

# **Erklärbarkeit und Visualisierung von Graph Codes mittels generativer KI**

Bachelorarbeit  
im Studiengang B.Sc. Informatik

eingereicht von  
**Jens Nathan Andreß**  
(Matrikelnummer 9763180)

angefertigt am  
**Lehrgebiet Multimedia  
und Internetanwendungen**

Betreuer:  
**Dr.-Ing. Stefan Wagenpfeil**

29. August 2023



# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b>	<b>i</b>
<b>Abbildungsverzeichnis</b>	<b>iii</b>
<b>Tabellenverzeichnis</b>	<b>iv</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Motivation . . . . .	3
1.2 Problembeschreibung . . . . .	5
1.2.1 Probleme in Bezug auf die Erklärbarkeit . . . . .	5
1.2.2 Probleme in Bezug auf die Integration . . . . .	8
1.2.3 Allgemeine Zielsetzung . . . . .	8
1.3 Forschungsfragen . . . . .	9
1.4 Methodik und Ziele . . . . .	9
1.5 Ansatz und Aufbau der Arbeit . . . . .	11
1.6 Arbeits- und Zeitplan . . . . .	13
1.7 Zusammenfassung . . . . .	14
<b>2 Stand der Wissenschaft und Technik</b>	<b>15</b>
2.1 FZ 1.1/O Erklärbarkeit von MMIR mittels generativer KI . . . . .	16
2.1.1 GMAF . . . . .	16
2.1.2 MMFG . . . . .	17
2.1.3 Graph Codes . . . . .	22
2.1.4 Generative KI . . . . .	29
2.1.5 Diskussion . . . . .	37
2.2 FZ 2.1/O Integration generativer KI in das GMAF . . . . .	47
2.2.1 Integrationsmöglichkeiten von Graph Codes . . . . .	47
2.2.2 Integrationsmöglichkeiten von Systemen generativer KI . . . . .	49
2.2.3 Diskussion . . . . .	52
2.3 Evaluierungsmethodiken . . . . .	54
2.4 Zusammenfassung . . . . .	57
2.4.1 Welche offenen Herausforderungen (OH) verbleiben? . . . . .	57
2.4.2 Gewonnene Erkenntnisse . . . . .	58
2.4.3 Getroffene Entscheidungen . . . . .	59
2.4.4 Behandlung offener Herausforderungen . . . . .	59
<b>3 Modellierung</b>	<b>62</b>
3.1 Konzeptions- und Modellierungsmethodik . . . . .	62
3.1.1 Rational Unified Process . . . . .	63
3.1.2 Unified Modeling Language . . . . .	64

3.2	FZ 1.2/TB Erklärbarkeit von MMIR mittels generativer KI . . . . .	66
3.2.1	Erklärbarkeit durch generative KI . . . . .	66
3.2.2	Anwendungsfälle . . . . .	66
3.2.3	Diskussion . . . . .	83
3.3	FZ 2.2/TB Integration generativer KI in das GMAF . . . . .	85
3.3.1	Transformation von Graph Codes . . . . .	85
3.3.2	Einbindung von Systemen generativer KI . . . . .	89
3.3.3	Diskussion . . . . .	93
3.4	Zusammenfassung . . . . .	94
3.4.1	Gewonnene Erkenntnisse . . . . .	94
3.4.2	Weiteres Vorgehen . . . . .	94
<b>4</b>	<b>Implementierung</b>	<b>97</b>
4.1	Grundlagen der Implementierung . . . . .	97
4.2	FZ 1.3/I Erklärbarkeit von MMIR mittels generativer KI . . . . .	99
4.3	FZ 2.3/I Integration generativer KI in das GMAF . . . . .	100
4.4	Zusammenfassung . . . . .	101
<b>5</b>	<b>Evaluierung</b>	<b>102</b>
<b>A</b>	<b>Anhang</b>	<b>103</b>

## Abbildungsverzeichnis

1.1	Begriffe im Rahmen von künstlicher Intelligenz. . . . .	2
1.2	Beispiel für eine generierte Erklärung . . . . .	4
1.3	Bauanleitung für ein Lego-Set [20]. . . . .	6
1.4	Beispiel für ein generiertes Vorschaubild . . . . .	6
1.5	HeadUp-Display in einem Kraftfahrzeug [24]. . . . .	7
1.6	Benutzerschnittstelle im <i>GMAF</i> zum Anpassen von Erklärungen. . . . .	7
1.7	Phasen der Problemlösung nach Nunamaker [25]. . . . .	10
1.8	Aufbau der Arbeit und Ansatz der Problemlösung nach Nunamaker [25]. . . . .	12
1.9	Arbeits- und Zeitplan. . . . .	13
2.1	Vereinfachte Übersicht des GMAFs (ohne UI / API) nach [22]. . . . .	17
2.2	Visualisierung eines MMFGs durch yEd. . . . .	19
2.3	Beispiel für einen <i>MMFG<sub>ex</sub></i> . . . . .	20
2.4	Wertungsmatrix des Beispiels <i>MMFG<sub>ex</sub></i> in a) Tabellenform und b) Form einer Matrix. . . . .	23
2.5	Codierte Wertungsmatrix bzw. Graph Code <i>GC<sub>ex</sub></i> zum entsprechenden Beispiel <i>MMFG<sub>ex</sub></i> in a) Tabellenform und b) in Form einer Matrix. c) Vom GMAF generierte Graph Code <i>GC<sub>ex</sub></i> . . . . .	24
2.6	Übersicht ausgewählter aktueller Systeme generativer KI nach [84]. . . . .	35
2.7	Beispiel für ein von Midjourney generiertes Bild anhand oben genannter Textbeschreibung [88]. . . . .	38
2.8	Beispiel für ein von Stable Diffusion generiertes Bild anhand oben genannter Textbeschreibung. . . . .	39
2.9	Beispiel für ein von DALL · E 2 generiertes Bild anhand oben genannter Textbeschreibung. . . . .	40
3.1	Modellierung zum Vorgehen des Rational Unified Process. . . . .	63
3.2	Übersicht über alle Anwendungsfälle. . . . .	67
3.3	Wireframe für die Grundbereiche der Benutzerschnittstelle. . . . .	72
3.4	Wireframe für den linken Arbeitsbereich (links) und den rechten Arbeitsbereich (rechts) der Benutzerschnittstelle. . . . .	73
3.5	Komplettes Wireframe für die Benutzerschnittstelle. . . . .	74
3.6	Mechanismus für den Anwendungsfall UC-1.1. . . . .	75
3.7	Mechanismus für den Anwendungsfall UC-1.2. . . . .	76
3.8	Mechanismus für den Anwendungsfall UC-1.3. . . . .	76
3.9	Mechanismus für den Anwendungsfall UC-1.5. . . . .	77
3.10	(i) Synchrone Nachricht, (ii) Asynchrone Nachricht, (iii) Antwort auf Nachricht, (iv) Selbstdelegation / Selbstaufruf. . . . .	79
3.11	Sequenzdiagramm für den Anwendungsfall UC-1.1. . . . .	80
3.12	Sequenzdiagramm für den Anwendungsfall UC-1.2. . . . .	81

3.13	Sequenzdiagramm für den Anwendungsfall UC-1.3.	82
3.14	Sequenzdiagramm für den Anwendungsfall UC-1.5.	83
3.15	Einfaches Klassendiagramm für die Erklärungspaneele.	86
3.16	Endpunkte der OpenAI Schnittstelle.	90
3.17	Erweiterung zum Sequenzdiagramm für den Anwendungsfall UC-1.3.	91
3.18	Mechanismus für den Anwendungsfall UC-1.8.	92
3.19	Sequenzdiagramm für den Anwendungsfall UC-1.8.	92
4.1	Hierarchie des Quellcodes.	98

## Tabellenverzeichnis

1.1	Aufschlüsselung der Forschungsfragen auf Forschungsziele . . . . .	11
2.1	Struktur dieses Kapitels. . . . .	15
2.2	$dict_{ex}$ vom Graph Code $GC_{ex}$ . . . . .	25
2.3	Offene Herausforderungen . . . . .	57
2.4	Tabelle zur Übersicht des aktuellen Arbeitstands. . . . .	60
3.1	Struktur dieses Kapitels. . . . .	62
3.2	Wesentliche Arbeitsergebnisse der Modellierung . . . . .	94
3.3	Tabelle zur Übersicht des aktuellen Arbeitstands. . . . .	95

# 1 Einleitung

Die Digitalisierung der Welt schreitet immer rasanter voran. Nicht zuletzt wird im Kontext der Digitalisierung auch von dem „Informationszeitalter“ und „Computarisierung“ gesprochen [1]. In einer zunehmend digitaleren Welt nimmt die Anzahl internethfähiger, digitaler Geräte, den sogenannten *Internet Of Things* (IoT)-Geräten und Anwendern derselben immer mehr zu [2]. Laut Statistiken belief sich die Zahl der Internetnutzer, den sogenannten „Onlinern“ [3], im Jahr 2022 auf rund 5.3 Milliarden. Im Vergleich zum letzten Jahrzehnt ist somit die Anzahl der Internetanwender innerhalb eines Jahrzehnts um rund 2.9 Milliarden auf beinahe das Doppelte angestiegen [3]. Auch die weltweite *Covid-19*-Pandemie hatte keinen hindernden Effekt auf diese Entwicklung [4]. Aufgrund der Pandemie waren viele Menschen dazu gezwungen mehr Zeit zu Hause zu verbringen, oder von zu Hause aus zu arbeiten. Dies hatte zur Folge, dass mehr Menschen sich digitalen Technologien und Lösungen zuwandten, um mittels Videokonferenzsystemen von zu Hause aus zu arbeiten, oder mittels sozialer Medien mit Freunden und Bekannten in Kontakt zu bleiben [4, 5]. Anhand der weltweiten Reaktionen auf die Pandemie wurde die Umstellung auf digitale Technologien um mehrere Jahre beschleunigt [4]. Die Verwendung dieser Technologien, wie z.B. den sozialen Medien, wie Instagram [6], Facebook [7], Snapchat [8] und weiteren führt zu enormen Datenmengen, die ohne Hilfsmittel für Menschen nicht mehr greifbar und überschaubar sind. So betrug die Datennutzung im Jahr 2022 schätzungsweise 100 000 Petabytes, von welchem Mobiltelefone einen Großteil ausmachten [9].

Sollen, unabhängig vom Anwendungsbereich, diese enormen Datenmengen und die daraus extrahierten Merkmale effektiv verarbeitet und gewonnene Erkenntnisse effektiv kommuniziert werden, so gilt es die Beziehung zwischen der Darstellung der identifizierten Merkmale und der für Menschen verständlichen Bedeutung zu berücksichtigen. Zwischen der Darstellung der identifizierten Merkmale und der für Menschen verständlichen Bedeutung besteht eine sogenannte semantische Lücke, die es zu überbrücken gilt [10]. Die semantische Lücke bezeichnet hierbei den Unterschied zwischen den jeweiligen Darstellungsformen. Die Darstellungsform der identifizierten Merkmale ist ihre maschinelle Darstellung in Form von Nullen und Einsen, auch binäre Darstellung genannt und die Darstellung der für Menschen verständliche Bedeutung kann in natürlicher Sprache erfolgen. Zwischen diesen Darstellungsformen liegt eine semantische Lücke, dessen Überbrückung in dieser Arbeit untersucht werden soll. Das Überbrücken dieser semantischen Lücke ist eine große Herausforderung und wird seit Jahren intensiv erforscht und ist, aufgrund immer breiterer Anwendung, Gegenstand zahlreicher Untersuchungen im Bereich des *Multimedia Information Retrieval*, kurz *MMIR* [10].

Das Thema dieser Arbeit ist im Bereich *MMIR* und künstlicher Intelligenz (KI) verortet. Das Themengebiet der künstlichen Intelligenz ist nicht neu und schon seit Jahrzehnten Gegenstand der Forschung und zahlreicher Untersuchungen [11]. Die Anwendungsmöglichkeiten von künstlicher Intelligenz sind weitreichend und haben zunehmend Einfluss auf unseren Alltag. Viele Aufgaben, die früher von Menschen erledigt wurden, oder nicht zu erledigen waren, werden bereits durch künstliche Intelligenz übernommen und automatisiert [12]. Weiterhin stellt künstliche Intelligenz für viele Unternehmen und Organisationen schon heute effektiv einen wirtschaftlichen Mehrwert dar [12]. Abb. 1.1 zeigt eine Übersicht wichtiger Begriffe im Bereich der künstlichen Intelligenz, sowie die Beziehung dieser Begriffe zueinander.

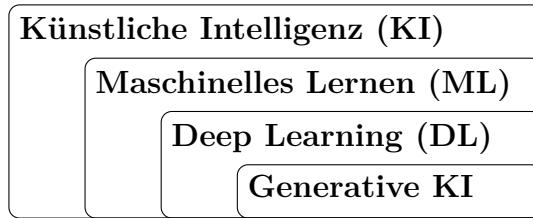


Abb. 1.1: Begriffe im Rahmen von künstlicher Intelligenz.

