Methodik sowie das SGAM Framework.
- **Eine Post-Survey** zur Bewertung der Systemnutzung und zur Erhebung von Feedback nach Abschluss des Experiments. Dies umfasst eine standardisierte Evaluation der Usability mittels des SUS [BKM08] sowie offene Fragen zu positiven und negativen Aspekten des Systems und potenziell fehlenden Funktionalitäten für die User Experience.

Beide Umfragen wurden mithilfe von SurveyJS [Sur25] als unterstützende Bibliothek realisiert und können im Online-Anhang<sup>3</sup> nachvollzogen werden. Das gesamte Vorgehen wurde mit drei Pilotteilnehmenden getestet und auf eine Gesamtdauer von ca. 60 Minuten geschätzt.

Für die Teilnahme am Experiment war es erforderlich, dass die Teilnehmenden zuvor sowohl mindestens ein IEC 62559 Use Case als auch SGAM Modell erstellt hatten. Zusätzlich wurden sie gebeten, ihre Erfahrung als „Einsteiger“, „Fortgeschrittene“ oder „Experte“ selbst einzuschätzen. Auf Basis dieser Selbsteinschätzung erfolgte die Gruppenzuweisung zur Kontroll- oder Testgruppe so, dass eine gleichmäßige Verteilung und eine ausgewogene Repräsentation der Erfahrungsniveaus sichergestellt wurden. Die Teilnehmenden waren zudem in eine Qualitätssicherungsphase eingebunden (vgl. Abbildung 7.2), in der sie die Qualität eines IEC 62559 Use Case und eines korrespondierenden SGAM Modells für das Anwendungsszenario „Verwendung von dynamischen Stromtarifen für EV-Ladestationen“ (dargestellt in Abschnitt 9.1.2) bewerten sollten. Dieser Anwendungsfall wurde gewählt, da er sowohl leicht verständlich als auch praxisrelevant ist, insbesondere im Kontext Elektromobilität, dynamischer Tarife und nachfrageseitiger Flexibilität [RSJ<sup>+</sup>23]. [KVN<sup>+</sup>26]

- **Identifikation von eingebauten Fehlern:** Im Evaluationsbeispiel „Verwendung von dynamischen Stromtarifen für EV-Ladestationen“ wurden ebenfalls die im Abschnitt 9.1.2 gezielten Fehler integriert. Aufgabe der Teilnehmenden war es, diese Fehler zu identifizieren und entweder direkt zu korrigieren oder über Kommentare zu markieren. [KVN<sup>+</sup>26]

<sup>3</sup> unter Evaluation/pre-evaluation-survey.ts bzw. Evaluation/post-evaluation-survey.ts

- **Identifikation von nicht-eingebauten (zusätzlicher) Fehler:** Da nicht davon ausgegangen werden kann, dass eine IEC 62559 Use Case-Beschreibung oder ein SGAM Modell ausschließlich die vorgegebenen Fehler enthält, hatten die Teilnehmenden die Möglichkeit, zusätzliche, nicht implementierte Fehler zu identifizieren. Diese wurden anschließend auf ihre Relevanz geprüft und (sofern valide) den in Kapitel 6 beschriebenen Qualitätsmerkmalen zugeordnet. [KVN<sup>+</sup>26]
- **Quantitative Datenanalyse:** Die erhobenen Daten wurden kodiert und mittels deskriptiver Statistik, Hypothesentests und Effektstärkeberechnungen analysiert. Für jede teilnehmende Person wurde eine Punktzahl berechnet, die die Leistung in Bezug auf die Fehler und Qualitätsmerkmale quantifiziert. Details zu den Verfahren sind in den jeweiligen Abschnitten zu den Ergebnissen beschrieben. [KVN<sup>+</sup>26]

Die Auswahl der Hypothesen leitet sich unmittelbar aus der übergeordneten Zielsetzung 2 ab (vgl. Abschnitt 1.2), den Mehrwert des Einsatzes von SGAAIRE als intelligentes Unterstützungssystem zur Verbesserung der Modellqualität systematisch zu evaluieren. Tabelle 9.5 fasst die betrachteten Hypothesen zusammen, diese adressieren wie folgt:

- **Hypothesen 1 und 2 zur Fehleridentifikation:**  $H1$  und  $H2$  untersuchen, inwiefern der Einsatz von SGAAIRE die Erkennung von Fehlern verbessert. Dabei wird sowohl die Identifikation von eingebauten Fehlern als auch die Entdeckung zusätzlicher, nicht-eingebauter Fehler betrachtet. Diese Hypothesen prüfen, ob und in welchem Ausmaß SGAAIRE zur Steigerung der Fehlererkennungsleistung beiträgt und ob es ggf. zu einer Verzerrung in der Erkennung nicht-eingebauter Fehler führt.
- **Hypothese 3 zur Qualitätsverbesserung.**  $H3$  fokussiert sich auf die zentrale Zielgröße der Untersuchung: die Verbesserung der Qualität von IEC 62559 Use Cases und SGAM Modellen. Da Qualität in mehreren Dimensionen operationalisiert (gemäß der Qualitätsmerkmale aus Kapitel 6) wird, verbindet diese Hypothese die reine Fehleridentifikation mit einem umfassenderen Qualitätsverständnis.
- **Hypothesen 4 und 5 zum Einfluss des Erfahrungsniveaus.** Bei  $H4$  und  $H5$  wird der Einfluss des (selbsteingeschätzten) Erfahrungsniveaus der Teilnehmenden berücksichtigt. Da Erfahrung den Zusammenhang zwischen Werkzeugunterstützung und Ergebnisqualität moderieren kann, zielen diese Hypothesen darauf ab, diesen Zusammenhang differenziert zu untersuchen. Dabei wird zum einen ein Vergleich der Qualitätsverbesserungen zwischen den verschiedenen Erfahrungsgruppen insgesamt ( $H4$ ) vorgenommen, zum anderen wird der Zusammenhang innerhalb der Gruppen mit und ohne SGAAIRE betrachtet ( $H5$ ).
- **Hypothese 6 zur Benutzerfreundlichkeit:**  $H6$  betrachtet die wahrgenommene Benutzerfreundlichkeit. Diese Dimension zeigt die Akzeptanz und Gebrauchstauglichkeit von SGAAIRE innerhalb des praktischen Einsatzes.

In den folgenden Unterabschnitten werden zunächst die Durchführung der RCT und die erhobenen Metadaten beschrieben. Anschließend werden die in Tabelle 9.5 for-

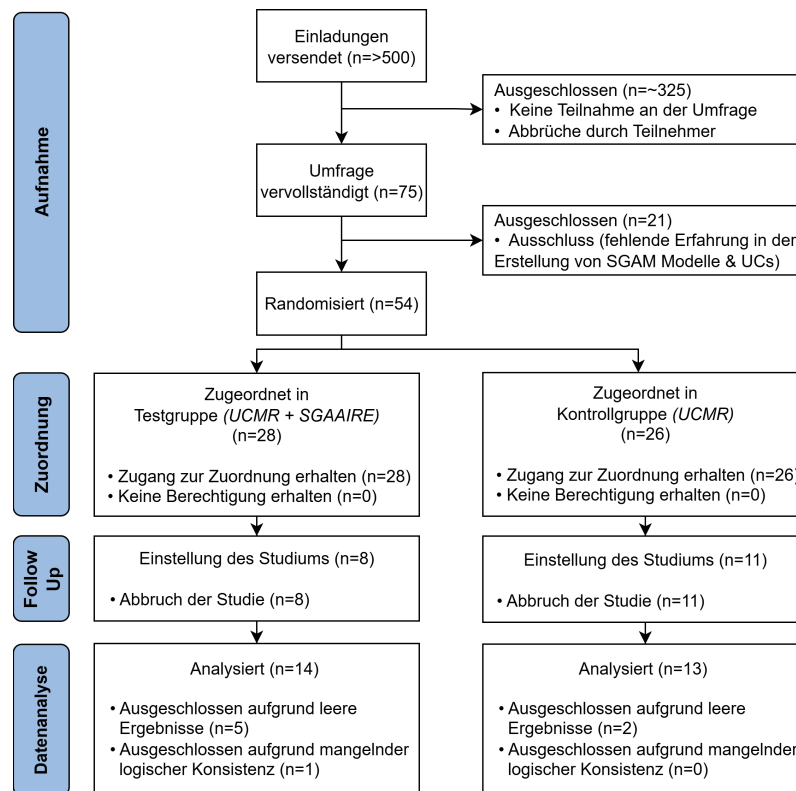
Tabelle 9.5: Übersicht der aufgestellten Hypothesen

ID	Nullhypothese $H_{x0}$	Alternativhypothese $H_{x1}$
H1	Es besteht kein Unterschied in der Anzahl der entdeckten eingebauten Fehler zwischen der Gruppe, die das UCMR-System mit SGAAIRE verwendet, und der Gruppe, die ausschließlich das UCMR-System nutzt.	Die Gruppe, die das UCMR-System in Kombination mit SGAAIRE verwendet, entdeckt mehr eingebaute Fehler als die Gruppe, die ausschließlich das UCMR-System nutzt.
H2	Es besteht kein Unterschied in der Anzahl der zusätzlich (nicht-eingebauten) entdeckten Fehler zwischen der Gruppe, die das UCMR-System in Kombination mit SGAAIRE verwendet, und der Gruppe, die ausschließlich das UCMR-System nutzt.	Die Gruppe, die das UCMR-System in Kombination mit SGAAIRE verwendet, entdeckt mehr zusätzliche (nicht-eingebaute) Fehler als die Gruppe, die ausschließlich das UCMR-System nutzt.
H3	Der Einsatz des UCMR-Systems in Kombination mit SGAAIRE führt zu keiner Verbesserung der Qualität im Vergleich zum alleinigen Einsatz des UCMR-Systems.	Der Einsatz des UCMR mit SGAAIRE führt zu einer Verbesserung der Qualität im Vergleich zum UCMR-System.
H4	Es existiert kein Unterschied in der Qualitätssteigerung zwischen den eingeschätzten Erfahrungsleveln.	Es existiert einen Unterschied in der Qualitätssteigerung zwischen den eingeschätzten Erfahrungsleveln.
H5	Das eingeschätzte Erfahrungslevel hat keinen Einfluss auf die Qualitätssteigerung, sowohl in der Gruppe UCMR mit SGAAIRE als auch in der Gruppe UCMR ohne SGAAIRE.	Das eingeschätzte Erfahrungslevel hat einen Einfluss auf die Qualitätssteigerung, sowohl in der Gruppe UCMR mit SGAAIRE als auch in der Gruppe UCMR ohne SGAAIRE.
H6	Es besteht kein Unterschied in der wahrgenommenen Benutzerfreundlichkeit zwischen der Gruppe, die das UCMR mit SGAAIRE verwendet, und der Gruppe, die ausschließlich das UCMR-System verwendet.	Die Gruppe, die das UCMR mit SGAAIRE verwendet, bewertet die Benutzerfreundlichkeit des Systems höher als die Gruppe, die ausschließlich das UCMR-System verwendet.

mulierten Hypothesen den jeweiligen Themenbereichen in den Abschnitten quantitativ adressiert. Dies umfasst eine deskriptive Auswertung zentraler Kennwerte, inferenzstatistische Hypothesentests sowie die Berechnung von Effektstärken.

### 9.2.1 Durchführung der randomisierten kontrollierten Studie

Das Experiment wurde im Rahmen eines Online-RCT asynchron im Zeitraum vom 7. Januar bis 7. Februar komplett anonymisiert und ohne zusätzliche Anreize durchgeführt. Für die Rekrutierung wurden über 500 Einladungen über verschiedene Kanäle verbreitet. Dazu zählten insbesondere die europäische BRIDGE Working Group (WG) Data Management, die Chairs der CEN-CENELEC-ETSI CG-SG sowie der Convenor der Set of Standards WG, ebenso wie Chair und Convenors der IEC SyC Smart Energy, die für die Verwaltung des IEC 62559 Use Case-Standards sowie des IEC 63200 (erweiterter SGAM-Standard) verantwortlich sind. Weitere Kanäle umfassten die int:net-Community, die Mailinglisten von OFFIS und des int:net-Projekts, LinkedIn, Weiterleitungen bereits eingeladener Personen sowie individuelle Kontakte, die beispielsweise im Rahmen früherer Studien ihr Interesse bekundet hatten.



Quelle: Übersetzt aus [KVN+26]

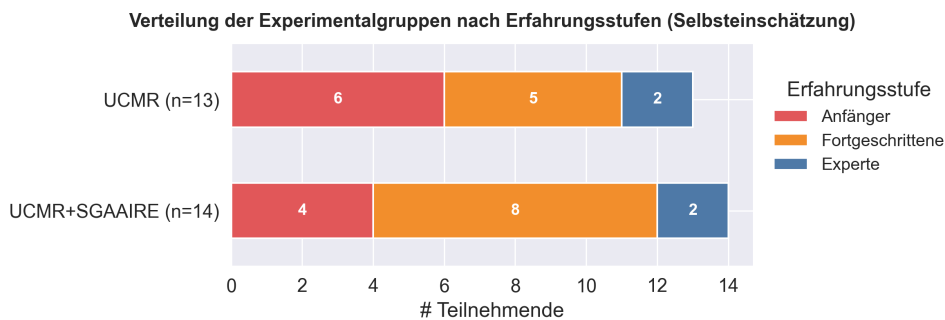
Abbildung 9.10: CONSORT-Diagramm zur Darstellung des Teilnehmerflusses

Einen Überblick über die Flüsse der Teilnehmenden bietet das CONSORT-Diagramm in Abbildung 9.10. Von den Eingeladenen füllten 75 Personen die Vorabumfrage (Pre-Survey) aus. Davon wurden 54 Teilnehmende zufällig den beiden Untersuchungsgruppen zugewiesen: 28 in die UCMR-Version mit Unterstützung durch das SGAAIRE-KI-System (Testgruppe) und 26 in die Baseline-Version von UCMR (Kontrollgruppe). Aufgrund von frühzeitigem Studienabbruch, leeren Abgaben sowie logischen Inkonsistenzen in den Antworten konnten in der finalen Auswertung 14 vollständige Datensätze in der Testgruppe und 13 in der Kontrollgruppe berücksichtigt werden.

Die Stichprobengröße von 27 Teilnehmenden ist im Kontext der Evaluierung eher gering, im Rahmen des Experiments aber als akzeptabel einzustufen. In das Experiment wurden Personen mit sehr unterschiedlichen Erfahrungsstufen einbezogen, von Studierenden über Gelegenheitsanwendende und Fachexperten bis hin zu Standardisierungsexperten aus dem Methodenkreis. Diese heterogene Zusammensetzung ermöglichte die Erhebung hochwertiger, praxisnaher Daten, die in den statistischen Analysen (vgl. entsprechende Abschnitte der Ergebnisse) klare Tendenzen aufzeigen. Zugleich verdeutlicht die Studie ein bekanntes Phänomen: Empirische Untersuchungen mit menschenbezogenen Evaluierungen im Requirements Engineering weisen weiterhin Verbesserungs-

potenzial auf, wie bereits [HS21] im Kontext des modellbasierten Systems Engineerings anmerkte und auch noch in aktuellen Arbeiten bestätigt wird [FFT<sup>+</sup>24, UL24]. Die Durchführung empirischer Studien, insbesondere randomisierter kontrollierter Experimente (RCTs), im Requirements Engineering ist methodisch anspruchsvoll. Sie erfordert sowohl domänenspezifisches Fachwissen als auch einen hohen zeitlichen Aufwand und wird zusätzlich dadurch erschwert, dass es keine objektive Grundwahrheit für „gute“ Anforderungen gibt. Der menschenzentrierte Charakter des Requirements Engineering verschärft diese Komplexität [FWK<sup>+</sup>17, AIUN18]. Die hohe Abbruchquote von 64 % (berechnet ab Umfrage vervollständigt) lässt sich dementsprechend darauf zurückführen, dass für die Bearbeitung ein überdurchschnittliches Maß an Fachwissen erforderlich war und die Bearbeitungsdauer von rund 60 Minuten ohne entsprechende Anreizsetzung eine zusätzliche Hürde darstellte. Eine weiterführende Einordnung erfolgt in Abschnitt 9.3.1 (Diskussion der Studien).

Die Verteilung der Erfahrung der Teilnehmenden ist in Abbildung 9.11 dargestellt. Unter den 27 gültigen Versuchspersonen wurden zehn als Einsteiger (sechs in der Kontrollgruppe, vier in der Testgruppe), dreizehn als fortgeschrittene Anwender (fünf in der Kontrollgruppe, acht in der Testgruppe) und vier als Experten (jeweils zwei in beiden Gruppen) klassifiziert.

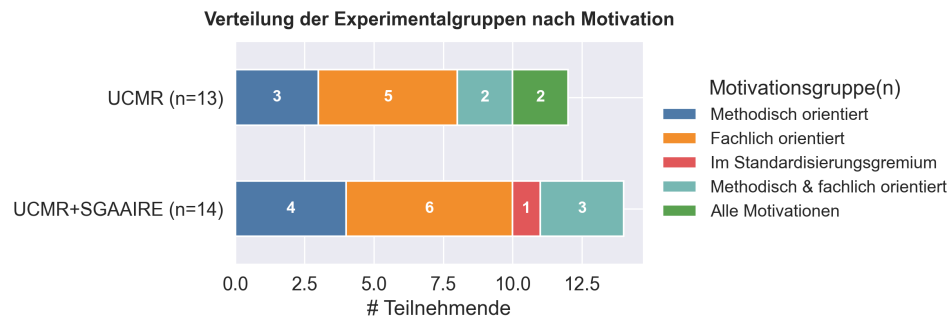


Quelle: Übersetzt aus [KVN<sup>+</sup>26]

Abbildung 9.11: Teilnehmeraufteilung (n=27) in Kontrollgruppe UCMR (n=13) und Testgruppe UCMR+SGAAIRE (n=14) nach Erfahrungsniveau

Darüber hinaus unterscheiden sich die Teilnehmenden in ihren Motivationstypen zur Beschäftigung mit der IEC 62559 Use Case Methodik und dem SGAM Framework. In Anlehnung an [SCUG21] lassen sie sich als methodenorientiert (Fokus auf methodisches Arbeiten mit Template und Framework) oder inhaltsorientiert (Beitrag über konkrete Anwendungsfälle ohne primären Fokus auf die Methodik) kategorisieren. Des Weiteren wurden über Teilnehmende aus Standardisierungsgremien (u. a. IEC TC 57 und IEC System Committee Smart Energy) erreicht. Eine Übersicht über die Motivationsorientierungen ist in Abbildung 9.12 dargestellt. In der Kontrollgruppe fanden sich drei methodenorientierte und fünf fachorientierte Teilnehmende, während die Testgruppe vier methoden- und sechs fachorientierte Personen umfasste. Darüber hinaus gaben

zwei Teilnehmende der Kontroll- und drei Teilnehmende der Testgruppe an, sowohl methoden- als auch fachorientiert motiviert zu sein. In der Testgruppe hatte eine Person die Motivation aus dem Standardisierungsgremium, während in der Kontrollgruppe zwei Personen allen drei Motivationsgruppen zugeordnet werden konnten.



Quelle: Übersetzt aus [KVN<sup>+</sup>26]

Abbildung 9.12: Teilnehmeraufteilung (n=27) in Kontrollgruppe UCMR (n=13) und Testgruppe UCMR+SGAAIRE (n=14) nach Motivation

An der abschließenden User Experience-/SUS-Umfrage (Post Survey) nahmen 35 Probandinnen und Probanden teil. Diese Zahl liegt über den analysierten Versuchspersonen aus dem Experiment (n=27). Ein möglicher Grund hierfür ist, dass einzelne Personen, die entweder nicht aktiv am Experiment beteiligt waren oder unzuverlässige Daten erzeugten, den Experimentteil übersprangen, jedoch dennoch an der abschließenden Umfrage teilgenommen haben. Die Ergebnisse wurden dennoch berücksichtigt, da die Teilnehmenden das System aus ihrer jeweiligen Perspektive getestet haben und somit ein breiteres Spektrum an Rückmeldungen in die Bewertung einfließen konnte.

## 9.2.2 Einfluss von SGAAIRE auf die Identifikation eingebauter Fehler

Ein Bestandteil dieser Studie ist die Untersuchung, welche der in einem IEC 62559 Use Case sowie dem zugehörigen SGAM Modell gezielt eingefügten Defizite von den Teilnehmenden erkannt wurden. Zu diesem Zweck wurden sowohl typische als auch gezielt eingefügte Fehler in das Evaluationsbeispiel integriert, wie in Abschnitt 9.1.2 erläutert. Die Teilnehmenden hatten die Aufgabe, diese Fehler zu identifizieren, zu kommentieren oder zu korrigieren. Zur Überprüfung wurden die Hypothese 2.1 (Nullhypothese,  $H_0$ ) sowie die entsprechende Alternativhypothese ( $H_1$ ) formuliert. [KVN<sup>+</sup>26]

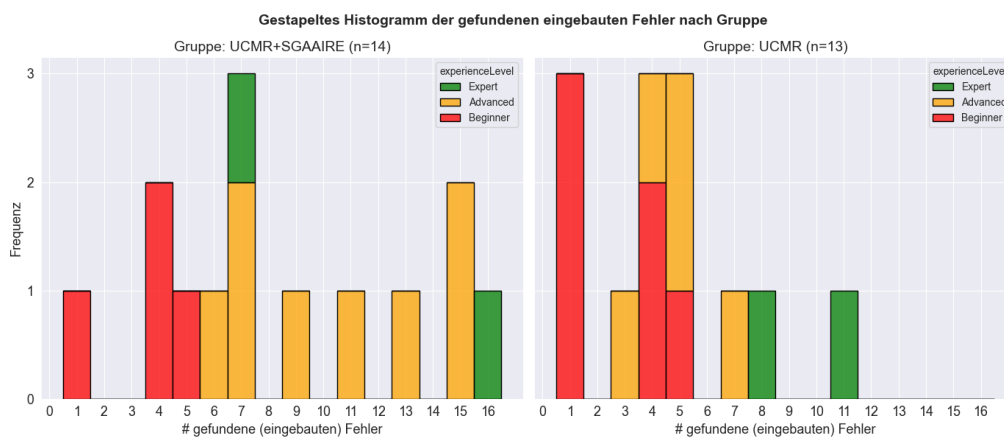
### Subhypothese 2.1 - Nullhypothese

$H_0$ : Es besteht kein Unterschied in der Anzahl der entdeckten eingebauten Fehler zwischen der Gruppe, die das UCMR-System mit SGAAIRE verwendet, und der Gruppe, die ausschließlich das UCMR-System nutzt.

### Subhypothese 2.1 - Alternativhypothese

$H_1$ : Die Gruppe, die das UCMR-System in Kombination mit SGAAIRE verwendet, entdeckte mehr eingebaute Fehler als die Gruppe, die ausschließlich das UCMR-System nutzt.

Die Verteilung der Ergebnisse ist in Abbildung 9.13 dargestellt. Die Interventionsgruppe (UCMR+SGAAIRE) erzielte dabei eine höhere durchschnittliche Trefferzahl ( $\mu_{SGAAIRE} = 8.41$ ) im Vergleich zur Kontrollgruppe (UCMR,  $\mu_{UCMR} = 4.27$ ). Zudem zeigte die Interventionsgruppe eine größere Varianz ( $\sigma_{SGAAIRE} = 4.82$ ) im Vergleich zur Kontrollgruppe ( $\sigma_{UCMR} = 2.88$ ). Auffällig ist außerdem, dass innerhalb der Gruppe, die das KI-basierte System nutzte, ein höherer Anteil an Teilnehmenden zehn oder mehr eingebaute Fehler identifizieren konnte. [KVN<sup>+</sup>26]



Quelle: Übersetzt aus [KVN<sup>+</sup>26]

Abbildung 9.13: Vergleich von Kontroll- und Testgruppe hinsichtlich der identifizierten eingebauten Fehler

Zur Prüfung der aufgestellten Hypothesen wurde ein *Mann-Whitney-U Signifikanztest* durchgeführt, der sich insbesondere für nicht normalverteilte Daten bei zwei unabhängigen Gruppen eignet. Der Test ergab einen Wert von  $U = 139.5$  bei einem p-Wert von  $p = 0.0195$ . Dieser liegt unterhalb des Signifikanzniveaus von  $\alpha = 0.05$ , wodurch die Nullhypothese  $H_0$  verworfen werden kann. Der berechnete Effekt von  $r = 0.77$  ist gemäß den Richtlinien nach Cohen als groß zu klassifizieren. [KVN<sup>+</sup>26]

### 9.2.3 Einfluss von SGAAIRE auf die Identifikation nicht-eingebauter Fehler

Im Zusammenhang mit dem Automation Bias ist es relevant zu prüfen, ob die Nutzung von SGAAIRE dazu führt, dass Teilnehmende weniger andere Fehler(arten) identifizieren, die nicht von SGAAIRE abgedeckt werden können. Dies würde bei einer gerin-

geren identifizierten Anzahl nicht-eingebauter Fehler zu einem solchen Bias hindeuten [VLH<sup>+</sup>23]. Zur Überprüfung wurden die Hypothese 2.2 (Nullhypothese,  $H_0$ ) sowie die entsprechende Alternativhypothese ( $H_1$ ) formuliert. [KVN<sup>+</sup>26]

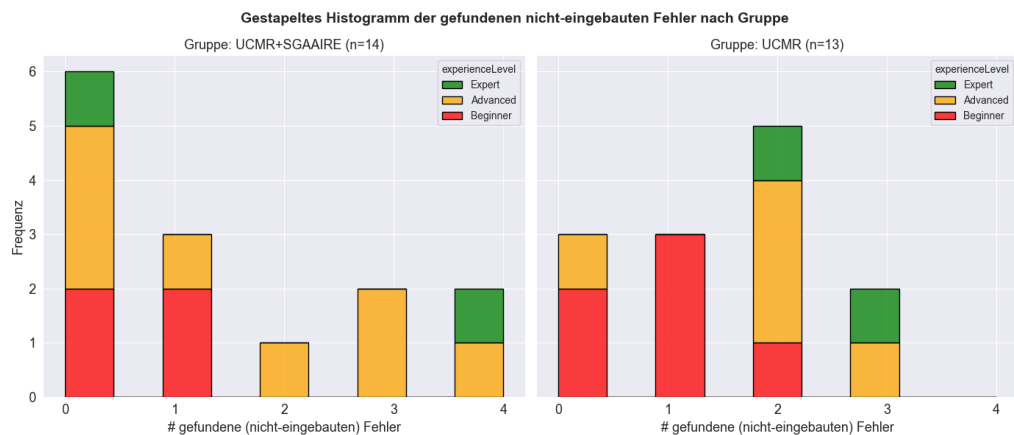
### Subhypothese 2.2 - Nullhypothese

$H_0$ : Es besteht kein Unterschied in der Anzahl der zusätzlich (nicht-eingebauten) entdeckten Fehler zwischen der Gruppe, die das UCMR-System in Kombination mit SGAAIRE verwendet, und der Gruppe, die ausschließlich das UCMR-System nutzt.

### Subhypothese 2.2 - Alternativhypothese

$H_1$ : Die Gruppe, die das UCMR-System in Kombination mit SGAAIRE verwendet, entdeckt mehr zusätzliche (nicht-eingebaute) Fehler als die Gruppe, die ausschließlich das UCMR-System nutzt.

Die Verteilung der Resultate ist in Abbildung 9.14 dargestellt. Der Mittelwert der identifizierten nicht-eingebauten Fehler in der Interventionsgruppe ( $\mu_{SGAAIRE} = 1.36$ ) liegt geringfügig unter dem Mittelwert der Kontrollgruppe ( $\mu_{UCMR} = 1.46$ ). Die Varianz in der Interventionsgruppe ( $\sigma_{SGAAIRE}^2 = 1.55$ ) ist jedoch höher als in der Kontrollgruppe ( $\sigma_{UCMR}^2 = 1.05$ ). Die Ergebnisse zeigen ein heterogenes Bild: Ein Teil der Teilnehmenden der Interventionsgruppe identifizierte vier oder mehr zusätzliche Fehler und übertraf damit alle Teilnehmenden der Kontrollgruppe. Gleichzeitig zeigte sich jedoch, dass eine größere Zahl von Probanden keine zusätzlichen Fehler identifizierte (drei in der Kontrollgruppe und sechs in der Interventionsgruppe). [KVN<sup>+</sup>26]



Quelle: Übersetzt aus [KVN<sup>+</sup>26]

Abbildung 9.14: Vergleich von Kontroll- und Testgruppe hinsichtlich der identifizierten nicht-eingebauten (zusätzlichen) Fehler

Zur Bewertung der Hypothesen wurde ein *Mann-Whitney-U-Test* durchgeführt. Der Test ergab einen Wert von  $U = 81$  mit einer Effektstärke von  $r = 0.45$ , die gemäß

Cohen als mittlerer Effekt einzuordnen ist. Der zugehörige p-Wert von  $p = 0.63$  liegt jedoch deutlich oberhalb des Signifikanzniveaus von  $\alpha = 0.05$ . Folglich kann die Nullhypothese  $H_0$  nicht verworfen werden. Auf Grundlage der vorliegenden Daten lässt sich somit kein signifikanter Unterschied feststellen, weder in der Richtung, dass SGAAIRE die Identifikation zusätzlicher Fehler unterstützt, noch in der Richtung, dass durch den Einsatz von SGAAIRE weniger Fehler erkannt werden. [KVN<sup>+</sup>26]

#### 9.2.4 Einfluss von SGAAIRE auf die Qualitätsverbesserung

Der Fokus von Forschungsfrage 2 liegt auf der Bewertung der Qualitätsverbesserungen, die durch den Einsatz intelligenter Systeme bei IEC 62559 Use Cases und SGAM Modellen erzielt werden können. Zu diesem Zweck wurden sowohl die in den vorherigen Abschnitten betrachteten eingebauten Fehler als auch die zusätzlich identifizierten Fehler in Qualitätswerte überführt. Für die eingebauten Fehler erfolgte die Überführung anhand der Qualitätsmerkmale des Demonstrationsbeispiels in Abschnitt 9.1.2. Die Berücksichtigung der nicht-eingebauten Fehler orientierte sich an der in Abschnitt 9.2 beschriebenen Vorgehensweise; deren Identifikation und Zuordnung im Datensatz sind in der begleitenden Excel-Tabelle<sup>4</sup> dokumentiert. Beide Fehlertypen wurden in die Beurteilung der Qualitätsverbesserung einbezogen. Um zu untersuchen, ob und in welchem Ausmaß SGAAIRE als KI-basiertes System zu einer Qualitätsverbesserung beiträgt, wurde die Hypothese 2.3 (Nullhypothese,  $H_0$ ) sowie die entsprechende Alternativhypothese ( $H_1$ ) formuliert. [KVN<sup>+</sup>26] [KVN<sup>+</sup>26]

##### **Subhypothese 2.3 - Nullhypothese**

$H_0$ : Der Einsatz des UCMR-Systems in Kombination mit SGAAIRE führt zu keiner Verbesserung der Qualität im Vergleich zum alleinigen Einsatz des UCMR-Systems.

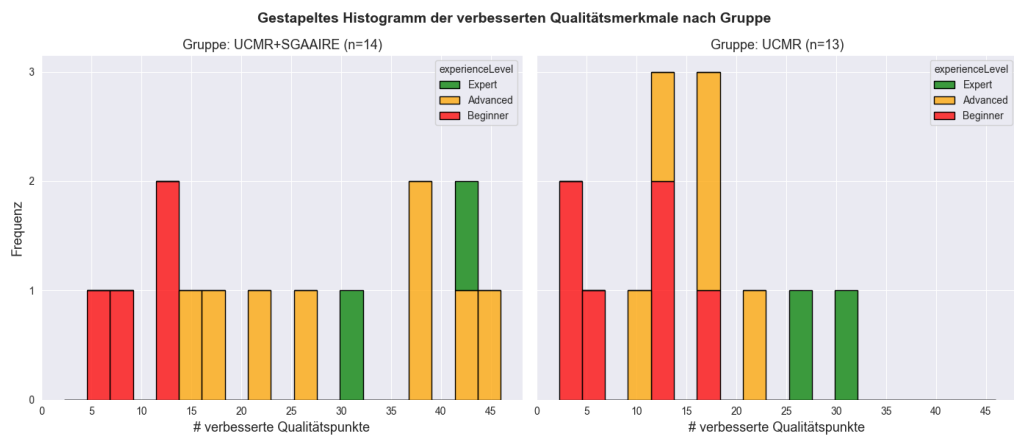
##### **Subhypothese 2.3 - Alternativhypothese**

$H_1$ : Der Einsatz des UCMR mit SGAAIRE führt zu einer Verbesserung der Qualität im Vergleich zum UCMR-System.

Abbildung 9.15 zeigt die Verteilung der Qualitätsverbesserungswerte zwischen Kontroll- und Interventionsgruppe. Es ist ein Verteilungseffekt zugunsten des KI-unterstützten Systems (UCMR+SGAAIRE) erkennbar. Der Mittelwert in dieser Gruppe liegt höher ( $\mu_{SGAAIRE} = 25.49$ ,  $\sigma_{SGAAIRE} = 14.33$ ) im Vergleich zur Kontrollgruppe ( $\mu_{UCMR} = 14.60$ ,  $\sigma_{UCMR} = 8.68$ ). Dies deutet darauf hin, dass die Nutzung von SGAAIRE durchschnittlich zu einer Qualitätssteigerung bei den Probanden führt (um 74,59 %). [KVN<sup>+</sup>26]

Zur statistischen Validierung dieser Beobachtung wurde ebenfalls ein *Mann-Whitney-U-Test* durchgeführt, der eine Teststatistik von  $U = 132$  und einen p-Wert von  $p =$

<sup>4</sup> vgl. Online-Anhang: Evaluation/evaluation.data.xlsx.



Quelle: Übersetzt aus [KVN<sup>+</sup>26]

Abbildung 9.15: Vergleich von Kontroll- und Testgruppe hinsichtlich der Qualitätsverbesserungen

0.0494 ergab. Da dieser Wert knapp unterhalb des Signifikanzniveaus von  $\alpha = 0.05$  liegt, kann die Nullhypothese  $H_0$  dennoch verworfen werden. Die berechnete Effektstärke von  $r = 0.73$  ist gemäß gängiger Richtlinien als groß zu klassifizieren und deutet auf einen positiven Einfluss von SGAIRE auf die Qualitätsverbesserung hin. [KVN<sup>+</sup>26]

Die in Abbildung 9.15 dargestellten Qualitätswerte repräsentieren die Gesamtsumme über alle betrachteten Qualitätsmerkmale. Um detaillierter zu untersuchen, wie die einzelnen Gruppen bei spezifischen Qualitätsattributen abschnitten, wurden die Werte zusätzlich getrennt nach Attributen betrachtet. Da die Verteilung der Fehlertypen im Datensatz nicht einheitlich war, wurde eine Min-Max Skalierung durchgeführt, um die Ergebnisse zu normalisieren. Abbildung 9.16 stellt die Resultate für ausgewählte Qualitätsattribute in Form einer Boxplot-Darstellung dar. [KVN<sup>+</sup>26]

Die Ergebnisse weisen darauf hin, dass SGAIRE im vorliegenden Anwendungsfall keine Nachteile gegenüber dem konventionellen UCMR-Ansatz bei den adressierten Qualitätsmerkmalen erkennen lässt. Vielmehr tragen die KI-basierten Empfehlungen insbesondere in den Dimensionen *Korrektheit*, *Konsistenz*, *Nachvollziehbarkeit* und *Eindeutigkeit* zu einer Qualitätssteigerung der IEC 62559 Use Cases und SGAM Modelle bei. In den Attributen *Erreichbarkeit*, *Relevanz* und *Vollständigkeit* sind ebenfalls höhere Mittelwerte erkennbar, die Verbesserungen fallen hier jedoch weniger stark aus. [KVN<sup>+</sup>26]

### 9.2.5 Einfluss der eingeschätzten Expertise

Die Abbildungen 9.13, 9.14 und 9.15 zeigen auf, dass Unterschiede zwischen den eingeschätzten Erfahrungsstufen bestehen. Vor diesem Hintergrund wurden zwei zentrale

Aspekte näher untersucht: Erstens, ob das Erfahrungsniveau einen messbaren Einfluss auf die Gesamtqualität ausübt, und zweitens, wie sich dieser Faktor im Kontext der Nutzung von SGAAIRE auswirkt.

### 9.2.5.1 Einfluss des Erfahrungsniveaus auf die Qualität

Zunächst ist quantitativ zu prüfen, ob ein statistischer Zusammenhang zwischen dem Erfahrungsniveau und der Qualitätssteigerung unabhängig von der Gruppenzuteilung besteht. Um zu untersuchen, ob sich die Gruppen „Anfänger“, „Fortgeschrittene“ und „Experten“ signifikant voneinander unterscheiden, wurde die Hypothese 2.4 (Nullhypothese,  $H_0$ ) sowie die entsprechende Alternativhypothese ( $H_1$ ) formuliert. [KVN<sup>+</sup>26]

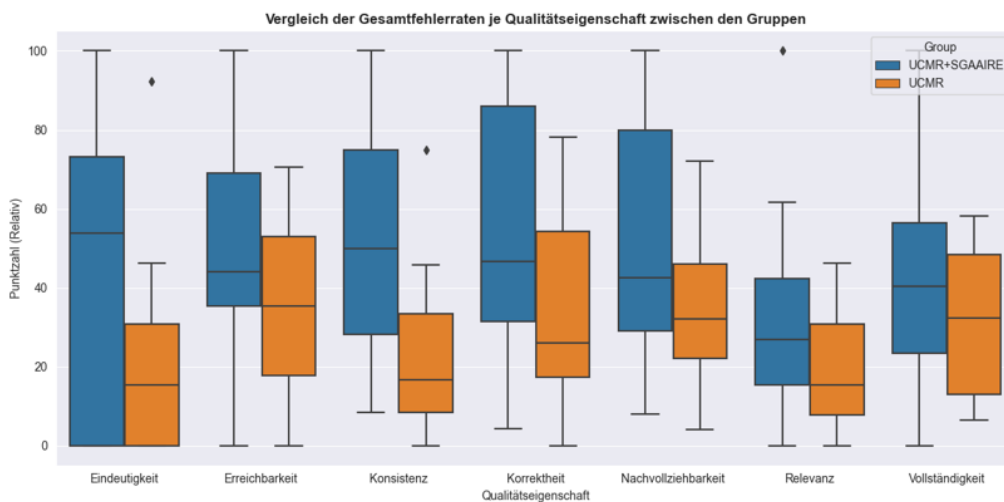
#### Subhypothese 2.4 - Nullhypothese

$H_0$ : Es existiert kein Unterschied in der Qualitätssteigerung zwischen den eingeschätzten Erfahrungsleveln.

#### Subhypothese 2.4 - Alternativhypothese

$H_1$ : Es existiert ein Unterschied in der Qualitätssteigerung zwischen den eingeschätzten Erfahrungsleveln.

Abbildung 9.15 zeigt die Anzahl der erreichten Qualitätsverbesserungspunkte in den einzelnen Erfahrungskategorien. Sowohl visuell als auch numerisch ist ein klarer Trend erkennbar: Die *Anfänger* trugen am wenigsten zur Qualitätsverbesserung bei ( $\mu_{\text{beg}} =$



Quelle: Übersetzt aus [KVN<sup>+</sup>26]

Abbildung 9.16: Vergleich von Kontroll- und Testgruppe hinsichtlich der Auswirkungen auf einzelne Qualitätsmerkmale (Min-Max Skalierung)

8.93,  $\sigma_{\text{beg}} = 4.94$ ), gefolgt von der Gruppe der *Fortgeschrittenen*, die ein breites Mittelfeld abbildet ( $\mu_{\text{adv}} = 25.12$ ,  $\sigma_{\text{adv}} = 12.27$ ). Die *Experten* leisteten den größten Beitrag zur Qualitätssteigerung ( $\mu_{\text{exp}} = 32.69$ ,  $\sigma_{\text{exp}} = 6.72$ ). [KVN<sup>+</sup>26]

Zur Überprüfung der Hypothese wurde ein *Kruskal-Wallis-Signifikanztest* eingesetzt, der Vergleiche über mehr als zwei Gruppen erlaubt und keine Normalverteilung der Daten voraussetzt. Der Test ergab eine Teststatistik von  $H = 15.11$  mit einem p-Wert von  $p = 0.0005$ , der deutlich unterhalb des Signifikanzniveaus von  $\alpha = 0.05$  liegt. Damit kann die Nullhypothese  $H_0$  verworfen werden. Die berechnete Effektstärke von  $\eta^2 = 0.56$  ist gemäß den Konventionen nach Cohen als groß einzuordnen. [KVN<sup>+</sup>26]

### 9.2.5.2 Einfluss des Erfahrungsniveaus auf die Qualität zwischen Kontroll- und Testgruppe

Die im vorherigen Abschnitt aufgeführten Ergebnisse beziehen sich auf den Gesamtdatensatz ohne die Betrachtung der Experimentalgruppen. Vor dem Hintergrund der Frage, welche Gruppen von einem KI-basierten System profitieren oder durch dieses benachteiligt sein könnten, wurde die Effizienz der Qualitätsverbesserung differenziert nach Erfahrungshintergrund untersucht. Hierfür wurde die Hypothese 2.5 (Nullhypothese,  $H_0$ ) sowie die entsprechende Alternativhypothese ( $H_1$ ) formuliert. [KVN<sup>+</sup>26]

#### Subhypothese 2.5 - Nullhypothese

$H_0$ : Das eingeschätzte Erfahrungslevel hat keinen Einfluss auf die Qualitätssteigerung, sowohl in der Gruppe UCMR mit SGAAIRE als auch in der Gruppe UCMR ohne SGAAIRE.

#### Subhypothese 2.5 - Alternativhypothese

$H_1$ : Das eingeschätzte Erfahrungslevel hat einen Einfluss auf die Qualitätssteigerung, sowohl in der Gruppe UCMR mit SGAAIRE als auch in der Gruppe UCMR ohne SGAAIRE.

Abbildung 9.15 zeigt die Anzahl der erreichten Qualitätsverbesserungspunkte für jedes Erfahrungsniveau, getrennt nach den Experimentalgruppen. Ein gleiches Muster wie in der vorangehenden Hypothese (vgl. Abschnitt 9.2.5) zeigt sich in beiden Gruppen. Mit zunehmendem Erfahrungsniveau nimmt auch die erzielte globale Qualitätsverbesserung zu. Auffällig ist jedoch, dass in der Interventionsgruppe eine deutliche Verteilungssverschiebung sichtbar wird, die darauf hindeutet, dass jede Erfahrungsgruppe von der Unterstützung durch das KI-System profitiert: [KVN<sup>+</sup>26]

- **Anfänger** profitieren am wenigsten von SGAAIRE, mit einer Verbesserung von etwa 5 % im Vergleich zum konventionellen UCMR-Ansatz ( $\mu_{\text{beg}}^{\text{SGAAIRE}} = 9.19$ ,  $\sigma_{\text{beg}}^{\text{SGAAIRE}} = 3.80$  vs.  $\mu_{\text{beg}}^{\text{UCMR}} = 8.75$ ,  $\sigma_{\text{beg}}^{\text{UCMR}} = 5.83$ ).
- **Fortgeschrittene** profitieren am stärksten von SGAAIRE, mit einem Anstieg von rund 95.09 % bei der Qualitätsverbesserung im Vergleich zum konventionellen UCMR-Ansatz ( $\mu_{\text{adv}}^{\text{SGAAIRE}} = 30.9$ ,  $\sigma_{\text{adv}}^{\text{SGAAIRE}} = 12.10$  vs.  $\mu_{\text{adv}}^{\text{UCMR}} = 15.85$ ,  $\sigma_{\text{adv}}^{\text{UCMR}} = 4.52$ ).

- **Experten** zeigen einen moderaten Nutzen von etwa 25.43 % durch SGAAIRE im Vergleich zum konventionellen UCMR-Ansatz ( $\mu_{\text{exp}}^{\text{SGAAIRE}} = 36.38$ ,  $\sigma_{\text{exp}}^{\text{SGAAIRE}} = 7.95$  vs.  $\mu_{\text{exp}}^{\text{UCMR}} = 29.00$ ,  $\sigma_{\text{exp}}^{\text{UCMR}} = 4.24$ ).

Zur Hypothesenprüfung wurde ein *Kruskal-Wallis-Signifikanztest* durchgeführt. Für die Interventionsgruppe ergab der Test eine Teststatistik von  $H = 8.14$  mit einem p-Wert von  $p = 0.0171$ , der unterhalb des Signifikanzniveaus von  $\alpha = 0.05$  liegt. Damit kann die Nullhypothese  $H_0$  verworfen werden. Die Effektstärke von  $\eta^2 = 0.56$  ist nach Cohen als groß einzuordnen. Für die Kontrollgruppe ergab der Kruskal-Wallis-Test eine Teststatistik von  $H = 6.67$  und einen p-Wert von  $p = 0.0356$ , ebenfalls unterhalb der Grenze von  $\alpha = 0.05$ . Somit kann auch hier die Nullhypothese  $H_0$  verworfen werden. Die zugehörige Effektstärke von  $\eta^2 = 0.46$  weist ebenfalls auf einen großen Effekt hin. [KVN<sup>+</sup>26]

### 9.2.6 Systembewertung der User Experience

Abschließend ist es relevant, den Einfluss des integrierten SGAAIRE auf die User Experience (UX) (dt. Benutzererlebnis) zu untersuchen. Hierbei sollte sich die UX nicht gegenüber der Basisversion des UCMR verschlechtern. Die Benutzerfreundlichkeit stellt eine Teilmenge des Benutzererlebnisses dar, welche quantitativ über den System Usability Scale (SUS)-Test erhoben werden kann [LS21]. In der abschließenden Umfrage wurde dieser SUS-Test integriert, um die Benutzerfreundlichkeit zu bewerten, sowie offene Fragen zum Benutzererlebnis, darunter: „Was hat Ihnen an dem System gefallen?“, „Was hat Ihnen an dem System missfallen?“ und „Haben Sie weitere Kommentare oder Verbesserungsvorschläge für das System?“. Im Rahmen der Auswertung des SUS-Tests wurden die folgenden Hypothesen zur Benutzerfreundlichkeit des Systems aufgestellt:

#### **Hypothese zur Benutzerfreundlichkeit - Nullhypothese**

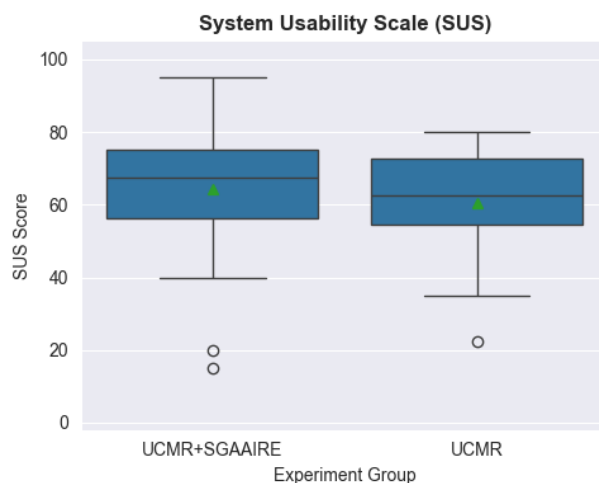
$H_0$ : Es besteht kein Unterschied in der wahrgenommenen Benutzerfreundlichkeit zwischen der Gruppe, die das UCMR mit SGAAIRE verwendet, und der Gruppe, die ausschließlich das UCMR-System verwendet.

#### **Hypothese zur Benutzerfreundlichkeit - Alternativhypothese**

$H_1$ : Die Gruppe, die das UCMR mit SGAAIRE verwendet, bewertet die Benutzerfreundlichkeit des Systems höher als die Gruppe, die ausschließlich das UCMR-System verwendet.

Tabelle 9.6: Statistische Kennzahlen der erreichten Punktzahlen im SUS-Test

	Count	Mean	Std	Min	25 %	50 %	75 %	Max
UCMR	16	60.47	15.79	22.5	54.38	62.5	72.5	80
UCMR + SGAAIRE	19	64.21	22.42	15	56.25	67.5	75	95



Quelle: Eigene Darstellung

Abbildung 9.17: Boxplot der erreichten Punktzahlen im SUS-Test

Dieser Abschnitt fokussiert sich nicht auf die Evaluierung der Basisversion des UCMR, sondern explizit auf die Version mit SGAAIRE. An der Erhebung über den SUS sowie die Fragen (vgl. Tabelle 9.7) haben die 35 Probanden teilgenommen (vgl. Abbildung 9.10), wobei 16 der Kontrollgruppe und 19 der Interventionsgruppe zugewiesen worden sind. Die Verteilungen der Punktzahlen können aus der Boxplot-Darstellung aus Abbildung 9.17 sowie aus Tabelle 9.6 entnommen werden. Hierbei zeigt sich, dass der Mittelwert als auch Median der Interventionsgruppe gegenüber der Kontrollgruppe höher ist, jedoch auch die Standardabweichung. Die Mittelwerte ( $\mu_{UCMR} = 60.47$  sowie  $\mu_{SGAAIRE} = 64.21$ ) lassen sich gemäß der Bewertungsskala nach [BKM09] im Mittel als „In Ordnung / OK“ einordnen. Dies deutet darauf hin, dass das System grundsätzlich als ausbaufähig, aber mit hoher Wahrscheinlichkeit als für die Inbetriebnahme geeignet betrachtet werden kann.

Um die Alternativhypothese zu überprüfen, wurde ein *einseitiger t-Test* durchgeführt, der ermittelt, ob die beobachteten Unterschiede in den Daten zufallsbedingt sein könnten [Stu08]. In der Auswertung von SUS-Tests ist der t-Test weit verbreitet, sofern die Daten annähernd normalverteilt sind [BKM08]. Dies ist im Rahmen der erhobenen Daten gegeben, weshalb der Einsatz eines parametrischen Tests hier gerechtfertigt ist. Zudem werden zwei unabhängige Gruppen miteinander verglichen, um signifikante Differenzen zwischen den Versuchsbedingungen zu evaluieren. Es wird erneut ein Fehler von  $\alpha = 0.05$  toleriert, was bedeutet, dass ein Ergebnis als statistisch signifikant gilt, wenn der p-Wert kleiner oder gleich ist. Der durchgeführte Test liefert jedoch einen p-Wert von  $p = 0.2896$ , was darauf hinweist, dass der beobachtete Effekt zufallsbedingt zustande gekommen ist. Daher wird die Alternativhypothese  $H_1$  verworfen.

Ebenfalls nur marginale Unterschiede lassen sich über die jeweiligen Boxplots zur Darstellung der Ergebnisse aus den zehn SUS-Fragen erkennen (vgl. Abbildung A.1, A.2 sowie Tabelle 9.7). Basierend auf der Annahme, dass die Punktzahlen normalverteilt sind und zwei unabhängige Gruppen betrachtet wurden, wurden für jede Frage *zweiseitige t-Tests* durchgeführt, um mögliche Unterschiede in den Verteilungen zu überprüfen. Bei keiner Frage konnte ein signifikanter Unterschied festgestellt werden, und es ist kein großer Unterschied zwischen den Gruppen erkennbar, weshalb davon ausgegangen werden muss, dass die beobachteten Mittelwerte zufällig zustande gekommen sind.

Tabelle 9.7: Vergleich der Mittelwerte, Standardabweichungen und p-Werte der beiden Gruppen

SUS-Frage	UCMR ( $\mu \pm \sigma$ )	UCMR + SGAAIRE ( $\mu \pm \sigma$ )	p-Wert
Ich denke, dass ich dieses System gerne häufiger nutzen würde.	2.25 $\pm$ 1.13	2.58 $\pm$ 1.17	0.40
Ich finde, dass das System unnötig komplex ist.	1.94 $\pm$ 1.00	1.58 $\pm$ 1.22	0.35
Ich finde, das System ist einfach zu benutzen.	2.75 $\pm$ 0.68	2.63 $\pm$ 0.96	0.68
Ich glaube, dass ich Unterstützung einer technischen Person benötigen würde, um das System nutzen zu können.	1.06 $\pm$ 1.34	1.47 $\pm$ 1.22	0.35
Ich finde, dass die verschiedenen Funktionen gut integriert sind.	2.50 $\pm$ 0.82	2.79 $\pm$ 0.79	0.29
Ich finde, das System hat zu viele Inkonsistenzen.	1.50 $\pm$ 1.03	1.26 $\pm$ 0.99	0.50
Ich kann mir vorstellen, dass die meisten Menschen schnell lernen, das System zu benutzen.	2.38 $\pm$ 1.15	2.79 $\pm$ 1.23	0.31
Ich finde, das System ist umständlich zu nutzen.	1.75 $\pm$ 1.18	1.84 $\pm$ 1.01	0.81
Ich fühle mich sicher im Umgang mit dem System.	2.38 $\pm$ 1.15	2.68 $\pm$ 1.16	0.44
Ich musste viele neue Dinge lernen, bevor ich das System benutzen konnte.	1.81 $\pm$ 1.17	1.63 $\pm$ 1.34	0.68

Positiv hervorgehoben wurde die Funktion in SGAAIRE, die den Nutzenden anzeigte, wo etwas fehlte, und dadurch eine gezielte Verbesserung ermöglichte. Ein Wunsch war es, visuelle Hinweise zu erhalten, um direkt zu sehen, an welcher Stelle fehlende Elemente ergänzt werden sollten, auch wenn die Umsetzung dieser Funktion als herausfordernd betrachtet wurde. Kritisch wurde der „schwebende SGAAIRE-Button“ wahrgenommen, die von einer Person als störend empfunden wurde. Außerdem gab es Unsicherheiten bezüglich der Bedeutung der Empfehlungen und Bewertungen des Systems. Nutzer wünschten sich mehr Klarheit darüber, wie diese Vorschläge interpretiert werden sollten und wie sie in die Arbeit integriert werden können. Insgesamt zeigt das Feedback, dass SGAAIRE gut genutzt wird, jedoch weiterhin Verbesserungspotential in der Benutzerführung und der Erklärung der Funktionen besteht.

### 9.3 Diskussion

In diesem Abschnitt werden die Ergebnisse der Dissertation kritisch reflektiert und in einen größeren wissenschaftlichen Kontext eingeordnet. Zunächst erfolgt die Analyse und Einordnung der empirischen Ergebnisse, insbesondere hinsichtlich der identifizierten Verbesserungen in den Qualitätsmerkmalen sowie deren Abgleich mit bestehenden wissenschaftlichen Arbeiten. Anschließend werden die Ergebnisse im Hinblick auf

die adressierte Forschungslücke und in dem Gesamtkontext der Dissertation betrachtet. Darüber hinaus werden die zentralen Limitationen der Arbeit erörtert, um eine fundierte Einschätzung der Aussagekraft sowie der Reichweite der empirischen Ergebnisse zu ermöglichen.

### 9.3.1 Einordnung der Studienergebnisse

Die Ergebnisse der Untersuchung verdeutlichen, dass der Einsatz KI-gestützter Systeme, repräsentiert durch SGAAIRE, im Durchschnitt zu einer höheren Qualität von IEC 62559 Use Cases und SGAM Modellen führt als die konventionelle Herangehensweise. So identifizierten die Teilnehmenden der Interventionsgruppe mit Unterstützung von SGAAIRE im Mittel 9,77 Fehler, während die Kontrollgruppe (nur mit UCMR) durchschnittlich 5,73 Fehler feststellte. Dies entspricht einer relativen Steigerung von 70,51 %. Darüber hinaus konnte eine generelle Qualitätsverbesserung von 75,59 % nachgewiesen werden. Besonders deutlich profitierten die Qualitätsmerkmale Konsistenz, Korrektheit, Nachvollziehbarkeit und Eindeutigkeit von der KI-Unterstützung. Auch Merkmale wie Zugänglichkeit, Relevanz und Vollständigkeit verbesserten sich, wenn auch in geringem Maße.

Die Literatur bestätigt diese Beobachtungen: [RDCC25] zeigen in einer umfassenden Übersicht, dass insbesondere Methoden des ML und der NLP positive Effekte auf die Merkmale Korrektheit, Nachvollziehbarkeit und Eindeutigkeit haben. Gleichzeitig weisen die Autoren darauf hin, dass insbesondere Konsistenz und Nachvollziehbarkeit trotz KI-Unterstützung weiterhin menschlicher Validierung bedürfen. Diese Einschätzung deckt sich mit den hier vorliegenden Ergebnissen. Zwar können KI-basierte Systeme wie SGAAIRE hochwertige, kontextbezogene Empfehlungen erzeugen, jedoch erfordert deren Nutzung eine Einordnung und Begründung, die vom Modellierungskontext abhängt [VLH<sup>+</sup>23]. Ein zentrales Risiko in diesem Zusammenhang ist der Automatisierungsbias, sprich die unkritische Übernahme von Systemempfehlungen. Um negative Effekte zu vermeiden, bleibt menschliches Urteilsvermögen essenziell [VLH<sup>+</sup>23].

Auffällig ist, dass besonders Fortgeschrittene mit 95,09 % den größten Mehrwert durch SGAAIRE erzielten. Interessanterweise entsprachen deren Verbesserungen in etwa denen von Experten ohne KI-Unterstützung. Dies deutet darauf hin, dass ein gewisses Maß an Vorerfahrung notwendig ist, um das Potenzial solcher Systeme effektiv ausschöpfen zu können. Gleichzeitig zeigt sich, dass auch Experten, entgegen den Erwartungen anderer Arbeiten [ICM<sup>+</sup>22], signifikant, wenn auch moderater (25,43 %), profitierten. Anfänger wiesen hingegen lediglich geringe Zugewinne von rund 5 % auf. Dies deutet darauf hin, dass grundlegendes Domänen- und Methodenwissen erforderlich ist, um KI-gestützte Empfehlungen adäquat zu bewerten und wirksam anzuwenden. Neben den Verbesserungen in der Modellqualität kann SGAAIRE darüber hinaus durch die Bereitstellung von Prozessunterstützungen Effizienzgewinne, beispielsweise in Form von Zeitersparnis ermöglichen. Eine vertiefte Untersuchung dieses Aspekts erfolgte im Rahmen der vorliegenden Arbeit jedoch nicht.

Ein weiterer Befund betrifft die Frage, ob das KI-System zu einer höheren Anzahl nicht-eingebauter (zusätzlicher) Fehler führt. Die Ergebnisse lassen keine eindeutige Tendenz erkennen, belegen jedoch, dass die Gesamtqualität nicht beeinträchtigt wurde. Eine unreflektierte Übernahme von Vorschlägen wurde im Fall von SGAAIRE durch den Einsatz von Konzepten aus dem Bereich XAI abgemildert [MST<sup>+</sup>24]. Konkret stellt das System sowohl Wahrscheinlichkeitswerte seiner Empfehlungen bereit, was Transparenz über die Entscheidungsgrundlage schafft, als auch globale Erklärungen zum Modellverhalten. Zusätzlich werden lokale Erklärungen auf Ebene einzelner Anwendungsfälle geliefert, sodass konkrete Empfehlungen im jeweiligen Kontext nachvollziehbar bleiben.

Die Ergebnisse legen ferner nahe, dass ein Lerneffekt bei den Nutzenden auftreten könnte. Eine belastbare Überprüfung dieses Effekts hätte jedoch wiederholte Durchläufe mit denselben Teilnehmenden erfordert, was aufgrund der Anonymisierungsrichtlinien im Studiendesign nicht realisierbar war. Besonders Fortgeschrittene, die ihre Interaktion mit dem System im Verlauf offenbar optimierten, könnten von einer solchen Wiederholung profitieren. Während Anfänger nur geringe Verbesserungen erzielten, deuten die Daten insgesamt darauf hin, dass Erfahrung und Anpassung notwendig sind, um das volle Potenzial der KI-Unterstützung auszuschöpfen. Zu betonen ist, dass die hier dargestellten Ergebnisse auf dem spezifischen Anwendungsfall „*Verwendung dynamischer Stromtarife für EV-Ladestationen*“ beruhen. Weitere Untersuchungen in anderen Anwendungsfällen sowie Langzeitstudien sind erforderlich, um die Generalisierbarkeit und Nachhaltigkeit der Effekte zu überprüfen.

### 9.3.2 Weiterführende Einordnung

Im Rahmen dieser Dissertation konnte erfolgreich SGAAIRE als Prototyp eines intelligenten Systems zur Unterstützung bei der Erstellung von IEC 62559-Use Cases und SGAM-Modellen entwickelt werden. Empirische Ergebnisse, wie im vorherigen Abschnitt dargestellt, belegen, dass die Nutzung von SGAAIRE zu einer Verbesserung der Qualität von IEC 62559 Use Cases und SGAM Modellen beitragen kann.

Ausgehend von der in Abschnitt 1.3 identifizierten Forschungslücke konnte mit SGAAIRE unter Nutzung KI-Verfahren die Qualitätsbewertung von Use Cases und SGAM-Modellen über die Konsistenzprüfung hinaus erweitert werden. Dabei wurden zusätzlich die Qualitätsdimensionen Korrektheit, Nachvollziehbarkeit und Eindeutigkeit adressiert, während Aspekte wie Zugänglichkeit, Relevanz und Vollständigkeit in geringerem Maße berücksichtigt wurden. Das System wurde im Rahmen zweier Demonstrationsbeispiele sowie durch Durchführung einer RCT mit unterschiedlichsten Interessensgruppen (Fach-, Methoden- sowie Standardisierungsexperten) mit diversen Disziplinen und persönlichen Hintergründen erprobt. Durch die Integration im UCMR sowie die offene API kann SGAAIRE auch weiterführend im praktischen Anforderungsmanagement in Energieprojekten eingesetzt werden. Ein Aspekt dieser Arbeit lag auch in der Untersuchung von Qualitätsmerkmalen und -Metriken für die Bewertung von IEC 62559 Use Cases und SGAM Modelle. Diese können sowohl direkt zur Qualitätssicherung- und Be-

wertung bei der Erstellung neuer IEC 62559 Use Cases und SGAM Modellen als auch indirekt über unterstützende Systeme wie SGAAIRE genutzt werden.

Die Ergebnisse verdeutlichen, dass insbesondere Sprachmodelle der dritten Generation, wie LMs (z. B. BERT, RoBERTa oder DeBERTa), in klassifikationsorientierten Aufgaben weiterhin eine hohe Leistungsfähigkeit aufweisen. Demgegenüber zeigen LLMs, wie GPT-4o(-mini) oder LLaMA, ihre Stärken vor allem in generativen Szenarien, stellen jedoch nicht in allen betrachteten Aufgabenfeldern die leistungsstärkste Lösung dar. Damit wird deutlich, dass auch zum Zeitpunkt dieser Dissertation eine differenzierte Betrachtung unterschiedlicher Modellgenerationen notwendig ist. Ähnliche Befunde berichten [BM24], die zeigen, dass feinabgestimmte kleinere LMs in Klassifikationsaufgaben bessere Ergebnisse erzielen als Zero-Shot-Ansätze mit LLMs. Während deren Untersuchung den Fokus auf Zero-Shot-Einstellungen legte, berücksichtigt diese Dissertation zusätzlich Few-Shot-Anpassungen, wodurch ein differenzierteres Bild entsteht. Dies impliziert, dass LLMs zum gegenwärtigen Zeitpunkt nicht als universelle Lösung für sämtliche Aufgaben im Bereich der NLP-basierten Künstlichen Intelligenz angesehen werden können, sondern in Abhängigkeit von der jeweiligen Problemstellung gezielt im Vergleich zu etablierten Verfahren zu bewerten sind.

### 9.3.3 Limitierungen

Als empirische Forschungsarbeit ist auch diese Dissertation mit bestimmten Limitationen behaftet. Diese betreffen sowohl die (multi-)methodische Herangehensweise als auch die inhaltliche Fokussierung der Arbeit. Eine transparente Darstellung dieser Einschränkungen ist notwendig, um die Aussagekraft der Ergebnisse angemessen einzuordnen und potenzielle Verzerrungen oder Grenzen in der Generalisierbarkeit zu reflektieren. Des Weiteren ermöglicht die kritische Diskussion der Limitationen die Identifikation von Ansatzpunkten für zukünftige Forschungsarbeiten, die auf den hier gewonnenen Erkenntnissen aufbauen und diese vertiefen oder erweitern können.

#### Anforderungs- und Kontextanalyse

In der Anforderungs- und Kontextanalyse nahmen Fach- und Methodenexperten im Rahmen einer Umfrage teil. Eine Limitation ist die geringe Zahl qualifizierter Teilnehmer, da von  $n = 48$  nur  $n = 21$  für den SGAM-Teil und  $n = 20$  für den IEC 62559 Use Case-Teil berechtigt waren. Diese Einschränkung war notwendig, um eine fundierte Repräsentation der tatsächlichen Methodenanwendenden zu gewährleisten und so hochwertigere Daten zu erhalten. Dies liegt vermutlich daran, dass die IEC 62559 Use Case-Methodik und das SGAM Framework zwar im Anforderungsmanagement energiewirtschaftlicher Projekte anerkannt sind, ihre konkrete Anwendung jedoch nur von einem kleinen Teil der Projektbeteiligten erfolgt. Nichtsdestotrotz konnten aus den qualitativen Daten wichtige Erkenntnisse gewonnen werden, und es ließ sich ein klares Bild der Anwendenden erfassen, was die Signifikanztests bei der Anforderungsanalyse aufzeigten.

Die aus der Umfrage systematisch erhobenen Anforderungen wurden anschließend mittels einer NWA transparent priorisiert und selektiert. Dieses Vorgehen trägt wesentlich zur Qualität, Relevanz und Nachvollziehbarkeit der identifizierten Anforderungen bei. Gleichwohl verbleibt eine theoretische Limitation: Die Gewichtung innerhalb der NWA kann Subjektivität nicht vollständig ausschließen [Wit79]. Auch die Festlegung des Schwellenwertes für die Anforderungsselektion basiert trotz transparenter Begründung auf einer forschungspraktischen Entscheidung und impliziert damit einen weiteren subjektiven Einflussfaktor. Dennoch stärkt die Kombination aus Struktur, Transparenz und interdisziplinärer Expertise die Nachvollziehbarkeit der Anforderungsanalyse.

## SGAAIRE

Die Entwicklung sowie die Erprobung von SGAAIRE konzentrierten sich auf den Prozess zur anwendungsfallgetriebenen Erstellung von SGAM-Modellen. Andere Erstellungsansätze wurden dabei nicht berücksichtigt, was jedoch nicht ausschließt, dass auch sie von SGAAIRE profitieren können. Laut der Kontexterhebung stellt dies zwar die in der Praxis gängigste Herangehensweise dar (vgl. [GDU17, FSBU20]), dennoch können alternative Herangehensweisen zu abweichenden Ergebnissen führen.

Auf technischer Ebene profitieren insbesondere LMs von möglichst großen und qualitativ hochwertigen Datenmengen. Da jedoch nur eine begrenzte Zahl an öffentlich verfügbaren IEC 62559 Use Cases und SGAM-Modellen existiert, könnte eine Erweiterung des Datenbestands die Qualität der KI-Modelle weiter steigern. Durch zusätzliche Einbeziehung weiterer Datenquellen, wie z. B. Übersichtsstandards und Referenzarbeiten (u. a. IEC Standards Map [IEC22] oder das ENTSO-E Harmonised Energy Market Role Model [ENT22]), sowie durch den Einsatz von Data Augmentation-Methoden konnte ein heterogener Datensatz realisiert werden.

Gleichzeitig ist zu berücksichtigen, dass das KI-System weiterhin False Positives und False Negatives erzeugen kann. Dies hängt sowohl von der Qualität der zugrunde liegenden IEC 62559 Use Cases und SGAM-Modelle ab als auch von der inhärent probabilistischen Natur von KI-Modellen, die prinzipiell keine hundertprozentige Korrektheit erreichen können. Allerdings ist dies im gegebenen Kontext nicht zwingend erforderlich. Durch die Integration lokaler und globaler Erklärbarkeitsansätze konnte dieser Aspekt zusätzlich adressiert werden.

## Randomisierte kontrollierte Studie

Trotz der Möglichkeit zur orts- und zeitunabhängigen Teilnahme am Online-Experiment war die effektive Stichprobengröße eingeschränkt. Von den ursprünglich  $n = 75$  Teilnehmenden konnten nach Ausschluss nicht berechtigter Personen, ungleichmäßiger Verteilung in den Experimentalgruppen sowie Dropouts lediglich  $n = 14$  in der Interventionsgruppe und  $n = 13$  in der Kontrollgruppe in die Auswertung einbezogen werden. Diese Fallzahlen erscheinen zunächst gering. Gleichwohl war die Stichprobengröße aus-

reichend, um statistisch signifikante Unterschiede nachzuweisen, wie die durchgeführten Hypothesentests belegen. Dies hängt möglicherweise damit zusammen, dass die Modellierung von SGAM-Modellen und IEC 62559 Use Cases zwar einen zentralen Bestandteil des Smart Grid Requirements Engineering bildet [USD<sup>+</sup>13], jedoch bislang weder fester Bestandteil der ingenieur- noch der informatikbezogenen Grundausbildung ist. Zudem dürfte die vergleichsweise lange Bearbeitungsdauer von etwa 60 Minuten, in Verbindung mit dem Fehlen extrinsischer Anreizsysteme (z. B. monetäre oder symbolische Belohnungen), zur erhöhten Abbruchquote beigetragen haben. Eine größere Stichprobe hätte insbesondere im Hinblick auf explorative Korrelationsanalysen zu einer höheren statistischen Robustheit und Verallgemeinerbarkeit der Ergebnisse beigetragen.

Ein weiterer Aspekt betrifft die Auswahl des Anwendungsfalls. Im Rahmen der Evaluation wurde ausschließlich der Anwendungsfall „Verwendung von dynamischen Stromtarifen für EV-Ladestationen“ untersucht, der aufgrund seiner potenziell hohen Zugänglichkeit und Verständlichkeit ausgewählt wurde. Andere Anwendungsfälle mit höherer Komplexität oder geringerem Alltagsbezug könnten jedoch andere Herausforderungen mit sich bringen und abhängig von der Vorerfahrung der Teilnehmenden zu abweichenden oder schwächeren Ergebnissen führen. Solche Unterschiede lassen sich unter anderem durch die jeweilige Erstellungsweise sowie die Struktur der zugrunde liegenden IEC 62559 Use Cases und SGAM Modelle erklären. Um die Aussagekraft der gewonnenen Erkenntnisse zu stärken, sollten künftige Studien weitere Anwendungsfälle untersuchen und im Rahmen praxisnaher Projekte als Langzeituntersuchungen durchgeführt werden.

## 9.4 Beiträge aus der Arbeit

Im Rahmen dieser Dissertation wurden Beiträge zur Weiterentwicklung sowie Nutzung von Methoden und Werkzeugen im Bereich des Smart Grid Requirements Engineering (in Hinblick der IEC 62559 Use Case-Methodik sowie dem SGAM Framework) geleistet. Diese umfassen sowohl wissenschaftliche als auch technische Lösungen, die auf die Verbesserung der Qualität von Use Case-Beschreibungen und SGAM Modellen abzielen und um Interoperabilitätsbemühungen zu verbessern. Die Kernergebnisse dieser Arbeit tragen dazu bei, den Stand der Forschung in der Energieinformatik zu erweitern und praktische Lösungen für Herausforderungen in der Modellierung und Qualitätssicherung anzubieten. Die folgenden Abschnitte beleuchten sowohl wissenschaftliche als auch technische Beiträge.

### 9.4.1 Wissenschaftliche Beiträge

Im Rahmen der Dissertation wurden einzelne Teilergebnisse publiziert und in den wissenschaftlichen Fachkreisen insbesondere in der Energieinformatik sowie Information System Research zur Diskussion gestellt. Dieser Abschnitt führt die wissenschaftlichen Publikationen mit Angaben zu den beteiligten Autoren, den Titeln der Beiträge, dem je-

weiligen Konferenz- oder Publikationsorgan sowie dem Monat/das Jahr der Präsentation auf. Es ergaben sich aus der Dissertation die folgenden (direkten) Publikationen:

- **R. Kuchenbuch, J. Schütz und J. Sauer**, „**Quality properties of IEC 62559 use cases and SGAM models**“, **12th DACH+ Energy Informatics, October 2023** (vgl. [KSS23]): Die Ergebnisse zur Untersuchung relevanter Qualitätsmerkmale von IEC 62559 Use Case-Beschreibungen und SGAM Modellen wurden im Rahmen dieses Papers veröffentlicht. Es wurden Kriterien identifiziert und strukturiert, die für die Bewertung und Optimierung der IEC 62559 Use Cases sowie SGAM Modelle eingesetzt werden können.
- **R. Kuchenbuch, P. Nale, L. Niemann, O. Werth und J. Sauer**, „**Practical Application of Energy-related Use Case Methodology and SGAM Framework: A Study of User Perspectives and Challenges**“, **13th DACH+ Energy Informatics, October 2024** (vgl. [KNN<sup>+</sup>24]): Der aktuelle Stand der Anwendung der IEC 62559 Use Case-Methodik und des SGAM Frameworks sowie die bestehenden Herausforderungen und typischen Arbeitsweisen in der Praxis wurden in diesem Paper wissenschaftlich und praxisorientiert evaluiert und veröffentlicht.
- **R. Kuchenbuch, L. Niemann, J. Schütz und J. Sauer**, „**Towards Intelligent Systems to Improve IEC 62559 Use Cases and Smart Grid Architecture Models Quality**“, **CAiSE 2024 Forum, June 2024**, (vgl. [KNSS24]): Die Ergebnisse zur Anforderungsanalyse und High-Level-Konzeption von KI-basierten Systemen zur Unterstützung des Modellierungs- und Qualitätssicherungsprozesses wurden in dieser Publikation vorgestellt. Funktionale und nicht-funktionale Anforderungen sowie deren Bedeutung für die Entwicklung solcher Systeme wurden analysiert und beschrieben.
- **R. Kuchenbuch, J. Sauer und S. Lehnhoff**, „**Smart Grid Assistive AI in Requirement Engineering: Improving the Modeling of Use Cases and Architecture Models with LLMs**“, **ACM e-Energy '25, June 2025** (vgl. [KLS25]): In dieser Publikation wurden die Ergebnisse zur detaillierten Umsetzung des SGAAIRE-Systems veröffentlicht. Das Paper zeigt die Demonstration des Systems anhand eines praxisnahen Fallbeispiels und hebt hervor, dass die erhobenen funktionalen Anforderungen mit bestehenden Daten und Methodiken realisiert werden konnten. Zudem wurde die *Effektivität* des Systems bei der Erkennung von integrierten Qualitätsdefiziten nachgewiesen.
- **R. Kuchenbuch, D. Vereno, L. Niemann, O. Werth, S. Lehnhoff und J. Sauer**, „**The Impact of Artificial Intelligence on Quality in Smart Grid Requirements Engineering**“, **IEEE Access, Oktober 2025** (vgl. [KVN<sup>+</sup>26]): In dieser Veröffentlichung werden Evaluationsexperimente des SGAAIRE-Systems vorgestellt. Der Fokus liegt auf dessen Einsatz durch Methoden- und Fachexperten, der Wahrnehmung seitens der Anwender sowie auf der Messung und Bewertung der Qualitätssteigerungen gegenüber der konventionellen Arbeitsweise.