Künstliche Intelligenz ist ein Überbegriff und bezeichnet den Ansatz, mit Maschinen intelligentes menschliches Verhalten nachzuahmen, um dadurch Probleme zu lösen [13]. Maschinelles Lernen ist ein Überbegriff für die Technologien, die eingesetzt werden, um künstliche Intelligenz zu erreichen [13]. Deep Learning ist eine Weiterentwicklung des maschinellen Lernens und macht Gebrauch von neuronalen Netzwerken, welche dem menschlichen Gehirn nachempfunden sind [13]. Generative KI ist ein besonderer Bereich der künstlichen Intelligenz und ein Sammelbegriff für auf künstlicher Intelligenz basierender Systeme, mit welcher neue und originelle Inhalte erstellt werden können [14].

Die enormen und komplexen Datenmengen, die mittlerweile im Internet generiert werden [9], verstärken das Potenzial von maschinellem Lernen, sowie aber auch ihre Notwendigkeit [15]. Noch bis vor Kurzem war maschinelles Lernen auf vorausschauende Modelle beschränkt, die nur in der Lage waren Beobachtungen vorzunehmen und Muster zu erkennen. Durch generative künstliche Intelligenz, wie ChatGPT-3.5 [16] oder Dall-E 2 [17], konnte ein Durchbruch erzielt werden, welcher ermöglicht neue Inhalte zu erzeugen [15].

In dieser Arbeit wird untersucht, ob und inwiefern generative KI verwendet werden kann, um die semantische Lücke zu schließen und inwieweit generative KI geeignet ist, präzise Erklärungen bzw. geeignete semantische Darstellungen aus identifizierten Merkmalen zu generieren.

## 1.1 Motivation

Bei dieser Arbeit handelt es sich um eine Bachelorarbeit im Forschungsumfeld von Dr.-Ing. Stefan Wagenpfeil am Lehrgebiet Multimedia und Internetanwendungen (MMIA) von Prof. Dr.-Ing. Matthias Hemmje. Schon vor der Bachelorarbeit bin ich über das von diesem Lehrgebiet angebotene Fachpraktikum Multimedia Information Retrieval mit dem Bereich *MMIR* in Kontakt gekommen. Da mich das Fachpraktikum und das Themengebiet des *MMIR* sehr interessiert und begeistert hat, schreibe ich nun in diesem Lehrgebiet meine Bachelorarbeit, um meine Fähigkeiten auszubauen und mein Wissen zu vertiefen.

Das Lehrgebiet MMIA hat zum Zwecke der Forschung im Bereich des *MMIR* das Forschungsprojekt *Generic Multimedia Analysis Framework*, kurz *GMAF* ins Leben gerufen. Dieses Forschungsprojekt wurde entwickelt, um aktuelle Missstände und Probleme des einfachen *MMIR* auf intelligente Art und Weise zu adressieren. Das *GMAF* hat dabei zum Ziel durch eine intelligente Erweiterung von *MMIR* semantische, erklärbare, für Menschen verständliche, effektive, effiziente, kompatible und integrierte Lösungen anzubieten [18, S. 20]. Diese Erweiterung wird *Smart Multimedia Information Retrieval*, kurz *Smart MMIR* genannt. *Smart MMIR* führt eine Reihe von Konzepten ein, die über klassisches *MMIR* hinaus gehen. Hierzu gehören erweiterte Metriken für die Berechnung von Ähnlichkeit oder Empfehlungen, Möglichkeiten der Merkmalsintegration und -fusion, sowie die Erklärbarkeit von Prozessen des *MMIR* und deren Inhalte. Unter Erklärbarkeit versteht man in diesem Zusammenhang für Menschen verständliche Antworten auf Fragestellungen, wie z.B. die Frage „Was ist die wichtigste Eigenschaft des Elements?“ oder „Was ist der Unterschied zwischen diesen zwei Bildern?“. Sollen Antworten auf diese Fragen auf für Menschen verständliche Art und Weise ausgedrückt werden, so müssen diese Antworten auf natürlicher Sprache aufbauen, oder prägnant visuell dargestellt werden. Ein Beispiel für eine generierte textuelle Erklärung, die auf natürlicher Sprache aufbaut, ist in Abb. 1.2 zu sehen.

Das *GMAF* bietet hierfür ein umfangreiches Konzept zur Generierung von textuellen Beschreibungen von *MMIR*-Inhalten in für Menschen lesbarer und verständlicher Form. Es ist allerdings wichtig anzumerken, dass dieses Konzept statistisch und rein mathematische Ansätze zur Generierung von natürlicher Sprache verfolgt, ohne dabei Methoden maschinellen Lernens zu verwenden. Diese Ansätze generieren jedoch nur textuelle Beschreibungen auf eine statische Art und Weise. Zudem können mit diesen Ansätzen keine anderen Formen, wie z.B. visuelle Beschreibungen, außer Text generiert werden. Es ist also erstrebenswert nach einer Lösung zu suchen, mit welcher man sich von diesen statischen Ansätzen lösen kann. Ein anderer Ansatz, der in den letzten Monaten stark an Popularität gewonnen hat, ist die generative KI. Generative KI scheint eine vielversprechende Alternative zu einem statischen, statistischen und rein mathematischen Ansatz zu sein.



The picture shows two primary objects. The first object is a dog, the second object is a woman. The dog is white and has black eyes. The woman is smiling and has blonde hair. The woman is wearing shoes, shorts, a top, and a watch. The shoes are blue. The shorts are blue. The top is black. The watch is an Apple Watch. In the background there is grass.

Abb. 1.2: Beispiel für eine generierte textuelle Beschreibung eines Bildes im GMAF.

Ließen sich mittels generativer KI tatsächlich *MMIR*-Prozesse erklären, so kann dies einen echten Mehrwert für Benutzer darstellen, da diese von präzisen Erklärungen oder visuellen Darstellungen profitieren würden. Beispiele für Mehrwert für Benutzer können sein:

- **Zeitersparnis.** Generative KI könnte anhand identifizierter Merkmale Datei(en) zusammenfassen und in kürzester Zeit präzise Beschreibungen der Datei(en) erstellen. Ein Benutzer müsste für diese Arbeit erheblich mehr Zeit aufwenden.
- **Überblick / Ordnung.** Generative KI könnte anhand identifizierter Merkmale in einer großen Kollektion an Dateien Zusammenhänge erkennen und dem Benutzer eine Beschreibung dieser Zusammenhänge bieten. Zusammenhänge können in diesem Fall Muster oder eine für den Benutzer versteckte bzw. nicht ersichtliche Ordnung innerhalb der Dateien sein.
- **Barrierefreiheit.** Es gibt bereits Programme, wie z.B. Screen Readers, die blinden Personen Beschreibungen von Bildern vorlesen. Nachteil ist, dass diese Programme abhängig von den im Bild eingebetteten Beschreibungen sind, und viele Bilder im Internet über keine eingebetteten Beschreibungen verfügen. Zusätzlich sind diese Beschreibungen meistens von Menschen verfasst [19]. Generative KI könnte anhand identifizierter Merkmale Beschreibungen solcher Bilder generieren. Diese Beschreibungen könnten in weiteren Schritten von anderen Programmen, oder selbst wieder generativer KI verwendet werden, um diese vorzulesen. Auf diese Weise könnte echte Barrierefreiheit erreicht werden.

## 1.2 Problembeschreibung

Das *GMAF* bietet bereits ein umfangreiches Konzept zur Erklärbarkeit von Prozessen des *MMIR*. Die Ansätze, die in diesem Konzept verfolgt werden, sind statisch, in ihrer Anwendung begrenzt und ihre Erklärbarkeit vergleichsweise rudimentär. In diesem Abschnitt werden diese Probleme angesprochen, umrissen und subsequent Problembeschreibungen definiert.

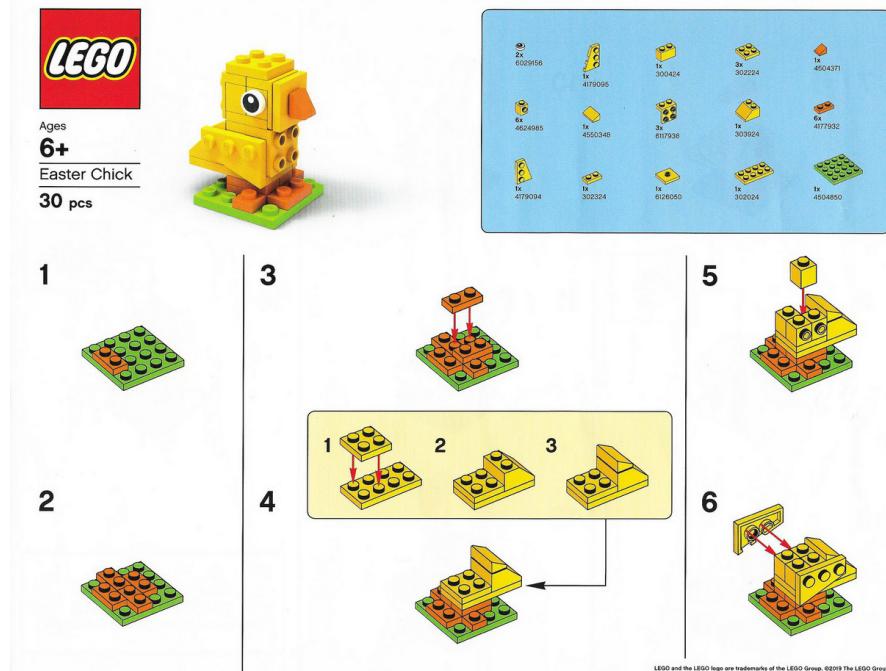
### 1.2.1 Probleme in Bezug auf die Erklärbarkeit

Das *GMAF* bietet bereits Verfahren zur Erklärbarkeit von *MMIR*-Prozessen. Die Ansätze, die mit diesen Verfahren umgesetzt werden, sind allerdings sehr statisch und basieren wiederum auf statistischen und rein mathematischen Ansätzen. Noch dazu bieten diese Verfahren nur die Erklärbarkeit von *MMIR*-Prozessen anhand von Text. Es ist allerdings wichtig anzumerken, dass Text nicht die einzige Form der Erklärbarkeit ist. Eine andere Form der Erklärbarkeit kann mittels Bilder erzielt werden. So lautet ein bekanntes Sprichwort: „*Ein Bild sagt mehr als 1000 Worte*“. Dieses Sprichwort will zum Ausdruck bringen, dass Bilder, im Vergleich zu Text, ebenfalls komplexe Sachverhalte und Wissen prägnant und in für Menschen verständlicher Form zum Ausdruck bringen können. Ein Beispiel hierfür wäre eine Bauanleitung zu einem Lego-Set. Ein Beispiel für eine Bauanleitung zum Bauen eines Osterkükens ist in Abb. 1.3 zu sehen. Die Bauanleitung veranschaulicht dabei die einzelnen Aufbauschritte durch fortwährende Nummerierung und Abbildungen.

Ein konkretes Anwendungsbeispiel für die Erklärbarkeit in Form von Bildern wäre im Rahmen des *MMIR* die Generierung von Vorschaubildern für eine Textdatei. Abb. 1.4 zeigt ein Beispiel für ein von Dall-E 2 [17] generiertes Vorschaubild für ein Gedicht über einen Schmetterling [21]. Anhand der in der Textdatei identifizierten Merkmale könnte eine generative KI, wie z.B Dall-E 2, ein Vorschaubild generieren, um somit den Inhalt des Textes visuell zu erklären. Ein Vorschaubild kann sich dann als sinnvoll erweisen, wenn viele Dateien zu verwalten sind und Benutzer nicht die Zeit haben diese einzeln zu durchforsten. Durch einen kurzen Blick auf das Vorschaubild könnte ein Benutzer die Datei einordnen und sich besser zurechtfinden.

Ein weiteres Anwendungsbeispiel wäre die Gemeinsamkeit/Differenz aus einer Kollektion an Bildern. Die als Gemeinsamkeit/Differenz identifizierten Merkmale könnten in diesem Beispiel ebenfalls konkret mittels eines Bildes kurz und prägnant visuell dargestellt werden. Diese Art der visuellen Erklärung kann sich dann als sinnvoll erweisen, wenn viele Bilder auf ihre Merkmale hin untersucht werden sollen. Ein Benutzer kann hier schnell den Überblick verlieren.

Statistische und rein mathematische Ansätze reichen für diese Beispiele nicht mehr aus. Ein anderer, vielversprechender Ansatz, um die Erklärbarkeit von Prozessen des



Haftungsausschluss: LEGO® und das LEGO Logo sind Marken der LEGO Unternehmensgruppe.

© 2023 Lego Gruppe. Diese Arbeit ist unabhängig und wurde von der LEGO Gruppe weder genehmigt, noch gesponsort.

Abb. 1.3: Bauanleitung für ein Lego-Set [20].

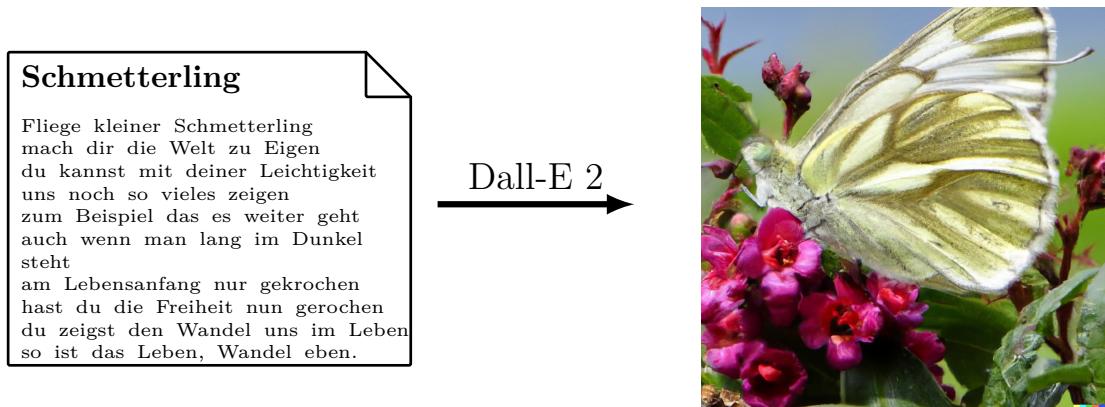


Abb. 1.4: Ein von Dall-E 2 generiertes Vorschaubild.

*MMIR* zu erreichen, liegt in generativer KI. Mittels generativer KI könnten anhand identifizierter Merkmale eigenständig neuartige Multimedia-Inhalte, wie Texte oder Bilder, erzeugt werden, die zur Erklärbarkeit von *MMIR*-Prozessen dienen könnten.

Je nach Anwendungsbereich, variieren die Anforderungen an Erklärungen allerdings erheblich. Somit ist ein weiterer wichtiger Aspekt die Anpassungsmöglichkeiten von Erklärungen in Abhängigkeit ihrer Anwendungsfälle. So wird *Smart MMIR* für die Erkennung sicherheitsrelevanter Zustände im Bereich der Automobilindustrie, auch Automotive genannt, eingesetzt [22]. Sicherheitsrelevante Zustände in einem Kraftfahrzeug sind Zustände, die für eine reibungslose Fahrt und Funktion sorgen, oder unterstützend und im Sinne des Autofahrers eingreifen oder im Ernstfall zum Schutz von Insassen und Außenstehenden beitragen [23]. Ein Beispiel für eine unterstützende Funktion kann in Form von HeadUp-Displays (HUD) gefunden werden und ist in Abb. 1.5 zu sehen. Das HeadUp-Display projiziert sicherheitrelevante Informationen, wie z.B. die Geschwindigkeit auf die Windschutzscheibe. Da im Bereich



Abb. 1.5: HeadUp-Display in einem Kraftfahrzeug [24].

Automotive schnelle Entscheidungen getroffen werden müssen, werden hier kurze, präzise Erklärungen benötigt. Es ist also in Bezug auf die Erklärbarkeit notwendig Eigenschaften, wie z.B. die Länge der Erklärung, anpassen zu können. Das *GMAF* bietet bereits eine Benutzerschnittstelle, die es ermöglicht Anpassungen von textuellen Erklärungen vorzunehmen. Diese Benutzerschnittstelle ist in Abb. 1.6 zu sehen und zeigt die Elemente, die zur Anpassung von textuellen Erklärungen vorgesehen sind. Diese Benutzerschnittstelle bietet allerdings nur grobe Möglichkeiten

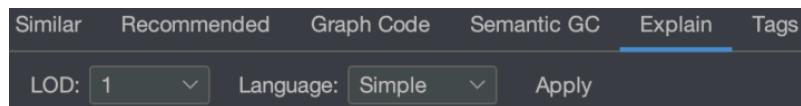


Abb. 1.6: Benutzerschnittstelle im *GMAF* zum Anpassen von Erklärungen.

zur Anpassung von textuellen Erläuterungen, wie die Anpassung der Komplexität der textuellen Beschreibung. Möglichkeiten zur Anpassungen der Länge der textuellen

Beschreibungen ermöglicht diese Schnittstelle jedoch nicht. Schlussendlich kann in Bezug auf die Erklärbarkeit von *MMIR*-Prozessen folgende Problembeschreibung formuliert werden:

**PB1:** Generative KI ist aktuell nicht für die Erklärbarkeit von *MMIR*-Prozessen nutzbar.

### 1.2.2 Probleme in Bezug auf die Integration

Damit generative KI-Systeme aus identifizierten Merkmalen präzise und für Menschen verständliche Erklärungen generieren können, müssen die identifizierten Merkmale als Eingabedaten in die jeweiligen Systeme überführt werden. Das *GMAF* bietet keine Möglichkeiten zur Anbindung von Systemen, die generative KI anbieten, sowie kein Modell zur Übersetzung von identifizierten Merkmalen in Eingabedaten für entsprechende Systeme. Folglich gibt es auch keine Modellierung, die die Anpassung der Eingabedaten ermöglicht, um die in Abschnitt 1.2.1 benannten Modifikationen von Erklärungen zu ermöglichen. Gelingt eine Integration von generativen KI-Systemen, so kann dies einen echten Mehrwert für Benutzer darstellen, sofern dadurch verbesserte und originellere Erklärungen von *MMIR*-Prozessen generiert werden. Besonders die Möglichkeit, je nach Anforderung die Eingabedaten anzupassen, kann für Benutzer von Nutzen sein. Schlussendlich kann in Bezug auf die Integration folgende Problembeschreibung formuliert werden:

**PB2:** Das *GMAF* bietet keine Integration generativer KI.

### 1.2.3 Allgemeine Zielsetzung

Aus den in Abschnitt 1.2.1 und Abschnitt 1.2.2 beschriebenen Problembereichen lassen sich zwei direkte Probleme erfassen: Die im *GMAF* implementierten Konzepte zur Erklärung von *MMIR*-Prozessen sind noch sehr statisch und bieten nur Erklärungen in Form von Text, sowie die Integration besserer Lösungen zur Erklärbarkeit von *MMIR*-Prozessen. Die Motivation dieser Arbeit besteht darin, dass die bereits implementierten Konzepte im *GMAF* in Bezug auf die Erklärbarkeit von *MMIR*-Prozessen noch großen Raum für Verbesserung aufweisen. Verbesserung wird sich durch das Anbinden von generativer KI versprochen, da generative KI in der Lage zu sein scheint auf kreative Weise neuartige Inhalte zu generieren.

Werden nun die Problembeschreibungen zusammengefasst, so wird das übergeordnete Ziel sichtbar:

*Die im GMAF implementierten Konzepte zur Erklärungen von MMIR-Prozessen sollen durch generative KI abgelöst werden, um bessere Erklärungen, sowie andere Formen von Erklärungen zu ermöglichen.*

Anhand der in diesem Abschnitt beschriebenen Problembeschreibungen und des allgemeinen Ziels können nun im nächsten Abschnitt 1.3 folgende Forschungsfragen formuliert werden. Im darauffolgenden Abschnitt 1.4 wird genauer auf die Methodik dieser Arbeit, sowie die konkreteren Ziele eingegangen.

### 1.3 Forschungsfragen

Um die im vorigen Abschnitt definierten Problembeschreibungen zu adressieren, werden folgende Forschungsfragen definiert:

- **FF1:** Kann generative KI für die Erklärbarkeit von *MMIR*-Prozessen genutzt werden?
- **FF2:** Wie lässt sich generative KI in das *GMAF* integrieren?

### 1.4 Methodik und Ziele

Um die in Abschnitt 1.2 identifizierten Probleme und subsequent die im vorigen Abschnitt 1.3 formulierten Forschungsfragen strukturiert beantworten zu können, wird diese Arbeit auf der vielfach bewährten Methodik von Nunamaker [25] aufbauen. Die Methodik nach Nunamaker teilt ein zu lösendes Problem anhand von Recherche und Entwicklung in vier Phasen der Problemlösung auf.

Diese vier Phasen lauten:

1. Beobachtungsphase
2. Theoriebildungsphase
3. Systementwicklungs - bzw. Implementierungsphase
4. Experimentphase

In der Beobachtungsphase werden im Rahmen einer Recherche Informationen gesammelt. In der Theoriebildungsphase werden Konzepte für Lösungen konzipiert und modelliert. Die Systementwicklungs bzw. Implementierungsphase beschäftigt sich mit der Entwicklung von Lösungen basierend auf den Theorien aus der Theoriebildungphase bzw. Beobachtungen aus der Beobachtungsphase, und welche in der letzten Phase, der Experimentphase evaluiert werden können. Diese Phasen sind eng miteinander verwoben und jede Phase mündet mit seinem Ergebnis in den anderen Phasen und trägt zu diesen bei. Durch die Methodik von Nunamaker wird sicher gestellt, dass jedes Forschungsziel eindeutig einer Problembeschreibung zugeordnet werden kann.

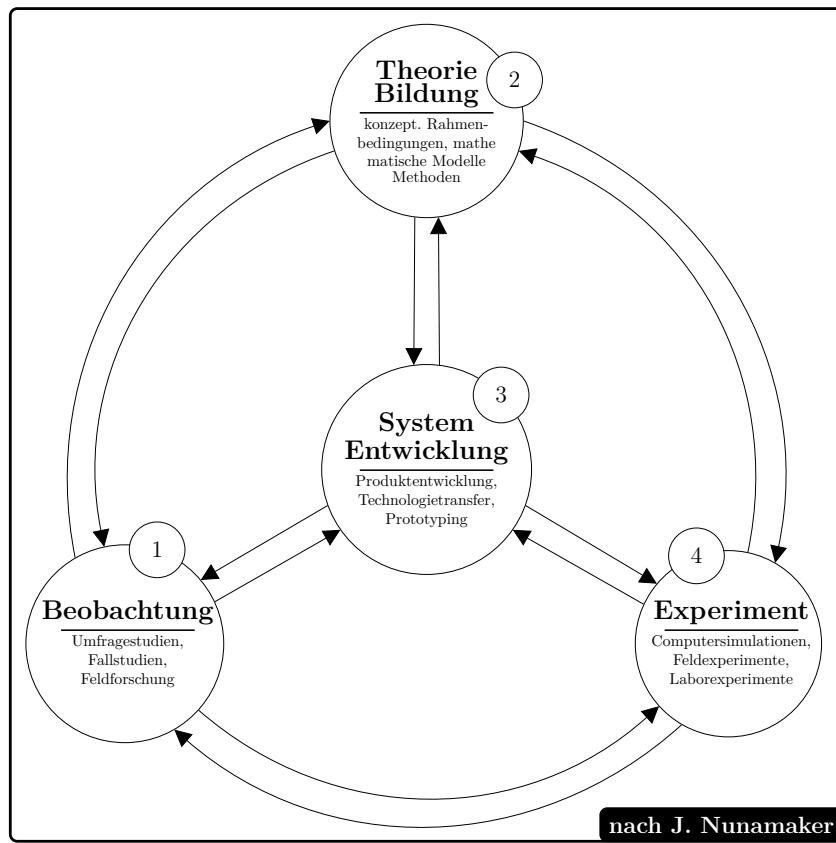


Abb. 1.7: Phasen der Problemlösung nach Nunamaker [25].

Die Anwendung der Methodik nach Nunamaker führt für jede Forschungsfrage zu jeweils vier Forschungszielen, die in jeweils einem Unterkapitel behandelt werden. Dies führt insgesamt zu 8 Forschungsaufgaben, die mit dieser Arbeit behandelt werden. Die anfängliche Planung der Forschungsziele ist in Tab. 1.1 einsehbar. Die Benennung der Forschungsaufgaben folgt einem einheitlichen Schema und ergeben sich aus dem Namen des jeweiligen Forschungsziels, erweitert durch einen Buchstaben als Anhängsel, z.B. „O“ für *Observation*/Beobachtung, „TB“ für *Theory Building*/Theoriebildung, „I“ für *Implementation*/Implementierung und „E“ für *Experiment*.

Tabelle 1.1: Aufschlüsselung der Forschungsfragen auf Forschungsziele

<b>Forschungsfragen</b>		
<b>FZ</b>	<b>Beschreibung</b>	<b>PB</b>
<b>Erklärbarkeit von MMIR mittels generativer KI</b>		
FZ 1.1/O	Recherche zur Erklärbarkeit von MMIR mittels generativer KI	PB1
FZ 1.2/TB	Modellierung zur Erklärbarkeit von MMIR mittels generativer KI	PB1
FZ 1.3/I	Implementierung zur Erklärbarkeit von MMIR mittels generativer KI	PB1
FZ 1.4/E	Experiment zur Erklärbarkeit von MMIR mittels generativer KI	PB1
<b>Integration generativer KI in das GMAF</b>		
FZ 2.1/O	Recherche zur Integration generativer KI in das GMAF	PB2
FZ 2.2/TB	Modellierung zur Integration generativer KI in das GMAF	PB2
FZ 2.3/I	Implementierung zur Integration generativer KI in das GMAF	PB2
FZ 2.4/E	Experiment zur Integration generativer KI in das GMAF	PB2

Um sicherzustellen, dass alle Forschungsziele durch den Inhalt dieser Arbeit abgedeckt werden, entspricht die Struktur dieses Dokuments direkt dem Ansatz nach Nunamaker. Der Ansatz und Aufbau der Arbeit wird im nächsten Abschnitt behandelt.