Neben den oben genannten direkten Beiträgen zur Dissertation wurde an weiteren wissenschaftlichen Arbeiten mitgewirkt, die einen **indirekten Bezug** zu den Forschungsergebnissen aufweisen. Die nachstehenden peer-reviewten Publikationen werden im Kontext dieser Dissertation eingeordnet:

- **R. Kuchenbuch, M. Azamat und J. Schütz**, „**Optimizing the Life-Span of (Smart) Transformers: A Review on Smart Services**“, **27th International Conference on Electricity Distribution, June 2023** (vgl. [KAS23]): Diese Arbeit dieser Publikation befasste sich mit der Optimierung von intelligenten Transformatoren mittels Smart Services. Im Rahmen dieser Dissertation wurden die veröffentlichten Ergebnisse als Demonstrationsbeispiel für SGAAIRE verwendet.
- **T. Strasser, E. Widl, R. Kuchenbuch, L. Lázaro-Elorriaga, B. Laraudogoitia, M. Ginocchi, T. Penthong, et al.** „**Towards Interoperability Testing of Smart Energy Systems – An Overview and Discussion of Possibilities**“, **14th Mediterranean Conference on Power Generation Transmission, Distribution and Energy Conversion, November 2024** (vgl. [SWK<sup>+</sup>24]): Dieser Beitrag befasst sich mit der Analyse bestehender Herangehensweisen, Konzepte, Methoden sowie Tools, die im Rahmen von Interoperabilitätstests eingesetzt werden. Die IEC 62559 Use Case-Methodik als auch das SGAM Framework werden ebenfalls im Rahmen von Interoperabilitätstests eingesetzt insb. in der Beschreibung von Technical Frameworks gemäß der IES Testing Methodology.
- **J. van Genuchten, J. Hurley, R. Kuchenbuch, L. Guitart und G. Lipari**, „**EMINENT, a Method for Assessing Interoperability Maturity**“, **28th International Conference on Electricity Distribution, June 2024** (vgl. [VGHK<sup>+</sup>25]): Diese Publikation befasst sich mit der Entwicklung eines Reifegradmodells für das Bewerten von Interoperabilitätsprozessen von energiebezogenen Systemen. Die IEC 62559 Use Case-Methodik und das SGAM Framework stellen etablierte Ansätze im Bereich des Smart Grids Requirements Engineering dar. Ihre Anwendung in Projekten trägt dazu bei, die Interoperabilitätsbestrebungen zu erhöhen. Zudem wurde das SGAM als Rahmenwerk zur Betrachtung der Interoperabilitäts Herausforderungen in EMINENT zur Hand gezogen.
- **D. Vereno, C. Neureiter, S. Eschlberger, M. Millaku, R. Kuchenbuch und M. Us-lar**, „**SGAM Toolbox Revisited: A Standards-Based Domain-Specific Modeling Language and Toolset**“, **IEEE Access, Juli 2025** (vgl. [VNE<sup>+</sup>25]): Diese Publikation stellt eine aktualisierte Version der SGAM Toolbox vor, die u.a. eine anwendungsfallgetriebene Modellierungsherangehensweise über die IEC 62559 Use Case-Methodik integriert. Zudem wird die Einbindung KI-basierter Unterstützung, insbesondere durch LLMs, vorgeschlagen, um das Verständnis der verschiedenen Stakeholder zu verbessern. Zudem leistet die Arbeit einen Beitrag zur Formalisierung der Darstellung von SGAM Modellen.
- **J. Van Genuchten, L. Guitart, C. Ayon Mac Gregor, R. Kuchenbuch und T. I. Strasser**, „**Redefining the concept of interoperability: A multidimensional**

**approach for the energy sector“**, Open Research Europe, August 2025 (vgl. [VGA<sup>+</sup>25]): Die Publikation aus dem int:net-Projekt analysiert unterschiedliche Interoperabilitätsdefinitionen und ordnet sie lexikal, temporal sowie im Kontext des SGAM ein. Besonders betont wird die mögliche Erweiterung um eine sechste Interoperabilitätsebene („Framework Layer“), welche Governance-Aspekte wie Standardisierung, Fördergeber und Technologiebereitsteller umfasst. Die Dissertation kann daran anknüpfen, diese Anforderungen aufgreifen und so die Anforderungslage weiter adressieren sowie für weiterführendes Training nutzbar machen.

- **T. I. Strasser, E. Widl, C. Ayon Mac Gregor, M. Ginocchi und R. Kuchenbuch**, „Mapping and Evolving Interoperability Testing in European Energy Systems: The int:net Perspective“, IEEE PES ISGT EUROPE 2025, Oktober 2025 (vgl. [SWG<sup>+</sup>25]): Aufbauend auf der in [SWK<sup>+</sup>24] dargestellten Studie wurde eine Analyse bestehender Interoperabilitätseinrichtungen durchgeführt. Dabei kann SGAAIRE als intelligente Bewertungs- und Assistenzkomponente innerhalb von Interoperabilitätsworkflows dienen und in Testszenarien auf Basis von IEC 62559 Use Cases und SGAM Modellen als Entscheidungs- und Hinweisgeber eingesetzt werden.

#### 9.4.2 Technische Beiträge

Neben den wissenschaftlichen Beiträgen der Dissertation sind auch technische Beiträge entstanden. Diese umfassen das KI-Artefakt „Smart Grid Assisting Artificial Intelligence in Requirements Engineering“, die Entwicklung eines Datensatzes bestehender IEC 62559 Use Cases und SGAM Modelle sowie den IEC 62559 Use Case-Serializer, die im Folgenden beschrieben werden.

##### KI-Artefakt „SGAAIRE“

Das KI-Artefakt „SGAAIRE“ stellt das primäre Element dieser Dissertation dar und bildet den zentralen Untersuchungsgegenstand der zugrunde liegenden Forschungsfragen. Im Rahmen der vorliegenden Arbeit wurde es konzipiert, implementiert sowie im Kontext der Forschungsfragen demonstriert und evaluiert. Die in Abschnitt 5.3 durch Fach- und Methodenexperten erhobenen funktionalen Anforderungen wurden systematisch im Artefakt umgesetzt. Tabelle 9.8 gibt einen Überblick über die jeweiligen Erfüllungsgrade der funktionalen Anforderungen. In der Tabelle sind zudem die Abweichungen zum Erfüllungsgrad beschrieben.

SGAAIRE unterliegt auch den in Abschnitt 5.2 beschriebenen nicht-funktionalen Anforderungen. Tabelle 9.9 gibt einen Überblick über den Erfüllungsgrad dieser Anforderungen und zeigt auf, wie die NFAs erfüllt bzw. gemessen wurden. Die Nutzung des KI-Artefakts „SGAAIRE“ ist sowohl direkt über die Hypertext Transfer Protocol (HTTP)-API<sup>5</sup> als auch indirekt über die OFFIS UCMR-Instanz<sup>6</sup> möglich. Die HTTP-API er-

<sup>5</sup> Öffentliche SGAAIRE API: [https://sgaaire.offis.de/api\\_SGAAIRE/](https://sgaaire.offis.de/api_SGAAIRE/)

<sup>6</sup> OFFIS UCMR Instanz (zugriffsbeschränkt, auf Anfrage verfügbar): <https://sgaaire.offis.de/>

Tabelle 9.8: Erfüllungsgrade der funktionalen Anforderungen

ID	Funktionale Anforderung	Erfüllungsgrad
RQ_UC.1	Das System soll vorschlagen, Elemente hinzuzufügen, zu entfernen oder zu ändern (z. B. Akteure, Informationsaustausch, Anforderungen, KPI, Schlüsselwörter usw.)  <i>Die Anforderung wird durch die Generierung von SGAM Modellen basierend auf bestehenden Informationen umgesetzt. Fehlende Daten werden durch das KI-Modell „M5: SGAM Proposer“ ergänzt und auf das IEC 62559 Use Case-Template übertragen. Dies betrifft jedoch nur Elemente der SGAM-Interoperabilitätsebenen. Empfehlungen für Akteure, Informationsobjekte und deren Kommunikation können generiert werden, jedoch nicht für Freitextfelder oder außerhalb des SGAM liegende Felder.</i>	(✓)
RQ_UC.2	Das System soll abweichende und nicht standardisierte Begriffe in Freitextfeldern erkennen	✓
RQ_UC.4	Das System sollte Anwendungsfälle in SGAM-Zonen und -Domänen klassifizieren	✓
RQ_UC.7	Das System soll auf Inkonsistenzen bei Akteuren, Informationsaustausch und Anforderungen hinweisen, die nicht zum Anwendungsfalltyp passen	✓
RQ_UC.9	Das System soll Akteure in SGAM-Zonen und -Domänen klassifizieren	✓
RQ_UC.11	Das System soll Scenario Services in der Step-by-Step Analysis vorschlagen	✓
RQ_SGAM.1	Das System soll Arbeitsvorschläge für SGAM Modelle basierend auf IEC 62559 Use Cases generieren	✓
RQ_SGAM.2	Das System soll SGAM Modelle auf Vollständigkeit, Redundanz, Einzigartigkeit, Korrektheit oder Verständlichkeit prüfen	✓
RQ_SGAM.3	Das System soll Inkonsistenzen innerhalb von SGAM Modellen erkennen und die Wahrscheinlichkeit anzeigen  <i>SGAAIRE erkennt Inkonsistenzen bei bestimmten Fehlerquellen. Neben kausalen Aspekten wie fehlenden Verbindungen (die keine KI benötigen) werden insbesondere die Klassifikation der Domänen und Zonen sowie der Abstraktionsgrad des SGAM Modells überprüft. Das „M5: SGAM Proposer“-Modell adressiert die Semantik. Eine Wahrscheinlichkeitsangabe ist jedoch nur bedingt über den Matching-Score möglich.</i>	(✓)

laubt es, einzelne Funktionen von SGAAIRE in der Entwicklung oder Nutzung von Drittanbieter-Tools und -Programmen zu integrieren. Darüber hinaus bietet sie Zugang zur direkten Inferenz der in SGAAIRE enthaltenen KI-Modelle. Die OFFIS UCMR-Instanz umfasst eine vollständige Integration von SGAAIRE und kann entsprechend der in dieser Dissertation beschriebenen Methodik für die direkte Bearbeitung und Analyse verwendet werden.

Tabelle 9.9: Erfüllungsgrade der nicht-funktionalen Anforderungen

ID	Nicht-funktionale Anforderung	Erfüllungsgrad
REQ_NFA.1	Qualitätsverbesserung durch KI  <i>Diese NFA steht in Verbindung mit Forschungsfrage 2 und wurde im Rahmen der Demonstration (vgl. Abschnitt 9.1) sowie der Evaluation (vgl. Abschnitt 9.2) behandelt. Die Bestätigung von Hypothese 2 führt zu dem Schluss, dass diese NFA als erfüllt angesehen werden kann.</i>	✓
REQ_NFA.2	Lose-gekoppelte Integration in Use Case Management Repository  <i>Die Architektur von SGAAIRE wurde gezielt so konzipiert, dass sie eine toolunabhängige Integration über eine standardisierte HTTP-API ermöglicht. Im Kontext des OFFIS UCMR erfolgt die Implementierung von SGAAIRE als modularer und gekapselter Service, wie in Abschnitt 8.1 detailliert beschrieben. Auf Anwendungsebene ist eine kontextsensitive Aktivierung oder Deaktivierung vorgesehen, die eine flexible Anpassung an individuelle Nutzungsszenarien erlaubt. Diese Funktionalität wird im Rahmen der Evaluation genutzt, um den benutzerspezifischen Zugriff auf SGAAIRE gemäß des Experiments zu regulieren.</i>	✓
REQ_NFA.3	Standalone Software System  <i>SGAAIRE wurde als selbstständiges System mithilfe einer extern zugänglichen HTTP-API entwickelt. Das System wurde als Docker-Image in einer virtualisierten Umgebung bereitgestellt werden und kann somit über Docker-Container instanziiert werden.</i>	✓
REQ_NFA.4	Standardisierte Serialisierung  <i>SGAAIRE sieht die Verwendung der offiziellen IEC 62559 Use Case XML-Serialisierung als auch das DISCERN SGAM-Format vor. Das DISCERN SGAM-Format stellt zum Zeitpunkt der Dissertation der de-facto Standard für SGAM Modelle dar.</i>	✓
REQ_NFA.5	Benutzerzentrierte Entscheidungsfreiheit  <i>Durch den Einsatz von kognitiven Zwangsfunktionen, wie der passiven Anzeige und der verzögerten Bereitstellung von Vorschlägen, wird eine unkritische Akzeptanz KI-generierter Entscheidungen verringert und die Reflexion der Benutzer gefördert [BMG21]. Insgesamt fördert das System kognitive Vielfalt, indem es Überlegungen und Neubewertungen in den Entscheidungsprozess einbindet, was durch Tests mit Zielgruppen validiert wurde [BNK<sup>+</sup>19]. Indirekt bestätigt wird die Erfüllung im Rahmen der Evaluierung über den SUS Test und den Freitext-Kommentaren.</i>	✓

## Datensatz von IEC 62559 Use Cases und SGAM Modellen

Die Erstellung des Datensatzes auf Basis bestehender IEC 62559 Use Cases und SGAM Modelle war Voraussetzung für das Training der KI-Modelle. Die methodische Herleitung des Datensatzes wurde bereits in Abschnitt 8.2 ausführlich beschrieben und geht detailliert auf die verwendeten Daten ein.

Die Datenpipeline zur Entwicklung des Datensatzes umfasst mehrere Schritte, wobei die Daten aus Gründen der Transparenz für jeden einzelnen Schritt bereitgestellt werden. Der finale Datensatz setzt sich aus den Rohdaten, erhoben aus der Literatur sowie aus bestehenden IEC 62559 Use Cases und SGAM Modellen, und den mittels Data Augmentation erzeugten Daten zusammen. Tabelle 9.10 zeigt die Anzahl der im Rahmen der Pipeline berücksichtigten Rohdaten auf, wobei ausschließlich öffentlich zugängliche Daten verwendet wurden. Tabelle 9.11 gibt einen Überblick über die be-

reinigten Daten, die nach Durchlaufen der Datenpipeline für das Modelltraining genutzt wurden. Um die Datengrundlage zu erweitern, kamen die in Abschnitt 8.2 beschriebenen Data Augmentation-Techniken zur Anwendung. Der Datensatz kann als Grundlage für technische sowie wissenschaftliche Projekte genutzt werden und ist über das separate GitLab Projekt<sup>7</sup> öffentlich zugänglich.

Tabelle 9.10: Im Datensatz berücksichtigte IEC 62559 Use Cases sowie SGAM Modelle (Quelldateien)

Art	Anzahl
<b>IEC 62559 Use Cases</b>	97
<i>davon PDF-Dateien</i>	60
<i>davon DOCX-Dateien</i>	11
<i>davon XML-Dateien</i>	26
<b>SGAM Modelle</b>	79
<i>davon PDF-Dateien</i>	79
<b>Literaturbezogene Daten</b>	7

Tabelle 9.11: Datensatz für das Trainieren der KI-Modelle

Model	Rohdaten	Data Augmentation	Datensatz (ingesamt)
<b>M1: SGAM Domänen</b>	664	605	1,269
<b>M2: SGAM Zonen</b>	664	411	1,075
<b>M3: SGAM Layer</b>	2,912	0	2,912
<b>M4: Service Step</b>	691	245	936
<b>M5: LoA Regressor</b>	2,433	0	2,433
<b>M6: Element Proposer</b>	1,624	0	1,624

## IEC 62559 Use Case Serializer

Der IEC 62559 Use Case Serializer wurde zur halbautomatischen Konvertierung von Word- und PDF-Anwendungsfällen nach dem IEC 62559-2 Use Case-Template (gemäß [IEC15]) in die IEC 62559-3 Use Case XML-Serialisierung (gemäß [IEC18]) entwickelt. Dieses Tool wurde im Rahmen der Erstellung des KI-Datensatzes entwickelt und ist ein integraler Bestandteil der Datenpipeline (vgl. Abschnitt 8.2). Der Serializer ist eine zustandslose Command-Line Interface (CLI)-Anwendung, die in Java 11 unter Verwendung des Spring Frameworks implementiert wurde. Die Serialisierung kann mithilfe einer Konfigurationsdatei angepasst werden, sodass individuelle Templateversionen berücksichtigt werden können. Der Quellcode sowie die Bedienungsanleitung für den IEC 62559 Use Case-Serializer sind über das SGAAIRE Git-Repository<sup>8</sup> unter dem Ordner Misc zugänglich.

<sup>7</sup> GitLab: [https://gitlab.offis.de/official/fbe-p-ucmr/sgaaire\\_dataset](https://gitlab.offis.de/official/fbe-p-ucmr/sgaaire_dataset)

<sup>8</sup> GitLab: <https://gitlab.offis.de/official/fbe-p-ucmr/SGAAIRE>

## 10 Zusammenfassung und Schluss

Dieses Kapitel fasst die wesentlichen Ergebnisse der Dissertation im Kontext des Forschungsgebiets zusammen. Dabei werden die im Rahmen der Arbeit gewonnenen Erkenntnisse zur Beantwortung der Forschungsfragen herangezogen. Anschließend wird im Ausblick der weitere Forschungsbedarf aufgezeigt.

### 10.1 Fazit

Im Rahmen dieser Dissertation wurden drei zentrale Zielsetzungen verfolgt: 1) die Konzeption und Implementierung eines Systems zur Unterstützung bei der Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen, 2) der Nachweis einer Qualitätssteigerung durch die Nutzung dieses Systems im Vergleich zu einer Erstellung ohne Unterstützung, sowie 3) der Entwurf von Qualitätsmerkmalen und Metriken für IEC 62559 Use Case-Beschreibungen und SGAM Modelle.

Diese Dissertation widmete sich der Frage, wie die Qualität von IEC 62559 Use Case-Beschreibungen und SGAM Modellen gesteigert werden kann. Hierfür wurde ein intelligentes System namens *Smart Grid Assisting Artificial Intelligence in Requirements Engineering (SGAIRE)* entworfen, implementiert und evaluiert. Ausgangspunkt war die Beobachtung, dass die Qualität dieser Artefakte eine zentrale Voraussetzung für das Schaffen von Interoperabilität in Smart Grid-Projekten darstellt, während gleichzeitig Fehler in frühen Phasen aufgrund unterschiedlicher Interpretationen durch heterogene Stakeholder häufig auftreten und mit erheblichen Kosten sowie Risiken verbunden sind.

#### Forschungsfrage 1

Wie können **intelligente Systeme** für Fachexperten zur Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen konzipiert werden?

Zur Beantwortung der ersten Forschungsfrage wurde das System SGAIRE im Rahmen der Design Science Research Methodology konzipiert und umgesetzt. Das zugrunde liegende Konzept basierte auf der Annahme, dass Methoden des NLP, insbesondere (L)LMs, in der Lage sind, textuelle Repräsentationen und Strukturen gezielt zu verarbeiten. Auf Basis einer Anforderungs- und Kontextanalyse mit Fach- und Methodenexperten sowie deren Auswertung durch eine Nutzwertanalyse wurden relevante funktionale Anforderungen abgeleitet. Diese Anforderungen wurden im Gesamtsystem durch den Einsatz von acht unterschiedlichen KI-Modellen realisiert.

Hierbei kamen verschiedene LMs wie BERT, RoBERTa, DeBERTa und GPT-2 sowie LLMs wie LLaMA 3.1 und GPT-4o(-mini) zum Einsatz, die jeweils feintrainiert, evaluiert und mit prozeduraler Logik in das Gesamtsystem integriert wurden. Im Rahmen zweier Demonstrationen sowie eines Experiments, an dem Fach- und Methodenexperten sowie Vertreter aus dem Standardisierungsumfeld beteiligt waren, konnte SGAIRE erfolgreich angewendet werden. Sowohl die funktionalen (vgl. Tabelle 9.8) als auch die

nicht-funktionalen Anforderungen (vgl. Tabelle 9.9) konnten nachweislich erfüllt werden. Damit konnte Hypothese 1 bestätigt und Forschungsfrage 1 abschließend beantwortet werden.

**Hypothese 1**

Intelligente Systeme, die auf natürliche Sprachverarbeitung setzen, können dazu verwendet werden, Fachexperten bei der Erstellung von IEC 62559 Use Cases und SGAM Modellen zu unterstützen, indem sie relevante Informationen aus verschiedenen Quellen extrahieren und strukturieren.

**Forschungsfrage 2.1**

Welche relevanten *Bewertungsmetriken für die Erfüllungsgrade von Qualitätsmerkmalen* können auf Basis der Grundsätze ordnungsmäßiger Modellierung (GoM) abgeleitet werden, um die Modellanforderungen von IEC 62559-2 Use Cases und SGAM Modellen zu erfüllen?

Zur Beantwortung der zweiten Forschungsfrage war es zunächst erforderlich, die Unterforschungsfrage 2.1 zu adressieren. Diese richtete sich auf die Identifikation relevanter Qualitätsmerkmale- und Metriken, um eine systematische Bewertung von IEC 62559 Use Cases und SGAM Modellen zu ermöglichen. Hierfür wurde eine Literaturrecherche durchgeführt, auf deren Grundlage Anforderungen an „gute“ IEC 62559 Use Cases sowie SGAM Modelle erhoben werden konnten. Darauf aufbauend wurden durch eine induktive Herleitung, unter Einbezug der GoMs sowie unter Berücksichtigung des ISO/IEC 25012 Standards, geeignete Qualitätsmerkmale abgeleitet. Die daraus entwickelten Metriken wurden anschließend über ein entsprechendes Mapping operationalisiert und im Rahmen der Demonstration angewendet.

Damit konnte gezeigt werden, dass die entwickelten Metriken in zweifacher Hinsicht von Nutzen sind: Zum einen können sie unmittelbar bei der Erstellung zukünftiger Beschreibungen oder Modelle eingesetzt werden, um deren Qualität bereits im Entstehungsprozess sicherzustellen. Zum anderen bieten sie eine belastbare Grundlage für die Implementierung unterstützender Systeme, die sowohl bestehende als auch neue IEC 62559 Use Cases und SGAM Modelle objektiv bewerten können, wie dies für SGAAIRE der Fall ist. Auf dieser Basis wurde Hypothese 3 bestätigt und die Unterforschungsfrage 2.1 abschließend beantwortet.

**Hypothese 3**

Bewertungsmetriken können als Grundlage dienen, um die Qualität von erstellten IEC 62559-2 Use Case-Beschreibungen und SGAM Modellen zu bewerten und somit deren Konformität mit den vorgegebenen Modellanforderungen zu überprüfen.

**Forschungsfrage 2**

Führt der Einsatz von **intelligenten Systemen** bei der Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen gegenüber der konventionellen Arbeitsweise zur **Qualitätsverbesserung**?

Auf Grundlage der entwickelten Qualitätsmerkmale und -metriken konnten die Unterforschungsfrage 2.2 sowie die zweite Forschungsfrage adressiert werden. Über die beiden durchgeführten Demonstrationen ließ sich zunächst eine Indikation für die Effektivität von SGAAIRE als intelligentes Unterstützungssystem aufzeigen. In beiden Anwendungsfällen wurden bei der Identifikation der eingebauten Fehler Metrikerwerte (F1-Werte) von über 85 % erzielt. Innerhalb des Experiments konnte SGAAIRE anhand des zweiten Demonstrationsbeispiels angewandt werden. Dabei zeigte sich, dass das System in der Interventionsgruppe eine aktive Unterstützung bei der Fehleridentifikation leistete. Für den betrachteten Anwendungsfall der „Verwendung von dynamischen Stromtarifen für EV-Ladestationen“ ergab sich eine Qualitätssteigerung von rund 74,58 %.

Eine differenzierte Betrachtung verdeutlichte, dass insbesondere Fortgeschrittene mit durchschnittlich 95,09 % eine deutliche Leistungssteigerung aufwiesen, während Experten mit 25,54 % lediglich eine moderate Verbesserung erzielten und Anfänger mit 5 % nur in sehr geringem Maße profitierten. Es lag die Vermutung nahe, dass durch den Effekt der Automatisierungsverzerrung die Probanden verstärkt dazu neigen, die Empfehlungen ungeprüft zu übernehmen und dadurch weniger zusätzliche Fehler entdecken. Ein statistisch signifikanter Unterschied konnte hierbei jedoch nicht festgestellt werden. Dies dürfte womöglich auf die transparente, sowohl lokale als auch globale Erklärung der Systemempfehlungen sowie auf die konsequente Verwendung passiver, nicht-imperativer Sprache in allen Vorschlägen zurückzuführen sein. Besonders stark profitierten die Qualitätsmerkmale Korrektheit, Nachvollziehbarkeit und Eindeutigkeit, während die Verbesserungen in den Bereichen Erreichbarkeit, Relevanz und Vollständigkeit vergleichsweise gering ausfielen. Die Ergebnisse konnten durch Hypothesentests untermauert werden, sodass Hypothese 2 angenommen, die Unterforschungsfrage 2.2 beantwortet und damit die zweite Forschungsfrage insgesamt adressiert werden konnte.

**Forschungsfrage 2.2**

Welche *Auswirkungen* hat der Einsatz von *intelligenten Systemen* im Hinblick auf die Erfüllungsgrade der Qualitätsmerkmale hergeleitet aus den Grundsätzen ordnungsmäßiger Modellierung (GoM) gemäß Modell-Anforderungen von IEC 62559-2 Use Cases und SGAM Modellen?

**Hypothese 2**

Der Einsatz von SGAAIRE bei der Erstellung von IEC 62559-2 Use Case-Beschreibungen und SGAM Modellen kann zu einer Verbesserung der Qualität beitragen.

## 10.2 Ausblick

Ausgehend von der Entwicklung von SGAAIRE bis hin zu dessen Evaluation zur Erhebung der Daten zur Beantwortung der Forschungsfragen, lassen sich verschiedene Ansatzpunkte für weiterführende Arbeiten ableiten. Neben einer Vertiefung durch Begleitstudien in praktischen Projekten bietet insbesondere die Nutzung von SGAAIRE sowie dessen API für die Entwicklung externer Werkzeuge zusätzliche Potenziale. Aufbauend auf den in dieser Dissertation gewonnenen Erkenntnissen können zukünftige Forschungsarbeiten in den folgenden Bereichen ansetzen.

### Berücksichtigung weiterer funktionaler Anforderungen und Designentscheidungen

Im Rahmen dieser Dissertation konnte nur eine Auswahl an funktionalen Anforderungen einbezogen werden. Diese wurden auf Basis einer Anforderungserhebung unter den Methodennutzenden identifiziert und mithilfe einer NWA priorisiert und selektiert. Durch die Selektionsproblematik blieben jedoch weitere Anforderungen unberücksichtigt, die potenziell Einfluss auf die Qualität der Ergebnisse haben könnten und somit zu abweichenden Resultaten führen würden. Dies betrifft auch die konkrete Umsetzung der berücksichtigten funktionalen Anforderungen, da Verzerrungen durch System- und Softwareentscheidungen in der empirischen Forschung nicht gänzlich zu vermeiden sind. Alternative Herangehensweisen könnten hier zu potenziellen Qualitätsverbesserungen beitragen.

### Stetige Dynamik im KI-Bereich

Zum Beginn dieses des Dissertationsvorhabens (September 2022) waren LLM-basierte Systeme wie ChatGPT noch nicht in der breiten Öffentlichkeit verfügbar. Dies änderte sich jedoch im Verlauf des Dissertationsprozesses und wurde in der Konzeption sowie in der Umsetzung dieser Arbeit berücksichtigt. Inzwischen erscheinen kontinuierlich weitere Alternativplattformen sowie modernere Modelle und Techniken. Da der Ansatz dieser Arbeit auf der Verwendung von (L)LMs basiert, könnten durch die Einbeziehung zukünftiger Modellgenerationen weitere Leistungssteigerungen erzielt werden. Darüber hinaus könnten zukünftige Arbeiten alternative Herangehensweisen untersuchen und insbesondere das Themenfeld des Prompt Engineerings adressieren, um auf Basis von KI-Plattformen und LLMs präzisere Empfehlungen für SGAM-Modelle und IEC 62559 Use Case-Beschreibungen zu generieren.

### Integration in andere Tools und Prozesse

Diese Dissertation untersuchte insbesondere den anwendungsfallgetriebenen Prozess zur Erstellung von SGAM-Modellen sowie die Erstellung von IEC 62559 Use Cases über

das Ausfüllen des IEC 62559-2 Use Case-Templates. Dabei wurde die gleichwertige Anwendung beider Methoden hervorgehoben, um sowohl statische als auch dynamische Aspekte der Systembeschreibung abzudecken. Dies entspricht auch Perspektiven in der Standardisierung durch das IEC System Committee Smart Energy sowie in Projekten wie TDX-Assist [LCR<sup>+</sup>19] und CoordiNet [GMIFM<sup>+</sup>23]. Aus der Kontextanalyse lässt sich jedoch ableiten, dass weitere Zielgruppen z. B. einen MBSE-Ansatz bevorzugen, etwa zur Erstellung von IEC 62559 Use Cases mit Modсарus [EDF] oder zur Modellierung von SGAM Modellen mit der SGAM Toolbox [NET<sup>+</sup>14]. Zudem kann der Fokus der Anwendenden auch auf einzelnen Artefakten liegen, wie etwa ausschließlich der IEC 62559 Use Case-Beschreibung oder dem SGAM-Modell, beispielsweise zur Durchführung von (Co-)Simulationen oder weiterführenden Interoperabilitätsanalysen. Die Ergebnisse dieser Dissertation können daher auch in andere Prozesse und Tools integriert werden, um eine erweiterte Qualitäts- und Modellierungsunterstützung zu bieten, um so die weiterführenden Analysen zu stützen. Zudem ist denkbar, dass sich aus der Nutzung alternativer Prozesse oder Werkzeuge neue Gegebenheiten ergeben, die wiederum zu abweichenden Ergebnissen führen können.

### Berücksichtigung von Entwicklungen im SGAM und der IEC 62559 Use Case-Methodik

Die Weiterentwicklung und Anwendung sowohl des SGAM als auch der IEC 62559 Use Case-Methodik ist weiterhin gegeben. So werden etwa zusätzliche Interoperabilitätsdisziplinen berücksichtigt, beispielsweise durch die Einbeziehung von Governance- und Regulierungsaspekten in einer erweiterten sechsten „Framework“-Ebene [KMH24] oder durch die Integration sozialer Dimensionen [PKML22]. Auch auf Domänen- und Zonenebene eröffnen sich neue Perspektiven, wie etwa im Electric Mobility Architecture Model (EMAM) [UG15]. Bei ausreichender Datenverfügbarkeit und entsprechender Akzeptanz könnte zudem ein erneutes Training der zugrunde liegenden KI-Modelle erfolgen, um diese Weiterentwicklungen adäquat abzubilden und deren Auswirkungen auf die Qualitätsmetriken zu erfassen.

### Übertragung und Analyse in anderen Sektoren

Die in dieser Arbeit erzielten Ergebnisse beziehen sich primär auf den Energiesektor im Kontext des SGAM. Perspektivisch ist jedoch eine Übertragung und Analyse auf weitere Referenzarchitekturmodelle anderer Sektoren denkbar. Beispiele hierfür sind das Reference Architecture Model for Industry 4.0 (RAMI 4.0) [HR24] für den Industriesektor, das Automotive Reference Architecture Model (RAMA) [PGBN22] für den Verkehrssektor sowie das International Data Spaces Reference Architecture Model (IDS RAM 4.0) [OSTL19] im Rahmen des domänenübergreifenden Data Spaces-Konzeptes.



## A Anhang

### A.1 Erklärung zum Online-Anhang und GitLab Projekten

Der Online-Anhang<sup>1</sup> enthält das konsolidierte Material der betrachteten Studien sowie ergänzende Unterlagen. Dazu zählen insbesondere die IEC 62559 Use Cases und die SGAM Modelle, welche in den Demonstrations- und Evaluationsbeispielen Anwendung finden. Das Online-Archiv ist in folgende Verzeichnisse gegliedert:

- **Online-Umfrage (Kapitel 5 und 7):** Enthält die Rohdaten der Befragungsteilnehmenden, den Aufbau des Fragebogens, eine Übersicht der demografischen Daten sowie das Auswertungsskript.
- **Nutzwertanalyse (Kapitel 5):** Beinhaltet die vollständige Nutzwertanalyse als Excel-Datei mit mehreren Tabellenblättern, die die einzelnen Zwischenschritte dokumentieren.
- **Demonstrationsbeispiele (Kapitel 9):** Beinhaltet für beide Fallstudien sowohl die Referenz- als auch die EvaluationsUse Case-Beschreibungen nach IEC 62559 sowie die zugehörigen SGAM Modelle. Zusätzlich ist dokumentiert, welche Fehler wie in die Beispiele integriert wurden.
- **Evaluationsergebnisse (Kapitel 9):** Umfasst die Rohdaten der Evaluation sowie die Struktur der Vor- und Nachbefragungen.

Darüber hinaus wurden im Rahmen der Dissertation zwei GitLab-Repositories in der institutseigenen OFFIS-Instanz eingerichtet:

- **GitLab-Repository zu SGAAIRE<sup>2</sup> (Kapitel 7 und 8):** Umfasst den vollständigen Quellcode des Systems, die zur Entwicklung der KI-Modelle verwendeten Jupyter Notebooks, die für das Training eingesetzten Daten, die Demonstrationsbeispiele, die trainierten Modelle, sämtliche in der Arbeit eingesetzten Skripte, weiterführende Evaluationsdaten sowie den im Rahmen der Dissertation entwickelten IEC 62559 Use Case-Serializer.
- **GitLab-Repository zum SGAAIRE-Datensatz<sup>3</sup> (Kapitel 7):** Enthält die Rohdaten sowie deren Aufbereitungsschritte zur Erstellung des finalen Trainingsdatensatzes.

---

<sup>1</sup> Verfügbar unter: [https://osf.io/y6c4n/?view\\_only=d1894be227324a44b468ebd8cb5da87f](https://osf.io/y6c4n/?view_only=d1894be227324a44b468ebd8cb5da87f)

<sup>2</sup> Intern verfügbar unter: <https://gitlab.offis.de/official/fbe-p-ucmr/SGAAIRE>

<sup>3</sup> Intern verfügbar unter: [https://gitlab.offis.de/official/fbe-p-ucmr/sgaaire\\_dataset](https://gitlab.offis.de/official/fbe-p-ucmr/sgaaire_dataset)

## A.2 Anwendungsfallbeschreibung gemäß IEC 62559-2: #1 Optimierung der Lebensdauer von (smarten) Transformatoren mithilfe von Smart Services

In diesem Abschnitt des Anhangs wird der Anwendungsfall des ersten Demonstrationsbeispiels „Optimierung der Lebensdauer von (smarten) Transformatoren mithilfe von Smart Services“ in seiner Referenzversion dargestellt. Der Anwendungsfall liegt in englischer Sprache vor.