## 1.5 Ansatz und Aufbau der Arbeit

Die Struktur dieser Arbeit ergibt sich 1:1 aus der Methodik von Nunamaker [25] und folgt seiner Modellierung. Ein Überblick auf die Struktur dieser Arbeit und Nunamakers Ansatz ist in Abb. 1.8. Die entsprechenden Forschungsziele werden, je nach Phase der Modellierung nach Nunamaker, umgeordnet und in die jeweiligen Kapitel dieser Arbeit umgruppiert.

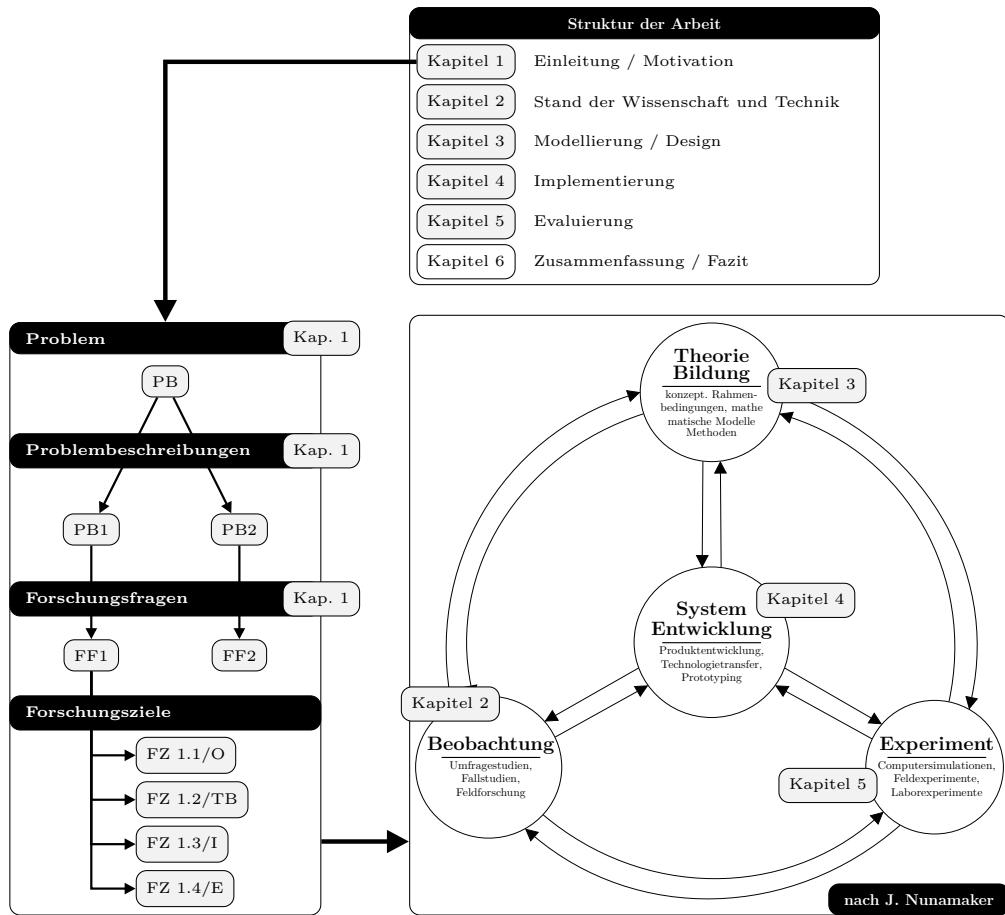


Abb. 1.8: Aufbau der Arbeit und Ansatz der Problemlösung nach Nunamaker [25].

Im Nachfolgenden werden die Kapitel jeweils einzelnen kurz beschrieben. Es wird genannt, welche Forschungsziele von welchem Typ abgearbeitet werden, und welche Aufgaben das entsprechende Kapitel jeweils übernimmt.

- „**Kapitel 2 - Stand der Wissenschaft und Technik**“ deckt alle Forschungsziele vom Typ *Beobachtung* ab, und wird der Reihenfolge nach den aktuellen Stand der Wissenschaft und Technik zusammenfassen, einführen und darstellen.
- „**Kapitel 3 - Modellierung**“ deckt alle Forschungsziele vom Typ *Theoriebildung* ab und beschreibt die Modellierung und das Design von Konzepten und Algorithmen zu vorgeschlagenen Problemlösungen.
- „**Kapitel 4 - Implementierung**“ beschreibt die Implementierung von Modellen und Konzepten und deckt alle Forschungsziele vom Typ *Systementwicklung bzw. Implementierung* ab.

- „Kapitel 5 - Experiment“ deckt alle Forschungsziele vom Typ *Experiment* ab und gibt eine detaillierte Beschreibung aller Ergebnisse.

Um in dieser Arbeit einen wissenschaftlich formalen und korrekten Prozess zur Problemlösung anzuwenden, wurde die Modellierung nach Nunamaker eingeführt. Dieser Modellierung folgend, wird eine klare Beziehung zwischen den Forschungszielen und den Forschungsfragen etabliert. Dadurch wird sichergestellt, dass jedes Ergebnis, das in dieser Arbeit dargestellt wird, in direktem Bezug zu einer entsprechenden Forschungsfrage und einer Problembeschreibung steht und umgekehrt. Des Weiteren folgt das Kapitel 3 „Modellierung“ dieser Arbeit dem Ansatz *User Centered System Design* nach Draper [26], welches den Benutzer in den Mittelpunkt sämtlicher Modellierung und Planung stellt. Die Modellierung und Planung wird auf der *Unified Modeling Language (UML)* [27] aufbauen.

## 1.6 Arbeits- und Zeitplan

Der Arbeitsplan ist dem Ansatz der Arbeit nach strukturiert, sodass sich ebenfalls eine 1:1 Korrespondenz zwischen dem Ansatz der Arbeit, der Gliederung der Arbeit und dem Arbeitsplan ergibt.

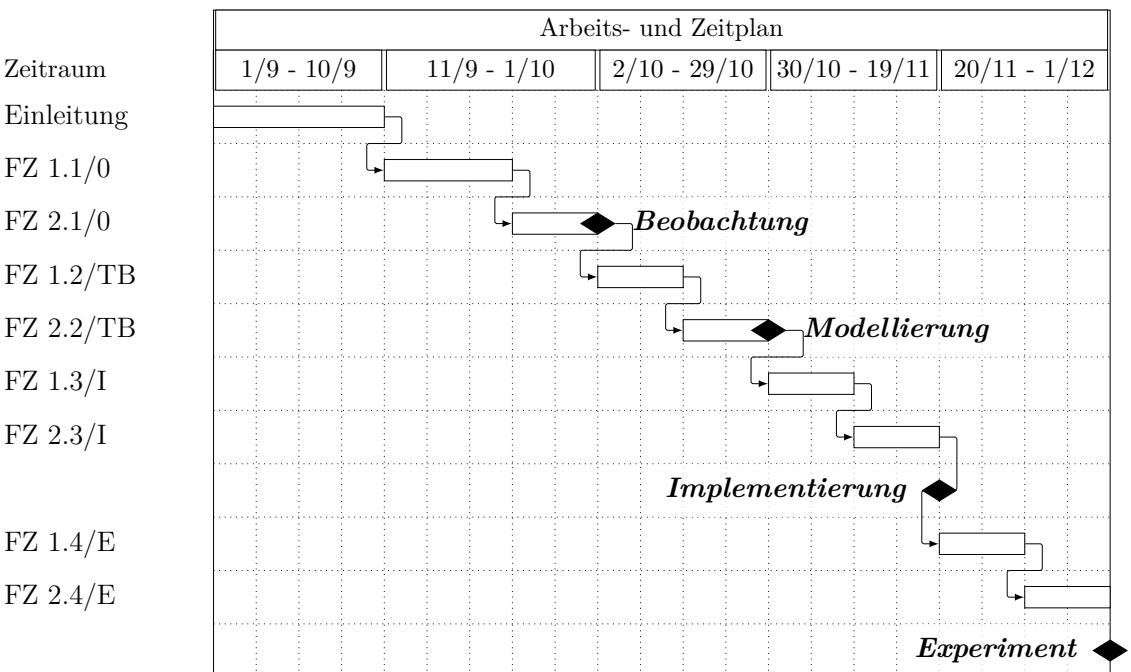


Abb. 1.9: Arbeits- und Zeitplan.

## 1.7 Zusammenfassung

Das erste Kapitel dieser Arbeit hat allgemein in die Motivation des Themas eingeführt. Eine kurze Beschreibung der aktuellen Entwicklung im Bereich Multimedia hat gezeigt, dass *MMIR* eine zunehmend wichtige Rolle spielt. In diesem Zuge spielt auch die Erklärbarkeit von *MMIR*-Prozessen eine immer wichtigere Rolle. Zwei Problembereiche konnten identifiziert werden, die mit dieser Arbeit adressiert werden. Im Bereich der Erklärbarkeit bietet das *GMAF* bereits umfangreiche Konzepte zur Erklärbarkeit von *MMIR*-Prozessen. Diese Konzepte sind aber sehr statisch und in ihrer Anwendung stark eingeschränkt. Hier gibt es allgemein großes Potenzial für Verbesserungen, die sich durch das Anbinden generativer KI versprochen wird. Im Bereich der Integration konnte festgestellt werden, dass das *GMAF* noch keine Anbindung für generative KI-Systeme bietet und hierfür eine geeignete Modellierung entwickelt werden muss.

Das Definieren von Problembeschreibungen, Forschungsfragen und Forschungszielen dient als formelle Grundlage für die Strukturierung dieser Arbeit. Basierend auf der Modellierung nach Nunamaker folgt diese Arbeit einem wissenschaftlich anerkannten Ansatz bei der Darstellung der Forschungsergebnisse. Um ein starkes Fundament für die Arbeit zu legen, wird im nächsten Kapitel ein detaillierter Überblick zum aktuellen Stand der Wissenschaft und Technik gegeben, in welchem die Ergebnisse der Forschungsziele vom Typ Beobachtung behandelt werden.

## 2 Stand der Wissenschaft und Technik

Die voranschreitende Digitalisierung, die Entwicklung des Internets, Mobiltelefone, sowie die sozialen Medien und die damit einhergehenden enormen Datenmengen haben zu einer enormen Zunahme an Wissen, Standards und auch Forschung im Bereich des *MMIR* geführt [18]. Im Bereich der Forschung spielt auch die Erklärbarkeit von *MMIR*-Prozessen eine zunehmend wichtige Rolle. In einschlägiger Literatur, wie z.B [18] wurde die Erklärbarkeit von *MMIR*-Prozessen bereits adressiert und es wurden Lösungsvorschläge und Konzepte entwickelt.

In diesem Kapitel wird, den Forschungszielen der Beobachtungsphase entsprechend, der aktuelle Stand der Wissenschaft und Technik bezogen auf ausgewählte Bereiche der Erklärbarkeit von Prozessen des *MMIR* beschrieben. Die in diesem Kapitel dargestellten Informationen wurden im Rahmen einer Literaturrecherche, speziell Publikationen, bezogen auf das GMAF und seine Technologien, sowie durch eine extensive Internetrecherche mit entsprechenden Stichworten zusammengetragen. Das Bezugssystem für diese Arbeit stellt das *GMAF* dar. Die für diese Arbeit relevanten Konzepte werden detailliert beschrieben, um so eine konkrete Vorlage, ein starkes Fundament, sowie eine scharfe Trennung zu den Phasen Modellierung, Implementierung und Experiment zu schaffen, die im weiteren Verlauf dieser Arbeit in den nachfolgenden Kapiteln Modellierung, Implementierung und Evaluierung folgen werden. Somit enthält dieses Kapitel alle nötigen Rechercheergebnisse, Beobachtungsergebnisse und damit Erkenntnisse, die der Beantwortung der Forschungsfragen laut Ansatz dienlich sein sollen. Der Methodik nach Nunamaker [25] folgend, behandelt dieses Kapitel alle Forschungsziele vom Typ Beobachtung, entsprechend der identifizierten Problembeschreibungen aus dem Abschnitt 1.2. In Abschnitt 2.1 wird auf die Erklärbarkeit von MMIR mittels generativer KI eingegangen, welches vorhandene Technologien im Bereich MMIR, sowie generativer KI beinhaltet. In Abschnitt 2.2 wird auf die Integration von generativer KI in das GMAF eingegangen, welches Technologien und technische Aspekte zur Integration beinhaltet. In Abschnitt 2.3 werden Methodiken zur Evaluierung vorgestellt, die später in Kapitel 5 im Rahmen von Experimenten Verwendung finden könnten. Abschließend wird in Abschnitt 2.4 eine Zusammenfassung der in diesem Kapitel erfassten Erkenntnisse und Ergebnisse gegeben. Dieser Strukturierung folgend, wird im nächsten Abschnitt die erste Problembeschreibung adressiert. Die Struktur dieses Kapitels ist in Tab. 2.1 abgebildet.

Tabelle 2.1: Struktur dieses Kapitels.

FZ	Kurze Beschreibung	Abschnitt
FZ 1.1/O	Recherche: Erklärbarkeit & generativer KI	Abschnitt 2.1

FZ	Kurze Beschreibung	Abschnitt
FZ 2.1/O	Recherche: Integration generativer KI	Abschnitt 2.2
Evaluierungsmeth. Zusammenfassung	Methodiken zur Evaluierung von Software Ergebnisse / Erkenntnisse	Abschnitt 2.3 Abschnitt 2.4

## 2.1 FZ 1.1/O Erklärbarkeit von MMIR mittels generativer KI

Dieser Abschnitt gibt einen Überblick über ausgewählte Ergebnisse und Erkenntnisse aus der Forschung im Bereich der Erklärbarkeit von MMIR und deckt verschiedene Aspekte ab, die sich auf die erste Problembeschreibung beziehen. Die Struktur dieses Abschnittes folgt einer logischen Abfolge. Zuerst werden bereits vorhandene Technologien im Bereich der Erklärbarkeit von MMIR vorgestellt und erläutert. Dies trifft besonders auf Technologien in Bezug auf das GMAF zu, da das GMAF das Bezugssystem dieser Arbeit darstellt. Im weiteren Verlauf wird dann auf generative KI eingegangen, Grundlagen werden erläutert, eine Auswahl an aktuellen Systemen generativer KI vorgestellt, sowie die Einschränkungen von generativer KI hervorgehoben.

### 2.1.1 GMAF

Das Generic Multimedia Analysis Framework (GMAF) [28] ist ein Framework, welches ausgewählte Technologien zum Verarbeiten verschiedener multimedialer Inhalte, wie Bilder, Videos oder Text nutzt und miteinander kombiniert [29, 30]. Beispiele für solche Technologien sind Verarbeitungssysteme zur Merkmalserkennung in Bildern (z.B zur Objekterkennung oder Erkennung von räumlichen Beziehungen mittels maschinellen Lernens). Das GMAF bietet eine flexible und erweiterbare Architektur, mit welcher auf einfache Art und Weise neue Algorithmen oder Schnittstellen zur Merkmalserkennung im Form von Plugins in das GMAF integriert werden können [30]. Die von den Plugins extrahierten Merkmalen werden vom GMAF in einer zentralen Datenstrukturen zusammengefasst und vereint [29, 22]. Abb. 2.1 zeigt eine vereinfachte Übersicht der Architektur des GMAFs und die grundlegenden Komponenten. Die Plugins im GMAF können als ein Zusammenschluss aus Anbietern von semantischen Informationen betrachtet werden [29]. Dieser Zusammenschluss produziert alle Daten, die notwendig sind, um ein Verzeichnis bzw. einen Index aus den semantischen Informationen zu generieren [29]. Allgemein kann jedes durch ein Plugin identifizierte Merkmal einen Indizierungsbegriff darstellen und somit Teil des Indexes sein [31]. Da für gewöhnlich diese Begriffe strukturelle und semantische Be-

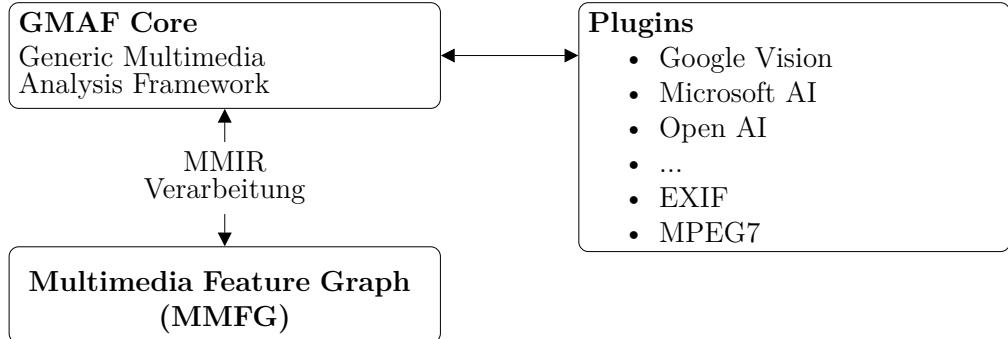


Abb. 2.1: Vereinfachte Übersicht des GMAFs (ohne UI / API) nach [22].

ziehungen zueinander aufweisen, bietet sich zur Darstellung dieser Merkmale und ihren Beziehungen zueinander eine auf Graphen basierende Struktur an [31]. Diese Struktur wird durch den Multimedia Feature Graphen (MMFG) realisiert, welcher die in ihm gespeicherten semantischen Informationen in einem einheitlichen Modell repräsentiert [31].

Der nächste Abschnitt beschreibt die Datenstruktur der MMFGs, welche im Rahmen des GMAF eine zentrale Rolle einnehmen. Zuerst wird die Grundstruktur der MMFGs beschrieben. Im weiteren Verlauf wird dann in Abschnitt 2.1.2.1 die Erklärbarkeit von MMFGs thematisiert. Darauffolgend wird in Abschnitt 2.1.2.2 die Verarbeitung von MMFGs angesprochen, welcher den Übergang zu einer weiteren wichtigen Technologie, den Graph Codes, schafft.

### 2.1.2 MMFG

Der Multimedia Feature Graph (MMFG) ist ein Resultat des GMAFs und stellt für Prozesse innerhalb des GMAFs eine zentrale gemeinsame Ressource dar [29]. Das grundlegende Konzept der MMFGs ist dabei eng verknüpft mit der Graphentheorie [29]. Genauer lässt sich ein MMFG als ein gewichteter und gerichteter Graph mit attribuierten Knoten und Kanten verstehen, der die aus Plugins identifizierten und extrahierten Merkmale multimedialer Inhalte in einer Graphenstruktur repräsentiert bzw. abbildet [29]. Des Weiteren ist ein MMFG in der Lage Verbindungen zu Teilen anderer MMFGs aufzubauen, wodurch eine rekursive Struktur aus MMFGs entsteht, die Schleifen und Teilgraphen besitzen kann [29]. Ein MMFG kann wie folgt definiert werden:

$$MMFG = (V, E)$$

Die Menge  $V$  repräsentiert die Knoten in einem MMFG und stellt die identifizierten Merkmale dar. Die Menge  $E$  stellt wiederum die Verbindungen bzw. Beziehungen zwischen den Knoten bzw. Merkmalen dar. Beide Mengen können anhand ihrer Attributierung in Unterklassen aufgeteilt werden. Die Knotenmenge  $V$  besteht aus fol-

genden Unterklassen:  $V \in \{n, r, e, w, cn, t, m, s, l, p\}$ . Die folgende Aufzählung wurde aus [29] entnommen und erläutert die Bedeutung dieser Unterklassen:

- **Nodes**  $n_1, \dots, n_n$  stellt wichtige Objekte, Aktivitäten oder Regionen in einer Datei dar.  $r$  ist die Wurzel des MMFGs und selbst wiederum ein Knoten.  $e$  bezeichnet externe Knoten, die externe Informationen darstellen. Externe Informationen können aus anderen MMFGs oder anderen externen Systemen, wie dem Semantic Web stammen und so in den MMFG eingebunden werden.
- **Weight**  $w_1, \dots, w_n$  stellt die Gewichtung bzw. Relevanz eines Knoten in einem besonderen Kontext dar. So können Objekte, abhängig vom Kontext und Gewichtung, unterschiedliche Bedeutungen in Bezug auf die gesamte Datei besitzen.
- **Childnodes**  $cn_1, \dots, cn_n$  sind Objekte, die in einer untergeordneten hierarchischen Beziehung zu anderen Objekten stehen. Diese Objekte entstehen durch wiederholte bzw. rekursive Anwendung des GMAFs auf bereits identifizierte Merkmale. Ein Beispiel für diese rekursive Anwendung ist  $Person \rightarrow Body \rightarrow Arm \rightarrow Watch$  und kann in Abb. 2.3 eingesehen werden.
- **TechnicalAttributes**  $t_1, \dots, t_n$  stellen nicht semantische Informationen eines Knoten dar. Beispiel für ein technisches Attribut kann die Abgrenzung eines Merkmals in einem Bild sein. Diese Abgrenzung wird in Form eines Rechtecks notiert ( $x: 468, y: 176, width: 28, height: 12$ ).
- **GeneralMetadata**  $m$  stellt die Metadaten der Multimedia-Datei dar. Mit Metadaten bezeichnet man Daten, die andere Informationen beschreiben und können durch EXIF oder MPEG7 identifiziert werden. Die Metadaten eines Buches könnten so z.B. der Name des Autors, Auflage oder das Erscheinungsjahr sein.
- **Location**  $l_1, \dots, l_n$  stellt Speicherorte dar. Dies kann der Speicherort der ursprünglichen Multimedia-Datei, aber auch von Kopien, unterschiedlichen Auflösungen oder Teilen der Datei sein.
- **SynonymInformation**  $s_1, \dots, s_n$  verweist auf Synonyme eines Begriffes.
- **Security and Privacy**  $p$  stellt die Sicherheitseigenschaften einer Datei dar. Beispiel für so eine Sicherheitseigenschaft ist eine Zugriffskontrollliste (engl. access control list). Auf diese Weise kann der Umgang mit vertraulichen Daten sichergestellt werden.

Auch die Kanten in  $E$  besitzen Unterklassen und können in folgende Unterklassen unterteilt werden:  $E \in \{cr, sr, sn, mm, al\}$ . Die folgende Aufzählung wurde aus [29] entnommen und erläutert, analog zu den Knoten, die Bedeutung der Unterklassen der Kanten:

- **Composition Relationship**  $cr_1, \dots, cr_n$  drückt das Verhältnis in dem zwei Objekte miteinander stehen aus. Beispiele für solche Beziehungen können „verbunden mit (engl. attached to)“, „ein Teil von (engl. part of)“ oder „bezogen auf (engl. related to)“ sein. Diese Beziehungen werden vom GMFA durch rekursives Berechnen der Abgrenzungen von Merkmalen und deren Distanzen zueinander berechnet.
  - **Sematic Relationship**  $sr_1, \dots, sr_n$  verbindet einen Knoten mit einem externen Knoten aus einem externen System, wie z.B. dem Semantic Web.
  - **SynonymLinks**  $sn_1, \dots, sn_n$  verweist auf Informationen zu Synonymen.
  - **MMFGLinks**  $mm_1, \dots, mm_n$  stellt Verbindungen zu anderen MMFGs dar.
  - **AssetLinks**  $al_1, \dots, al_n$  zeigt auf Knoten, die Speicherorte darstellen.

Abb. 2.2 zeigt die Visualisierung eines MMFGs durch yEd. Dieser MMFG besitzt 185 Knoten und 222 Kanten. Aufgrund des steigenden Level Of Detail (LOD) vieler multimedialer Inhalte, wie z.B. durch höhere Auflösung von Bildern, steigt die Anzahl der Knoten und Kanten in MMFGs drastisch [29]. In der Praxis können MMFGs tausende Knoten und Kanten aufweisen [22]. In diesen Maßstäben ist eine Visualisierung eines MMFGs auf sinnvolle Weise kaum noch möglich. Das Hauptaugenmerk der MMFGs liegt nicht darauf für Menschen lesbar, darstellbar oder verständlich zu sein, sondern anhand ihrer Graphenstruktur geeignete Verarbeitung der Merkmale zu ermöglichen [29].

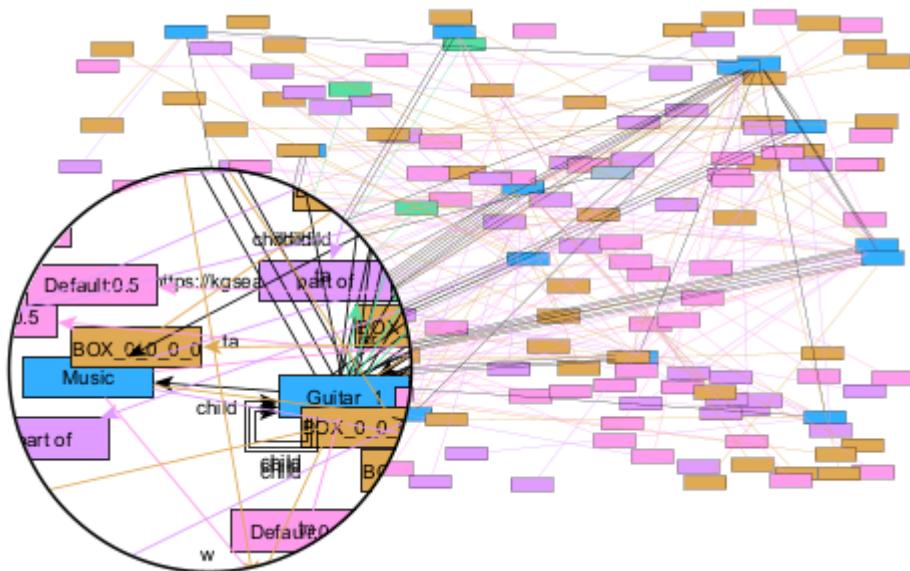


Abb. 2.2: Visualisierung eines MMFGs durch yEd.