### 1. Description of the use case

Tabelle A.1: Anwendungsfall #1 - Sektion „1.1 Name of the use case“

Use case identification		
ID	Area Domain(s)/Zone(s)	Name of use case
SSST	Transmission, Distribution, DER, Customer Premises / Operation, Station, Field	Optimizing the life-span of (smart) transformers

Tabelle A.2: Anwendungsfall #1 - Sektion „1.2 Version management“

Version management				
Version No.	Date	Name of author(s)	Changes	Approval status
0.1	9th November 2023	R. Kuchenbuch	Initial	Draft

Tabelle A.3: Anwendungsfall #1- Sektion „1.3 Scope and objectives of use case“

Scope and objectives of use case	
<b>Scope</b>	The scope of the use case relates to transformers. As transformers are to be used by Transmission System Operators (TSOs), Distribution System Operators (DSOs), Distributed Energy Resources (DER) and Customer Premises (CP), the proposed Smart Service system (used for improving the life-span) should refer generically to all these domains.
<b>Objective(s)</b>	<b>O1:</b> Development of a smart service (including a dashboard visualization) to improve the service life of (smart) transformers
<b>Related business case(s)</b>	

Tabelle A.4: Anwendungsfall #1 - Sektion „1.4 Narrative of use case“

Narrative of use case	
<b>Short description</b>	The aim of this Use Case is to optimize the service life of (smart) transformers via an improved maintenance environment. The aspects relating to Condition Monitoring (CM), Predictive Maintenance (PdM) and Asset Management (AM) are considered as Smart Services in the national research and development project STIM and already being used in practice in other domains to improve both maintenance and health monitoring. By developing a dedicated smart service (SSST standing for Smart Service for Smart Transformers) including a dashboard visualization; the aim is to increase the service life of smart transformers based on the previous mentioned Smart Services by adjusting processes for the maintenance.
<b>Complete description</b>	<p>This Use Case aims to enhance the longevity and performance of smart transformers through the implementation of an advanced maintenance framework. It falls under the scope of the national research and development project STIM, where the concepts of Condition Monitoring (CM), Predictive Maintenance (PdM), and Asset Management (AM) are categorized as Smart Services. These concepts have already proven their effectiveness in various domains by improving maintenance practices and health monitoring. In this particular project, a dedicated smart service, referred to as SSST (Smart Service for Smart Transformers), is being developed, complete with a comprehensive dashboard visualization. The primary objective is to extend the service life of smart transformers by refining maintenance processes based on the aforementioned Smart Services.</p> <p>The SSST encompasses several critical elements, including the Baseline Assets List, Health Status Evaluation, Components Residual Life Estimation, Failure Detection, Failure Diagnosis, Remaining Useful Life Estimation, Failure Prognosis, Pattern Recognition, and Lifecycle Cost Estimation. Additionally, it offers management decision-making support, which can lead to actionable insights. The system is designed to generate warning messages and present all relevant data on a user-friendly Data Visualization Dashboard. For the realization of the named components, sensor values are to be collected, which are continuously evaluated and assessed using Machine Learning (ML)/Artificial Intelligence (AI) methods and techniques, such as Artificial Neural Networks (ANN) or Transformers (in the sense of ML) within SSST. Furthermore, the SSST incorporates an analytical component, which yields valuable reports such as Transformer Health Assessment, Failure Analysis Report, Essential Maintenance Tips, Transformer Lifetime Prediction, Risk Assessment, and Alarm Reports. These reports provide stakeholders with vital information to make informed decisions, ensure the transformers' optimal health, and extend their operational life. Overall, the SSST system is poised to enhance the maintenance and management of smart transformers, ultimately enhancing their efficiency and reliability.</p>

Tabelle A.5: Anwendungsfall #1 - Sektion „1.5 Key performance indicators (KPI)“

Key performance indicators			
ID	Name	Description	Reference to mentioned use case objectives
KPI.1	Increased Service life	The service life of the (smart) transformer should increase by at least 5 years compared to conventional ones	O1
KPI.2	Established transformers using SSST	10 (Smart) Transformers are using SSST on field	O1

Tabelle A.6: Anwendungsfall #1 - Sektion „1.6 Use case conditions“

Use case conditions	
<b>Assumptions</b>	-/-
<b>Prerequisites</b>	- Required number of transformers provided for evaluation in the field by organizations

Tabelle A.7: Anwendungsfall #1 - Sektion „1.7 Further information to the use case for classification / mapping“

Classification information	
<b>Relation to other use cases</b>	-/-
<b>Level of depth</b>	Low Level Use Case
<b>Prioritisation</b>	Nice to have
<b>Generic, regional or national relation</b>	Generic
<b>Nature of the use case</b>	Technical Use Case
<b>Further keywords for classification</b>	Monitoring, Sensors, Smart Services, Transformers

Tabelle A.8: Anwendungsfall #1 - Sektion „1.8 General remarks“

General remarks
<ul style="list-style-type: none"> <li>The use case envisions the usage of the sensor values and metadata most relevant to the analysis acc. [Ref. 1].</li> </ul>

## 2. Diagrams

Tabelle A.9: Anwendungsfall #1 - Sektion „2 Diagram(s) of use case“

Diagram(s) of use case

### 3. Technical details

Tabelle A.10: Anwendungsfall #1 - Sektion „3.1 Actors (Grouping: Primary Actors)“

Actors			
Grouping		Group description	
Primary actors			
Actor name	Actor type	Actor description	Further information specific to this use case
(Smart) Transformer	Device	A transformer is an essential electrical device used in energy-related systems and applications. It is designed to facilitate the efficient transfer of electrical energy between two or more circuits by means of electromagnetic induction. [...]	
Sensors / Components	Device	Decentralized microcontroller systems, equipped with sensor connectivity (for measuring parameters such as temperature and humidity), capable of collecting and transmitting valuable data.	
SSST System	System	The Smart Service for Smart Transformers is the concrete system related to the objective of this use case. Purpose of the SSST is the support of the stakeholders (in particular technical personal but also decision-makers) to inform about the condition of the transformer, the next maintenance points and about critical/unusual data. Included (AI) modules and complex algorithms.	
SSST Dashboard	System	A loosely-coupled dashboard specifically for the SSST System which consists of the access management.	

Tabelle A.11: Anwendungsfall #1 - Sektion „3.1 Actors (Grouping: Further Actors)“

Actors			
Grouping		Group description	
Further Actors			
Actor name	Actor type	Actor description	Further information specific to this use case
Transmission System Operator (TSO)	Role	A Transmission System Operator (TSO) is a fundamental entity in the energy sector responsible for overseeing and managing the high-voltage electricity transmission grid. [...]	
Distributed System Operator (DSO)	Role	A Distributed System Operator (DSO) is a critical entity in the energy sector responsible for managing and maintaining the efficient and reliable distribution of electrical power within a specific geographic area or network. [...]	
Distributed Energy Resources (DER)	Role	Distributed Energy Resources (DER) are decentralized energy generation and storage systems typically located near the point of energy consumption. [...]	
Enterprises	Role	An enterprise refers to a formal organization or business entity established for the purpose of conducting commercial, industrial, or service-related activities. [...]	

Tabelle A.12: Anwendungsfall #1 - Sektion „3.2 References“

References						
No.	Reference type	Reference	Status	Impact on use case	Originator	Link
Ref 1.	Research paper	Optimizing the Life-Span of (Smart) Transformers: A Review on Smart Services	Published	Basic concept of the system	René Kuchenbuch (OFFIS e.V.), Mana Azamat (OFFIS e.V.) and Johann Schütz (OFFIS e.V.)	<a href="http://dx.doi.org/10.1049/icp.2023.0819">http://dx.doi.org/10.1049/icp.2023.0819</a>

4. Step by step analysis of use case

Tabelle A.13: Anwendungsfall #1 - Sektion „4.1 Overview of scenarios“

Scenario conditions						
No.	Scenario name	Scenario description	Primary actor	Triggering event	Pre-condition	Post-condition
1	Data collection, compilation and management	The collection and management of data is a key component of the data-driven system. The data is necessary to train the (ML/AI-) modules and develop/test the algorithms on the one hand and to measure the actual status of the (smart) transformer and provide this data on the other.	Sensors / Components	Periodic (Streaming data)	Availability of the SS-ST system	Available / stored compiled data within the SSST system
2	Detection of unusual/critical conditions	Within the use case, critical and unusual conditions are to be identified and (if available) reported to the users.	SSST System	Periodic (Timer)	Compiled data available	In the event of critical and unusual messages, a notification should be sent to the involved instance
3	Provision of the compiled data	The compiled data should be made available to users.	Transmission System Operator (TSO), Distributed System Operator (DSO), Distributed Energy Resources (DER), Enterprises	On request	Compiled data available	Compiled data accessible for the respective stakeholders

Tabelle A.14: Anwendungsfall #1 - Sektion „4.2 Scenario (1)“

Scenario							
Scenario name:			No. 1: Data collection, compilation and management				
Step No.	Event	Name of process/ activity	Description of process/ activity	Service	Information producer (actor)	Information receiver (actor)	Inf. exchanged (IDs)
1.1	Periodic (Streaming data)	Creating raw data frame from	A data frame is created based on the raw data from the sensors, which are standardized/normalized for the data types (e.g. temperature as a floating point number in C°)	CREATE	Sensors / Component		
1.2	–	Transferring raw data to SSST System	Transmission of raw data to the SSST system for further processing and provision	REPORT	Sensors / Components	SSST System	INF-1
1.3	–	Compile data	Raw data are processed using (AI) modules and algorithms	EXECUTE	SSST System		
1.4	–	Storing data into a RDMS	Raw data as well as the compiled are to be stored within an RDMS	EXECUTE	SSST System		
1.5	–	Commit data to the dashboard	The data is provided to the dashboard (in the sense of a stream)	REPORT	SSST System	SSST Dashboard	INF-2

Tabelle A.15: Anwendungsfall #1 - Sektion „4.2 Scenario (2)“

Scenario							
Scenario name:			No. 2: Detection of unusual/critical conditions				
Step No.	Event	Name of process/ activity	Description of process/ activity	Service	Information producer (actor)	Information receiver (actor)	Inf. exchanged (IDs)
2.1	Periodic (Timer)	Verification of the data frame	The available data is checked for unusual and critical conditions.	EXECUTE	SSST System		
2.2	In case of critical or unusual condition	Sending the warning message and action code	If an error occurs, a high-priority warning message should be sent to the users (DSO, TSO, DER or CP), as well as a message to the SSST dashboard. In addition, a message should be sent to the (Smart) Transformer in the event of a catastrophic state, which can react on this basis (e.g. hard shutdown).	REPORT	SSST System	(Smart) Transformer, SSST Dashboard, Transmission System Operator (TSO), Distributed System Operator (DSO), Distributed Energy Resources (DER), Enterprises	INF-3

Tabelle A.16: Anwendungsfall #1 - Sektion „4.2 Scenario (3)“

Scenario							
Scenario name:			No. 3: Provision of the compiled data				
Step No.	Event	Name of process/ activity	Description of process/ activity	Service	Information producer (actor)	Information receiver (actor)	Inf. exchanged (IDs)
3.1	On request	Request visualization	Users (in the sense of a DSO, TSO, DER actor or company) would like to use the dashboard	GET	Transmission System Operator (TSO), Distributed System Operator (DSO), Distributed Energy Resources (DER), Enterprises	SSST Dashboard	
3.2	–	Data provision	The compiled data are made available to the stakeholders (acc. to the request)	REPORT	SSST Dashboard	Transmission System Operator (TSO), Distributed System Operator (DSO), Distributed Energy Resources (DER), Enterprises	INF-2

## 5. Information exchanged

Tabelle A.17: Anwendungsfall #1 - Sektion „5. Information Exchanged“

Information exchanged			
Information exchanged ID	Name of information exchanged	Description of information exchanged	Req. IDs
INF-1	Raw data	Raw data from e.g. sensor data or other measurements or key indicators.	
INF-2	Compiled data	Processed data that is additionally handled via the SSST (e.g. results of AI).	
INF-3	Warning message with error code and details	A data packet including a human-readable message and additional flags for the error code and other technical values, for an interpretation at a technical point.	

## 6. Requirements (optional)

Tabelle A.18: Anwendungsfall #1 - Sektion „6. Requirements (optional)“

Requirements (optional)		
Categories ID	Category name for requirements	Category description
Requirement ID	Requirement name	Requirement description

## 7. Common Terms and Definitions

Tabelle A.19: Anwendungsfall #1 - Sektion „7. Common terms and definitions“

Common Terms and Definitions	
Term	Definition
SSST	Smart Service for Smart Transformer

## 8. Custom information (optional)

Tabelle A.20: Anwendungsfall #1 - Sektion „8. Custom information (optional)“

Custom information (optional)		
Key	Value	Refers to section

### A.3 Anwendungsfallbeschreibung gemäß IEC 62559-2: #2 Verwendung von dynamischen Stromtarifen für EV-Ladestationen

In diesem Abschnitt des Anhangs wird der Anwendungsfall des zweiten Demonstrationsbeispiels „Verwendung von dynamischen Stromtarifen für EV-Ladestationen“ in seiner Referenzversion dargestellt. Der Anwendungsfall liegt in englischer Sprache vor.

#### 1. Description of the use case

Tabelle A.21: Anwendungsfall #2 - Sektion „1.1 Name of the use case“

Use case identification		
ID	Area Domain(s)/Zone(s)	Name of use case
ToUT.EV	Enterprise, Operation, Market / DER, Distribution, Customer Premises	Time of Use Tariff for Electric Vehicle Charging Stations

Tabelle A.22: Anwendungsfall #2 - Sektion „1.2 Version management“

Version management				
Version No.	Date	Name of author(s)	Changes	Approval status
0.1	16th September 2024	R. Kuchenbuch	Initial creation	Draft

Tabelle A.23: Anwendungsfall #2 - Sektion „1.3 Scope and objectives of use case“

Scope and objectives of use case	
<b>Scope</b>	This Use Case provides an incentive-based mechanism for load shifting. For this, the Energy Service Providers (ESPs) or the DSO may send incentive tables to the customer energy manager (CEM) of buildings. The CEM may manage the connected appliances within their flexibility (which depends on the capabilities of the appliances) to optimize the power consumption or production according to the communicated incentives, e.g., price, CO <sub>2</sub> emission, etc.
<b>Objective(s)</b>	<b>Operation of Electric Vehicle Charging Stations with Time of Use Tariff:</b> The objective of this use case is to optimize the operation of electric vehicle (EV) charging stations by utilizing an incentive table with unique tiers that depend on consumed or produced power. This method aims to reduce operational costs, provide cost savings for EV users, and support grid stability.
<b>Related business case(s)</b>	

Tabelle A.24: Anwendungsfall #2 - Sektion „1.4 Narrative of use case“

Narrative of use case	
<b>Short description</b>	This Use Case provides an incentive-based mechanism for load shifting for Electric Vehicles (EV).
<b>Complete description</b>	<p>The electricity grid is more and more challenged by energy consuming appliances such as electric vehicles (EVs), heat pumps, etc. on the one hand and transition to regenerative energy production on the other hand. To enable this, energy demand and energy production need to be harmonized. Energy Service Providers (ESPs) or Distribution System Operators (DSOs) need to time-shift the energy demand of consumers to manage over and underload scenarios in the distribution grid. As many appliances are not capable or willing to be shifted directly, a mechanism to trigger the appliances to shift their demand themselves is needed.</p> <p>This Use Case provides an incentive-based mechanism for load shifting. For this, the ESP or the DSO may send incentive tables to the customer energy manager (CEM) of buildings. The CEM may manage the connected appliances within their flexibility (which depends on the capabilities of the appliances) to optimize the power consumption or production according to the communicated incentives, e.g., price, CO<sub>2</sub> emission, etc.</p> <p>Within this Use Case the following Actors are used:</p> <ul style="list-style-type: none"> <li>• „Energy Broker“ as optimization instance at the customer’s ESP,</li> <li>• „Delivery Broker“ for the DSO’s or the ESP’s role in this Use Case (depending on the overall system setup),</li> <li>• „CEM“ for the energy manager at the customer’s building.</li> </ul> <p>Depending on the contracts of the customer with the ESP and the DSO and the regulations in the country they are located, either both Actors send individual incentive tables to the CEM or only the ESP (see section 2.2.1). Each incentive table with a tariff can have multiple tariff levels (called „tiers“) which have one or more incentives (e.g., absolute price, CO<sub>2</sub> emission, etc.). Based on the incentives, the CEM may calculate where to plan the consumption and production cycles of its managed appliances (see section 2.2.4). At least, the absolute price SHALL be included in the incentive table.</p> <p>Within this Use Case, a time-based incentive information approach is used to model incentive tables. Subsequently, these are called „incentive-based incentive tables“.</p>

Tabelle A.25: Anwendungsfall #2 - Sektion „1.5 Key performance indicators (KPI)“

Key performance indicators			
ID	Name	Description	Reference to mentioned use case objectives
KPI.1	Establishing Time of Use Tariff for three different EV Charging Stations	Establish three distinct providers for E-Vehicle charging stations.	Operation of Electric Vehicle Charging Stations with Time of Use Tariff

Tabelle A.26: Anwendungsfall #2 - Sektion „1.6 Use case conditions“

Use case conditions	
<b>Assumptions</b>	The CEM operates one or more EV charging stations.
<b>Prerequisites</b>	The communication infrastructure between the Energy Service Provider (ESP), the Distribution System Operator (DSO), and the Customer Energy Manager (CEM) must be established and compliant with standardized protocols.

Tabelle A.27: Anwendungsfall #2 - Sektion „1.7 Further information to the use case for classification / mapping“

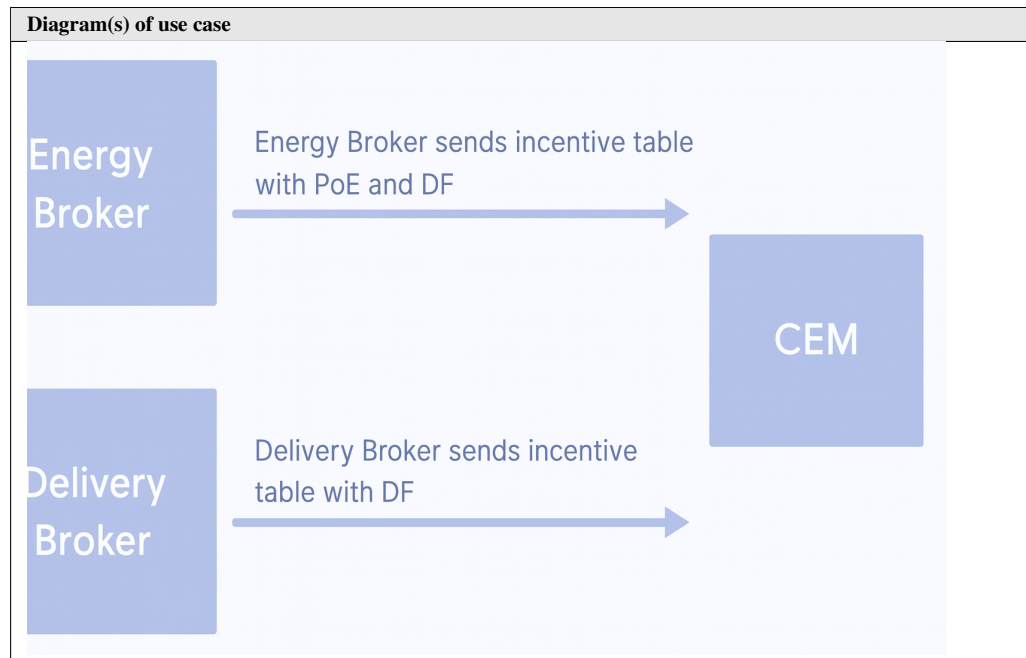
Classification information	
<b>Relation to other use cases</b>	Load shifting for Heat Pumps (associate) / Integration of Renewable Energy Sources (include) / Flexibility Management by Customer Energy Manager (associate)
<b>Level of depth</b>	High Level Use Case
<b>Prioritisation</b>	Mandatory
<b>Generic, regional or national relation</b>	Generic
<b>Nature of the use case</b>	Technical Use Case
<b>Further keywords for classification</b>	Dynamic Tariffs, Electric Vehicles, Smart Grid, ICT

Tabelle A.28: Anwendungsfall #2 - Sektion „1.8 General remarks“

General remarks
<ul style="list-style-type: none"> <li>• This use case focuses on EV charging with incentive-based tariffs and does not claim to be exhaustive regarding other demand response scenarios.</li> <li>• The structure can be reused for future EV charging scenarios such as vehicle-to-grid (V2G) integration.</li> </ul>

## 2. Diagrams

Tabelle A.29: Anwendungsfall #2 - Sektion „2 Diagram(s) of use case“



## 3. Technical details

Tabelle A.30: Anwendungsfall #2 - Sektion „3.1 Actors (Grouping: Application-related Actors)“

Actors			
Grouping		Group description	
Application-related Actors			
Actor name	Actor type	Actor description	Further information specific to this use case
Charging Station	Internal	Charging stations for Electric Vehicles (EVs) are infrastructure facilities that supply electric energy for recharging EV batteries. They come in various types, including Level 1 (slow charging), Level 2 (moderate charging), and DC Fast Charging (rapid charging), catering to different charging needs and speeds. [...]	
Electric Vehicle	Internal	Electric Vehicles (EVs) play a significant role in various applications, particularly through the aggregation of multiple EVs to provide enhanced services. In public transportation, the aggregation of electric buses improves fleet management and reduces emissions in cities. In logistics, aggregating EV fleets streamlines delivery operations and cuts fuel costs. [...]	

Tabelle A.31: Anwendungsfall #2 - Sektion „3.1 Actors (Grouping: General)“

Actors			
Grouping		Group description	
General			
Actor name	Actor type	Actor description	Further information specific to this use case
Energy Broker	External	The Actor Energy Broker submits tariff information to the building. The tariff may include the delivery fee and the price of energy (PoE+DF) or only the latter (PoE).	
Energy Service Provider	External	An Energy Service Provider (ESP) is a company that offers various services to improve energy efficiency and manage energy use for consumers. These services include energy audits, energy management, renewable energy solutions, demand response, and energy procurement. [...]	
Customer Energy Manager	Internal	The Actor CEM manages the devices in the building in order to shift the load to times of lower costs (e.g., monetary costs, CO <sub>2</sub> emission, etc.).	
Delivery Broker	External	The Actor Delivery Broker sends tariff information into the building including the delivery fee only. This incentive is used by the Delivery Broker to motivate a CEM to shift the building's energy demand in times of lower costs that typically correlates with more available energy in the electricity grid.	

Tabelle A.32: Anwendungsfall #2 - Sektion „3.2 References“

References						
No.	Reference type	Reference	Status	Impact on use case	Originator	Link
[IEC 62746-2]	Standard	IEC 62746-2	Published	Description of the Time of Use Tariff Use Case	IEC	<a href="https://www.vde-verlag.de/iec-normen/221696/iec-tr-62746-2-2015.html">https://www.vde-verlag.de/iec-normen/221696/iec-tr-62746-2-2015.html</a>
[IEC 63380]	Standard	IEC 63380	Published	Relevant for the operation of solar charging stations	IEC	<a href="https://www.dke.de/de/normen-standards/dokument?id=7157429&amp;type=dke%7Cdokument">https://www.dke.de/de/normen-standards/dokument?id=7157429&amp;type=dke%7Cdokument</a>

## 4. Step by step analysis of use case

Tabelle A.33: Anwendungsfall #2 - Sektion „4.1 Overview of scenarios“

Scenario conditions						
No.	Scenario name	Scenario description	Primary actor	Triggering event	Pre-condition	Post-condition
1	Transfer of Incentive Schemes to Electric Vehicles Charging Stations	In this scenario, the Energy Broker is responsible for sending incentive-based tables to the Customer Energy Manager (CEM). These tables contain information about the price of energy and/or the delivery fee. There can be different incentive tables for the consumption of power and the production of energy.	Energy Broker	The Actor Energy Broker wants to update one or more incentive tables on the Actor CEM.	The incentive table(s) of the Actor CEM are outdated.	The Actor CEM has new incentive table(s).

Tabelle A.34: Anwendungsfall #2 - Sektion „4.2 Scenario“

Scenario								
Scenario name:		Transfer of Incentive Schemes to Electric Vehicles Charging Stations						
Step No.	Event	Name of process/ activity	Description of process/ activity	Service	Information producer (actor)	Information receiver (actor)	Information exchanged (IDs)	Requirements R-ID
1a	The Actor Energy Broker wants to update one or more incentive tables on the Actor CEM.	Sending incentive table with PoE and/or DF		REPORT	Energy Broker	Customer Energy Manager	Incentive-based incentive table for energy production/-consumption with price of energy and/or delivery fee	REQ_2
1b	The Actor Delivery Broker wants to update one or more incentive tables on the Actor CEM.	Sending incentive table with DF		REPORT	Delivery Broker	Customer Energy Manager	Incentive-based incentive table for energy production/-consumption with price of energy and/or delivery fee	REQ_2
2	–	Sending pricing information to the charging station	Is required for displaying prices, calculating consumption and production prices, etc.	REPORT	Customer Energy Manager	Charging Station	Incentive-based incentive table for energy production/-consumption with price of energy and/or delivery fee	REQ_2
3	–	Loading process according to IEC 63380	Data exchange according to IEC 63380	EXECUTE	Charging Station	Electric Vehicle	E-Vehicle charging data	REQ_1

## 5. Information exchanged

Tabelle A.35: Anwendungsfall #2 - Sektion „5. Information Exchanged“

Information exchanged			
Information exchanged ID	Name of information exchanged	Description of information exchanged	Requirements IDs
1	E-Vehicle charging data	Consumption related-data of charging processes on e-vehicles.	REQ_1
2	Incentive-based incentive table for energy production/consumption with price of energy and/or delivery fee	Information about the incentive values per time-period for the price of energy and delivery fee for the consumption of energy according to Common Information Model (CIM).	REQ_2

## 6. Requirements (optional)

Tabelle A.36: Anwendungsfall #2 - Sektion „6. Requirements (optional)“

Requirements (optional)		
Categories ID	Category name for requirements	Category description
–	Envisioned Standards	Standards that are intended for communication but also for the exchange of information.
Requirement ID	Requirement name	Requirement description
REQ_1	IEC 63380	Local Charging station management systems; Standard that describes the communication and information exchange of charging stations.
REQ_2	Common Information Model	CIM communication/information exchange according to the defined CIM Profile.

## 7. Common Terms and Definitions

Tabelle A.37: Anwendungsfall #2 - Sektion „7. Common terms and definitions“

Common Terms and Definitions	
Term	Definition
DF	Delivery Fee
PoE	Price of Energy

## 8. Custom information (optional)

Tabelle A.38: Anwendungsfall #2 - Sektion „8. Custom information (optional)“

Custom information (optional)		
Key	Value	Refers to section
EV.charging.modes	Charging modes considered: Level 1 (slow), Level 2 (normal), DC Fast Charging.	3.1 Actors

## A.4 Anhang zur Nutzwertanalyse

In diesem Abschnitt werden die zur NWA verwendeten weiterführenden Tabellen dargestellt. Sie umfassen die zugrunde liegenden Bewertungsmaßstäbe, die ermittelten Gewichtungen sowie die Skalierungen und Zielgrößen der Kriterien. Zudem wird die Berechnungsdatei zur NWA beschrieben.

### A.4.1 Tabellen zur NWA

Tabelle A.39: Vorgesehene Skala für den Kriterienvergleich, orientiert an [Saa87]

Intensitätswert	Definition	Beschreibung
1	Gleiche Wichtigkeit	Beide zu vergleichende Kriterien haben eine gleiche Wichtigkeit
2	Größere Wichtigkeit	Ein Kriterium zeigt durch Begründung oder Erfahrungen gegenüber einem anderem Kriterium eine moderat-höhere Wichtigkeit
4	Deutlich-größere Wichtigkeit	Ein Kriterium zeigt durch Begründung oder Erfahrungen gegenüber einem anderem Kriterium eine wesentlich-höhere Wichtigkeit
6	Erhebliche-größere Wichtigkeit	Ein Kriterium zeigt durch Begründung oder Erfahrungen gegenüber einem anderem Kriterium eine erhebliche höhere Wichtigkeit
3, 5	Zwischenstufen	Sofern keine klare Klassifikation in den Definitionen möglich ist (als Kompromiss)

Tabelle A.40: Aufgestellte Bewertungsmatrix der Kriterien

	als	$k_1$	$k_2$	$k_3$	$k_4$	$k_5$	$k_6$
wichtiger							
[k1] Auswirkung auf die Qualität		1	2	4	4	$0.\overline{16}$	0.25
[k2] Weitere Mehrwerte/Folgen		0.5	1	2	4	$0.\overline{16}$	0.25
[k3] Zeitlicher Aufwand		0.25	0.5	1	0.5	$0.\overline{16}$	0.25
[k4] Realisierbarkeit		0.25	0.25	2	1	$0.\overline{16}$	0.25
[k5] Mehrwert von KI-Algorithmen		6	6	6	6	1	2
[k6] Interoperabilität mit dem UCMR		4	4	4	4	0.5	1

Tabelle A.41: Gewichte und Rangordnung der Kriterien für die zusätzliche funktionalen Anforderungen

Kriterien	Gewicht $\omega_i$	Rang
[k3] Zeitlicher Aufwand	0.0549	4
[k4] Realisierbarkeit	0.0640	3
[k5] Mehrwert von KI-Algorithmen	0.5472	1
[k6] Interoperabilität mit dem UCMR	0.3337	2

Tabelle A.42: Verwendeter Wertebereiche und Beschreibung der Messung der Kriterien

Krit.	Datenart der Messung	Mögliche Werte	Beschreibung
[k1]	Numerisch	$[1, 5] \in \mathbb{R}$	Mittelwerte aus der Umfrage
[k2]	Kategorisch	Sehr positiv Eher positiv  Neutral Eher negativ Sehr negativ	Sehr positive Kommentare und/oder signifikante Mehrwerte erkennbar Eher positive Kommentare und/oder weitere (kleinere) Mehrwerte erkennbar  Keine Kommentare oder ausgewogenes Meinungsbild Eher negative Kommentare bzw. geringer Mehrwert Sehr negative Kommentare bei der Einschätzung der Anforderung
[k3]	Kategorisch	Gering  Mittel  Hoch	Geringe Kosten zur Realisierung (schätzungsweise $< 2$ Wochen/Vollzeit zur Realisierung) Mittlere Kosten zur Realisierung (schätzungsweise $\geq 2$ Wochen, aber $< 1$ Monat/Vollzeit zur Realisierung) Hohe Kosten zur Realisierung (schätzungsweise $> 1$ Monat/Vollzeit zur Realisierung)
[k4]	Kategorisch	Hoch Mittel Gering	Realistisch Realistisch, jedoch hohe Komplexität gegeben Unrealistisch
[k5]	Kategorisch	Hoch  Mittel  Gering  Keine	Hoher Mehrwert durch Einsatz von KI-Verfahren (z. B. Keine deterministische Lösung vorhanden, hohe Komplexität in der Anwendung gegeben, KI als Enabler) Mittlerer Mehrwert durch Einsatz von KI (z. B. KI könnte bessere Ergebnisse erzielen als Prozesse) Geringer Mehrwert durch Einsatz von KI (z. B. Vereinfachung einfacher Prozesse mittels KI) Kein Mehrwert durch Einsatz von KI (z. B. direkter Algorithmus oder einfache Regel ist besser als KI)
[k6]	Kategorisch	Ja  Ja, aber [...]  Nein	Implementierung der Anforderung im UCMR zeigt niedrige Komplexität auf Implementierung der Anforderung im UCMR zeigt Komplexität auf, da grundlegende Änderungen oder Erweiterungen notwendig sind Implementierung der Anforderung ist nicht möglich oder sieht zu hohe Aufwände für grundlegende Änderungen oder Erweiterungen vor

Tabelle A.43: Skalierte Zielgrößen der Kriterien

Kriterium	Mögliche Werte	Zielgrößen
[k1] Auswirkungen auf die Qualität	$[1, 5] \in \mathbb{R}$	$\frac{x-1}{5} \in [0, 1]$
[k2] Weitere Mehrwerte/Folgen	Sehr positiv Eher positiv Neutral Eher negativ Sehr negativ	1 0.75 0.5 0.25 0
[k3] Zeitlicher Aufwand	Gering Mittel Hoch	1 0.5 0
[k4] Realisierbarkeit	Hoch Mittel Gering	1 0.5 0
[k5] Mehrwert von KI-Algorithmen	Hoch Mittel Gering Keine	1 0.66 0.33 0
[k6] Interoperabilität mit dem UCMR	Ja Ja, aber [...] Nein	1 0.5 0

#### A.4.2 Berechnungsdatei zur Nutzwertanalyse

Im Online-Anhang (vgl. Abschnitt A.1) befindet sich die Datei zur durchgeführten Nutzwertanalyse. Die bereitgestellte Datei umfasst sämtliche quantitativen Daten, Bewertungsmatrizen sowie die zugrunde liegenden Gewichtungen, die für die präzise Ermittlung der Nutzwerte der einzelnen Optionen im Abschnitt 5.3.1 herangezogen wurden. Die daraus resultierenden Nutzwerte gewährleisten eine nachvollziehbare und methodisch fundierte Entscheidungsgrundlage, die unmittelbar auf den zuvor definierten Kriterien und deren Gewichtungen basiert.

Bitte beachten Sie, dass die relevante Datei *Nutzwertanalyse.xlsx* im Excel-Format vorliegt und mithilfe eines entsprechenden Tabellenverarbeitungsprogramm geöffnet werden kann. Die Datei *Nutzwertanalyse.xlsx* ist in den folgenden Tabellen strukturiert:

**Tabelle 1: Kriterienselektion** Diese Tabelle enthält die verwendeten Kriterien sowie detaillierte Informationen über die angewandten Messungsarten, Wertebereiche, Zielwerte und Beschreibungen für jedes einzelne Kriterium.

**Tabelle 2: Skala Wertevergleich** In dieser Tabelle wird die Kriterienskala aufgeführt, die für den Vergleich der verschiedenen Kriterien verwendet wurde. Es werden die Skalenwerte und die entsprechenden Bedeutungen erläutert.

**Tabelle 3: Gewichteberrechnung (AHP)** Die Tabelle beschreibt den Prozess der Gewichtebestimmung für die Kriterien mittels einer Dominanzmatrix. Anschließend

wird die Berechnung der Gewichtungen gemäß der Analytic Hierarchy Process Methode dargestellt.

**Tabelle 4: Nutzwertanalyse** In dieser Tabelle werden die Nutzwerte für die einzelnen Optionen berechnet. Die Tabelle enthält die Werte, Zielwerte und Zielwertgrößen, die für die Berechnung der Nutzwerte verwendet wurden.

**Tabelle 5: Nutzwerte (für den CSV-Export)** Diese Tabelle enthält die speziell für den CSV-Export berechneten Nutzwerte. Die Daten in dieser Tabelle dienen als Grundlage für die anschließende Anforderungsselektion.

## A.5 Weitere Darstellungen aus den Demonstrationen

In diesem Abschnitt werden die von SGAAIRE identifizierten zusätzlichen Fehler sowie Falschfehler tabellarisch dargestellt.

Tabelle A.44: Fallbeispiel A: Zusätzliche identifizierte Fehler sowie False Positives

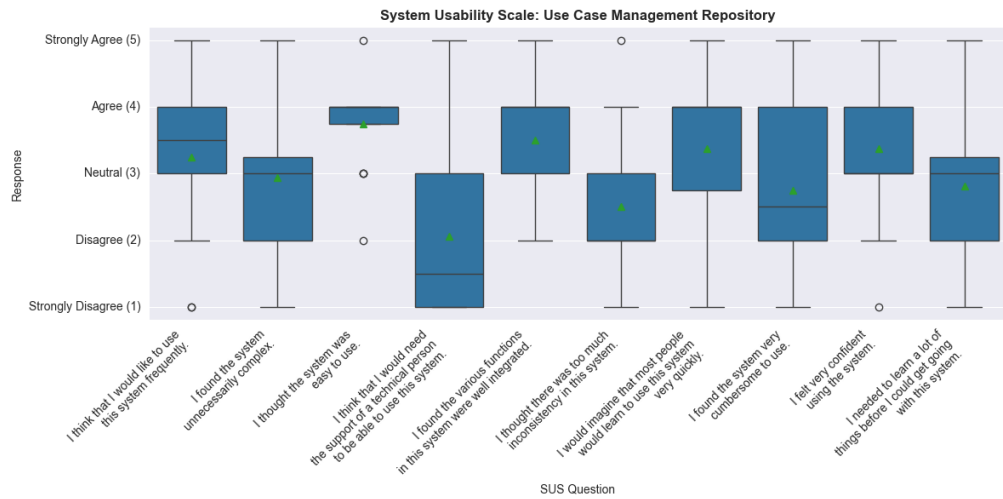
	Beschreibung	Betroffene Qualitätsmerkmale
<b>Zusätzliche identifizierte Fehler</b>	(Weitere „True Positives“)	
Acronyme „STIM“, „UC“, „SSST“ nicht eingeführt	Die genannten Akronyme wurden im Anwendungsfall nicht definiert.	Nachvollziehbarkeit (++) Vollständigkeit (+) Zugänglichkeit (+)
Akteur „(Smart) Transformer“ redundant	Der (smarte) Transformator wurde in der Schritt-für-Schritt-Analyse nicht verwendet. Da der Anwendungsfall auf Systemebene betrachtet wird, werden einzelne Teilkomponenten detaillierter untersucht.	Relevanz (++) Eindeutigkeit (+) Zugänglichkeit (+)
Schreibfehler „STIMrelated“	Es handelt sich um einen Schreibfehler, der durch das Akronym-Modul identifiziert wurde.	Zugänglichkeit (+)
„TSO“, „DSO“ und „DER“ als Operation klassifiziert	Im Referenzmodell wurden alle Instanzen als Enterprise klassifiziert. Dies ist im Rahmen der Modellierungsfreiheit zulässig, jedoch kann argumentiert werden, dass insbesondere die Operation-Zone mit den Daten arbeitet. Beide Zonen können ausgewählt werden.	Korrektheit (++) Erreichbarkeit (+)
Standardempfehlung „CIM“ anstelle von „HTTP“	SGAAIRE macht hier einen spezifischeren Vorschlag, nämlich die Verwendung von CIM. Für den Anwendungsfall muss geprüft werden, welche Daten über CIM übertragen werden können. Gegebenenfalls ist die Entwicklung eines separaten Profils erforderlich.	Nachvollziehbarkeit (++) Korrektheit (?) Erreichbarkeit (+)
<b>Falschfehler</b>	(„False Positives“)	
Fehlerhafte Domänenklassifikation „SSST Dashboard“	Das SSST Dashboard konnte keiner Domäne eindeutig zugeordnet werden.	Korrektheit (—) Nachvollziehbarkeit (—)

Tabelle A.45: Fallbeispiel B: Zusätzliche identifizierte Fehler sowie False Positives

	Beschreibung	Betroffene Qualitätsmerkmale
<b>Zusätzliche identifizierte Fehler</b>	(Weitere „True Positives“)	
Acronyme „DC“ nicht eingeführt	Das genannte Akronym wurden im Anwendungsfall nicht definiert.	Nachvollziehbarkeit (++) Vollständigkeit (+) Zugänglichkeit (+)
<b>Falschfehler</b>	(„False Positives“)	
Vorschlag für „Undefined/Proprietary Communication“	Die angegebene Kommunikation ist zu unpräzise und irrelevant. In dem Fall wäre der Einsatz von CIM-basierter Kommunikation auf Basis der vorliegenden Informationen naheliegender.	Relevanz (—) Nachvollziehbarkeit (—)

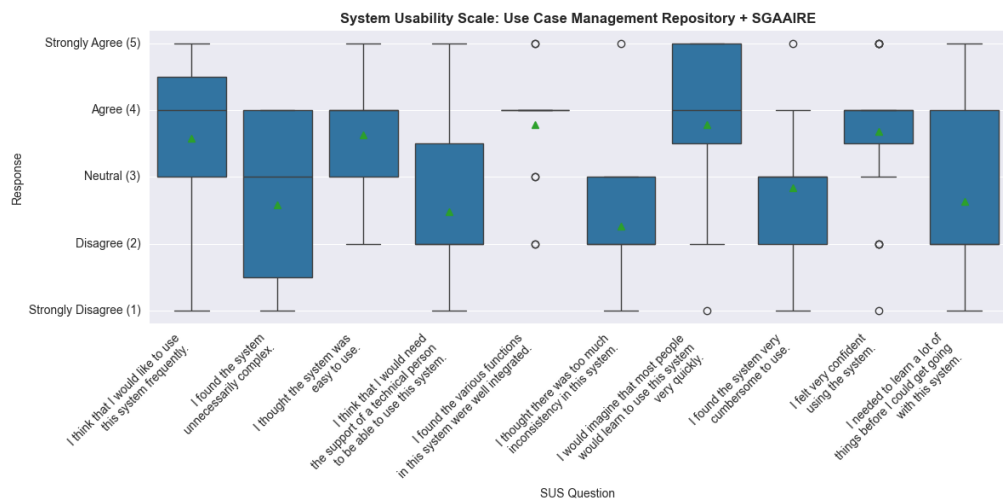
### A.6 Weitere Darstellungen aus der Evaluation

In diesem Abschnitt werden die im Rahmen der Evaluation erhobenen SUS-Daten in Form von Boxplots dargestellt. Die Abbildungen zeigen die Verteilung der Bewertungen für die einzelnen Fragen des SUS-Tests innerhalb der UCMR- sowie der KI-Gruppe.



Quelle: Eigene Darstellung

Abbildung A.1: Bewertung der einzelnen Fragen im SUS-Test für die UCMR-Gruppe



Quelle: Eigene Darstellung

Abbildung A.2: Bewertung der einzelnen Fragen im SUS-Test für die KI-Gruppe

## A.7 Anwendungsfallbeschreibungen zu SGAAIRE

In diesem Abschnitt werden die im Rahmen von SGAAIRE definierten Anwendungsfälle dargestellt. Sie beschreiben die zentralen Funktionen und Interaktionen des Systems sowie deren Bezug zu den zugrundeliegenden IEC 62559 Use Case-Beschreibungen und SGAM-Modellen. Die einzelnen Tabellen folgen dem etablierten Template nach [Coc01] und enthalten sowohl übergeordnete als auch spezialisierte Unteranwendungsfälle, die den modularen Aufbau von SGAAIRE abbilden.

Tabelle A.46: Anwendungsfallbeschreibung #1 - Bereitstellen von Vorschlägen für IEC 62559 Use Case-Beschreibungen

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #1</b>	Bereitstellen von Vorschläge für IEC 62559 Use Case-Beschreibungen
Verwendungskontext	Zur Verbesserung der Modellierung von IEC 62559 Use Case-Beschreibungen und deren Qualität.
Rahmen	Erstellung und Qualitätsmanagement von IEC 62559 Use Case-Beschreibungen
Level	Zusammenfassung
Primärer Akteur	Fach- oder Methodenexperten
<b>STAKEHOLDER &amp; INTERES-SIERTE</b>	<i>Fach- oder Methodenexperten:</i> Nutzen SGAAIRE zur Verbesserung der Modellierung bzw. der Qualität von IEC 62559 Use Case-Beschreibung <i>UCMR:</i> Integration des Anwendungsfalls in dem Tool
Voraussetzungen	Bestimmte Inhalte in der IEC 62559 Anwendungsfallbeschreibung (Konkretisierung der Unteranwendungsfälle)
Erfolgsbedingung	Vorgeschlagene Aspekte für eine IEC 62559 Use Case-Beschreibung (gemäß der Unteranwendungsfälle)
Gescheiterter Endschutz	Bei fehlenden Informationen wird eine Warnmeldung ausgegeben; Generelle Informationen zum Nachvollziehen gegeben, ansonsten werden keine Vorschläge bereitgestellt
Auslöser	1. Die nutzende Person möchte direkte Vorschläge erhalten 2. Die nutzende Person erhält basierend auf den Eingaben innerhalb des integrierten UCMR indirekte Vorschläge
<b>BESCHREIBUNG</b>	1. Die nutzende Person stellt direkt bzw. indirekt eine Anfrage für die Bereitstellung von Vorschlägen 2. SGAAIRE identifiziert basierend auf dem aktuellen Stand der IEC 62559 Use Case-Beschreibung mögliche Vorschläge und leitet diese den zugehörigen KI-Modulen (Zugehörige Unteranwendungsfälle) weiter 3. Die KI-Module von SGAAIRE übernehmen die Bestimmung von Vorschlägen 4. Die Vorschläge werden der nutzenden Person bereitgestellt
<b>ERWEITERUNGEN</b>	1a. Die nutzende Person stellt syntaktisch und semantisch falsche Anfragen: Stellt die Anfrage trotzdem 2a. Bei der Überprüfung der möglichen KI-Optionen lassen sich syntaktisch oder semantisch inkorrekte Inhalte der eingehenden Anfrage erkennen: Warnmeldung bereitstellen und KI-Modulaufruf überspringen 3a. Technische Fehler sind entstanden: Warnmeldung bereitstellen
<b>VARIATIONEN</b>	Keine

Tabelle A.47: Anwendungsfallbeschreibung #1.1 - SGAM-Domänen und -Zonen für Akteure, Komponenten sowie IEC 62559 Use Cases vorschlagen

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #1.1</b>	SGAM-Domänen und -Zonen für Akteure, Komponenten sowie IEC 62559 Use Cases vorschlagen
Verwendungskontext Level	– Primäre Aufgabe
Voraussetzungen	<i>Für die UC-Klassifikation:</i> Auflistung der Akteure, vollständige Beschreibung und Angabe des UC-Typs. <i>Für die Akteurs- und Komponentenklassifikation:</i> Name und/oder Beschreibung des Akteurs.
Erfolgsbedingung	Vorschläge für SGAM-Zonen und -Domänen.
<b>BESCHREIBUNG</b>	<ol style="list-style-type: none"> <li>1. Die nutzende Person stellt direkt oder indirekt eine Anfrage zur Bereitstellung von Vorschlägen.</li> <li>2. Überprüfung der aufgeführten Voraussetzungen und anschließender Aufruf der SGAM Klassifikationsmodelle.</li> <li>3. Klassifikation innerhalb der SGAM-Zonen und -Domänen.</li> <li>4. Die Vorschläge werden der nutzenden Person bereitgestellt.</li> </ol>
<b>ERWEITERUNGEN</b>	<ol style="list-style-type: none"> <li>1a. Die nutzende Person stellt syntaktisch oder semantisch fehlerhafte Anfragen: Die Anfrage wird dennoch verarbeitet.</li> <li>2a. Die Anfrage ist fehlerhaft oder die Voraussetzungen sind nicht erfüllt: Warnmeldung bereitstellen und Aufruf des KI-Moduls überspringen.</li> <li>3a. Es treten technische Fehler auf: Warnmeldung bereitstellen.</li> </ol>

Tabelle A.48: Anwendungsfallbeschreibung #1.2 - Erkennen von abweichenden und nicht standardisierten Termen

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #1.2</b>	Erkennen von abweichenden und nicht standardisierten Termen
Verwendungskontext Level	– Primäre Aufgabe
Voraussetzungen Erfolgsbedingung	Ausgefüllte Freitextfelder. Erkennung von abweichenden und nicht standardisierten Termen.
<b>BESCHREIBUNG</b>	<ol style="list-style-type: none"> <li>1. Die nutzende Person stellt direkt oder indirekt eine Anfrage zur Bereitstellung von Vorschlägen.</li> <li>2. Überprüfung der aufgeführten Voraussetzungen und anschließender Aufruf des KI-Moduls zur Erkennung abweichender und nicht standardisierter Terme gemäß der Use Case-Template-Sektion 7. <i>Common Terms and Definitions Templates</i>.</li> <li>3. Erkennung der abweichenden und nicht standardisierten Terme gemäß der Use Case-Template-Sektion 7. <i>Common Terms and Definitions Templates</i>.</li> <li>4. Die Vorschläge werden der nutzenden Person bereitgestellt.</li> </ol>
<b>ERWEITERUNGEN</b>	<ol style="list-style-type: none"> <li>1a. Die nutzende Person stellt syntaktisch oder semantisch fehlerhafte Anfragen: Die Anfrage wird dennoch verarbeitet.</li> <li>2a. Die Anfrage ist fehlerhaft oder die Voraussetzungen sind nicht erfüllt: Warnmeldung bereitstellen und Aufruf des KI-Moduls überspringen.</li> <li>3a. Es treten technische Fehler auf: Warnmeldung bereitstellen.</li> </ol>

Tabelle A.49: Anwendungsfallbeschreibung #1.3 - Vorschläge zum Hinzufügen, Entfernen und Bearbeiten von Elementen

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #1.3</b>	Vorschläge zum Hinzufügen, Entfernen und Bearbeiten von Elementen
Verwendungskontext Level	– Primäre Aufgabe
Voraussetzungen Erfolgsbedingung	– Erkennung relevanter Elemente sowie Vorschläge zum Entfernen oder Bearbeiten basierend auf der Beschreibung.
<b>BESCHREIBUNG</b>	<ol style="list-style-type: none"> <li>1. Die nutzende Person stellt direkt oder indirekt eine Anfrage zur Bereitstellung von Vorschlägen.</li> <li>2. Überprüfung der aufgeführten Voraussetzungen und anschließender Aufruf des KI-Moduls zur Erkennung relevanter Elemente.</li> <li>3. Identifikation redundanter oder aktualisierbarer Elemente.</li> <li>4. Die Vorschläge werden der nutzenden Person bereitgestellt.</li> </ol>
<b>ERWEITERUNGEN</b>	<ol style="list-style-type: none"> <li>1a. Die nutzende Person stellt syntaktisch oder semantisch fehlerhafte Anfragen: Die Anfrage wird dennoch verarbeitet.</li> <li>2a. Die Anfrage ist fehlerhaft oder die Voraussetzungen sind nicht erfüllt: Warnmeldung bereitstellen und Aufruf des KI-Moduls überspringen.</li> <li>3a. Es treten technische Fehler auf: Warnmeldung bereitstellen.</li> </ol>

Tabelle A.50: Anwendungsfallbeschreibung #1.4 - Service Steps in der Step-by-Step Analysis vorschlagen

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #1.4</b>	Service Steps in der Step-by-Step Analysis vorschlagen
Verwendungskontext Level	– Primäre Aufgabe
Voraussetzungen Erfolgsbedingung	Prozessbeschreibung in einem Step sowie die beteiligten Akteure. Vorschlag eines Service Steps gemäß Standard.
<b>BESCHREIBUNG</b>	<ol style="list-style-type: none"> <li>1. Die nutzende Person stellt direkt oder indirekt eine Anfrage zur Bereitstellung von Vorschlägen.</li> <li>2. Überprüfung der aufgeführten Voraussetzungen und anschließender Aufruf des KI-Moduls zur Klassifikation des Service Steps.</li> <li>3. Klassifikation des Service Steps.</li> <li>4. Die Vorschläge werden der nutzenden Person bereitgestellt.</li> </ol>
<b>ERWEITERUNGEN</b>	<ol style="list-style-type: none"> <li>1a. Die nutzende Person stellt syntaktisch oder semantisch fehlerhafte Anfragen: Die Anfrage wird dennoch verarbeitet.</li> <li>2a. Die Anfrage ist fehlerhaft oder die Voraussetzungen sind nicht erfüllt: Warnmeldung bereitstellen und Aufruf des KI-Moduls überspringen.</li> <li>3a. Es treten technische Fehler auf: Warnmeldung bereitstellen.</li> </ol>

Tabelle A.51: Anwendungsfallbeschreibung #1.5 - Erkennen von Inkonsistenzen bei Akteuren, Informationsaustauschen und Anforderungen gegenüber dem Use Case-Typ

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #1.5</b>	Erkennen von Inkonsistenzen bei Akteuren, Informationsaustauschen und Anforderungen gegenüber dem Use Case-Typ
Verwendungskontext Level	– Primäre Aufgabe
Voraussetzungen Erfolgsbedingung	Das Feld <i>Use Case-Typ</i> enthält eine der folgenden Informationen: BUC, SUC oder HLUC. Erkennung von Akteuren, Informationsaustauschen und Anforderungen, die nicht dem jeweiligen Use Case-Typ entsprechen.
<b>BESCHREIBUNG</b>	<ol style="list-style-type: none"> <li>1. Die nutzende Person stellt direkt oder indirekt eine Anfrage zur Bereitstellung von Vorschlägen.</li> <li>2. Überprüfung der aufgeführten Voraussetzungen und anschließender Aufruf des KI-Moduls zur Inkonsistenzerkennung.</li> <li>3. Durchführung der Inkonsistenzerkennung durch das KI-Modul.</li> <li>4. Bereitstellung der Vorschläge an die nutzende Person.</li> </ol>
<b>ERWEITERUNGEN</b>	<ol style="list-style-type: none"> <li>1a. Die nutzende Person stellt syntaktisch oder semantisch fehlerhafte Anfragen: Die Anfrage wird dennoch verarbeitet.</li> <li>2a. Die Anfrage ist fehlerhaft oder die Voraussetzungen sind nicht erfüllt: Warnmeldung bereitstellen und Aufruf des KI-Moduls überspringen.</li> <li>3a. Es treten technische Fehler auf: Warnmeldung bereitstellen.</li> </ol>

Tabelle A.52: Anwendungsfallbeschreibung #2 - Bereitstellen von Vorschlägen für SGAM Modelle

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #2</b>	Bereitstellen von Vorschlägen für SGAM Modelle
Verwendungskontext	Zur Verbesserung der Modellierung von SGAM Modellen und deren Qualität.
Rahmen	Erstellung und Qualitätsmanagement von SGAM Modellen
Level	Zusammenfassung
Primärer Akteur	Fach- oder Methodenexperten
<b>STAKEHOLDER &amp; INTER-ESSIERTE</b>	<i>Fach- oder Methodenexperten:</i> Nutzen SGAAIRE zur Verbesserung der Modellierung bzw. der Qualität von SGAM Modellen. <i>UCMR:</i> Integration des Anwendungsfalls in das Tool. <i>SGAM Add-on:</i> Erweiterung des UCMR zur Bereitstellung der CRUD-Operationen für die Realisierung der Anforderungen.
Voraussetzungen	Bestimmte Inhalte in der IEC 62559 -Anwendungsfallbeschreibung und in den SGAM Modellen (Konkretisierung der Unteranwendungsfälle).
Erfolgsbedingung	Erfolgreiche Bereitstellung vorgeschlagener Aspekte für SGAM Modelle (gemäß den Unteranwendungsfällen).
Gescheiterter Endzustand	Bei fehlenden Informationen wird eine Warnmeldung ausgegeben; allgemeine Hinweise zum Nachvollziehen werden bereitgestellt, ansonsten keine Vorschläge.
Auslöser	Die nutzende Person möchte direkte Vorschläge gemäß den Unteranwendungsfällen erhalten.
<b>BESCHREIBUNG</b>	<ol style="list-style-type: none"> <li>1. Die nutzende Person stellt direkt oder indirekt eine Anfrage zur Bereitstellung von Vorschlägen.</li> <li>2. SGAAIRE identifiziert auf Basis der aktuellen IEC 62559 Use Case-Beschreibung bzw. der SGAM Modelle mögliche Vorschläge und leitet diese an die zugehörigen KI-Module (Unteranwendungsfälle) weiter.</li> <li>3. Die KI-Module von SGAAIRE generieren die Vorschläge.</li> <li>4. Die Vorschläge werden der nutzenden Person bereitgestellt.</li> </ol>
<b>ERWEITERUNGEN</b>	<ol style="list-style-type: none"> <li>1a. Die nutzende Person stellt syntaktisch oder semantisch fehlerhafte Anfragen: Die Anfrage wird dennoch verarbeitet.</li> <li>2a. Bei der Überprüfung der möglichen KI-Optionen werden syntaktisch oder semantisch inkorrekte Inhalte der eingehenden Anfrage erkannt: Warnmeldung bereitstellen und Aufruf des KI-Moduls überspringen.</li> <li>3a. Es treten technische Fehler auf: Warnmeldung bereitstellen.</li> </ol>
<b>VARIATIONEN</b>	Keine.

Tabelle A.53: Anwendungsfallbeschreibung #2.1 - Generieren von SGAM Arbeitsmodellen

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #2.1</b>	Generieren von SGAM Arbeitsmodellen
Verwendungskontext Level	– Primäre Aufgabe
Voraussetzungen	Ausgefüllte IEC 62559 Use Case-Beschreibung.
<b>BESCHREIBUNG</b>	<ol style="list-style-type: none"> <li>1. Die nutzende Person stellt eine Anfrage zur Generierung von SGAM Arbeitsmodellen.</li> <li>2. Überprüfung der aufgeführten Voraussetzungen und anschließender Aufruf des KI-Moduls zur Generierung.</li> <li>3. Generierung von SGAM Arbeitsmodellen basierend auf der Beschreibung.</li> <li>4. Die Vorschläge werden der nutzenden Person bereitgestellt.</li> </ol>
<b>ERWEITERUNGEN</b>	<ol style="list-style-type: none"> <li>1a. Die nutzende Person stellt syntaktisch oder semantisch fehlerhafte Anfragen: Die Anfrage wird dennoch verarbeitet.</li> <li>2a. Die Anfrage ist fehlerhaft oder die Voraussetzungen sind nicht erfüllt: Warnmeldung bereitstellen und Aufruf des KI-Moduls überspringen.</li> <li>3a. Es treten technische Fehler auf: Warnmeldung bereitstellen.</li> </ol>

Tabelle A.54: Anwendungsfallbeschreibung #2.2 - Überprüfen der Konsistenz von SGAM Modellen gegenüber den IEC 62559 Use Case-Beschreibungen

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #2.2</b>	Überprüfen der Konsistenz von SGAM Modellen gegenüber den IEC 62559 Use Case-Beschreibungen
Verwendungskontext Level	– Primäre Aufgabe
Voraussetzungen	Ausgefüllte IEC 62559 Use Case-Beschreibung und SGAM Modell.
<b>BESCHREIBUNG</b>	<ol style="list-style-type: none"> <li>1. Die nutzende Person stellt eine Anfrage zur Konsistenzprüfung von SGAM Modellen auf Basis einer IEC 62559 Use Case-Beschreibung.</li> <li>2. Überprüfung der aufgeführten Voraussetzungen und anschließender Aufruf des KI-Moduls zur SGAM/UC-Konsistenzprüfung.</li> <li>3. Überprüfung der SGAM Modelle basierend auf der IEC 62559 Use Case-Beschreibung.</li> <li>4. Die Ergebnisse werden der nutzenden Person bereitgestellt.</li> </ol>
<b>ERWEITERUNGEN</b>	<ol style="list-style-type: none"> <li>1a. Die nutzende Person stellt syntaktisch oder semantisch fehlerhafte Anfragen: Die Anfrage wird dennoch verarbeitet.</li> <li>2a. Die Anfrage ist fehlerhaft oder die Voraussetzungen sind nicht erfüllt: Warnmeldung bereitstellen und Aufruf des KI-Moduls überspringen.</li> <li>3a. Es treten technische Fehler auf: Warnmeldung bereitstellen.</li> </ol>

Tabelle A.55: Anwendungsfallbeschreibung #2.3 - Identifikation von Inkonsistenzen innerhalb von SGAM Modellen

Quelle: Template basierend auf [Coc01]

<b>ANWENDUNGSFALL #2.3</b>	Identifikation von Inkonsistenzen innerhalb von SGAM Modellen
Verwendungskontext Level	– Primäre Aufgabe
Voraussetzungen	Ausgefülltes SGAM Modell.
<b>BESCHREIBUNG</b>	<ol style="list-style-type: none"> <li>1. Die nutzende Person stellt eine Anfrage zur Konsistenzprüfung von SGAM Modellen.</li> <li>2. Überprüfung der aufgeführten Voraussetzungen und anschließender Aufruf des KI-Moduls zur SGAM-Konsistenzprüfung.</li> <li>3. Überprüfung der SGAM Modelle auf interne Konsistenz.</li> <li>4. Die Ergebnisse werden der nutzenden Person bereitgestellt.</li> </ol>
<b>ERWEITERUNGEN</b>	<ol style="list-style-type: none"> <li>1a. Die nutzende Person stellt syntaktisch oder semantisch fehlerhafte Anfragen: Die Anfrage wird dennoch verarbeitet.</li> <li>2a. Die Anfrage ist fehlerhaft oder die Voraussetzungen sind nicht erfüllt: Warnmeldung bereitstellen und Aufruf des KI-Moduls überspringen.</li> <li>3a. Es treten technische Fehler auf: Warnmeldung bereitstellen.</li> </ol>



## Abkürzungen

<b>AAL</b>	Ambient Assisted Living
<b>AD</b>	Acronym Disambiguation
<b>AHP</b>	Analytic Hierarchy Process
<b>AId</b>	Acronym Identification
<b>AUC</b>	Area Under the Curve
<b>API</b>	Application Programming Interfaces
<b>BIO</b>	Beginning-Inside-Outside
<b>BERT</b>	Bidirectional Encoder Representations from Transformers
<b>BUC</b>	Business Use Case
<b>CEM</b>	Customer Energy Manager
<b>CFF</b>	Cognitive Forcing Function
<b>CGMES</b>	Common Grid Model Exchange Standard
<b>CI</b>	Consistence Index
<b>CIM</b>	Common Information Model
<b>CLM</b>	Causal Language Modelling
<b>CRUD</b>	Create Read Update Delete
<b>CR</b>	Consistency Ratio
<b>CS</b>	Constituent System
<b>DER</b>	Distributed Energy Resources
<b>DNN</b>	Deep Neural Network
<b>DL</b>	Deep Learning
<b>DSO</b>	Distribution System Operator
<b>DSRM</b>	Design Science Research Methodology
<b>DSRF</b>	Design Science Research Framework
<b>EA</b>	Enterprise Architect
<b>ELMo</b>	Embeddings from Language Model
<b>ENTSO-E</b>	European Network of Transmission System Operators for Electricity
<b>ESP</b>	Energy Service Provider
<b>EV</b>	Electric Vehicle
<b>FLAN-T5</b>	Fine-tuned Language Net Text-to-Text Transfer Transformer
<b>GoM</b>	Grundsätze ordnungsmäßiger Modellierung
<b>GPT</b>	Generative Pretrained Transformer

---

<b>GWAC</b>	GridWise Architecture Council
<b>HCD</b>	Human-Centred Design
<b>HTTP</b>	Hypertext Transfer Protocol
<b>HLUC</b>	High-Level Use Case
<b>IDE</b>	Integrated Development Environment
<b>IdiNA</b>	Intelligente Digitalisierung der Energieversorgung zur Optimierung des Netzbetriebs und zur Erhöhung der Akzeptanz
<b>IKT</b>	Informations- und Kommunikationstechnologien
<b>IOP</b>	Interoperabilität
<b>IS</b>	Information Systems
<b>JAXB</b>	Jakarta XML Binding
<b>KI</b>	Künstliche Intelligenz
<b>KNN</b>	Künstliches neuronales Netz
<b>KPI</b>	Key Performance Indicator
<b>LM</b>	Language Model
<b>LLaMA</b>	Large Language Model Meta AI
<b>LLM</b>	Large Language Model
<b>LLUC</b>	Low-Level Use Case
<b>ML</b>	Machine Learning
<b>MLM</b>	Masked Language Modeling
<b>MBSE</b>	Model-Based Systems Engineering
<b>MDD</b>	Modellgetriebene Softwareentwicklung
<b>MSE</b>	Mean Squared Error
<b>MAE</b>	Mean Absolute Error
<b>NER</b>	Named Entity Recognition
<b>NFA</b>	Nicht-funktionale Anforderungen
<b>NIST</b>	National Institute of Standards and Technology
<b>NSP</b>	Next Sentence Prediction
<b>NLG</b>	Natural Language Generation
<b>NLU</b>	Natural Language Understanding
<b>NLP</b>	Natural Language Processing
<b>NWA</b>	Nutzwertanalyse
<b>ROC</b>	Receiver Operating Characteristic
<b>RAG</b>	Retrieval Augmented Generation

---

<b>RDS</b>	Reference Designation System
<b>RegEx</b>	Regular Expression
<b>RI</b>	Random Consistency Index
<b>RoBERTa</b>	Robustly Optimized BERT Pretraining Approach
<b>DeBERTa</b>	Decoding-enhanced BERT with Disentangled Attention
<b>SGAAIRE</b>	Smart Grid Assisting Artificial Intelligence in Requirements Engineering
<b>SGAM</b>	Smart Grid Architecture Model
<b>SGD</b>	Stochastic Gradient Descent
<b>SIA</b>	Seamless Integration Architecture
<b>SoC</b>	System-of-Concern
<b>SoS</b>	System-of-Systems
<b>SysML</b>	Systems Modeling Language
<b>SUC</b>	System Use Case
<b>SUS</b>	System Usability Scale
<b>TC</b>	Technical Committee
<b>TUC</b>	Technical Use Case
<b>TSO</b>	Transmission System Operator
<b>UCMR</b>	Use Case Management Repository
<b>UCR</b>	Use Case Repository
<b>UML</b>	Unified Modeling Language
<b>UX</b>	User Experience
<b>VM</b>	Virtuelle Maschine
<b>XSD</b>	XML Schema Definition
<b>XAI</b>	eXplainable Artificial Intelligence
<b>TL</b>	Transfer Learning
<b>RCT</b>	Randomized Controlled Trial
<b>CLI</b>	Command-Line Interface
<b>RNN</b>	Rekurrentes neuronales Netz
<b>LSTM</b>	Long Short-Term Memory
<b>WG</b>	Working Group



## Abbildungen

1.1	Darstellung möglicher Akteure im Smart Grid . . . . .	1
1.2	Ausschnitt eines SGAM Modells mit potenziellen Fehlern . . . . .	2
1.3	Abhängigkeiten zwischen IEC 62559 Use Cases, SGAM Modellen und weiterführenden Tätigkeiten . . . . .	3
1.4	Erstellungsprozess von IEC 62559 Use Cases und SGAM Modellen . . . . .	4
2.1	Visualisierung der Konkretisierung des DSRF für das Dissertationsvorhaben . . . . .	15
2.2	Vorgehensmodell der Arbeit . . . . .	18
2.3	Zuordnung der Forschungsmethoden zu den Forschungsfragen . . . . .	18
3.1	Unterschiedliche Arten der Expertise bei der Erstellung von IEC 62559 Use Cases sowie SGAM Modellen . . . . .	23
3.2	GridWise Architecture Council Interoperability Context-Setting Framework . . . . .	26
3.3	Konzeptionelles Modell von Architekturbeschreibungen gemäß ISO/IEC/IE-EE 42010 . . . . .	28
3.4	Aufbau der IEC 62559 Standard Serie . . . . .	30
3.5	Zusammenfassung von Use Cases in einem Cluster . . . . .	36
3.6	Rahmenwerk zur Definition des XML-basierten Datenaustauschformats des IEC 62559-3 Standards . . . . .	37
3.7	Generelles Design des Use Case Management Repository . . . . .	38
3.8	Screenshot des OFFIS Use Case Management Repository . . . . .	39
3.9	Screenshot des BRIDGE Use Case Repository . . . . .	39
3.10	Smart Grid Architecture Model (SGAM) . . . . .	40
3.11	Mapping der Interoperabilitätskategorien aus dem SGAM Framework und GWAC Interoperability Context-Setting Framework . . . . .	41
3.12	SGAM Grid Plane: Domänen und Zonen . . . . .	43
3.13	Kategorisierung der verschiedenen Abstraktionsgrade über die SGAM Interoperabilitätsebenen . . . . .	45
3.14	Beispielhaftes Mapping des IEC 62559 Use Case-Template und dem SGAM Framework mit Berücksichtigung der Abstraktionsgrade . . . . .	46
3.15	Modellierung von SGAM Modelle anhand von IEC 62559-2 Use Cases . . . . .	47
4.1	(Binäres) Klassifikationsproblem . . . . .	51
4.2	Regressionsproblem . . . . .	51
4.3	Vereinfachte Abbildung eines (tiefen) KNN . . . . .	54
4.4	Aktivierungsfunktionen: Sigmoid, Tanh und ReLU . . . . .	55

4.5	Beispielhafte Darstellung von semantischer Nähe und Distanz im eingebetteten Vektorraum . . . . .	59
4.6	Architektur eines Transformers . . . . .	61
4.7	Beispiel für Masked Language Modeling (MLM) mit BERT als Modell	65
4.8	Beispiel für Tokenklassifikation mit BIO-Tagging . . . . .	72
4.9	Retrieval Augmented Generation (RAG) . . . . .	81
5.1	Auswirkungen der Ergebnisse aus Kapitel 4 auf das Forschungsvorhaben	83
5.2	Flussdiagramm der Durchführung der Studienteilnehmenden . . . . .	86
5.3	Exemplarische Codierung für "Herausforderungen bei der Erstellung von IEC 62559 Use Cases" . . . . .	89
5.4	Boxplot der Resonanzwerte der erhobenen funktionalen Anforderungen für IEC 62559 Use Cases (n=18) und SGAM Modelle (n=19) . . . . .	94
5.5	Prozess zur Herleitung der funktionalen Anforderungen . . . . .	97
5.6	Verwendete Schritte der Nutzwertanalyse . . . . .	98
5.7	Selektion der funktionalen Anforderungen für SGAAIRE über den Grenzwert $\Theta = 0.7$ . . . . .	107
6.1	Auswirkungen der Ergebnisse aus Kapitel 6 auf das Forschungsvorhaben	109
6.2	Visualisierung des Prozesses zur Herleitung der Bewertungsmetriken . .	110
6.3	Vorliegendes Modellverständnis und Herausforderungen im Kontext . .	116
7.1	Auswirkungen der Ergebnisse aus Kapitel 7 auf das Forschungsvorhaben	121
7.2	Zugrundeliegender Prozess der Erstellung von IEC 62559 Use Case-Beschreibungen und SGAM Modellen . . . . .	123
7.3	Schematische Darstellung des V-Modells . . . . .	127
7.4	Systemkontextmodell von SGAAIRE . . . . .	128
7.5	Systemarchitektur von SGAAIRE . . . . .	130
7.6	UML-Anwendungsfalldiagramm für die Interaktionen mit SGAAIRE für die IEC 62559 Use Case-bezogenen funktionalen Anforderungen . .	132
7.7	Sequenzdiagramm für die Interaktionen mit SGAAIRE für die IEC 62559 Use Case-bezogenen funktionalen Anforderungen . . . . .	134
7.8	Sequenzdiagramm für die Interaktionen mit SGAAIRE für die SGAM-bezogenen funktionalen Anforderungen . . . . .	135
7.9	Visualisierung des Konzeptes für die Verwendung von Acronym Identification (AId) und Acronym Disambiguation (AD) für das KI-Modul zur Erkennung von abweichenden und nicht-standardisierten Termen . . . .	139
7.10	Beispiele von Inkonsistenzen zw. SGAM Modell und IEC 62559 UC . .	141
7.11	Konzeptuelle Zuordnung der Sektionen des IEC 62559 Use Case-Templates im Rahmen des KI-basierten Mappings . . . . .	143

7.12	Beispiele zur Elementableitung für die IEC 62559 Use Case-Methodik mit Hilfe des SGAM Frameworks . . . . .	147
7.13	Modellübersicht von SGAAIRE . . . . .	148
7.14	Einordnung des Akteurs <i>EV</i> im SGAM . . . . .	149
7.15	Einordnung des Standards <i>CIM</i> im SGAM . . . . .	149
7.16	Beispielhafte Klassifikation einer Aktivität im Feld „Service“ im IEC 62559 Use Case-Template . . . . .	150
7.17	Mögliche Elementkombinationen innerhalb eines SGAM Modells . . . . .	151
7.18	Beispielhafte Anwendung des Modells für Erhalten der Vorschläge für die Kommunikationsebene . . . . .	151
7.19	Einordnung zweier Systeme nach Abstraktionsgrad . . . . .	152
8.1	Auswirkungen der Ergebnisse aus Kapitel 8 auf das Forschungsvorhaben	155
8.2	Übersicht der SGAAIRE API . . . . .	160
8.3	Screenshot zur Illustration des IEC 62559 Use Case-Hub . . . . .	161
8.4	Screenshot zur Illustration eines Service Step-Vorschlag für ein IEC 62559 Use Case . . . . .	162
8.5	Screenshot zur Illustration des SGAM Qualitätschecks . . . . .	163
8.6	Screenshot zur Generierung von SGAM Modellen . . . . .	164
8.7	Datenpipeline für die Überführung in einheitliche XML-serialisierte IEC 62559 Use Cases sowie SGAM Modelle . . . . .	167
8.8	Domänenkonvexität . . . . .	179
8.9	Zonenkonvexität . . . . .	179
9.1	Auswirkungen der Ergebnisse aus Kapitel 9 auf das Forschungsvorhaben	191
9.2	SGAM Modell (Referenz) für Fallbeispiel A . . . . .	193
9.3	SGAM Modell (Evaluation) für Fallbeispiel A . . . . .	194
9.4	Fallbeispiel A: Konfusionsmatrix zur Fehleridentifikation durch die Anwendung von SGAAIRE . . . . .	197
9.5	SGAM Modell (Referenz) für Fallbeispiel B . . . . .	198
9.6	SGAM Modell (Evaluation) für Fallbeispiel B . . . . .	199
9.7	Fallbeispiel B: Konfusionsmatrix zur Fehleridentifikation durch die Anwendung von SGAAIRE . . . . .	201
9.8	Vorgehensweise zur Erhebung, Kodierung und Analyse der Evaluationsdaten . . . . .	202
9.9	Aufbau der vorliegenden RCT . . . . .	203
9.10	CONSORT-Diagramm zur Darstellung des Teilnehmerflusses . . . . .	206
9.11	Teilnehmeraufteilung (n=27) in Kontrollgruppe UCMR (n=13) und Testgruppe UCMR+SGAAIRE (n=14) nach Erfahrungsniveau . . . . .	207

---

9.12	Teilnehmeraufteilung (n=27) in Kontrollgruppe UCMR (n=13) und Testgruppe UCMR+SGAAIRE (n=14) nach Motivation . . . . .	208
9.13	Vergleich von Kontroll- und Testgruppe hinsichtlich der identifizierten eingebauten Fehler . . . . .	209
9.14	Vergleich von Kontroll- und Testgruppe hinsichtlich der identifizierten nicht-eingebauten (zusätzlichen) Fehler . . . . .	210
9.15	Vergleich von Kontroll- und Testgruppe hinsichtlich der Qualitätsverbesserungen . . . . .	212
9.16	Vergleich von Kontroll- und Testgruppe hinsichtlich der Auswirkungen auf einzelne Qualitätsmerkmale (Min-Max Skalierung) . . . . .	213
9.17	Boxplot der erreichten Punktzahlen im SUS-Test . . . . .	216
A.1	Bewertung der einzelnen Fragen im SUS-Test für die UCMR-Gruppe . . . . .	258
A.2	Bewertung der einzelnen Fragen im SUS-Test für die KI-Gruppe . . . . .	258