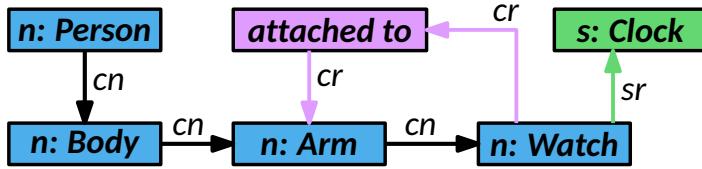


Abb. 2.3: Beispiel für einen  $MMFG_{ex}$ .

Abb. 2.3 zeigt einen simplen  $MMFG_{ex}$ . Die Knoten *Person*, *Body*, *Arm*, *Watch* repräsentieren identifizierte Merkmale, *Clock* ist ein Synonym für *Watch* und *attached to* verweist auf eine Beziehung zwischen *Arm* und *Watch*.

### **Erklärbarkeit von MMFGs**

Um Erklärbarkeit von MMFGs zu erreichen, muss die technische Darstellung der Merkmale in eine für Menschen verständliche Bedeutung überführt werden. Das Überbrücken dieser Lücke ist nicht in einem einzigen Schritt machbar. Es benötigt mehrere Schritte um diese Lücke zu schließen. In einem ersten Zwischenschritt muss die Lücke zwischen der technischen Darstellung der Merkmale und ihrer semantischen Darstellung überbrückt werden. Hierzu wird ein formaler und für Maschinen lesbarer Ansatz zur Erweiterung von MMFGs benötigt [32]. In verwandten Arbeiten, wie [32], wurden umfangreiche Konzepte entwickelt und beschrieben, die in der Lage sind durch Erweiterungen von MMFGs Erklärbarkeit zu erreichen. Das Überbrücken der Lücke zwischen der technischen Darstellung der Merkmale und ihrer semantischen Darstellung gelingt mit einer semantischen Erweiterung. Eine semantische Erweiterung von MMFGs kann durch das Anbinden eines externen Informationssystems, wie z.B. dem Semantic Web [33], erreicht werden. Durch das Definieren und Anwenden einer formalen kontextfreien Phrasenstruktur-Grammatik [34] in Kombination mit einer semantischen Erweiterung sind die Darstellungen von MMFGs nicht nur für Menschen lesbar, sondern es können auch für Menschen verständliche, sowie lesbare und auf natürlicher Sprache basierende Ausdrücke generiert werden. Diese Erweiterungen ermöglichen formale semantische Darstellungen und schaffen die Grundlage für auf natürlicher Sprache basierende Erklärungen. Diese Erweiterungen von MMFGs zu formalen semantischen Darstellungen führen zu Semantischen Multimedia Feature Graphen (SMMFGs) [32], sowie zu Erklärbaren Semantischen Multimedia Feature Graphen (ESMMFGs) [32].

Jeder auf natürlicher Sprache basierende Ausdruck, der mit solch einer Grammatik generiert wird, ist inhaltlich *korrekt*, da es nur die ursprünglichen identifizierten Merkmale in einer formalen und für Menschen lesbaren Form darstellt [32]. Auf diese Weise kann jedes Element eines MMFG bzw. identifizierte Merkmal struktu-

rell und semantisch dargestellt werden [32]. Des Weiteren bieten sich anhand der Grammatik und ihrer Produktionsregeln unbegrenzte Möglichkeiten zur Konstruktion natürlicher Ausdrücke [32]. Durch Anpassungen der Produktionsregeln können Applikationen individuell die Konstruktion natürlicher Ausdrücke bestimmen [32]. Anhand dieser Techniken können bereits gute Ergebnisse bzw. Erklärungen erzielt werden. Besonders textbasierte Dokumente können mit dieser Herangehensweise erklärt werden, da die Struktur des MMFGs der Struktur des Dokuments folgt [32]. Dies gilt allerdings nicht für andere Multimedia-Dateien, wie z.B. Bildern. Für diese Dateitypen bleibt die Reihenfolge der konstruierten Sätze zufällig [32].

Wichtig anzumerken ist, dass diese Konzepte ohne maschinelles Lernen auskommen und einen rein mathematischen und statistischen Ansatz verfolgen, um Erklärungen anhand identifizierter Merkmale zu generieren [32].

### ***Verarbeitung von MMFGs***

Der steigende Detaillierungsgrad (LOD) vieler multimedialer Inhalte führt zu einem Anstieg in der Anzahl an Elementen in einem MMFG [30]. So können in der Praxis MMFGs tausende Knoten und Kanten enthalten [29]. Ein höherer Detaillierungsgrad und eine höhere Anzahl an Elementen in einem MMFG stärkt zwar die Wirksamkeit der Ergebnisse von Berechnungen eines MMFGs, wie z.B der Berechnung der Ähnlichkeit zweier MMFGs, sorgt aber allgemein auch für eine komplexere Graphenstruktur [31]. Diese Eigenschaft von MMFGs birgt eine Reihe von Problemen. Zum einen ist die Umsetzung aktueller Algorithmen, wie z.B. zur Bestimmung der Ähnlichkeit von MMFGs, auf rechenschwachen Systemen, wie z.B Smartphones, mit dieser Anzahl an Knoten und Kanten aufgrund Speicherplatz und Laufzeitbegrenzungen nicht einfach zu handhaben und zum anderen sind aktuelle Algorithmen, die auf dem Durchlaufen eines Graphen basieren, wie z.B die Bestimmung der Ähnlichkeit zweier MMFGs, bei solch einer Anzahl an Knoten und Kanten nicht mehr effizient und zudem sehr rechenaufwendig [31].

Mathematisch betrachtet können Graphen durch ihre Adjanzenzmatrix dargestellt werden [31]. Basierend auf einer Adjanzenzmatrix können weitere Umformungen und Berechnungen durchgeführt werden [31]. Ein Beispiel hierfür ist die Umwandlung eines Graphen in einen n-dimensionalen Vektorraum. Auf diesem Vektorraum können dann wiederum weitere Berechnungen durchgeführt werden. Speziell für gewichtete Graphen, zu welchen auch per Definition MMFGs zählen, werden in der Regel Matrizen angewandt [31]. Diese Matrizen erweitern eine Adjanzenzmatrix um die Gewichtung der Kanten. Im weiteren Verlauf dieser Arbeit werden diese Matrizen auch Wertungsmatrizen (engl. valuation matrices) genannt. Basierend auf diesen Darstellungen können mathematische Konzepte angewandt werden. Ein Beispiel für so ein mathematisches Konzept ist die Anwendung der Eigenwert-Methode, durch

welche eine auf Merkmale bezogene Ähnlichkeitsberechnung von Adjazenzmatrizen durchgeführt werden kann [31]. Die Eigenwert-Methode reduziert eine quadratische Matrix  $M$  auf eine einfache rationale Zahl  $\lambda$  unter der Annahme  $\exists \vec{v} : M \cdot \vec{v} = \lambda \cdot \vec{v}$  [31]. Auf diese Weise können Ähnlichkeitsberechnungen auf  $\lambda$  anstatt der gesamten Matrix  $M$  durchgeführt werden [31]. Allerdings weisen viele dieser Algorithmen eine polynomiale Laufzeit von  $\mathcal{O}((n + e)^2)$  oder gar exponentielle Laufzeit auf (hierbei verweist  $n$  auf die Anzahl der Knoten und  $e$  auf die Anzahl der Kanten) [31]. Um effektive und effiziente Verarbeitungen, wie z.B. Abfragen, für große Datenstrukturen und Sammlungen multimedialer Inhalte, wie einem MMFG, zu ermöglichen, müssen effizientere Algorithmen, wie zur Bestimmung der Ähnlichkeit von MMFGs, die mit komplexeren Graphenstrukturen umgehen können, eingesetzt werden [31].

Im nächsten Abschnitt wird eine Projektion eines Graphen in einen zweidimensionalen Vektorraum, sowie entsprechende Algorithmen vorgestellt. Diese Projektion stellt eine Erweiterung der Wertungsmatrix eines Graphen dar und ermöglicht statt einer exponentiellen Komplexität der Laufzeit eine lineare Laufzeitkomplexität von  $\mathcal{O}(n + e)$  und unterstützt zudem ein höheren Detaillierungsgrad von MMFGs [31]. Das Konzept dieser Projektion wird im weiteren Verlauf der Arbeit als Graph Code [35] bezeichnet.

### 2.1.3 Graph Codes

In diesem Abschnitt werden die mathematischen und algorithmischen Konzepte der Graph Codes beschrieben. Zuerst werden die Grundlagen von Graph Codes beschrieben. Diese Grundlagen umfassen des Weiteren die in Abschnitt 2.1.3.1 beschriebene Codierung von Graph Codes, sowie die in Abschnitt 2.1.3.2 beschriebenen technischen Eigenschaften, wie Vokabular und Wörterbücher von Graph Codes. In Abschnitt 2.1.3.3 wird dann die Ähnlichkeit von Graph Codes behandelt.

Graph Codes sind eine Projektion eines MMFG in einen zweidimensionalen Raum [35]. Dabei basieren Graph Codes auf Operationen auf der Adjazenzmatrix eines Graphen und können allgemein, per Prinzip, auf jede Art von Graphen angewendet werden. Allerdings wurden Graph Codes spezifisch für MMIR entwickelt. Graph Codes werden eingesetzt, um MMFGs zu transformieren und in einer Form darzustellen, die schnelleres und effizienteres MMIR ermöglicht [31]. Grundlage dieser Umformung ist ein Algorithmus zur Codierung eines MMFG. Dieser Algorithmus nimmt die Wertungsmatrix  $WM$  eines MMFGs als Eingabe [35]. Um die grundlegenden Konzepte von Graph Codes vorzustellen, wird das Beispiel  $MMFG_{ex}$  aus Abb. 2.3 als Basis benutzt. Eine Wertungsmatrix enthält jeweils eine Zeile und Spalte für einen Knoten in einem MMFG. Aufgrund dessen ist eine Wertungsmatrix auch eine quadratische Matrix. Eine gerichtete Kante zwischen zwei Knoten  $n_1$  und  $n_2$  wird in der Wertungsmatrix durch ihre Gewichtung oder durch den Wert 1 im Eintrag  $(n_1, n_2)$  der Wertungsmatrix  $WM$  dargestellt.

Die Knoten *Person*, *Body*, *Arm*, *Watch*, *Clock*, *attached to* des Beispiels  $MMFG_{ex}$  werden in der Diagonalen der Wertungsmatrix ebenfalls durch den Wert 1 dargestellt. Dies führt zu der in Abb. 2.4 dargestellten Wertungsmatrix.

	Person	Body	Arm	Watch	Clock	attached to	b)
Person	1	1	0	0	0	0	$WM(MMFG_{ex}) =$
Body	0	1	1	0	0	0	$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$
Arm	0	0	1	1	0	0	
Watch	0	0	0	1	1	1	
Clock	0	0	0	0	1	0	
attached to	0	0	1	0	0	1	

Abb. 2.4: Wertungsmatrix des Beispiels  $MMFG_{ex}$  in a) Tabellenform und b) Form einer Matrix.

### Codierung von Graph Codes

Graph Codes wenden eine Funktion  $f_{enc}$  zur Codierung von Informationen in einem MMFG an. Diese Funktion codiert die Informationen zu den Typen der Knoten oder Kanten und ihren entsprechenden Attributen. Hierzu werden für alle Einträge bzw. Felder in der Wertungsmatrix  $WM$ , die nicht null sind, numerische Werte berechnet. Die Funktion  $f_{enc}$  zur Codierung dieser Informationen kann von Applikationen je nach ihren jeweiligen Anforderungen angepasst werden [35]. Eine beispielhafte Codierung kann wie folgt lauten: Knoten = 1, Blatt = 2, Kindsbeziehung = 3, Synonymbeziehung = 4, Beziehung = 5, Knoten (räumliche Beziehung) = 6, Knoten (Synonym) = 7. Angewandt auf das Beispiel  $MMFG_{ex}$  führt diese Codierung zu folgender in Abb. 2.5 a-b) dargestellten codierten Wertungsmatrix  $WM_{enc}$  bzw. Graph Code  $GC_{ex}$ . Die Farben der Einträge in Abb. 2.5 entsprechen der Färbung der Knoten aus Abb. 2.3. Die dazugehörigen Beziehungen sind ebenfalls entsprechend gefärbt. Abb. 2.5 c) hingegen zeigt den durch das GMAF generierten Graph Code  $GC_{ex}$  des Beispiels  $MMFG_{ex}$ . Es sei angemerkt, dass die Unterschiede zwischen dem Graph Code  $GC_{ex}$  aus Abb. 2.5 a-b) und Abb. 2.5 c) durch die vom GMAF eigens definierte Funktion  $f_{enc}$  zu Stande kommen. Des Weiteren inkludiert das GMAF die Synonym-Knoten *Clock* und *attached to* nicht in seiner Abbildung. Die Beziehung *attached to* wird im GMAF durch den Eintrag (*Arm*, *Watch*) im Graph Code

$GC_{ex}$  dargestellt. Hier wird eine feinere Differenzierung der räumlichen Beziehung vorgenommen und durch den Wert 17 dargestellt. Durch komplexere Funktionen zur Codierung können Attribute, Gewichtung oder auch andere Informationen in einem MMFG codiert und durch beliebige Zahlen dargestellt werden, welche im weiteren Verlauf einer Verarbeitung gezielt genutzt werden können.