## Tabellenverzeichnis

1.1	Vergleich der wesentlichen verwandten Arbeiten . . . . .	10
2.1	Einordnung des Vorgehensmodells in den Aufbau der Arbeit . . . . .	19
2.2	Aufführung beispielhafter Definitionen für Künstliche Intelligenz . . . . .	20
2.3	Zielsetzungen von KI nach [Win92] . . . . .	20
5.1	Übersicht über die ungefilterten Metadaten zur Methodennutzung . . . . .	88
5.2	Übersicht über die demografischen Informationen . . . . .	88
5.3	Übersicht aus den funktionalen Anforderungen, hergeleitet aus den Modell- anforderungen (vgl. Abschnitt 6.2) . . . . .	93
5.4	Übersicht der von den Umfrageteilnehmenden vorgeschlagenen funktio- nalen Anforderungen . . . . .	95
5.5	Gewichte und Rangordnung der Kriterien . . . . .	103
5.6	Stochastisch von [Saa87] erzeugte RI-Werte mit n=500 zufällig gene- rierte Bewertungsmatrizen . . . . .	103
5.7	Bestimmte Werte für CR, CI, $\lambda_{max}$ und RI auf Basis der vierten Iteration	103
5.8	Ausprägungswerte der Alternativen . . . . .	105
5.9	Aufführung der erzielten (Teil-)Nutzwerte der gegenübergestellten Al- ternativen . . . . .	106
5.10	Übersicht der selektierten funktionalen Anforderungen . . . . .	107
6.1	Betrachtete Literatur zur Anforderungserhebung von IEC 62559 Use Case-Beschreibungen sowie SGAM Modellen . . . . .	112
6.2	Anforderungen für IEC 62559 Use Case-Beschreibungen (M_UC_ID) . . . . .	113
6.3	Anforderungen für SGAM Modelanforderungen (M_SGAM_ID) . . . . .	114
6.4	Anforderungen für generische Modelanforderungen (M_G_ID) . . . . .	114
6.5	Abgeleitete Qualitätsdimensionen und deren Metriken für IEC 62559 Use Case-Beschreibungen und SGAM Modelle nach dem Anforderungs- mapping . . . . .	119
7.1	Einschätzung der identifizierten Herausforderungen auf die funktionalen Anforderungen . . . . .	124
7.2	Referenzen der Anwendungsfallbeschreibungen sowie Abbildungen für den Use Case-Teil . . . . .	133
8.1	Übersicht der extrahierten IEC 62559 Use Cases . . . . .	165
8.2	Übersicht der extrahierten SGAM Modelle . . . . .	165
8.3	Datenquellen für die KI-Modelle (M1-M6) . . . . .	170
8.4	Übersicht der Datensätze und deren Aufteilung für die KI-Modelle M1–M6	171

8.5	Überblick der verwendeten Sprachmodelle (LMs) . . . . .	174
8.6	Überblick der verwendeten großen Sprachmodelle (LLMs) . . . . .	174
8.7	Auswahl der Hyperparameter beim Fine-Tuning . . . . .	174
8.8	Modellergebnisse für M1: SGAM Domänen Klassifikation . . . . .	177
8.9	Modellergebnisse für M2: SGAM Zonen Klassifikation . . . . .	181
8.10	Modellergebnisse für M3: SGAM IOP Klassifikation . . . . .	182
8.11	Modellergebnisse für M4: Service Step Klassifikation . . . . .	183
8.12	Modellergebnisse für M5: SGAM Element Proposer . . . . .	184
8.13	Modellergebnisse für M6: LoA Regression . . . . .	186
8.14	Modellergebnisse für M7a & b: Acronym Identification und -Disambiguation	189
9.1	Fallbeispiel A: Beschreibung der eingebetteten Fehler im IEC 62559 Use Case . . . . .	195
9.2	Fallbeispiel A: Beschreibung der eingebetteten Fehler im SGAM Modell	196
9.3	Fallbeispiel B: Beschreibung der eingebetteten Fehler im IEC 62559 Use Case . . . . .	200
9.4	Fallbeispiel B: Beschreibung der eingebetteten Fehler im SGAM Modell	200
9.5	Übersicht der aufgestellten Hypothesen . . . . .	205
9.6	Statistische Kennzahlen der erreichten Punktzahlen im SUS-Test . . . . .	215
9.7	Vergleich der Mittelwerte, Standardabweichungen und p-Werte der bei- den Gruppen . . . . .	217
9.8	Erfüllungsgrade der funktionalen Anforderungen . . . . .	226
9.9	Erfüllungsgrade der nicht-funktionalen Anforderungen . . . . .	227
9.10	Im Datensatz berücksichtigte IEC 62559 Use Cases sowie SGAM Mo- delle (Quelldateien) . . . . .	228
9.11	Datensatz für das Trainieren der KI-Modelle . . . . .	228
A.1	Anwendungsfall #1 - Sektion „1.1 Name of the use case“ . . . . .	236
A.2	Anwendungsfall #1 - Sektion „1.2 Version management“ . . . . .	236
A.3	Anwendungsfall #1 - Sektion „1.3 Scope and objectives of use case“ . . . . .	236
A.4	Anwendungsfall #1 - Sektion „1.4 Narrative of use case“ . . . . .	237
A.5	Anwendungsfall #1 - Sektion „1.5 Key performance indicators (KPI)“ . . . . .	237
A.6	Anwendungsfall #1 - Sektion „1.6 Use case conditions“ . . . . .	238
A.7	Anwendungsfall #1 - Sektion „1.7 Further information to the use case for classification / mapping“ . . . . .	238
A.8	Anwendungsfall #1 - Sektion „1.8 General remarks“ . . . . .	238
A.9	Anwendungsfall #1 - Sektion „2 Diagram(s) of use case“ . . . . .	238
A.10	Anwendungsfall #1 - Sektion „3.1 Actors (Grouping: Primary Actors)“ . . . . .	239
A.11	Anwendungsfall #1 - Sektion „3.1 Actors (Grouping: Further Actors)“ . . . . .	239
A.12	Anwendungsfall #1 - Sektion „3.2 References“ . . . . .	240

A.13 Anwendungsfall #1 - Sektion „4.1 Overview of scenarios“ . . . . .	241
A.14 Anwendungsfall #1 - Sektion „4.2 Scenario (1)“ . . . . .	242
A.15 Anwendungsfall #1 - Sektion „4.2 Scenario (2)“ . . . . .	243
A.16 Anwendungsfall #1 - Sektion „4.2 Scenario (3)“ . . . . .	243
A.17 Anwendungsfall #1 - Sektion „5. Information Exchanged“ . . . . .	244
A.18 Anwendungsfall #1 - Sektion „6. Requirements (optional)“ . . . . .	244
A.19 Anwendungsfall #1 - Sektion „7. Common terms and definitions“ . . . . .	244
A.20 Anwendungsfall #1 - Sektion „8. Custom information (optional)“ . . . . .	244
A.21 Anwendungsfall #2 - Sektion „1.1 Name of the use case“ . . . . .	245
A.22 Anwendungsfall #2 - Sektion „1.2 Version management“ . . . . .	245
A.23 Anwendungsfall #2 - Sektion „1.3 Scope and objectives of use case“ . . . . .	245
A.24 Anwendungsfall #2 - Sektion „1.4 Narrative of use case“ . . . . .	246
A.25 Anwendungsfall #2 - Sektion „1.5 Key performance indicators (KPI)“ . . . . .	246
A.26 Anwendungsfall #2 - Sektion „1.6 Use case conditions“ . . . . .	247
A.27 Anwendungsfall #2 - Sektion „1.7 Further information to the use case for classification / mapping“ . . . . .	247
A.28 Anwendungsfall #2 - Sektion „1.8 General remarks“ . . . . .	247
A.29 Anwendungsfall #2 - Sektion „2 Diagram(s) of use case“ . . . . .	248
A.30 Anwendungsfall #2 - Sektion „3.1 Actors (Grouping: Application-related Actors)“ . . . . .	248
A.31 Anwendungsfall #2 - Sektion „3.1 Actors (Grouping: General)“ . . . . .	249
A.32 Anwendungsfall #2 - Sektion „3.2 References“ . . . . .	249
A.33 Anwendungsfall #2 - Sektion „4.1 Overview of scenarios“ . . . . .	250
A.34 Anwendungsfall #2 - Sektion „4.2 Scenario“ . . . . .	250
A.35 Anwendungsfall #2 - Sektion „5. Information Exchanged“ . . . . .	251
A.36 Anwendungsfall #2 - Sektion „6. Requirements (optional)“ . . . . .	251
A.37 Anwendungsfall #2 - Sektion „7. Common terms and definitions“ . . . . .	251
A.38 Anwendungsfall #2 - Sektion „8. Custom information (optional)“ . . . . .	252
A.39 Vorgesehene Skala für den Kriterienvergleich, orientiert an [Saa87] . . . . .	253
A.40 Aufgestellte Bewertungsmatrix der Kriterien . . . . .	253
A.41 Gewichte und Rangordnung der Kriterien für die zusätzliche funktiona- len Anforderungen . . . . .	253
A.42 Verwendeter Wertebereiche und Beschreibung der Messung der Kriterien . . . . .	254
A.43 Skalierte Zielgrößen der Kriterien . . . . .	255
A.44 Fallbeispiel A: Zusätzliche identifizierte Fehler sowie False Positives . . . . .	257
A.45 Fallbeispiel B: Zusätzliche identifizierte Fehler sowie False Positives . . . . .	257
A.46 Anwendungsfallbeschreibung #1 - Bereitstellen von Vorschlägen für IEC 62559 Use Case-Beschreibungen . . . . .	259

---

A.47 Anwendungsfallbeschreibung #1.1 - SGAM-Domänen und -Zonen für Akteure, Komponenten sowie IEC 62559 Use Cases vorschlagen . . . .	260
A.48 Anwendungsfallbeschreibung #1.2 - Erkennen von abweichenden und nicht standardisierten Termen . . . . .	260
A.49 Anwendungsfallbeschreibung #1.3 - Vorschläge zum Hinzufügen, Entfernen und Bearbeiten von Elementen . . . . .	261
A.50 Anwendungsfallbeschreibung #1.4 - Service Steps in der Step-by-Step Analysis vorschlagen . . . . .	261
A.51 Anwendungsfallbeschreibung #1.5 - Erkennen von Inkonsistenzen bei Akteuren, Informationsaustauschen und Anforderungen gegenüber dem Use Case-Typ . . . . .	262
A.52 Anwendungsfallbeschreibung #2 - Bereitstellen von Vorschlägen für SGAM Modelle . . . . .	263
A.53 Anwendungsfallbeschreibung #2.1 - Generieren von SGAM Arbeitsmodellen . . . . .	264
A.54 Anwendungsfallbeschreibung #2.2 - Überprüfen der Konsistenz von SGAM Modellen gegenüber den IEC 62559 Use Case-Beschreibungen . . . . .	264
A.55 Anwendungsfallbeschreibung #2.3 - Identifikation von Inkonsistenzen innerhalb von SGAM Modellen . . . . .	265

## Literatur

- [AABW19] ALPAR, Paul ; ALT, Rainer ; BENSBERG, Frank ; WEIMANN, Peter: *Anwendungsorientierte Wirtschaftsinformatik: Strategische Planung, Entwicklung und Nutzung von Informationssystemen*. Wiesbaden : Springer Fachmedien Wiesbaden, 2019. <http://dx.doi.org/10.1007/978-3-658-25581-7>. <http://dx.doi.org/10.1007/978-3-658-25581-7>. – ISBN 9783658255800 9783658255817
- [AAZG<sup>+</sup>23] AL-ASHWAL, Fahmi Y. ; ZAWIAH, Mohammed ; GHARAIBEH, Lobna ; ABU-FARHA, Rana ; BITAR, Ahmad N.: Evaluating the Sensitivity, Specificity, and Accuracy of ChatGPT-3.5, ChatGPT-4, Bing AI, and Bard Against Conventional Drug-Drug Interactions Clinical Tools. In: *Drug, Healthcare and Patient Safety* 15 (2023), September, 137–147. <http://dx.doi.org/10.2147/DHPS.S425858>. – DOI 10.2147/DHPS.S425858
- [AGA23] ARORA, Chetan ; GRUNDY, John ; ABDELRAZEK, Mohamed: Advancing Requirements Engineering through Generative AI: Assessing the Role of LLMs. (2023). <http://dx.doi.org/10.48550/ARXIV.2310.13976>. – DOI 10.48550/ARXIV.2310.13976
- [AHN19] ANANDARAJAN, Murugan ; HILL, Chelsey ; NOLAN, Thomas: *Practical Text Analytics: Maximizing the Value of Text Data*. 1st ed. 2019. Cham : Springer International Publishing : Imprint: Springer, 2019 (Advances in Analytics and Data Science 2). – ISBN 9783319956633
- [AIUN18] AMBREEN, Talat ; IKRAM, Naveed ; USMAN, Muhammad ; NIAZI, Mahmood: Empirical research in requirements engineering: trends and opportunities. In: *Requirements Engineering* 23 (2018), März, Nr. 1, 63–95. <http://dx.doi.org/10.1007/s00766-016-0258-2>. – DOI 10.1007/s00766-016-0258-2. – ISSN 0947-3602, 1432-010X
- [Alj18] ALJOHANI, Tawfiq M.: *Analysis of the Smart Grid as a System of Systems*. <http://dx.doi.org/10.48550/arXiv.1810.11453>. Version: Oktober 2018. – arXiv:1810.11453 [eess]
- [ALN23] ABDULKADIROV, Ruslan ; LYAKHOV, Pavel ; NAGORNOV, Nikolay: Survey of Optimization Algorithms in Modern Neural Networks. In: *Mathematics* 11 (2023), Mai, Nr. 11, 2466. <http://dx.doi.org/10.3390/math11112466>. – DOI 10.3390/math11112466. – ISSN 2227-7390
- [AMA21] AGUILAR MADRID, Ernesto ; ANTONIO, Nuno: Short-Term Electricity Load Forecasting with Machine Learning. In: *Information*

- 12 (2021), Januar, Nr. 2, 50. <http://dx.doi.org/10.3390/info12020050>. – DOI 10.3390/info12020050. – ISSN 2078–2489
- [AS10] AKERKAR, Rajendra ; SAJJA, Priti: *Knowledge-based systems*. Sudbury, Mass : Jones and Bartlett, 2010. – ISBN 9780763776473. – OCLC: ocn320804456
- [ASIM14] ACHIMUGU, Philip ; SELAMAT, Ali ; IBRAHIM, Roliana ; MAHRIN, Mohd N.: A systematic literature review of software requirements prioritization research. In: *Information and Software Technology* 56 (2014), Juni, Nr. 6, 568–585. <http://dx.doi.org/10.1016/j.infsof.2014.02.001>. – DOI 10.1016/j.infsof.2014.02.001. – ISSN 09505849
- [ATCG<sup>+</sup>19] ANABY-TAVOR, Ateret ; CARMELI, Boaz ; GOLDBRAICH, Esther ; KANTOR, Amir ; KOUR, George ; SHLOMOV, Segev ; TEPPER, Naama ; ZWERDLING, Naama: *Not Enough Data? Deep Learning to the Rescue!* <http://dx.doi.org/10.48550/arXiv.1911.03118>. Version: November 2019. – arXiv:1911.03118
- [BA05] BERANDER, Patrik ; ANDREWS, Anneliese: Requirements Prioritization. Version: 2005. [http://dx.doi.org/10.1007/3-540-28244-0\\_4](http://dx.doi.org/10.1007/3-540-28244-0_4). In: AURUM, Aybüke (Hrsg.) ; WOHLIN, Claes (Hrsg.): *Engineering and Managing Software Requirements*. Berlin/Heidelberg : Springer-Verlag, 2005. – DOI 10.1007/3-540-28244-0\_4. – ISBN 9783540250432, 69–94
- [BADRDS<sup>+</sup>20] BARREDO ARRIETA, Alejandro ; DÍAZ-RODRÍGUEZ, Natalia ; DEL SER, Javier ; BENNETOT, Adrien ; TABIK, Siham ; BARBADO, Alberto ; GARCIA, Salvador ; GIL-LOPEZ, Sergio ; MOLINA, Daniel ; BENJAMINS, Richard ; CHATILA, Raja ; HERRERA, Francisco: Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. In: *Information Fusion* 58 (2020), Juni, 82–115. <http://dx.doi.org/10.1016/j.inffus.2019.12.012>. – DOI 10.1016/j.inffus.2019.12.012. – ISSN 15662535
- [Bal16] BALSLEV, Henrik: The Reference Designation System (RDS) A common naming convention for systems and their elements. In: *INCOSE International Symposium* 26 (2016), Juli, Nr. 1, S. 1639–1656. <http://dx.doi.org/10.1002/j.2334-5837.2016.00251.x>. – DOI 10.1002/j.2334-5837.2016.00251.x. – ISSN 23345837
- [Bec78] BECHMANN, Arnim: *Nutzwertanalyse, Bewertungstheorie und Planung*. 1. Aufl. Bern Stuttgart : Haupt, 1978 (Beiträge zur Wirtschaftspolitik 29). – ISBN 9783258026947

- [BECV08] BANA E COSTA, Carlos A. ; VANSNICK, Jean-Claude: A critical analysis of the eigenvalue method used to derive priorities in AHP. In: *European Journal of Operational Research* 187 (2008), Juni, Nr. 3, 1422–1428. <http://dx.doi.org/10.1016/j.ejor.2006.09.022>. – DOI 10.1016/j.ejor.2006.09.022. – ISSN 03772217
- [BHA<sup>+</sup>22] BOMMASANI, Rishi ; HUDSON, Drew A. ; ADELI, Ehsan ; ALTMAN, Russ ; ARORA, Simran ; ARX, Sydney v. ; BERNSTEIN, Michael S. ; BOHG, Jeannette ; BOSSELUT, Antoine ; BRUNSKILL, Emma ; BRYNJOLFSSON, Erik ; BUCH, Shyamal ; CARD, Dallas ; CASTELLON, Rodrigo ; CHATTERJI, Niladri ; CHEN, Annie ; CREEL, Kathleen ; DAVIS, Jared Q. ; DEMSZKY, Dora ; DONAHUE, Chris ; DOUMBOUYA, Moussa ; DURMUS, Esin ; ERMON, Stefano ; ETCHEMENDY, John ; ETHAYARAJH, Kawin ; FEI-FEI, Li ; FINN, Chelsea ; GALE, Trevor ; GILLESPIE, Lauren ; GOEL, Karan ; GOODMAN, Noah ; GROSSMAN, Shelby ; GUHA, Neel ; HASHIMOTO, Tatsunori ; HENDERSON, Peter ; HEWITT, John ; HO, Daniel E. ; HONG, Jenny ; HSU, Kyle ; HUANG, Jing ; ICARD, Thomas ; JAIN, Saahil ; JURAFSKY, Dan ; KALLURI, Pratyusha ; KARAMCHETI, Siddharth ; KEELING, Geoff ; KHANI, Fereshte ; KHATTAB, Omar ; KOH, Pang W. ; KRASS, Mark ; KRISHNA, Ranjay ; KUDITIPUDI, Rohith ; KUMAR, Ananya ; LADHAK, Faisal ; LEE, Mina ; LEE, Tony ; LESKOVEC, Jure ; LEVENT, Isabelle ; LI, Xiang L. ; LI, Xuechen ; MA, Tengyu ; MALIK, Ali ; MANNING, Christopher D. ; MIRCHANDANI, Suvir ; MITCHELL, Eric ; MUNYIKWA, Zanele ; NAIR, Suraj ; NARAYAN, Avaniika ; NARAYANAN, Deepak ; NEWMAN, Ben ; NIE, Allen ; NIEBLES, Juan C. ; NILFOROSHAN, Hamed ; NYARKO, Julian ; OGUT, Giray ; ORR, Laurel ; PAPADIMITRIOU, Isabel ; PARK, Joon S. ; PIECH, Chris ; PORTELANCE, Eva ; POTTS, Christopher ; RAGHUNATHAN, Aditi ; REICH, Rob ; REN, Hongyu ; RONG, Frieda ; ROOHANI, Yusuf ; RUIZ, Camilo ; RYAN, Jack ; RÉ, Christopher ; SADIGH, Dorsa ; SAGAWA, Shiori ; SANTHANAM, Keshav ; SHIH, Andy ; SRINIVASAN, Krishnan ; TAMKIN, Alex ; TAORI, Rohan ; THOMAS, Armin W. ; TRAMÈR, Florian ; WANG, Rose E. ; WANG, William ; WU, Bohan ; WU, Jiajun ; WU, Yuhuai ; XIE, Sang M. ; YASUNAGA, Michihiro ; YOU, Jiakuan ; ZAHARIA, Matei ; ZHANG, Michael ; ZHANG, Tianyi ; ZHANG, Xikun ; ZHANG, Yuhui ; ZHENG, Lucia ; ZHOU, Kaitlyn ; LIANG, Percy: *On the Opportunities and Risks of Foundation Models*. <http://dx.doi.org/10.48550/arXiv.2108.07258>. Version: Juli 2022. – arXiv:2108.07258
- [Bis06] BISHOP, Christopher M.: *Pattern recognition and machine learning*. New York : Springer, 2006 (Information science and statistics). – ISBN 9780387310732

- [BKM08] BANGOR, Aaron ; KORTUM, Philip T. ; MILLER, James T.: An Empirical Evaluation of the System Usability Scale. In: *International Journal of Human-Computer Interaction* 24 (2008), Juli, Nr. 6, 574–594. <http://dx.doi.org/10.1080/10447310802205776>. – DOI 10.1080/10447310802205776. – ISSN 1044–7318, 1532–7590
- [BKM09] BANGOR, Aaron ; KORTUM, Philip ; MILLER, James: Determining what individual SUS scores mean: adding an adjective rating scale. In: *J. Usability Studies* 4 (2009), Mai, Nr. 3, S. 114–123
- [BM24] BUCHER, Martin Juan J. ; MARTINI, Marco: *Fine-Tuned 'Small' LLMs (Still) Significantly Outperform Zero-Shot Generative AI Models in Text Classification*. <http://dx.doi.org/10.48550/arXiv.2406.08660>. Version: August 2024. – arXiv:2406.08660
- [BMG21] BUÇINCA, Zana ; MALAYA, Maja B. ; GAJOS, Krzysztof Z.: To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. In: *Proc. ACM Hum.-Comput. Interact.* 5 (2021), April, Nr. CSCW1, 188:1–188:21. <http://dx.doi.org/10.1145/3449287>. – DOI 10.1145/3449287
- [BMN09] BACHER, Peder ; MADSEN, Henrik ; NIELSEN, Henrik A.: Online short-term solar power forecasting. In: *Solar Energy* 83 (2009), Oktober, Nr. 10, 1772–1783. <http://dx.doi.org/10.1016/j.solener.2009.05.016>. – DOI 10.1016/j.solener.2009.05.016. – ISSN 0038092X
- [BMR<sup>+</sup>20] BROWN, Tom B. ; MANN, Benjamin ; RYDER, Nick ; SUBBIAH, Melanie ; KAPLAN, Jared ; DHARIWAL, Prafulla ; NEELAKANTAN, Arvind ; SHYAM, Pranav ; SASTRY, Girish ; ASKELL, Amanda ; AGARWAL, Sandhini ; HERBERT-VOSS, Ariel ; KRUEGER, Gretchen ; HENIGHAN, Tom ; CHILD, Rewon ; RAMESH, Aditya ; ZIEGLER, Daniel M. ; WU, Jeffrey ; WINTER, Clemens ; HESSE, Christopher ; CHEN, Mark ; SIGLER, Eric ; LITWIN, Mateusz ; GRAY, Scott ; CHESS, Benjamin ; CLARK, Jack ; BERNER, Christopher ; MCCANDLISH, Sam ; RADFORD, Alec ; SUTSKEVER, Ilya ; AMODEI, Dario: *Language Models are Few-Shot Learners*. <http://dx.doi.org/10.48550/arXiv.2005.14165>. Version: Juli 2020. – arXiv:2005.14165
- [BNK<sup>+</sup>19] BANSAL, Gagan ; NUSHI, Besmira ; KAMAR, Ece ; LASECKI, Walter S. ; WELD, Daniel S. ; HORVITZ, Eric: Beyond Accuracy: The Role of Mental Models in Human-AI Team Performance. In: *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 7 (2019), Oktober, 2–11. <http://dx.doi.org/10.1609/hcomp.v7i1.5285>. – DOI 10.1609/hcomp.v7i1.5285. – ISSN 2769–1349, 2769–1330

- [BP21] BELLE, Vaishak ; PAPANTONIS, Ioannis: Principles and Practice of Explainable Machine Learning. In: *Frontiers in Big Data* 4 (2021), Juli. <http://dx.doi.org/10.3389/fdata.2021.688969>. – DOI 10.3389/fdata.2021.688969. – ISSN 2624–909X
- [BRS95] BECKER, Jörg ; ROSEMAN, Michael ; SCHÜTTE, Reinhard: Grundsätze ordnungsmäßiger Modellierung. In: *Wirtschaftsinf.* 37 (1995), Nr. 5, S. 435–445
- [CCG21] CELIKYILMAZ, Asli ; CLARK, Elizabeth ; GAO, Jianfeng: *Evaluation of Text Generation: A Survey*. <http://dx.doi.org/10.48550/arXiv.2006.14799>. Version: Mai 2021. – arXiv:2006.14799
- [CEN12] CEN-CENELEC-ETSI: *Smart Grid Reference Architecture*. [https://www.cencenelec.eu/media/CEN-CENELEC/AreasOfWork/CEN-CENELEC\\_Topics/Smart%20Grids%20and%20Meters/Smart%20Grids/reference\\_architecture\\_smartgrids.pdf](https://www.cencenelec.eu/media/CEN-CENELEC/AreasOfWork/CEN-CENELEC_Topics/Smart%20Grids%20and%20Meters/Smart%20Grids/reference_architecture_smartgrids.pdf). Version: 2012
- [CEN14] CEN-CENELEC-ETSI: *SGAM User Manual - Applying, testing & refining the Smart Grid Architecture Model (SGAM)*. [https://syc-se.iec.ch/wp-content/uploads/2019/10/SGCG\\_Methodology\\_SGAMUserManual.pdf](https://syc-se.iec.ch/wp-content/uploads/2019/10/SGCG_Methodology_SGAMUserManual.pdf). Version: 2014
- [CHL<sup>+</sup>22] CHUNG, Hyung W. ; HOU, Le ; LONGPRE, Shayne ; ZOPH, Barret ; TAY, Yi ; FEDUS, William ; LI, Yunxuan ; WANG, Xuezhi ; DEGHANI, Mostafa ; BRAHMA, Siddhartha ; WEBSON, Albert ; GU, Shixiang S. ; DAI, Zhuyun ; SUZGUN, Mirac ; CHEN, Xinyun ; CHOWDHARY, Aakanksha ; CASTRO-ROS, Alex ; PELLAT, Marie ; ROBINSON, Kevin ; VALTER, Dasha ; NARANG, Sharan ; MISHRA, Gaurav ; YU, Adams ; ZHAO, Vincent ; HUANG, Yanping ; DAI, Andrew ; YU, Hongkun ; PETROV, Slav ; CHI, Ed H. ; DEAN, Jeff ; DEVLIN, Jacob ; ROBERTS, Adam ; ZHOU, Denny ; LE, Quoc V. ; WEI, Jason: *Scaling Instruction-Finetuned Language Models*. <http://dx.doi.org/10.48550/arXiv.2210.11416>. Version: Dezember 2022. – arXiv:2210.11416 [cs]
- [Cho20] CHOWDHARY, K. R.: Natural Language Processing. Version: 2020. [http://dx.doi.org/10.1007/978-81-322-3972-7\\_19](http://dx.doi.org/10.1007/978-81-322-3972-7_19). In: CHOWDHARY, K.R. (Hrsg.): *Fundamentals of Artificial Intelligence*. New Delhi : Springer India, 2020. – DOI 10.1007/978-81-322-3972-7\_19. – ISBN 9788132239727, 603–649
- [Coc01] COCKBURN, Alistair: *Writing effective use cases*. Boston : Addison-Wesley, 2001 (The Crystal series for software development). – ISBN 9780201702255

- [DAZ<sup>+</sup>23a] DIRECTORATE-GENERAL FOR ENERGY (EUROPEAN COMMISSION) ; ALACREU, Lola ; ZAMBRANO, Alberto ; PONS, Lucas ; MIRZ, Markus: *Action#1: set up a use case repository 2.0 : Data Management Working Group : May 2022*. LU : Publications Office of the European Union, 2023 <https://data.europa.eu/doi/10.2833/329067>. – ISBN 9789276995371
- [DAZ<sup>+</sup>23b] DIRECTORATE-GENERAL FOR ENERGY (EUROPEAN COMMISSION) ; ALACREU, Lola ; ZAMBRANO, Alberto ; PONS, Lucas ; MIRZ, Markus: *Action#1: set up a use case repository 2.0 : Data Management Working Group : May 2022*. LU : Publications Office of the European Union, 2023. <http://dx.doi.org/10.2833/329067>. <http://dx.doi.org/10.2833/329067>. – ISBN 9789276995371
- [DCLT19] DEVLIN, Jacob ; CHANG, Ming-Wei ; LEE, Kenton ; TOUTANOVA, Kristina: *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. <http://dx.doi.org/10.48550/arXiv.1810.04805>. Version: Mai 2019. – arXiv:1810.04805 [cs]
- [Dev22] DEVELOPERS, TensorFlow: TensorFlow. In: *Zenodo* (2022). <http://dx.doi.org/10.5281/zenodo.4758419>. – DOI 10.5281/zenodo.4758419
- [DG06] DAVIS, Jesse ; GOADRICH, Mark: The relationship between Precision-Recall and ROC curves. In: *Proceedings of the 23rd international conference on Machine learning - ICML '06*. Pittsburgh, Pennsylvania : ACM Press, 2006. – ISBN 9781595933836, 233–240
- [DGH24] DALAVI, Farzad ; GOLSHAN, Mohamad Esmail H. ; HATZIARGYRIOU, Nikos D.: A review on topology identification methods and applications in distribution networks. In: *Electric Power Systems Research* 234 (2024), September, 110538. <http://dx.doi.org/10.1016/j.epsr.2024.110538>. – DOI 10.1016/j.epsr.2024.110538. – ISSN 0378–7796
- [DIN15] DIN DEUTSCHES INSTITUT FÜR NORMUNG E V: *DIN EN ISO 9000:2015-11, Qualitätsmanagementsysteme - Grundlagen und Begriffe*. <http://dx.doi.org/10.31030/2325650>. Version: 2015
- [DIN20] DIN DEUTSCHES INSTITUT FÜR NORMUNG E V: *DIN EN ISO 9241-210:2020-03, Ergonomie der Mensch-System-Interaktion - Teil 210: Menschzentrierte Gestaltung interaktiver Systeme; Deutsche Fassung*. Beuth Verlag GmbH, 2020. <http://dx.doi.org/10.31030/3104744>. <http://dx.doi.org/10.31030/3104744>
- [DTSOD23] DURANGO, María C. ; TORRES-SILVA, Ever A. ; OROZCO-DUQUE, Andrés: Named Entity Recognition in Electronic Health Records:

- A Methodological Review. In: *Healthcare Informatics Research* 29 (2023), Oktober, Nr. 4, S. 286–300. <http://dx.doi.org/10.4258/hir.2023.29.4.286>. – DOI 10.4258/hir.2023.29.4.286. – ISSN 2093–3681
- [EBB<sup>+</sup>19] ESTEBSARI, Abouzar ; BARBIERATO, Luca ; BAHMANYAR, Ali-reza ; BOTTACCIOLI, Lorenzo ; MACII, Enrico ; PATTI, Edoardo: A SGAM-Based Test Platform to Develop a Scheme for Wide Area Measurement-Free Monitoring of Smart Grids under High PV Penetration. In: *Energies* 12 (2019), April, Nr. 8, 1417. <http://dx.doi.org/10.3390/en12081417>. – DOI 10.3390/en12081417. – ISSN 1996–1073
- [EDF] EDF: *MODSARUS*© *USER GUIDE*. <https://cimug.ucaiug.org/Meetings/eu2018/2018%20Ljubljana%20Presentations/CIMU%20Track%203/EDF%20MODSARUS%20UserGuide.pdf>
- [EHP<sup>+</sup>24] EHRMANN, Maud ; HAMDY, Ahmed ; PONTES, Elvys L. ; ROMANELLO, Matteo ; DOUCET, Antoine: Named Entity Recognition and Classification in Historical Documents: A Survey. In: *ACM Computing Surveys* 56 (2024), Februar, Nr. 2, 1–47. <http://dx.doi.org/10.1145/3604931>. – DOI 10.1145/3604931. – ISSN 0360–0300, 1557–7341
- [Ele25] ELECTRIC POWER RESEARCH INSTITUTE: *EPRI Use Case Repository*. <https://smartgrid.epri.com/repository/repository.aspx>. Version: Juli 2025
- [ENT22] ENTSO-E: *Harmonised Electricity Role Model*. <https://www.entsoe.eu/data/cim/role-models/>. Version: 2022
- [EU12] ENGLERT, Heiko ; USLAR, Mathias: Europäisches architekturmodell für smart grids-methodik und anwendung der ergebnisse der arbeitsgruppe referenzarchitektur des eu normungsmandats m/490. In: *Ta-gungsband VDE-Kongress*, 2012
- [Eur20] EUROPÄISCHE KOMMISSION, GENERALDIREKTION KOMMUNIKATIONSNETZE, INHALTE UND TECHNOLOGIEN: *WEISSBUCH Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen*. <https://eur-lex.europa.eu/legal-content/DE/ALL/?uri=CELEX:52020DC0065>. Version: 2020
- [FFT<sup>+</sup>24] FRATTINI, Julian ; FUCCI, Davide ; TORKAR, Richard ; MONTGOMERY, Lloyd ; UNTERKALMSTEINER, Michael ; FISCHBACH, Jan-nik ; MENDEZ, Daniel: Applying bayesian data analysis for causal inference about requirements quality: a controlled experiment.