	Person	Body	Arm	Watch	Clock	attached to	
Person	1	3	0	0	0	0	$GC_{ex} =$
Body	0	1	3	0	0	0	$WM_{enc}(MMFG_{ex}) =$
Arm	0	0	1	3	0	0	$\begin{pmatrix} 1 & 3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 4 & 5 \\ 0 & 0 & 0 & 0 & 7 & 0 \\ 0 & 0 & 5 & 0 & 0 & 6 \end{pmatrix}$
Watch	0	0	0	2	4	5	
Clock	0	0	0	0	7	0	
attached to	0	0	5	0	0	6	

Abb. 2.5: Codierte Wertungsmatrix bzw. Graph Code  $GC_{ex}$  zum entsprechenden Beispiel  $MMFG_{ex}$  in a) Tabellenform und b) in Form einer Matrix. c) Vom GMAF generierte Graph Code  $GC_{ex}$ .

Bei der Berechnung der Ähnlichkeit von Graph Codes müssen diese miteinander verglichen werden. Damit Graph Codes verglichen werden können, muss der Vergleich dieser Graph Codes auf einer genaustens definierten Metrik für Ähnlichkeit beruhen. Diese Metrik muss die mathematischen Aspekte eines Vergleichs von Matrizen und auch die semantischen Aspekte von Graph Codes berücksichtigen [35]. Im nächsten Abschnitt 2.1.3.2 werden weitere grundlegende Definitionen eingeführt und definiert, die zur Definition einer Metrik zur Ähnlichkeit von Graph Codes notwendig sind und eine Berechnung der Ähnlichkeit von Graph Codes unabhängig von den Typen der Knoten und Kanten ermöglicht.

### Vokabular und Wörterbuch eines Graph Codes

Die Menge der Knoten  $V$  in einem MMFG, welche die identifizierten Merkmale darstellen, können als eindeutige Bezeichner für das Vokabular bzw. die Gesamtheit aller Wörter, welche einen MMFG ausmachen, betrachtet werden. Die Menge dieser Bezeichner wird mit  $FVT_{MMFG} = \{fvt_1, \dots, fvt_n\}$  notiert und stellt somit die Elemente des „Wörterbuchs“ eines Graph Codes dar. Damit ein Bezeichner  $fvt_i$  eines Eintrages in einem Graph Code eindeutig identifiziert werden kann, wird ein Vektor

$dict_{GC}$  definiert. Dieser Vektor bietet eine geordnete Reihenfolge für die Menge an Bezeichner  $FVT_{MMFG}$  und weist jedem Bezeichner eine eindeutig definierte Position zu. Die Ordnung dieser Bezeichner im Vektor  $dict_{GC}$  kann durch eine beliebige Strategie, wie z.B. durch Breitensuche, Tiefensuche oder manuell, erfolgen. Durch den Vektor  $dict_{GC}$  kann jedes Feld in der Diagonalen des entsprechenden Graph Codes einem Eintrag im Wörterbuch, dargestellt durch  $dict_{GC} = (fvt_1, \dots, fvt_n)$ , zugeordnet werden [35]. Angewandt auf den Graph Code  $GC_{ex}$  ergibt sich für die Menge an Bezeichnern  $FVT_{ex} = \{Person, Body, Arm, Watch, Clock, attached to\}$ , sowie der Vektor  $dict_{ex} = (Person, Body, Arm, Watch, Clock, attached to)$ . Die Reihenfolge der Bezeichner in  $FVT_{ex}$  ist unerheblich und ist für gewöhnlich zufällig.

Anders als  $FVT_{ex}$  weist  $dict_{ex}$  eine spezifische Reihenfolge auf, die in Tab. 2.2 dargestellt wird.

Index $i$	1	2	3	4	5	6
$fvt_i$	Person	Body	Arm	Watch	Clock	attached to

Tabelle 2.2:  $dict_{ex}$  vom Graph Code  $GC_{ex}$ .

Beim Vergleichen von Graph Codes ist es wichtig nur die Einträge äquivalenter Merkmale in den Diagonalen der Graph Codes miteinander zu vergleichen. Da jeder Graph Code einen individuellen Vektor  $dict_{GC}$  mit individueller Ordnung besitzt, unterscheiden sich die Wörterbücher der Graph Codes für gewöhnlich voneinander  $dict_{GC_1} \neq dict_{GC_2}$ . Äquivalente Merkmale können durch die Schnittmenge der Bezeichner bestimmt werden:  $FVT_{1 \cap 2} = \{fvt_1, \dots, fvt_n\} = V_{MMFG_1} \cap V_{MMFG_2}$ . Analog kann dieses Prinzip auch auf die Wörterbücher von Graph Codes angewendet werden:  $dict_{1 \cap 2} = dict_{GC_1} \cap dict_{GC_2}$ . Aufgrund der unterschiedlichen Ordnungen der Wörterbücher von Graph Codes können diese nicht einfach miteinander verglichen werden, da die Einträge in den zu vergleichenden Graph Codes unterschiedliche Bezeichner darstellen. Um diesem Problem zu begegnen und die Ordnung von Graph Codes anzupassen, gibt es eine Äquivalenzfunktion  $f_{equ}(M_{\cap})$ , die die Matrix der Schnittmenge der Wertematrizen bzw. das Wörterbuch so transformiert, dass das entsprechende Wörterbuch nach dem Vektor bzw. Wörterbuch  $dict_{\cup}$  geordnet ist [35].

Konträr zur Schnittmenge kann auch eine Vereinigung von Graph Codes erfolgen. Eine Vereinigung kann dann erfolgen, wenn eine Applikation eine Sammlung an MMFGs enthält. Im Falle der Vereinigung von Graph Codes werden die Menge der Bezeichner, sowie die Wörterbücher der Graph Codes miteinander vereint. Die Gesamtheit aller Bezeichner  $FVT_{Coll}$  kann als Vereinigung aller Bezeichner in der Sammlung an MMFGs definiert werden und als Vereinigung aller Wörterbücher dargestellt werden  $dict_{\cup} : FVT_{Coll} = \bigcup_{i=1}^n FVT_{MMFG_i}$  mit  $\forall i, j < n : dict_{\cup} = dict_i \times dict_j$ . Um die Vereinigung der Wörterbücher zu erstellen, wird jedes Wörter-

buch  $dict_i$  durchlaufen und, sofern noch nicht im Wörterbuch  $dict_{\cup}$  vorhanden, jeder individueller Bezeichner hinzugefügt. Wird ein  $dict_{\cup}$  aus der vollständigen Sammlung an Graph Codes berechnet, so kann  $dict_{\cup}$  als ein globales Wörterbuch betrachtet werden, in welchem jeder einzigartige Bezeichner eine einzigartige Position besitzt [35].

Die in diesem Abschnitt beschriebenen Definitionen ermöglichen eine effiziente und effektive Verarbeitung von Informationen eines Graph Codes. Ein Beispiel für so eine Verarbeitung ist die Berechnung der Ähnlichkeit zweier Graph Codes, die auf einem Vergleich dieser beiden basiert. Wie bereits erwähnt, muss der Vergleich dieser Graph Codes auf einer genaustens definierten Metrik beruhen. Im nächsten Abschnitt werden diese Metriken vorgestellt.

### **Ähnlichkeit von Graph Codes**

In diesem Abschnitt wird eine Metrik zur Berechnung der Ähnlichkeit bzw. zum Vergleich von Graph Codes beschrieben. Die in diesem Abschnitt vorgestellten Metriken wurden aus [35] entnommen und ermöglichen die Quantifizierung der Eigenschaften von Graph Codes. Diese Metrik muss die Zeilen, Spalten und die Einträge, die die Knoten und Kanten darstellen, berücksichtigen. D.h. diese Metrik zur Berechnung der Ähnlichkeit von Graph Codes muss auf Matrizen anwendbar sein, dessen Zeilen und Spalten die Knoten und Kanten eines MMFGs darstellen. Damit ein Vergleich dieser Einträge möglich ist, müssen die Einträge vom selben Typ sein. Da die Wörterbücher von Graph Codes für gewöhnlich unterschiedlich sind, ist es zudem wichtig die korrekten Einträge in einer Matrix zu vergleichen [35]. Es kann zwischen drei Typen an Einträgen differenziert werden:

- Knoten (stellt Merkmale dar)
- Kanten (Beziehungen zwischen Knoten bzw. Merkmalen)
- Kantentyp (Typ der Beziehung zwischen Knoten bzw. Merkmalen)

Für jeden Typ eines Eintrages kann jeweils eine eigene Metrik definiert werden:

- Metrik für Merkmale  $M_F$
- Metrik für Merkmalsbeziehungen  $M_{FR}$
- Metrik für Beziehungstypen  $M_{RT}$

Diese Metriken können wiederum als Tripel in einer Metrik  $M_{GC} = (M_F, M_{FR}, M_{RT})$  zusammengefasst werden und bilden die gesamte Metrik, die die Ähnlichkeit von Graph Codes darstellen soll. Diese Metriken werden im Folgenden vorgestellt und beschrieben.

Die **Metrik für Merkmale**  $M_F$  berechnet die Ähnlichkeit von Graph Codes anhand der Schnittmenge der Bezeichner in einem Wörterbuch. Die Metrik  $M_F$  ist definiert als das Verhältnis zwischen der Kardinalität der Schnittmenge von Bezeichnern  $dict_{\cap}$  der Wörterbücher  $dict_i$  und  $dict_j$  der Graph Codes  $GC_i$  und  $GC_j$  und der Kardinalität des Wörterbuchs  $dict_i$  vom Graph Code  $GC_i$ . Die Formel für diese Metrik lautet  $M_F(GC_i, GC_j) = \frac{|dict_{\cap}|}{|dict_i|}$ . Hierbei bezeichnet  $|\vec{v}|$  die Kardinalität eines Vektors, sprich die Anzahl der Elemente in einem Vektor. Anders formuliert, ist der Wert der Ähnlichkeit höher, desto mehr Merkmale zwei Graph Codes gemeinsam haben. Es ist wichtig anzumerken, dass diese Metrik nur Knoten bzw. Merkmale betrachtet, unabhängig von den Beziehungen zwischen diesen Merkmalen [35].

Die **Metrik für Merkmalsbeziehungen**  $M_{FR}$  ist die Grudlage für die Berechnung der Ähnlichkeit von Beziehungen zwischen Merkmalen. Beziehungen in einem Graph Code werden durch Einträge dargestellt, die nicht in der Diagonalen der Matrix vorkommen und deren Wert größer null ist. Dabei bezieht die Berechnung dieser Metrik nur Einträge aus der Schnittmenge  $M_{\cap}$  der Matrizen ein, die Beziehungen mit demselben Ausgangs- und Zielknoten darstellen. Die Berechnung dieser Metrik basiert auf den nicht in der Diagonalen vorkommenden Einträgen der Adjanzenzmatrix  $AM(M_{\cap})$ . Die Metrik  $M_{FR}$  ist definiert als das Verhältnis zwischen der Summe aller nicht in der Diagonalen vorkommenden Einträge und der Kardinalität aller nicht in der Diagonalen vorkommenden Einträge. Die Formel für diese Metrik lautet  $M_{FR}(GC_i, GC_j) = \frac{\sum_{i,j} AM(M_i \cap M_j) - n}{|AM(M_{\cap})| - n}$ . Die Metrik  $M_{FR}$  stellt somit das Verhältnis zwischen der Zahl der Einträge, die Beziehungen darstellen und der Anzahl der äquivalenten und in der Schnittmenge vorkommenden Einträge, die Beziehungen darstellen, von Graph Codes. Somit zählt die Metrik  $M_{FR}$  die Anzahl aller Beziehungen zwischen Ausgangs- und Zielknoten. Die Metrik berücksichtigt jedoch nicht die Äquivalenz der Typen der Beziehungen. Die Äquivalenz der Typen von Beziehungen wird von der nächsten Metrik, der Metrik für Beziehungstypen berücksichtigt [35].