- In: *Empirical Software Engineering* 30 (2024), November, Nr. 1, 29. <http://dx.doi.org/10.1007/s10664-024-10582-1>. – DOI 10.1007/s10664-024-10582-1. – ISSN 1573-7616
- [FGP19] FRANZL, Gerald ; GOTTSCHALK, Marion ; PASTEKA, Richard: The IES Cookbook - Enabling interoperability the IES way - Edition 0.8. (2019). <http://dx.doi.org/10.13140/RG.2.2.13780.86409>. – DOI 10.13140/RG.2.2.13780.86409
- [FGW<sup>+</sup>21] FENG, Steven Y. ; GANGAL, Varun ; WEI, Jason ; CHANDAR, Sarath ; VOSOUGHI, Soroush ; MITAMURA, Teruko ; HOVY, Eduard: *A Survey of Data Augmentation Approaches for NLP*. <http://dx.doi.org/10.48550/arXiv.2105.03075>. Version: Dezember 2021. – arXiv:2105.03075
- [Fin94] FINKELSTEIN, A.: Requirements engineering: a review and research agenda. In: *Proceedings of 1st Asia-Pacific Software Engineering Conference*. Tokyo, Japan : IEEE Comput. Soc. Press, 1994. – ISBN 9780818669606, 10–19
- [FSBU20] FALLER, Sebastian ; SCHÜTZ, Johann ; BOGENSPERGER, Alexander ; USLAR, Mathias: Anwendungshilfe SGAM: Smart Grid Use Cases modellieren. (2020). <http://dx.doi.org/10.13140/RG.2.2.24347.95525>. – DOI 10.13140/RG.2.2.24347.95525
- [FWK<sup>+</sup>17] FERNÁNDEZ, D. M. ; WAGNER, S. ; KALINOWSKI, M. ; FELDERER, M. ; MAFRA, P. ; VETRÒ, A. ; CONTE, T. ; CHRISTIANSSON, M.-T. ; GREER, D. ; LASSENIUS, C. ; MÄNNISTÖ, T. ; NAYABI, M. ; OIVO, M. ; PENZENSTADLER, B. ; PFAHL, D. ; PRIKLADNICKI, R. ; RUHE, G. ; SCHEKELMANN, A. ; SEN, S. ; SPINOLA, R. ; TUZCU, A. ; VARA, J. L. I. ; WIERINGA, R.: Naming the pain in requirements engineering. In: *Empirical Software Engineering* 22 (2017), Oktober, Nr. 5, 2298–2338. <http://dx.doi.org/10.1007/s10664-016-9451-7>. – DOI 10.1007/s10664-016-9451-7. – ISSN 1573-7616
- [GBC16] GOODFELLOW, Ian ; BENGIO, Yoshua ; COURVILLE, Aaron: *Deep learning*. Cambridge, Mass : The MIT press, 2016 (Adaptive computation and machine learning). – ISBN 9780262035613
- [GBD92] GEMAN, Stuart ; BIENENSTOCK, Elie ; DOURSAT, René: Neural Networks and the Bias/Variance Dilemma. In: *Neural Computation* 4 (1992), Januar, Nr. 1, 1–58. <http://dx.doi.org/10.1162/neco.1992.4.1.1>. – DOI 10.1162/neco.1992.4.1.1. – ISSN 0899-7667, 1530–888X
- [GDM<sup>+</sup>25] GARDAZI, Nadia M. ; DAUD, Ali ; MALIK, Muhammad K. ; BUKHARI, Amal ; ALSAHFI, Tariq ; ALSHEMAIMRI, Bader: BERT appli-

- cations in natural language processing: a review. In: *Artificial Intelligence Review* 58 (2025), März, Nr. 6, 166. <http://dx.doi.org/10.1007/s10462-025-11162-5>. – DOI 10.1007/s10462-025-11162-5. – ISSN 1573-7462
- [GDU17] GOTTSCHALK, Marion ; DELFS, Christina ; USLAR, Mathias: *The Use Case and Smart Grid Architecture Model Approach: The IEC 62559-2 Use Case Template and the SGAM applied in various domains*. 1st ed. 2017. Cham : Springer International Publishing : Imprint: Springer, 2017 (SpringerBriefs in Energy). – ISBN 9783319492292
- [GFF<sup>+</sup>18] GOTTSCHALK, Marion ; FRANZL, Gerald ; FROHNER, Matthias ; PASTEKA, Richard ; USLAR, Mathias: From Integration Profiles to Interoperability Testing for Smart Energy Systems at Connectathon Energy. In: *Energies* 11 (2018), Dezember, Nr. 12, 3375. <http://dx.doi.org/10.3390/en11123375>. – DOI 10.3390/en11123375. – ISSN 1996-1073
- [GMIFM<sup>+</sup>23] GARCÍA-MUÑOZ, Fernando ; IVANOVA, Anzhelika ; FARRÉ MONTANÉ, Jordi ; SERRANO, Manel ; CORCHERO, Cristina: A DSO flexibility platform based on an optimal congestion management model for the European CoordiNet project. In: *Electric Power Systems Research* 221 (2023), August, 109386. <http://dx.doi.org/10.1016/j.epsr.2023.109386>. – DOI 10.1016/j.epsr.2023.109386. – ISSN 0378-7796
- [GMR<sup>+</sup>18] GUIDOTTI, Riccardo ; MONREALE, Anna ; RUGGIERI, Salvatore ; TURINI, Franco ; PEDRESCHI, Dino ; GIANNOTTI, Fosca: *A Survey Of Methods For Explaining Black Box Models*. <http://dx.doi.org/10.48550/arXiv.1802.01933>. Version: Juni 2018. – arXiv:1802.01933
- [Gri18] GRINBERG, Miguel: *Flask Web Development*. Ö'Reilly Media, Inc., 2018. – ISBN 9781491991695. – Google-Books-ID: cVIPDwAAQ-BAJ
- [GS15] GOTTSCHALK, Marion ; SAUER, Jürgen: Towards identifying an approach for consistency checks to smart grid descriptions. In: *2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST)*, 2015, S. 380-385
- [GTP15] GINEVIČIUS, Romualdas ; TRISHCH, Halyna ; PETRAŠKEVIČIUS, Vladislavas: Quantitative assessment of quality management systems' processes. In: *Economic Research-Ekonomiska Istraživanja* 28 (2015), Januar, Nr. 1, 1096-1110. <http://dx.doi.org/10.1080/1331677X.2015.1087676>. – DOI 10.1080/1331677X.2015.1087676. – ISSN 1331-677X, 1848-9664

- [GU15] GOTTSCHALK, Marion ; USLAR, Mathias: Supporting the Development of Smart Cities using a Use Case Methodology. In: *Proceedings of the 24th International Conference on World Wide Web*, 2015, S. 541–545
- [GXG<sup>+</sup>24] GAO, Yunfan ; XIONG, Yun ; GAO, Xinyu ; JIA, Kangxiang ; PAN, Jinliu ; BI, Yuxi ; DAI, Yi ; SUN, Jiawei ; WANG, Meng ; WANG, Haofen: *Retrieval-Augmented Generation for Large Language Models: A Survey*. <http://dx.doi.org/10.48550/arXiv.2312.10997>. Version: März 2024. – arXiv:2312.10997
- [Hay09] HAYKIN, Simon S.: *Neural networks and learning machines*. 3. ed. New York Munich : Prentice-Hall, 2009. – ISBN 9780131471399
- [HGC23] HE, Pengcheng ; GAO, Jianfeng ; CHEN, Weizhu: *DeBERTaV3: Improving DeBERTa using ELECTRA-Style Pre-Training with Gradient-Disentangled Embedding Sharing*. <http://dx.doi.org/10.48550/arXiv.2111.09543>. Version: März 2023. – arXiv:2111.09543 [cs]
- [HJS22] HALEEM, Abid ; JAVAID, Mohd ; SINGH, Ravi P.: An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. In: *BenchCouncil Transactions on Benchmarks, Standards and Evaluations 2* (2022), Oktober, Nr. 4, 100089. <http://dx.doi.org/10.1016/j.tbench.2023.100089>. – DOI 10.1016/j.tbench.2023.100089. – ISSN 2772–4859
- [HKP12] HAN, Jiawei ; KAMBER, Micheline ; PEI, Jian: *Data mining: concepts and techniques*. 3rd ed. Amsterdam Boston : Elsevier/Morgan Kaufmann, 2012 (Morgan Kaufmann series in data management systems). – ISBN 9780123814791
- [HLGC21] HE, Pengcheng ; LIU, Xiaodong ; GAO, Jianfeng ; CHEN, Weizhu: *DeBERTa: Decoding-enhanced BERT with Disentangled Attention*. <http://dx.doi.org/10.48550/arXiv.2006.03654>. Version: Oktober 2021. – arXiv:2006.03654 [cs]
- [HMPR04] HEVNER ; MARCH ; PARK ; RAM: Design Science in Information Systems Research. In: *MIS Quarterly* 28 (2004), Nr. 1, 75. <http://dx.doi.org/10.2307/25148625>. – DOI 10.2307/25148625. – ISSN 02767783
- [HMW<sup>+</sup>20] HARRIS, Charles R. ; MILLMAN, K. J. ; WALT, Stéfan J. d. ; GOMMERS, Ralf ; VIRTANEN, Pauli ; COURNAPEAU, David ; WIESER, Eric ; TAYLOR, Julian ; BERG, Sebastian ; SMITH, Nathaniel J. ; KERN, Robert ; PICUS, Matti ; HOYER, Stephan ; KERKWIJK, Marten H. ; BRETT, Matthew ; HALDANE, Allan ; RÍO, Jaime F. ; WIEBE,

- Mark ; PETERSON, Pearu ; GÉRARD-MARCHANT, Pierre ; SHEPPARD, Kevin ; REDDY, Tyler ; WECKESSER, Warren ; ABBASI, Hameer ; GOHLKE, Christoph ; OLIPHANT, Travis E.: Array programming with NumPy. In: *Nature* 585 (2020), September, Nr. 7825, 357–362. <http://dx.doi.org/10.1038/s41586-020-2649-2>. – DOI 10.1038/s41586-020-2649-2. – ISSN 1476-4687
- [Hou15] HOULE, Sherilyn: An Introduction to the Fundamentals of Randomized Controlled Trials in Pharmacy Research. In: *The Canadian Journal of Hospital Pharmacy* 68 (2015), Februar, Nr. 1. <http://dx.doi.org/10.4212/cjhp.v68i1.1422>. – DOI 10.4212/cjhp.v68i1.1422. – ISSN 1920-2903, 0008-4123
- [HR24] HANKEL, Martin ; REXROTH, Bosch: Reference Architectural Model Industrie 4.0 (RAMI 4.0). 2 (2024), April, Nr. 2, S. 4–9
- [HS21] HENDERSON, Kaitlin ; SALADO, Alejandro: Value and benefits of model-based systems engineering (MBSE): Evidence from the literature. In: *Systems Engineering* 24 (2021), Januar, Nr. 1, 51–66. <http://dx.doi.org/10.1002/sys.21566>. – DOI 10.1002/sys.21566. – ISSN 1098-1241, 1520-6858
- [HSD<sup>+</sup>04] HASKINS, Bill ; STECKLEIN, Jonette ; DICK, Brandon ; MORONEY, Gregory ; LOVELL, Randy ; DABNEY, James: 8.4.2 Error Cost Escalation Through the Project Life Cycle. In: *INCOSE International Symposium* 14 (2004), Juni, Nr. 1, 1723–1737. <http://dx.doi.org/10.1002/j.2334-5837.2004.tb00608.x>. – DOI 10.1002/j.2334-5837.2004.tb00608.x. – ISSN 23345837
- [HZL<sup>+</sup>24] HOU, Xinyi ; ZHAO, Yanjie ; LIU, Yue ; YANG, Zhou ; WANG, Kailong ; LI, Li ; LUO, Xiapu ; LO, David ; GRUNDY, John ; WANG, Haoyu: *Large Language Models for Software Engineering: A Systematic Literature Review*. <http://dx.doi.org/10.48550/arXiv.2308.10620>. Version: April 2024. – arXiv:2308.10620
- [ICM<sup>+</sup>22] INKPEN, Kori ; CHAPPIDI, Shreya ; MALLARI, Keri ; NUSHI, Besmira ; RAMESH, Divya ; MICHELUCCI, Pietro ; MANDAVA, Vani ; VEPŘEK, Libuše H. ; QUINN, Gabrielle: *Advancing Human-AI Complementarity: The Impact of User Expertise and Algorithmic Tuning on Joint Decision Making*. <http://dx.doi.org/10.48550/arXiv.2208.07960>. Version: August 2022. – arXiv:2208.07960
- [IEC15] IEC TC 8: Use case methodology - Part 2: Definition of the templates for use cases, actor list and requirements list / International Organization for Standardization. 2015 (IEC 62559-2:2015). – Standard

- [IEC16] IEC TC 57: Power systems management and associated information exchange - Part 1: Reference architecture / International Electrotechnical Commission. 2016 (IEC TR 62357-1:2016). – Standard
- [IEC18] IEC TC 8: Use case methodology - Part 3: Definition of use case template artefacts into an XML serialized format / International Organization for Standardization. 2018 (IEC 62559-3:2018). – Standard
- [IEC19] IEC TC 8: Use case methodology - Part 1: Concept and processes in standardization / International Organization for Standardization. 2019 (IEC 62559-1:2019). – Standard
- [IEC21a] IEC TC 57: Communication networks and systems for power utility automation - Part 1: Introduction and overview / International Electrotechnical Commission. 2021 (IEC TR 61850-1:2013). – Standard
- [IEC21b] IEC TC 57: Energy management system application program interface (EMS-API) - Part 301: Common information model (CIM) base / International Electrotechnical Commission. 2021 (IEC 61970-301:2020). – Standard
- [IEC21c] IEC TC 57: Energy management system application program interface (EMS-API) - Part 600-1: Common Grid Model Exchange Standard (CGMES) - Structure and rules / International Electrotechnical Commission. 2021 (IEC 61970-600-1:2021). – Standard
- [IEC22] IEC: *Smart Grid Standards Map*. <https://www.iec.ch/sycsmartenergy>. Version: 2022
- [IN15] ISO/TC 176/SC 1 ; NA 147-00-01 AA: Qualitätsmanagementsysteme – Grundlagen und Begriffe / International Organization for Standardization and Deutsche Institut für Normierung. 2015 (DIN EN ISO 9000:2015). – Standard. – Standard
- [IPK21] IMAMBI, Sagar ; PRAKASH, Kolla B. ; KANAGACHIDAMBARESAN, G. R.: PyTorch. Version: 2021. [http://dx.doi.org/10.1007/978-3-030-57077-4\\_10](http://dx.doi.org/10.1007/978-3-030-57077-4_10). In: PRAKASH, Kolla B. (Hrsg.) ; KANAGACHIDAMBARESAN, G. R. (Hrsg.): *Programming with TensorFlow: Solution for Edge Computing Applications*. Cham : Springer International Publishing, 2021. – DOI 10.1007/978-3-030-57077-4\_10. – ISBN 9783030570774, 87–104
- [ISO08] ISO/IEC: Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Data quality model / International Organization for Standardization. 2008 (25012). – Standard

- 
- [ISO11] ISO/IEC JTC 1/SC 7: Systems and software engineering — Architecture description / International Organization for Standardization. 2011 (ISO/IEC/IEEE 42010:2011). – Standard
- [ISO12] ISO/IEC JTC 1: Object Management Group Unified Modeling Language / International Organization for Standardization and International Electrotechnical Commission. 2012 (ISO/IEC 19505-2:2012). – Standard
- [ISO19] ISO/IEC JTC 1/SC 7: Systems and software engineering — System of systems (SoS) considerations in life cycle stages of a system / International Organization for Standardization. 2019 (ISO/IEC/IEEE 21839:2019). – Standard
- [Jai22] JAIN, Shashank M.: Hugging Face. Version: 2022. [http://dx.doi.org/10.1007/978-1-4842-8844-3\\_4](http://dx.doi.org/10.1007/978-1-4842-8844-3_4). In: JAIN, Shashank M. (Hrsg.): *Introduction to Transformers for NLP: With the Hugging Face Library and Models to Solve Problems*. Berkeley, CA : Apress, 2022. – DOI 10.1007/978-1-4842-8844-3\_4. – ISBN 9781484288443, 51–67
- [JC07] JAMSEN, J ; CORLEY, Kevin: E-survey methodology. In: *Handbook of research on electronic surveys and measurements*. IGI Global, 2007, S. 1–8
- [Jet25] JETBRAINS: *JetBrains: Essential tools for software developers and teams*. <https://www.jetbrains.com/>. Version: 2025
- [JLF<sup>+</sup>24] JI, Ziwei ; LEE, Nayeon ; FRIESKE, Rita ; YU, Tiezheng ; SU, Dan ; XU, Yan ; ISHII, Etsuko ; BANG, Yejin ; CHEN, Delong ; DAI, Wenliang ; CHAN, Ho S. ; MADOTTO, Andrea ; FUNG, Pascale: *Survey of Hallucination in Natural Language Generation*. <http://dx.doi.org/10.48550/arXiv.2202.03629>. Version: Juli 2024. – arXiv:2202.03629
- [JMNR09] JURAFSKY, Dan ; MARTIN, James H. ; NORVIG, Peter ; RUSSELL, Stuart J.: *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition*. Second Edition, Pearson International Edition. Upper Saddle River, NJ : Prentice Hall, Pearson Education International, 2009 (Prentice Hall series in artificial intelligence). – ISBN 9780135041963
- [KA18] KALRA, Vaishali ; AGGARWAL, Rashmi: Importance of Text Data Pre-processing & Implementation in RapidMiner, 2018, 71–75
- [KAS23] KUCHENBUCH, R. ; AZAMAT, M. ; SCHÜTZ, J.: Optimizing the lifespan of (smart) transformers: a review on smart services. In: *27th In-*

- ternational Conference on Electricity Distribution (CIRED 2023)*. Rome, Italy : Institution of Engineering and Technology, 2023. – ISBN 9781839538551, 3479–3483
- [KJH<sup>+</sup>19] KOWSARI ; JAFARI MEIMANDI ; HEIDARYSAFA ; MENDU ; BARNES ; BROWN: Text Classification Algorithms: A Survey. In: *Information* 10 (2019), April, Nr. 4, 150. <http://dx.doi.org/10.3390/info10040150>. – DOI 10.3390/info10040150. – ISSN 2078–2489
- [KKKS23] KHURANA, Diksha ; KOLI, Aditya ; KHATTER, Kiran ; SINGH, Sukhdev: Natural language processing: state of the art, current trends and challenges. In: *Multimedia Tools and Applications* 82 (2023), Januar, Nr. 3, 3713–3744. <http://dx.doi.org/10.1007/s11042-022-13428-4>. – DOI 10.1007/s11042-022-13428-4. – ISSN 1573–7721
- [KLS25] KUCHENBUCH, René ; LEHNHOFF, Sebastian ; SAUER, Jürgen: Smart Grid Assistive AI in Requirement Engineering: Improving the Modeling of Use Cases and Architecture Models with LLMs. In: *Proceedings of the 16th ACM International Conference on Future and Sustainable Energy Systems*. Rotterdam Netherlands : ACM, Juni 2025. – ISBN 9798400711251, 495–504
- [KLZ24] KONG, Xiangwei ; LIU, Shujie ; ZHU, Luhao: Toward Human-centered XAI in Practice: A survey. In: *Machine Intelligence Research* 21 (2024), August, Nr. 4, 740–770. <http://dx.doi.org/10.1007/s11633-022-1407-3>. – DOI 10.1007/s11633-022-1407-3. – ISSN 2731–5398
- [KM14] KUCKARTZ, Udo ; MCWHERTOR, Anne: *Qualitative text analysis: a guide to methods, practice & using software*. Los Angeles : SAGE, 2014. – ISBN 9781446267752 9781446267745. – OCLC: ocn875376801
- [KMH24] KARG, Ludwig ; MUGICA, Mainer S. ; HUARTE, Joseba J.: *Deliverable 4.3: Engagement Towards Interoperability in Governance (v 1.0)*. <https://intnet.eu/resources/technical-resources>. Version: 2024
- [KNN<sup>+</sup>24] KUCHENBUCH, René ; NALE, Pranit ; NIEMANN, Laura ; WERTH, Oliver ; SAUER, Jürgen: Practical Application of Energy-related Use Case Methodology and SGAM Framework: A Study of User Perspectives and Challenges. In: *Energy Informatics* (2024)
- [KNSS24] KUCHENBUCH, René ; NIEMANN, Laura ; SCHÜTZ, Johann ; SAUER, Jürgen: Towards Intelligent Systems to Improve IEC 62559 Use Cases and Smart Grid Architecture Models Quality. In: ISLAM, Shareeful (Hrsg.) ; STURM, Arnon (Hrsg.): *Intelligent Information Systems*.

- 
- Cham : Springer Nature Switzerland, 2024. – ISBN 9783031610004, S. 38–46
- [KOH<sup>+</sup>23] KRUEGER, Carsten ; OTTE, Marcel ; HOLLY, Stefanie ; RATHJEN, Saskia ; WELLSSOW, Arlena ; LEHNHOFF, Sebastian: Redispatch 3.0 – Congestion Management for German Power Grids – Considering Controllable Resources in Low-Voltage Grids. In: *ETG Congress 2023*, 2023, 1–7
- [Kot04] KOTHARI, C. R.: *Research methodology: methods & techniques*. 2nd rev. ed. New Delhi : New Age International (P) Ltd., 2004. – ISBN 9788122424881. – OCLC: 395725716
- [Kot20] KOTARU, Venkata K.: *Angular for Material Design: Leverage Angular Material and TypeScript to Build a Rich User Interface for Web Apps*. Berkeley, CA : Apress, 2020. <http://dx.doi.org/10.1007/978-1-4842-5434-9>. <http://dx.doi.org/10.1007/978-1-4842-5434-9>. – ISBN 9781484254332 9781484254349
- [Kra16] KRAMER, Oliver: Scikit-Learn. Version: 2016. [http://dx.doi.org/10.1007/978-3-319-33383-0\\_5](http://dx.doi.org/10.1007/978-3-319-33383-0_5). In: KRAMER, Oliver (Hrsg.): *Machine Learning for Evolution Strategies*. Cham : Springer International Publishing, 2016. – DOI 10.1007/978-3-319-33383-0\_5. – ISBN 9783319333830, 45–53
- [KSS23] KUCHENBUCH, René ; SCHÜTZ, Johann ; SAUER, Jürgen: Quality properties of IEC 62559 use cases and SGAM models. In: *Energy Informatics* 6 (2023), Oktober, Nr. S1, 38. <http://dx.doi.org/10.1186/s42162-023-00280-5>. – DOI 10.1186/s42162-023-00280-5. – ISSN 2520-8942
- [Kur99] KURZWEIL, Ray: *The age of intelligent machines*. 3. print. Cambridge, Mass : MIT Press, 1999. – ISBN 9780262610797
- [KVN<sup>+</sup>26] KUCHENBUCH, René A. ; VERENO, Dominik ; NIEMANN, Laura ; WERTH, Oliver ; LEHNHOFF, Sebastian ; SAUER, Jürgen: The Impact of Artificial Intelligence on Quality in Smart Grid Requirements Engineering. In: *IEEE Access* 14 (2026), 11788–11804. <http://dx.doi.org/10.1109/ACCESS.2026.3654443>. – DOI 10.1109/ACCESS.2026.3654443. – ISSN 2169-3536
- [Kü14] KÜHNAPFEL, Jörg B.: *Nutzwertanalysen in Marketing und Vertrieb*. Wiesbaden : Springer Gabler, 2014 (essentials). – ISBN 9783658055097
- [LBH15] LECUN, Yann ; BENGIO, Yoshua ; HINTON, Geoffrey: Deep learning. In: *Nature* 521 (2015), Mai, Nr. 7553, 436–444. [http:](http://)

- [//dx.doi.org/10.1038/nature14539](https://dx.doi.org/10.1038/nature14539). – DOI 10.1038/nature14539. – ISSN 1476–4687
- [LCR<sup>+</sup>19] LAMBERT, Eric ; CANTENOT, Jérôme ; REIS, Francisco ; SULJANOVIC, Nermin ; SIMÃO, Tiago ; PETROVIC, Nejc ; TAYLOR, Gareth ; MORAIS, Hugo: *Standards assessment of Business Use Cases proposed in TDX-ASSIST*. AIM, 2019 <https://cired-repository.org/handle/20.500.12455/153>. – ISBN 9782960241501
- [LHC22] LI, Bohan ; HOU, Yutai ; CHE, Wanxiang: *Data Augmentation Approaches in Natural Language Processing: A Survey*. <http://dx.doi.org/10.48550/arXiv.2110.01852>. Version: Juni 2022. – arXiv:2110.01852
- [Liu21] LIU, Bin: ”Weak AI is Likely to Never Become ”Strong AI”, So What is its Greatest Value for us? (2021). <http://dx.doi.org/10.48550/ARXIV.2103.15294>. – DOI 10.48550/ARXIV.2103.15294
- [LOG<sup>+</sup>19] LIU, Yinhan ; OTT, Myle ; GOYAL, Naman ; DU, Jingfei ; JOSHI, Mandar ; CHEN, Danqi ; LEVY, Omer ; LEWIS, Mike ; ZETTMAYER, Luke ; STOYANOV, Veselin: *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. <http://dx.doi.org/10.48550/arXiv.1907.11692>. Version: Juli 2019. – arXiv:1907.11692 [cs]
- [LPL<sup>+</sup>22] LI, Qian ; PENG, Hao ; LI, Jianxin ; XIA, Congying ; YANG, Renyu ; SUN, Lichao ; YU, Philip S. ; HE, Lifang: A Survey on Text Classification: From Traditional to Deep Learning. In: *ACM Transactions on Intelligent Systems and Technology* 13 (2022), April, Nr. 2, 1–41. <http://dx.doi.org/10.1145/3495162>. – DOI 10.1145/3495162. – ISSN 2157–6904, 2157–6912
- [LS21] LEWIS, James R. ; SAURO, Jeff: USABILITY AND USER EXPERIENCE: DESIGN AND EVALUATION. Version: 1, September 2021. <http://dx.doi.org/10.1002/9781119636113.ch38>. In: SALVENDY, Gavriel (Hrsg.) ; KARWOWSKI, Waldemar (Hrsg.): *HANDBOOK OF HUMAN FACTORS AND ERGONOMICS*. 1. Wiley, September 2021. – DOI 10.1002/9781119636113.ch38. – ISBN 9781119636083 9781119636113, 972–1015
- [LTZ<sup>+</sup>24] LI, Junyi ; TANG, Tianyi ; ZHAO, Wayne X. ; NIE, Jian-Yun ; WEN, Ji-Rong: Pre-Trained Language Models for Text Generation: A Survey. In: *ACM Computing Surveys* 56 (2024), Oktober, Nr. 9, 1–39. <http://dx.doi.org/10.1145/3649449>. – DOI 10.1145/3649449. – ISSN 0360–0300, 1557–7341

- [LVHD23] LENNON, Breffn  ; VELASCO-HERREJ N, Paola ; DUNPHY, Niall P.: Operationalizing participation: Key obstacles and drivers to citizen energy community formation in Europe’s energy transition. In: *Science Talks* 5 (2023), M rz, 100104. <http://dx.doi.org/10.1016/j.sctalk.2022.100104>. – DOI 10.1016/j.sctalk.2022.100104. – ISSN 2772–5693
- [Mag01] MAGUIRE, Martin: Methods to support human-centred design. In: *International Journal of Human-Computer Studies* 55 (2001), Oktober, Nr. 4, 587–634. <http://dx.doi.org/10.1006/ijhc.2001.0503>. – DOI 10.1006/ijhc.2001.0503. – ISSN 10715819
- [MCCD13] MIKOLOV, Tomas ; CHEN, Kai ; CORRADO, Greg ; DEAN, Jeffrey: *Efficient Estimation of Word Representations in Vector Space*. <http://dx.doi.org/10.48550/arXiv.1301.3781>. Version: September 2013. – arXiv:1301.3781
- [McK11] MCKINNEY, Wes: pandas: a Foundational Python Library for Data Analysis and Statistics. (2011). <https://www.semanticscholar.org/paper/pandas%3A-a-Foundational-Python-Library-for-Data-and-McKinney/1a62eb61b2663f8135347171e30cb9dc0a8931b5>
- [MCL<sup>+</sup>18] MARKIDIS, Stefano ; CHIEN, Steven Wei D. ; LAURE, Erwin ; PENG, Ivy B. ; VETTER, Jeffrey S.: NVIDIA Tensor Core Programmability, Performance & Precision. In: *2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 2018, 522–531
- [Med00] MEDSKER, Larry R. (Hrsg.): *Recurrent neural networks: design and applications*. Boca Raton, FL : CRC Press, 2000 (The CRC Press international series on computational intelligence). – ISBN 9780849371813
- [Mer14] MERKEL, D.: Docker: lightweight Linux containers for consistent development and deployment. In: *Linux Journal* (2014), M rz. <https://www.semanticscholar.org/paper/Docker%3A-lightweight-Linux-containers-for-consistent-Merkel/875d90d4f66b07f90687b27ab304e04a3f666fc2>
- [Mic25] MICROSOFT: *Visual Studio Code - Code Editing. Redefined*. <https://code.visualstudio.com/>. Version: 2025
- [MMRS06] MCCARTHY, John ; MINSKY, Marvin L. ; ROCHESTER, Nathaniel ; SHANNON, Claude E.: A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. In: *AI magazine* 27 (2006), Nr. 4, S. 12–12

- [Mol22] MOLNAR, Christoph: *Interpretable machine learning: a guide for making black box models explainable*. Second edition. Munich, Germany : Christoph Molnar, 2022. – ISBN 9798411463330
- [MPR17] MU, Enrique ; PEREYRA-ROJAS, Milagros: *Practical Decision Making: An Introduction to the Analytic Hierarchy Process (AHP) Using Super Decisions V2*. 1st ed. 2017. Cham : Springer International Publishing : Imprint: Springer, 2017 (SpringerBriefs in Operations Research). – ISBN 9783319338613
- [MRS08] MANNING, Christopher D. ; RAGHAVAN, Prabhakar ; SCHUETZE, Hinrich: *Introduction to information retrieval*. New York : Cambridge University Press, 2008. – ISBN 9780521865715. – OCLC: ocn190786122
- [MS95] MARCH, Salvatore T. ; SMITH, Gerald F.: Design and natural science research on information technology. In: *Decision Support Systems* 15 (1995), Dezember, Nr. 4, 251–266. [http://dx.doi.org/10.1016/0167-9236\(94\)00041-2](http://dx.doi.org/10.1016/0167-9236(94)00041-2). – DOI 10.1016/0167-9236(94)00041-2. – ISSN 0167-9236
- [MST<sup>+</sup>24] MORRISON, Katelyn ; SPITZER, Philipp ; TURRI, Violet ; FENG, Michelle ; KÜHL, Niklas ; PERER, Adam: *The Impact of Imperfect XAI on Human-AI Decision-Making*. <http://dx.doi.org/10.48550/arXiv.2307.13566>. Version: Mai 2024. – arXiv:2307.13566
- [Mur14] MURPHY, Kevin P.: *Machine Learning - A Probabilistic Perspective*. Cambridge : MIT Press, 2014 (Adaptive Computation and Machine Learning). – ISBN 9780262018029 9780262305242
- [NBAKK17] NAZIR, Farhana ; BUTT, Wasi H. ; ANWAR, Muhammad W. ; KHAN KHATTAK, Muazzam A.: The Applications of Natural Language Processing (NLP) for Software Requirement Engineering - A Systematic Literature Review. In: KIM, Kuinam (Hrsg.) ; JOUKOV, Nikolai (Hrsg.): *Information Science and Applications 2017*, Springer, 2017 (Lecture Notes in Electrical Engineering). – ISBN 9789811041549, S. 485–493
- [NBJ<sup>+</sup>24] NAIMI, Lahbib ; BOUZIANE, El M. ; JAKIMI, Abdeslam ; SAADANE, Rachid ; CHEHRI, Abdellah: Automating Software Documentation: Employing LLMs for Precise Use Case Description. In: *Procedia Computer Science* 246 (2024), Januar, 1346–1354. <http://dx.doi.org/10.1016/j.procs.2024.09.568>. – DOI 10.1016/j.procs.2024.09.568. – ISSN 1877-0509
- [NET<sup>+</sup>14] NEUREITER, Christian ; ENGEL, Dominik ; TREFKE, Jörn ; SANTODOMINGO, Rafael ; ROHJANS, Sebastian ; USLAR, Mathias: Towards

- consistent smart grid architecture tool support: From use cases to visualization. In: *IEEE PES Innovative Smart Grid Technologies, Europe*, 2014, S. 1–6. – ISSN: 2165-4824
- [NEU16] NEUREITER, Christian ; ENGEL, Dominik ; USLAR, Mathias: Domain Specific and Model Based Systems Engineering in the Smart Grid as Prerequisite for Security by Design. In: *Electronics* 5 (2016), Mai, Nr. 2, 24. <http://dx.doi.org/10.3390/electronics5020024>. – DOI 10.3390/electronics5020024. – ISSN 2079–9292
- [NNN24] NAIK, Dishita ; NAIK, Ishita ; NAIK, Nitin: *Decoder-Only Transformers: The Brains Behind Generative AI, Large Language Models and Large Multimodal Models*. <http://dx.doi.org/10.36227/techrxiv.173198819.91727188/v1>. Version: November 2024
- [NUEL16] NEUREITER, Christian ; USLAR, Mathias ; ENGEL, Dominik ; LASTRO, Goran: A standards-based approach for domain specific modelling of smart grid system architectures. In: *2016 11th System of Systems Engineering Conference (SoSE)*, 2016, S. 1–6
- [OAA+24] OPENAI ; ACHIAM, Josh ; ADLER, Steven ; AGARWAL, Sandhini ; AHMAD, Lama ; AKKAYA, Ilge ; ALEMAN, Florencia L. ; ALMEIDA, Diogo ; ALTENSCHMIDT, Janko ; ALTMAN, Sam ; ANADKAT, Shyamal ; AVILA, Red ; BABUSCHKIN, Igor ; BALAJI, Suchir ; BALCOM, Valerie ; BALTESCU, Paul ; BAO, Haiming ; BAVARIAN, Mohammad ; BELGUM, Jeff ; BELLO, Irwan ; BERDINE, Jake ; BERNADETT-SHAPIRO, Gabriel ; BERNER, Christopher ; BOGDONOFF, Lenny ; BOIKO, Oleg ; BOYD, Madelaine ; BRAKMAN, Anna-Luisa ; BROCKMAN, Greg ; BROOKS, Tim ; BRUNDAGE, Miles ; BUTTON, Kevin ; CAI, Trevor ; CAMPBELL, Rosie ; CANN, Andrew ; CAREY, Brittany ; CARLSON, Chelsea ; CARMICHAEL, Rory ; CHAN, Brooke ; CHANG, Che ; CHANTZIS, Fotis ; CHEN, Derek ; CHEN, Sully ; CHEN, Ruby ; CHEN, Jason ; CHEN, Mark ; CHESS, Ben ; CHO, Chester ; CHU, Casey ; CHUNG, Hyung W. ; CUMMINGS, Dave ; CURRIER, Jeremiah ; DAI, Yunxing ; DECAREAUX, Cory ; DEGRY, Thomas ; DEUTSCH, Noah ; DEVILLE, Damien ; DHAR, Arka ; DOHAN, David ; DOWLING, Steve ; DUNNING, Sheila ; ECOFFET, Adrien ; ELETI, Atty ; ELOUNDOU, Tyna ; FARHI, David ; FEDUS, Liam ; FELIX, Niko ; FISHMAN, Simón P. ; FORTE, Juston ; FULFORD, Isabella ; GAO, Leo ; GEORGES, Elie ; GIBSON, Christian ; GOEL, Vik ; GOGINENI, Tarun ; GOH, Gabriel ; GONTIJO-LOPES, Rapha ; GORDON, Jonathan ; GRAFSTEIN, Morgan ; GRAY, Scott ; GREENE, Ryan ; GROSS, Joshua ; GU, Shixiang S. ; GUO, Yufei ; HALLACY, Chris ; HAN, Jesse ; HARRIS, Jeff ; HE, Yuchen ; HEATON, Mike ; HEIDECHE, Johannes ; HESSE, Chris ; HICKEY, Alan ; HICKEY, Wade ; HOESCHELE, Peter ; HOUGHTON, Brandon ; HSU, Kenny ; HU, Shengli ; HU, Xin ; HUIZINGA, Joost ;

JAIN, Shantanu ; JAIN, Shawn ; JANG, Joanne ; JIANG, Angela ; JIANG, Roger ; JIN, Haozhun ; JIN, Denny ; JOMOTO, Shino ; JONN, Billie ; JUN, Heewoo ; KAFTAN, Tomer ; KAISER, Łukasz ; KAMALI, Ali ; KANITSCHIEDER, Ingmar ; KESKAR, Nitish S. ; KHAN, Tabarak ; KILPATRICK, Logan ; KIM, Jong W. ; KIM, Christina ; KIM, Yongjik ; KIRCHNER, Jan H. ; KIROS, Jamie ; KNIGHT, Matt ; KOKOTAJLO, Daniel ; KONDRACIUK, Łukasz ; KONDRICH, Andrew ; KONSTANTINIDIS, Aris ; KOSIC, Kyle ; KRUEGER, Gretchen ; KUO, Vishal ; LAMPE, Michael ; LAN, Ikai ; LEE, Teddy ; LEIKE, Jan ; LEUNG, Jade ; LEVY, Daniel ; LI, Chak M. ; LIM, Rachel ; LIN, Molly ; LIN, Stephanie ; LITWIN, Mateusz ; LOPEZ, Theresa ; LOWE, Ryan ; LUE, Patricia ; MAKANJU, Anna ; MALFACINI, Kim ; MANNING, Sam ; MARKOV, Todor ; MARKOVSKI, Yaniv ; MARTIN, Bianca ; MAYER, Katie ; MAYNE, Andrew ; MCGREW, Bob ; MCKINNEY, Scott M. ; MCLEAVEY, Christine ; MCMILLAN, Paul ; MCNEIL, Jake ; MEDINA, David ; MEHTA, Aalok ; MENICK, Jacob ; METZ, Luke ; MISHCHENKO, Andrey ; MISHKIN, Pamela ; MONACO, Vinnie ; MORIKAWA, Evan ; MOSSING, Daniel ; MU, Tong ; MURATI, Mira ; MURK, Oleg ; MÉLY, David ; NAIR, Ashvin ; NAKANO, Reiichiro ; NAYAK, Rajeev ; NEELAKANTAN, Arvind ; NGO, Richard ; NOH, Hyeonwoo ; OUYANG, Long ; O'KEEFE, Cullen ; PACHOCKI, Jakub ; PAINO, Alex ; PALERMO, Joe ; PANTULIANO, Ashley ; PARASCANDOLO, Giambattista ; PARISH, Joel ; PARPARITA, Emy ; PASSOS, Alex ; PAVLOV, Mikhail ; PENG, Andrew ; PERELMAN, Adam ; PERES, Filipe de Avila B. ; PETROV, Michael ; PINTO, Henrique Ponde de O. ; MICHAEL ; POKORNY ; POKRASS, Michelle ; PONG, Vitchyr H. ; POWELL, Tolly ; POWER, Alethea ; POWER, Boris ; PROEHL, Elizabeth ; PURI, Raul ; RADFORD, Alec ; RAE, Jack ; RAMESH, Aditya ; RAYMOND, Cameron ; REAL, Francis ; RIMBACH, Kendra ; ROSS, Carl ; ROTSTED, Bob ; ROUSSEZ, Henri ; RYDER, Nick ; SALTARELLI, Mario ; SANDERS, Ted ; SANTURKAR, Shibani ; SASTRY, Girish ; SCHMIDT, Heather ; SCHNURR, David ; SCHULMAN, John ; SELSAM, Daniel ; SHEPPARD, Kyla ; SHERBAKOV, Toki ; SHIEH, Jessica ; SHOKER, Sarah ; SHYAM, Pranav ; SIDOR, Szymon ; SIGLER, Eric ; SIMENS, Maddie ; SITKIN, Jordan ; SLAMA, Katarina ; SOHL, Ian ; SOKOLOWSKY, Benjamin ; SONG, Yang ; STAUDACHER, Natalie ; SUCH, Felipe P. ; SUMMERS, Natalie ; SUTSKEVER, Ilya ; TANG, Jie ; TEZAK, Nikolas ; THOMPSON, Madeleine B. ; TILLET, Phil ; TOOTOONCHIAN, Amin ; TSENG, Elizabeth ; TUGGLE, Preston ; TURLEY, Nick ; TWOREK, Jerry ; URIBE, Juan Felipe C. ; VALLONE, Andrea ; VIJAYVERGIYA, Arun ; VOSS, Chelsea ; WAINWRIGHT, Carroll ; WANG, Justin J. ; WANG, Alvin ; WANG, Ben ; WARD, Jonathan ; WEI, Jason ; WEINMANN, C. J. ; WELIHINDA, Akila ; WELINDER, Peter ; WENG, Jiayi ; WENG, Lili-an ; WIETHOFF, Matt ; WILLNER, Dave ; WINTER, Clemens ; WOL-

- RICH, Samuel ; WONG, Hannah ; WORKMAN, Lauren ; WU, Sherwin ; WU, Jeff ; WU, Michael ; XIAO, Kai ; XU, Tao ; YOO, Sarah ; YU, Kevin ; YUAN, Qiming ; ZAREMBA, Wojciech ; ZELLERS, Rowan ; ZHANG, Chong ; ZHANG, Marvin ; ZHAO, Shengjia ; ZHENG, Tianhao ; ZHUANG, Juntang ; ZHUK, William ; ZOPH, Barret: *GPT-4 Technical Report*. <http://dx.doi.org/10.48550/arXiv.2303.08774>. Version: März 2024. – arXiv:2303.08774
- [OBS12] OVERHAGE, Sven ; BIRKMEIER, Dominik Q. ; SCHLAUDERER, Sebastian: Quality Marks, Metrics, and Measurement Procedures for Business Process Models. 4 (2012), Nr. 5, 229–246. <https://doi.org/10.1007/s12599-012-0230-8>. – ISSN 1867–0202
- [Ogu19] OGUZ, Yasin: *Konsistenzprüfung im Anforderungsmanagement mithilfe von Text Mining*. Oldenburg, Carl von Ossietzky Universität Oldenburg, Diplomarbeit, 2019. – unveröffentlichte Abschlussarbeit
- [OSTL19] OTTO, B. ; STEINBUSS, S. ; TEUSCHER, A. ; LOHMANN, S.: IDS Reference Architecture Model / Zenodo. Version: April 2019. <http://dx.doi.org/10.5281/ZENODO.5105529>. 2019. – Forschungsbericht. – Version Number: Version 3.0
- [PBG23] PATIL, Rajvardhan ; BOIT, Sorio ; GUDIVADA, Venkat ; NANDIGAM, Jagadeesh: A Survey of Text Representation and Embedding Techniques in NLP. In: *IEEE Access* 11 (2023), 36120–36146. <http://dx.doi.org/10.1109/ACCESS.2023.3266377>. – DOI 10.1109/ACCESS.2023.3266377. – ISSN 2169–3536
- [PBVDCN21] POURAN BEN VEYSEH, Amir ; DERNONCOURT, Franck ; CHANG, Walter ; NGUYEN, Thien H.: MadDog: A Web-based System for Acronym Identification and Disambiguation. In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*. Online : Association for Computational Linguistics, 2021, 160–167
- [Pet19] PETERS, Christian: *Konzeption und Implementierung eines Expertensystems für die Unterstützung bei der Erstellung von Use Cases nach IEC 62559*. Oldenburg, Carl von Ossietzky Universität Oldenburg, Diplomarbeit, 2019. – unveröffentlichte Abschlussarbeit
- [PGBN22] POLANEC, Katharina ; GROSS, Jounes-Alexander ; BRANKOVIC, Boris ; NEUREITER, Christian: Evolution of the Automotive Reference Architecture Model towards a Domain-Specific Systems Engineering Approach. In: *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2022, 1–4
- [PKML22] PAUSTIAN, Sabrina ; KÖHLKE, Julia ; MATTES, Jannika ; LEHNHOFF, Sebastian: The (still unexplored) social side of smart grid development:

- towards a social layer for the smart grid architecture model (SGAM). In: *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*. New York, NY, USA : Association for Computing Machinery, Juni 2022 (e-Energy '22). – ISBN 978–1–4503–9397–3, 521–531
- [PKT06] POURBEIK, P. ; KUNDUR, P.S. ; TAYLOR, C.W.: The anatomy of a power grid blackout - Root causes and dynamics of recent major blackouts. In: *IEEE Power and Energy Magazine* 4 (2006), September, Nr. 5, 22–29. <http://dx.doi.org/10.1109/MPAE.2006.1687814>. – DOI 10.1109/MPAE.2006.1687814. – ISSN 1540–7977
- [PNI<sup>+</sup>18] PETERS, Matthew E. ; NEUMANN, Mark ; IYYER, Mohit ; GARDNER, Matt ; CLARK, Christopher ; LEE, Kenton ; ZETTLEMOYER, Luke: Deep Contextualized Word Representations. In: WALKER, Marilyn (Hrsg.) ; JI, Heng (Hrsg.) ; STENT, Amanda (Hrsg.): *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. New Orleans, Louisiana : Association for Computational Linguistics, Juni 2018, 2227–2237
- [PR15] POHL, Klaus ; RUPP, Chris: *Requirements engineering fundamentals: a study guide for the certified professional for requirements engineering exam, foundation level, IREB compliant*. Second edition. Santa Barbara, CA : Rocky Nook, 2015. – ISBN 9781937538774
- [PRT22] PONELAT, Joshua S. ; ROSENSTOCK, Lukas L. ; TAM, Tony: *Designing APIs with Swagger and OpenAPI*. Shelter Island : Manning Publications, 2022. – ISBN 9781617296284
- [PSM14] PENNINGTON, Jeffrey ; SOCHER, Richard ; MANNING, Christopher: GloVe: Global Vectors for Word Representation. In: MOSCHITTI, Alessandro (Hrsg.) ; PANG, Bo (Hrsg.) ; DAELEMANS, Walter (Hrsg.): *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar : Association for Computational Linguistics, Oktober 2014, 1532–1543
- [PTRC07] PEFFERS, Ken ; TUUNANEN, Tuure ; ROTHENBERGER, Marcus ; CHATTERJEE, Samir: A Design Science Research Methodology for Information Systems Research. In: *Journal of Management Information Systems* 24 (2007), Dezember, Nr. 3, 45–77. <http://dx.doi.org/10.2753/MIS0742-1222240302>. – DOI 10.2753/MIS0742-1222240302. – ISSN 0742–1222
- [RAS<sup>+</sup>24] RETZLAFF, Carl O. ; ANGERSCHMID, Alessa ; SARANTI, Anna ; SCHNEEBERGER, David ; RÖTTGER, Richard ; MÜLLER, Heimo ; HOLZINGER, Andreas: Post-hoc vs ante-hoc explanations: xAI

- design guidelines for data scientists. In: *Cognitive Systems Research* 86 (2024), August, 101243. <http://dx.doi.org/10.1016/j.cogsys.2024.101243>. – DOI 10.1016/j.cogsys.2024.101243. – ISSN 1389–0417
- [RD15] RIEGEL, Norman ; DOERR, Joerg: A Systematic Literature Review of Requirements Prioritization Criteria. In: FRICKER, Samuel A. (Hrsg.) ; SCHNEIDER, Kurt (Hrsg.): *Requirements Engineering: Foundation for Software Quality*. Cham : Springer International Publishing, 2015 (Lecture Notes in Computer Science). – ISBN 9783319161013, S. 300–317
- [RDCC25] ROSADO DA CRUZ, António M. ; CRUZ, Estrela F.: Machine Learning Techniques for Requirements Engineering: A Comprehensive Literature Review. In: *Software* 4 (2025), Juni, Nr. 3, 14. <http://dx.doi.org/10.3390/software4030014>. – DOI 10.3390/software4030014. – ISSN 2674–113X
- [Reu14] REUTER, Martin: *Technischer und wirtschaftlicher Vergleich von Herstellungsverfahren bei der Entwicklung von Kunststoffhohlkörpern in Automobilanwendungen*, Duisburg, Essen, Universität Duisburg-Essen, Diss., 2013, PhD Thesis, 2014
- [RF08] RUDAS, Imre J. ; FODOR, János: Intelligent systems. In: *International Journal of Computers, Communications & Control* 3 (2008), Nr. 3, S. 132–138
- [RML24] RAKOTOSON, Loïc ; MASSIP, Sylvain ; LALEYE, Fréjus A. A.: Science Checker Reloaded: A Bidirectional Paradigm for Transparency and Logical Reasoning. (2024). <http://dx.doi.org/10.48550/ARXIV.2402.13897>. – DOI 10.48550/ARXIV.2402.13897
- [RN16] RUSSELL, Stuart J. ; NORVIG, Peter: *Artificial intelligence: a modern approach*. Third edition, Global edition. Boston Columbus Indianapolis New York San Francisco Upper Saddle River Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto Delhi Mexico City Sao Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo : Pearson, 2016 (Prentice Hall series in artificial intelligence). – ISBN 9781292153964
- [Ros13] ROSEMANN, Michael: *Komplexitätsmanagement in Prozeßmodellen: Methodenspezifische Gestaltungsempfehlungen für die Informationsmodellierung*. Springer-Verlag, 2013. – ISBN 9783322992314. – Google-Books-ID: hyPQBgAAQBAJ
- [RPGB17] RANGLES, Bernadette M. ; PASQUETTO, Irene V. ; GOLSHAN, Milena S. ; BORGMAN, Christine L.: Using the Jupyter Notebook as a

- Tool for Open Science: An Empirical Study. In: *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 2017, 1–2
- [RSAM21] REDDY SHABAD, Prem K. ; ALRASHIDE, Abdulmueen ; MOHAMMED, Osama: Anomaly Detection in Smart Grids using Machine Learning. In: *IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society*, 2021, 1–8. – ISSN: 2577-1647
- [RSJ<sup>+</sup>23] REIF, Valerie ; STRASSER, Thomas I. ; JIMENO, Joseba ; FARRE, Marjolaine ; GENEST, Oliver ; GYRARD, Amélie ; MCGRANAGHAN, Mark ; LIPARI, Gianluca ; SCHÜTZ, Johann ; USLAR, Mathias ; VOGEL, Sebastian ; BYTYQI, Arsim ; DORNMAIR, Rita ; CORUSA, Andreas ; ROY, Gaurav ; PONCI, Ferdinanda ; DOGNINI, Alberto ; MONTI, Antonello: Towards an interoperability roadmap for the energy transition. In: *e & i Elektrotechnik und Informationstechnik* 140 (2023), August, Nr. 5, 478–487. <http://dx.doi.org/10.1007/s00502-023-01144-2>. – DOI 10.1007/s00502-023-01144-2. – ISSN 0932-383X, 1613–7620
- [RSR<sup>+</sup>23] RAFFEL, Colin ; SHAZEER, Noam ; ROBERTS, Adam ; LEE, Katherine ; NARANG, Sharan ; MATENA, Michael ; ZHOU, Yanqi ; LI, Wei ; LIU, Peter J.: *Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer*. <http://dx.doi.org/10.48550/arXiv.1910.10683>. Version: September 2023. – arXiv:1910.10683 [cs, stat]
- [RTL09] REFAEILZADEH, Payam ; TANG, Lei ; LIU, Huan: Cross-Validation. Version: 2009. [http://dx.doi.org/10.1007/978-0-387-39940-9\\_565](http://dx.doi.org/10.1007/978-0-387-39940-9_565). In: *Encyclopedia of Database Systems*. Springer, Boston, MA, 2009. – DOI 10.1007/978-0-387-39940-9\_565. – ISBN 9780387399409, 532–538
- [RWC<sup>+</sup>19] RADFORD, Alec ; WU, Jeff ; CHILD, R. ; LUAN, D. ; AMODEI, Dario ; SUTSKEVER, I.: *Language Models are Unsupervised Multitask Learners*, 2019
- [Saa87] SAATY, R.W.: The analytic hierarchy process—what it is and how it is used. In: *Mathematical Modelling* 9 (1987), Nr. 3-5, 161–176. [http://dx.doi.org/10.1016/0270-0255\(87\)90473-8](http://dx.doi.org/10.1016/0270-0255(87)90473-8). – DOI 10.1016/0270-0255(87)90473-8. – ISSN 02700255
- [Sch12] SCHREIER, Margrit: *Qualitative Content Analysis in Practice*. SAGE Publications Ltd, 2012. – ISBN 9781937538774
- [Sch14] SCHREIER, Margrit: Ways of Doing Qualitative Content Analysis: Disentangling Terms and Terminologies. In: *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 15 (2014), Jan.,

- Nr. 1. <http://dx.doi.org/10.17169/fqs-15.1.2043>. – DOI 10.17169/fqs-15.1.2043
- [Sch15] SCHMIDHUBER, Jürgen: Deep learning in neural networks: An overview. In: *Neural Networks* 61 (2015), Januar, 85–117. <http://dx.doi.org/10.1016/j.neunet.2014.09.003>. – DOI 10.1016/j.neunet.2014.09.003. – ISSN 08936080
- [SCUG21] SCHÜTZ, J. ; CLAUSEN, M. ; USLAR, M. ; GÓMEZ, J. M.: IEC 62559-2 USE CASE TEMPLATE-BASED SMART GRID ARCHITECTURE ANALYTICS. In: *CIRE2021 - The 26th International Conference and Exhibition on Electricity Distribution* Bd. 2021, 2021, S. 2935–2939
- [SDCW20] SANH, Victor ; DEBUT, Lysandre ; CHAUMOND, Julien ; WOLF, Thomas: *DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter*. <http://dx.doi.org/10.48550/arXiv.1910.01108>. Version: Februar 2020. – arXiv:1910.01108 [cs]
- [SH03] SCHWARTZ, Ariel S. ; HEARST, Marti A.: A simple algorithm for identifying abbreviation definitions in biomedical text. In: *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* (2003), S. 451–462. – ISSN 2335–6928
- [SJ19] STANČIN, I. ; JOVIĆ, A.: An overview and comparison of free Python libraries for data mining and big data analysis. In: *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2019, S. 977–982. – ISSN: 2623-8764
- [SKKK24] SCHMITT, Beate ; KOSSLERS, Sebastian ; KUCHENBUCH, René ; KUNG, Antonio: *Deliverable 4.1: Analysis of standards environment relevant for int:net*. <https://intnet.eu/resources/technical-resources>. Version: 2024
- [SLCD<sup>+</sup>21] SAMOILI, Sofia ; LÓPEZ COBO, Montserrat ; DELIPETREV, Blagoj ; MARTÍNEZ-PLUMED, Fernando ; GÓMEZ, Emilia ; DE PRATO, Giuditta: *AI watch: defining artificial intelligence 2.0: towards an operational definition and taxonomy for the AI landscape*. Luxembourg : Publications Office of the European Union, 2021 (JRC technical reports 30873). – ISBN 9789276426486
- [SM21] SINGH, Sushant ; MAHMOOD, Ausif: The NLP Cookbook: Modern Recipes for Transformer Based Deep Learning Architectures. In: *IEEE Access* 9 (2021), 68675–68702. <http://dx.doi.org/10.1109/ACCESS.2021.3077350>. – DOI 10.1109/ACCESS.2021.3077350. – ISSN 2169–3536

- [SMGS20] SOWMYA V. B ; MAJUMDER, Bodhisattwa ; GUPTA, Anju ; SURANA, Harshit: *Practical natural language processing: a comprehensive guide to building real-world NLP systems*. First edition. Sebastopol, CA : O'Reilly Media, 2020. – ISBN 9781492054054. – OCLC: on1125266646
- [SNT04] STOLBERG, Harald O. ; NORMAN, Geoffrey ; TROP, Isabelle: Randomized Controlled Trials. In: *American Journal of Roentgenology* 183 (2004), Dezember, Nr. 6, 1539–1544. <http://dx.doi.org/10.2214/ajr.183.6.01831539>. – DOI 10.2214/ajr.183.6.01831539. – ISSN 0361–803X, 1546–3141
- [Som12] SOMMERVILLE, Ian: *Software engineering*. 9., aktualisierte Auflage. München Harlow Amsterdam : Pearson, 2012 (Always learning). – ISBN 9783868940992
- [SSPA<sup>+</sup>12] SIDI, Fatimah ; SHARIAT PANAHY, Payam H. ; AFFENDEY, Lilly S. ; JABAR, Marzanah A. ; IBRAHIM, Hamidah ; MUSTAPHA, Aida: Data quality: A survey of data quality dimensions. In: *2012 International Conference on Information Retrieval & Knowledge Management, IEE-EE*, 2012. – ISBN 9781467310918 9781467310901, 300–304
- [Sta25] STACK OVERFLOW: *2024 Stack Overflow Developer Survey*. <https://survey.stackoverflow.co/2024/#>. Version: 2025
- [Stu08] STUDENT: The Probable Error of a Mean. In: *Biometrika* 6 (1908), März, Nr. 1, 1. <http://dx.doi.org/10.2307/2331554>. – DOI 10.2307/2331554. – ISSN 00063444
- [SUG<sup>+</sup>14] SANTODOMINGO, R. ; USLAR, M. ; GORING, A. ; GOTTSCHALK, M. ; NORDSTROM, L. ; SALEEM, A. ; CHENINE, M.: SGAM-based methodology to analyse Smart Grid solutions in DISCERN European research project. In: *2014 IEEE International Energy Conference (ENERGYCON)*, IEEE, 2014. – ISBN 9781479924493, S. 751–758
- [SUG<sup>+</sup>16] SANTODOMINGO, Rafael ; USLAR, Mathias ; GOTTSCHLAK, Marion ; GOERING, Andre ; NORDSTRÖM, Lars ; VALDENMAIER, Georgii: The DISCERN tool support for knowledge sharing in large Smart Grid projects. In: *CIREN Workshop 2016*, 2016, 1–4
- [Sur25] SURVEYJS: *SurveyJS - JavaScript Libraries for Surveys and Forms*. <https://surveyjs.io/>. Version: 2025
- [SWG<sup>+</sup>25] STRASSER, Thomas I. ; WIDL, Edmund ; GREGOR, Carlos Ayon M. ; GINOCCHI, Mirko ; KUCHENBUCH, René: Mapping and Evolving Interoperability Testing in European Energy Systems: The int.net Perspective, arXiv, Oktober 2025. – arXiv:2510.19460

- [SWK<sup>+</sup>24] STRASSER, Thomas I. ; WIDL, Edmund ; KUCHENBUCH, René A. ; LÁZARO-ELORRIAGA, Laura ; LARAUDOGOITIA, Borja T. ; GINOCCHI, Mirko ; PENTHONG, Thanakorn ; PONCI, Ferdinanda ; GYRARD, Amelie ; KUNG, Antonio ; GREGOR, Carlos A. M. ; MONTERO, Carmen G. ; ALGABA, Eduardo R.: *Towards Interoperability Testing of Smart Energy Systems – An Overview and Discussion of Possibilities*. <http://dx.doi.org/10.48550/arXiv.2411.03369>. Version: November 2024. – arXiv:2411.03369
- [SWM17] SAMEK, Wojciech ; WIEGAND, Thomas ; MÜLLER, Klaus-Robert: *Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models*. <http://dx.doi.org/10.48550/arXiv.1708.08296>. Version: August 2017. – arXiv:1708.08296
- [SyC18] SYC SMART ENERGY: Use case methodology - Part 4: Best practices in use case development for IEC standardization processes and some examples for application outside standardization / International Organization for Standardization. 2018 (IEC SRD 62559-4:2020). – Standard
- [Tah24] TAHA, Kamal: Text Regression Analysis: A Review, Empirical, and Experimental Insights. In: *IEEE Access* 12 (2024), 137333–137344. <http://dx.doi.org/10.1109/ACCESS.2024.3446765>. – DOI 10.1109/ACCESS.2024.3446765. – ISSN 2169–3536
- [TGD13] TREFKE, Jörn ; GONZÁLEZ, José M. ; DÄNEKAS, Christian: IEC/PAS 62559-Based Use Case Management for Smart Grids. Version: 2013. [http://dx.doi.org/10.1007/978-3-642-34916-4\\_3](http://dx.doi.org/10.1007/978-3-642-34916-4_3). In: USLAR, Mathias (Hrsg.) ; SPECHT, Michael (Hrsg.) ; DÄNEKAS, Christian (Hrsg.) ; TREFKE, Jörn (Hrsg.) ; ROHJANS, Sebastian (Hrsg.) ; GONZÁLEZ, José M. (Hrsg.) ; ROSINGER, Christine (Hrsg.) ; BLEIKER, Robert (Hrsg.): *Standardization in Smart Grids: Introduction to IT-Related Methodologies, Architectures and Standards*. Springer, 2013 (Power Systems). – DOI 10.1007/978-3-642-34916-4\_3. – ISBN 9783642349164, S. 39–57
- [TLI<sup>+</sup>23] TOUVRON, Hugo ; LAVRIL, Thibaut ; IZACARD, Gautier ; MARTINET, Xavier ; LACHAUX, Marie-Anne ; LACROIX, Timothée ; ROZIÈRE, Baptiste ; GOYAL, Naman ; HAMBRO, Eric ; AZHAR, Faisal ; RODRIGUEZ, Aurelien ; JOULIN, Armand ; GRAVE, Edouard ; LAMPLE, Guillaume: *LLaMA: Open and Efficient Foundation Language Models*. <http://dx.doi.org/10.48550/arXiv.2302.13971>. Version: Februar 2023. – arXiv:2302.13971
- [TYJB24] TONG, Xie ; YUWEI, Wan ; JUNTAO, Fang ; BRAM, Hox: *UNSW-MasterAI/EnergyBERT · Hugging Face*. <https://github.com/HuggingFace/transformers>

//huggingface.co/UNSW-MasterAI/EnergyBERT.  
Version: 2024

- [UG15] USLAR, Mathias ; GOTTSCHALK, Marion: Extending the sgam for electric vehicles. In: *International ETG Congress 2015; Die Energiewende-Blueprints for the new energy age*, VDE, 2015, S. 1–8
- [UL24] UMAR, Muhammad A. ; LANO, Kevin: Advances in automated support for requirements engineering: a systematic literature review. In: *Requirements Engineering* 29 (2024), Juni, Nr. 2, 177–207. <http://dx.doi.org/10.1007/s00766-023-00411-0>. – DOI 10.1007/s00766-023-00411-0. – ISSN 1432-010X
- [URMG<sup>+</sup>11] USLAR, M. ; ROHJANS, S. ; M. GONZÁLEZ, J. ; SPECHT, M. ; TREFKE, J.: Das Standardisierungsumfeld im Smart Grid – Roadmap und Outlook. In: *e & i Elektrotechnik und Informationstechnik* 128 (2011), April, Nr. 4, 135–140. <http://dx.doi.org/10.1007/s00502-011-0818-6>. – DOI 10.1007/s00502-011-0818-6. – ISSN 1613-7620
- [URN<sup>+</sup>19] USLAR, Mathias ; ROHJANS, Sebastian ; NEUREITER, Christian ; PRÖSTL ANDRÉN, Filip ; VELASQUEZ, Jorge ; STEINBRINK, Cornelius ; EFTHYMIIOU, Venizelos ; MIGLIAVACCA, Gianluigi ; HORSMANHEIMO, Seppo ; BRUNNER, Helfried ; STRASSER, Thomas: Applying the Smart Grid Architecture Model for Designing and Validating System-of-Systems in the Power and Energy Domain: A European Perspective. 12 (2019), Nr. 2, S. 258. <http://dx.doi.org/10.3390/en12020258>. – DOI 10.3390/en12020258. – ISSN 1996-1073
- [URS14] USLAR, Mathias ; ROSINGER, Christine ; SCHLEGEL, Stefanie: Security by Design for the Smart Grid: Combining the SGAM and NISTIR 7628. In: *2014 IEEE 38th International Computer Software and Applications Conference Workshops*, 2014, S. 110–115
- [USD<sup>+</sup>13] USLAR, Mathias ; SPECHT, Michael ; DANEKAS, Christian ; TREFKE, Jorn ; ROHJANS, Sebastian ; GONZALEZ VAZQUEZ, José M. ; ROSINGER, Christine ; BLEIKER, Robert: *Standardization in smart grids: introduction to IT-related methodologies, architectures and standards*. Berlin ; New York : Springer, 2013 (Power systems). – ISBN 978-3-642-34915-7. – OCLC: ocn842074536
- [USS<sup>+</sup>21] UECKER, Denis ; SCHWARZ, Reinhard ; SOROKOS, Ioannis ; ALZUGHBI, Hanna ; USLAR, Mathias ; HANNA, Sebastian ; ROSINGER, Christine: *Addressing quality properties in use case descriptions: adding assurance to a use case process*. Fraunhofer Verlag, 2021. – ISBN 9783839617243

- [VDN<sup>+</sup>20] VEYSEH, Amir Pouran B. ; DERNONCOURT, Franck ; NGUYEN, Thien H. ; CHANG, Walter ; CELI, L.: Acronym Identification and Disambiguation Shared Tasks for Scientific Document Understanding. In: *ArXiv* (2020), Dezember. <https://www.semanticscholar.org/paper/Acronym-Identification-and-Disambiguation-Shared-Veyseh-Dernoncourt/19036f1d59e53ae4da3153863a5324089945b37f>
- [VDTN20] VEYSEH, Amir Pouran B. ; DERNONCOURT, Franck ; TRAN, Quan H. ; NGUYEN, Thien H.: *What Does This Acronym Mean? Introducing a New Dataset for Acronym Identification and Disambiguation*. <https://arxiv.org/abs/2010.14678v1>. Version: Oktober 2020
- [VGA<sup>+</sup>25] VAN GENUCHTEN, Joep ; GUITART, Laia ; AYNON MAC GREGOR, Carlos ; KUCHENBUCH, René ; STRASSER, Thomas I.: Redefining the concept of interoperability: A multidimensional approach for the energy sector. In: *Open Research Europe* 5 (2025), August, 250. <http://dx.doi.org/10.12688/openreseurope.20755.1>. – DOI 10.12688/openreseurope.20755.1. – ISSN 2732–5121
- [VGHK<sup>+</sup>25] VAN GENUCHTEN, Joep ; HURLEY, Jemima ; KUCHENBUCH, Rene ; GUITART, Laia ; LIPARI, Gianluca: EMINENT, a method for assessing interoperability maturity. In: *IET Conference Proceedings* 2024 (2025), Januar, Nr. 5, 628–632. <http://dx.doi.org/10.1049/icp.2024.2117>. – DOI 10.1049/icp.2024.2117. – ISSN 2732–4494
- [VHN23] VERENO, Dominik ; HARB, Jonas ; NEUREITER, Christian: Paving the Way for Reinforcement Learning in Smart Grid Co-simulations. In: MASCI, Paolo (Hrsg.) ; BERNARDESCHI, Cinzia (Hrsg.) ; GRAZIANI, Pierluigi (Hrsg.) ; KODDENBROCK, Mario (Hrsg.) ; PALMIERI, Maurizio (Hrsg.): *Software Engineering and Formal Methods. SEFM 2022 Collocated Workshops*. Cham : Springer International Publishing, 2023. – ISBN 9783031262364, S. 242–257
- [VLH<sup>+</sup>23] VERED, Mor ; LIVNI, Tali ; HOWE, Piers Douglas L. ; MILLER, Tim ; SONENBERG, Liz: The effects of explanations on automation bias. In: *Artificial Intelligence* 322 (2023), September, 103952. <http://dx.doi.org/10.1016/j.artint.2023.103952>. – DOI 10.1016/j.artint.2023.103952. – ISSN 0004–3702
- [VNE<sup>+</sup>25] VERENO, Dominik ; NEUREITER, Christian ; ESCHLBERGER, Simon ; MILLAKU, Mergim ; KUCHENBUCH, René ; USLAR, Mathias: SGAM Toolbox Revisited: A Standards-Based Domain-Specific Modeling Language and Toolset. In: *IEEE Access* (2025), 1–1. <http://dx.doi.org/10.1109/ACCESS.2025.3586722>. – DOI 10.1109/ACCESS.2025.3586722. – ISSN 2169–3536

- [VSP<sup>+</sup>17] VASWANI, Ashish ; SHAZEER, Noam ; PARMAR, Niki ; USZKOREIT, Jakob ; JONES, Llion ; GOMEZ, Aidan N. ; KAISER, Łukasz ; POLOSUKHIN, Illia: Attention is All you Need. In: *Advances in Neural Information Processing Systems* Bd. 30, Curran Associates, Inc., 2017
- [VWU23] VEITH, Eric M. S. P. ; WELLSOW, Arlena ; USLAR, Mathias: Learning new attack vectors from misuse cases with deep reinforcement learning. In: *Frontiers in Energy Research* 11 (2023), März. <http://dx.doi.org/10.3389/fenrg.2023.1138446>. – DOI 10.3389/fenrg.2023.1138446. – ISSN 2296–598X
- [WBC10] WATSON ; BOUDREAU ; CHEN: Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community. In: *MIS Quarterly* 34 (2010), Nr. 1, 23. <http://dx.doi.org/10.2307/20721413>. – DOI 10.2307/20721413. – ISSN 02767783
- [WH07] WILDE, Thomas ; HESS, Thomas: Forschungsmethoden der Wirtschaftsinformatik: Eine empirische Untersuchung. In: *WIRTSCHAFTSINFORMATIK* 49 (2007), August, Nr. 4, 280–287. <http://dx.doi.org/10.1007/s11576-007-0064-z>. – DOI 10.1007/s11576-007-0064-z. – ISSN 0937–6429, 1861–8936
- [WHA<sup>+</sup>08] WIDERGREN, Steven ; HARDIN, Dave ; AMBROSIO, Ron ; DRUMMOND, R. ; GUNTHER, Erich ; GILCHRIST, Grant ; COHEN, David: GridWise Interoperability Context-Setting Framework. (2008), Januar
- [Win92] WINSTON, Patrick H.: *Artificial intelligence*. 3rd ed. Reading, Mass : Addison-Wesley Pub. Co, 1992. – ISBN 9780201533774
- [Wit79] WITTE, Hermann: Kritische Aspekte der Nutzwertanalyse. In: *Jahrbuch für Sozialwissenschaft* 30 (1979), Nr. 1, 135–146. <http://dx.doi.org/10.2307/20714317>. – DOI 10.2307/20714317. – ISSN 0075–2770
- [WLMD10] WIDERGREN, S ; LEVINSON, A ; MATER, J ; DRUMMOND, R: Smart grid interoperability maturity model. In: *IEEE PES General Meeting*. Minneapolis, MN : IEEE, Juli 2010. – ISBN 978–1–4244–6549–1, 1–6
- [WSM<sup>+</sup>19] WANG, Alex ; SINGH, Amanpreet ; MICHAEL, Julian ; HILL, Felix ; LEVY, Omer ; BOWMAN, Samuel R.: *GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding*. <http://dx.doi.org/10.48550/arXiv.1804.07461>. Version: Februar 2019. – arXiv:1804.07461
- [XUD<sup>+</sup>19] XU, Feiyu ; USZKOREIT, Hans ; DU, Yangzhou ; FAN, Wei ; ZHAO, Dongyan ; ZHU, Jun: Explainable AI: A Brief Survey on History,

- Research Areas, Approaches and Challenges. In: TANG, Jie (Hrsg.) ; KAN, Min-Yen (Hrsg.) ; ZHAO, Dongyan (Hrsg.) ; LI, Sujian (Hrsg.) ; ZAN, Hongying (Hrsg.): *Natural Language Processing and Chinese Computing*. Cham : Springer International Publishing, 2019. – ISBN 9783030322366, S. 563–574
- [Zan14] ZANGEMEISTER, Christof: *Nutzwertanalyse in der Systemtechnik: eine Methodik zur multidimensionalen Bewertung und Auswahl von Projektalternativen*. 5. Auflage 2014 (erweitert). Winnemark : Zangemeister & Partner, 2014. – ISBN 978-3-923264-00-1
- [Zay24] ZAYYANU, Zaki M.: Revolutionising Translation Technology: A Comparative Study of Variant Transformer Models - BERT, GPT, and T5. In: *Computer Science & Engineering: An International Journal* 14 (2024), Juni, Nr. 3, 15–27. <http://dx.doi.org/10.5121/cseij.2024.14302>. – DOI 10.5121/cseij.2024.14302. – ISSN 22313583
- [Zik12] ZIKOPOULOS, Paul: *Understanding big data: analytics for enterprise class Hadoop and streaming data*. New York : McGraw-Hill, 2012. – ISBN 9780071790536. – OCLC: ocn760977087
- [ZKW+20] ZHANG, Tianyi ; KISHORE, Varsha ; WU, Felix ; WEINBERGER, Kilian Q. ; ARTZI, Yoav: *BERTScore: Evaluating Text Generation with BERT*. <http://dx.doi.org/10.48550/arXiv.1904.09675>. Version: Februar 2020. – arXiv:1904.09675 [cs]
- [ZLX+23] ZHOU, Chunting ; LIU, Pengfei ; XU, Puxin ; IYER, Srini ; SUN, Jiao ; MAO, Yuning ; MA, Xuezhe ; EFRAT, Avia ; YU, Ping ; YU, Lili ; ZHANG, Susan ; GHOSH, Gargi ; LEWIS, Mike ; ZETTLEMOYER, Luke ; LEVY, Omer: *LIMA: Less Is More for Alignment*. <http://dx.doi.org/10.48550/arXiv.2305.11206>. Version: Mai 2023. – arXiv:2305.11206
- [ZQD+21] ZHUANG, Fuzhen ; QI, Zhiyuan ; DUAN, Keyu ; XI, Dongbo ; ZHU, Yongchun ; ZHU, Hengshu ; XIONG, Hui ; HE, Qing: A Comprehensive Survey on Transfer Learning. In: *Proceedings of the IEEE* 109 (2021), Januar, Nr. 1, 43–76. <http://dx.doi.org/10.1109/JPROC.2020.3004555>. – DOI 10.1109/JPROC.2020.3004555. – ISSN 1558-2256
- [ZZL+25] ZHAO, Wayne X. ; ZHOU, Kun ; LI, Junyi ; TANG, Tianyi ; WANG, Xiaolei ; HOU, Yupeng ; MIN, Yingqian ; ZHANG, Beichen ; ZHANG, Junjie ; DONG, Zican ; DU, Yifan ; YANG, Chen ; CHEN, Yushuo ; CHEN, Zhipeng ; JIANG, Jinhao ; REN, Ruiyang ; LI, Yifan ; TANG, Xinyu ; LIU, Zikang ; LIU, Peiyu ; NIE, Jian-Yun ; WEN, Ji-Rong: *A Sur-*

---

*vey of Large Language Models*. <http://dx.doi.org/10.48550/arXiv.2303.18223>. Version: März 2025. – arXiv:2303.18223

---

## Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen verwendet, sowie keine unzulässige fremde Hilfe in Anspruch genommen habe. Zudem wurde die Arbeit nicht unter unkenntlichem Einsatz generativer KI erbracht. Ich habe erlaubterweise KI-basierte Werkzeuge (z. B. ChatGPT, DeepL) ausschließlich zur Verbesserung von Sprache und Grammatik eingesetzt. Sofern nicht kenntlich gemacht, stammen sämtliche inhaltlichen Entscheidungen, wissenschaftlichen Argumentationen und Implementierungen vollständig von mir. Außerdem versichere ich, dass ich die allgemeinen Prinzipien wissenschaftlichen Arbeitens und Veröffentlichens, wie sie in den Leitlinien guter wissenschaftlicher Praxis der Carl von Ossietzky Universität Oldenburg festgelegt sind, befolgt habe.

Weiterhin erkläre ich, dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Mir ist bekannt, dass eine eidesstattliche Versicherung eine nach den §§ 156, 161 Strafgesetzbuch (StGB) strafbewertete Bestätigung der Richtigkeit meiner Erklärung ist. Mir sind die strafrechtlichen Folgen einer unrichtigen, d.h. nicht den Tatsachen entsprechenden und unvollständigen Erklärung, d.h. das Verschweigen der wesentlichen Tatsachen bekannt.

Ort, Datum: \_\_\_\_\_

Vorname, Nachname: \_\_\_\_\_

Unterschrift (handschriftlich): \_\_\_\_\_