Die **Metrik für Beziehungstypen**  $M_{RT}$  basiert auf der Metrik für Merkmalsbeziehungen. Die Codierungsfunktion  $f_{enc}$  codiert für verschiedene Typen von Beziehungen unterschiedliche Werte. Eine Ähnlichkeit von Typen von Beziehungen setzt aber voraus, dass die Einträge, die die Beziehungen darstellen, vom selben Typ sind, damit sie verglichen werden können. Diese Metrik führt die Berechnungen nicht mehr anhand der Adjanzenzmatrix der zu vergleichenden Graph Codes aus, sondern anhand der Schnittmenge  $M_{\cap}$  der codierten Wertungsmatrizen. Bei dieser Berechnung wird die Subtraktion auf äquivalenten Einträgen der zu vergleichenden Graph Codes durchgeführt. Ist das Ergebnis dieser Subtraktion gleich null, so ist der Typ der Beziehung äquivalent. Sind alle äquivalenten Einträge gleich, so ist die Summe der Subtraktionen dieser Einträge zueinander gleich null. Im Vergleich zur Metrik  $M_{FR}$ , die nur Berechnungen anhand einer Adjanzenzmatrix durchführt, verwendet

diese Metrik zusätzlich den Wert der Codierung, welche den Typ der Beziehung darstellt, um festzustellen, ob die verglichenen Typen äquivalent sind. Die Metrik  $M_{RT}$  ist definiert als das Verhältnis zwischen der Summe aller nicht in der Diagonalen vorkommenden Einträge und der Kardinalität dieser Einträge. Die Formel für diese Metrik lautet  $M_{RT}(GC_i, GC_j) = \frac{\sum_{i,j}^{n,i \neq j} (|M_{\cap i} - M_{\cap j}|)}{|M_{\cap i} - n|}$  [35].

### ***Erklärbarkeit von Graph Codes***

Analog zur Erklärbarkeit von MMFGs in Abschnitt 2.1.2.1, muss die technische Darstellung der Merkmale in eine für Menschen verständliche Bedeutung überführt werden. Auch hier gilt es in einem ersten Schritt die Lücke zwischen der technischen Darstellung und ihrer semantischen Darstellung zu überbrücken. Die semantische Darstellung von Merkmalen wurde in Abschnitt 2.1.2.1 bereits durch das Anbinden von externen Informationssystemen erreicht und resultierte in SMMFGs. Das Anwenden der Algorithmen von Graph Codes auf SMMFGs ermöglicht die semantische Darstellung von Merkmalen in Graph Codes und führt zu Semantischen Graph Codes (SGCs) [32]. Während das Wörterbuch eines Graph Codes die Menge der Bezeichner der Merkmale darstellt, stellt das Wörterbuch eines Semantischen Graph Codes die semantische Darstellung der Bedeutung dieser Bezeichner dar. Analog zu MMFGs können auch für Graph Codes, durch das Definieren und Anwenden einer formalen kontextfreien Phrasenstruktur-Grammatik in Kombination mit Semantischen Graph Codes, für Menschen lesbare, verständliche und auf natürlicher Sprache basierende Ausdrücke generiert werden. Die Erweiterung von Graph Codes um eine formale kontextfreie Grammatik führt zu Erklärbaren Semantischen Graph Codes (ESGCs) [30].

In Bezug auf Graph Codes ist wichtig anzumerken, dass Graph Codes nicht einen allgemeinen Ersatz bzw. eine allgemeine Alternative zu MMFGs darstellen und nur in bestimmten Fällen, wie z.B. beim Vergleich von MMFGs zur Bestimmung der Ähnlichkeit, eine Verbesserung bieten [35]. Des Weiteren werden Erklärungen von Graph Codes bislang nur anhand statistischer und rein mathematischer Vorangehensweisen erzeugt. Es existiert keine Möglichkeit oder Untersuchung Systeme generativer KI zur Erklärbarkeit von Graph Codes zu nutzen. Dies führt zur ersten offenen Herausforderung.

#### **Offene Herausforderung OH 1.1 - Erklärbarkeit durch generative KI**

Einsatz von generativer KI als Ersatz bzw. Alternative zur statistischen und rein mathematischen Vorgehensweise um Erklärungen zu generieren.

Die offene Herausforderung **OH 1.1** wird im Rahmen einer Modellierung, sowie Implementierung in jeweils FZ 1.2/TB und FZ 1.3/I angesprochen und untersucht.

#### 2.1.4 Generative KI

Generative künstliche Intelligenz (GenKI, engl. generative AI (GenAI)) ist ein Typ künstlicher Intelligenz, die es ermöglicht, anhand vorhandener Informationen und Vorgaben, eine große Vielfalt an neuen Inhalten zu erstellen. Beispiele für diese Inhalte können Bilder, Videos, Audiodateien, Texte oder anderweitige synthetische Daten sein [36]. Synthetische Daten bezeichnet Daten, die nicht auf natürliche, sondern auf künstliche Art und Weise durch eine KI erstellt wurden, um echte Daten zu erweitern, oder zu ersetzen [37]. Damit eine generative KI synthetische Daten erzeugen kann, wird das Modell der KI mit Hilfe eines Algorithmus anhand echter Daten so trainiert, dass es in der Lage ist, die ursprüngliche Struktur und die Eigenschaften, wie Muster, Zusammenhänge und statistische Merkmale, der echten Daten zu reproduzieren [38, 39]. Generative KI wird daher als generativ bezeichnet, da sie auf Grundlage bereits vorhandener Datensätze etwas Neues bzw. etwas bisher nicht Dagewesenes erzeugt. Diese Eigenschaft hebt generative KI wesentlich von diskriminativer KI ab, welche einzig und allein zwischen verschiedenen Eingaben differenzieren kann [40].

Ein Anwendungsbeispiel für eine diskriminative KI wäre eine simple binäre Klassifikation von Bildern und ob ein Bild eine Katze darstellt bzw. enthält, oder nicht. Mit genügend Daten kann das diskriminative Modell einer KI so trainiert werden, dass es bestimmte Eigenschaften, wie Formen und Texturen, die einer Katze gleichen, in einem Bild erkennen kann. Anhand erkannter Merkmale kann das Modell eine Vorhersage treffen, ob in einem Bild eine Katze zu sehen ist, oder nicht. Eine wichtige Voraussetzung für das Training dieses Modells ist es, dass jede Beobachtung in den Trainingsdaten manuell mit einer Bezeichnung versehen ist. Im Fall einer binären Klassifikation sind die Bilder, die Katzen darstellen bzw. enthalten, mit der Bezeichnung 1 versehen, alle anderen Bilder mit der Bezeichnung 0. Durch das Training erlernt das diskriminative Modell beide Gruppen zu unterscheiden.

Die diskriminative Modellierung ist daher gleichbedeutend mit dem überwachten maschinellen Lernen (engl. supervised machine learning) [41]. In der Tat wird ein Großteil der diskriminativen Modelle für überwachtes maschinelles Lernen verwendet [42]. Diskriminative KI bildet also die Eingabe eines mit Bezeichnungen versehenen Datensatzes auf eine Ausgabe ab. Im Vergleich dazu benötigt das Training eines generativen Modells für gewöhnlich keine explizit mit Bezeichnungen versehenen Datensätze und kann somit mit dem nicht überwachten maschinellen Lernen (engl. unsupervised machine learning) gleichgesetzt werden. Das Training auf einem mit Bezeichnungen versehenen Datensatz ist damit allerdings nicht ausgeschlossen und kann auch Verwendung finden [41].

Es ist wichtig anzumerken, dass generative KI, die auf der Methode des Deep Learnings aufbaut, nicht neu ist und bereits in den 60ern in Form von Chatbots vorge-

stellte wurde [36]. Die erste Instanz eines solchen Chatbots hieß „ELIZA“ [43] und wurde 1966 am MIT entwickelt und hatte zur Aufgabe Gespräche mit einem Psychotherapeuten zu simulieren [40]. Durch das Aufkommen von künstlichen neuronalen Netzwerken (engl. articial neural networks (ANN)) und entsprechender Algorithmen in den 80ern und 90ern, sowie durch bessere Datenlage und Rechenkapazitäten wurde die Methode des Deep Learning praxistauglicher. Die in 2014 von Ian Goodfellow vorgestellte Technik der generativen gegnerischen Netzwerke (engl. generative adversarial networks (GAN)) führte zu einem Durchbruch und ermöglichte es generativer KI überzeugend wirkende neuartige Inhalte zu generieren [44].

### ***Grundlegende Techniken und Konzepte***

In diesem Abschnitt wird genauer auf wichtige Techniken und Technologien eingegangen, die in aktuellen generativen KI-Systemen Verwendung finden. Die Techniken und Technologien werden kurz beschrieben und erläutert. Zuerst werden die allgemeinen Techniken, wie Machine Learning oder Deep Learning vorgestellt bzw. rekapituliert und weiter wichtige Konzepte im Bereich der generativen KI angesprochen. Die in diesem Abschnitt vorgestellten Techniken bzw. Technologien sind dabei so geordnet, dass sie einer logischen Abfolge in Bezug auf ihre Abhängigkeiten zueinander folgen.

**Machine Learning** (z.Dt. maschinelles Lernen) ist ein Teilbereich der KI [45] und hat zum Ziel die Art und Weise, wie Menschen lernen zu imitieren [13]. Maschinelles Lernen ermöglicht es Systemen, Muster und Zusammenhänge automatisch und selbstständig aus Daten und Erfahrungen zu erlernen und zu verbessern, ohne dabei explizit programmiert zu sein [46]. Damit Machine Learning funktioniert und Muster und Zusammenhänge erlernt werden können, muss ein Datensatz zum Training vorhanden sein [46]. Im Bereich der KI spielen (**un**)-**supervised machine learning** (z.Dt. (un)-überwachtes maschinelles Lernen) eine zentrale Rolle [45]. Überwachtes Lernen zeichnet sich dadurch aus, dass Algorithmen mit Hilfe von Daten trainiert werden, die im Trainingsdatensatz manuell mit Bezeichnungen versehen wurden [46]. Daher ist überwachtes Lernen mit einem relativ hohen menschlichen Aufwand verbunden [46]. Weit verbreitete Problemstellungen bzw. Aufgaben, in denen überwachtes Lernen eingesetzt werden, sind die Klassifikation und Regression [47]. Unüberwachtes Lernen zeichnet sich hingegen dadurch aus, dass es ohne die Einflussnahme von Menschen Gemeinsamkeiten, Unterschiede, Gruppen und Muster in Daten erkennen kann [46]. Unüberwachtes Lernen hat, aufgrund seiner Fähigkeit Muster eigenständig erlernen zu können, viele Anwendungsfälle und kann besonders für die Bildanalyse, wie Objekterkennung, oder für die Erkennung von Auffälligkeiten in Daten Verwendung finden [48].

**Deep Learning** (z.Dt. tiefes Lernen) ist eine besondere Form des maschinellen Ler-

nens und ist in der Lage große Datenmengen zu verarbeiten [49]. Im Mittelpunkt des tiefen Lernens stehen **Artificial Neural Networks (ANNs)** (z.Dt. künstliche neurale Netzwerke (KNN)), die ebenfalls ein Teilgebiet des maschinellen Lernens sind [50]. Künstliche neuronale Netzwerke sind Algorithmen, deren Struktur und Verhalten nach dem biologischen Vorbild des menschlichen Gehirns und seiner Neuronen modelliert ist [49, 51]. Neuronale Netzwerke ermöglichen somit das Erlernen von komplexen Mustern und sind in der Lage genaue Vorhersagen zu treffen. Da tiefes Lernen sehr rechenintensiv ist, kann es Monate dauern, bis Vorhersagen und Entscheidungen gute Ergebnisse liefern [51]. Ein sehr bekanntes Beispiel für ein neuronales Netzwerk ist der Suchalgorithmus von Google [50]. Bekannte Modelle tiefen Lernens im Bereich der generativen KI sind **Generative Adversarial Networks (GANs)** [52] und **Variational Auto Encoders (VAEs)** [53].

**Large Language Models (LLMs)** (z.Dt. großes Sprachmodell) sind große generative Sprachmodelle und stellen einen Teilbereich der KI dar [54]. Eine sehr prominente Architektur von großen Sprachmodellen sind Transformer [54]. Die Architektur von Transformern basiert auf neuronalen Netzwerken und zeichnet sich dadurch aus, dass es Kontext und Bedeutung durch Beobachtung der Beziehungen zwischen Wörtern erlernen kann [54]. Transformer wurden zuerst von Google in „Attention Is All You Need.“ [55] vorgestellt und führen zwei neue Techniken ein: positionelle Kodierung und Selbstaufmerksamkeit [54]. Große Sprachmodelle werden meist durch unüberwachtes Lernen auf riesigen Datenmengen trainiert und weisen in der Praxis Milliarden an Parameter auf [54]. Als solches erlernen LLMs die Muster der Statistiken und Wahrscheinlichkeitsverteilungen innerhalb der Trainingsdaten und optimieren mit diesen die Genauigkeit ihrer Vorhersagen anhand ihrer Anwendungsfälle [56]. Anwendungsfälle von Transformern sind vielfältig und können das z.B. Erzeugen, Übersetzen oder Zusammenfassen von Text umfassen [54].

**Explainable AI (xAI)** (z.Dt. erklärbare KI) befasst sich mit der Erklärbarkeit von KI. Erklärbare KI beschäftigt sich u.a. mit der Frage „Wie kommt eine KI zu einem Ergebnis“ und versucht die inneren Abläufe in einer KI zu beleuchten [57]. Das generelle Problem mit KI ist, dass die Algorithmen wie neuronale Netzwerke und tiefes Lernen sogenannte „Black-Boxes“ sind [57]. Zwar ist die Funktionsweise dieser Algorithmen, sowie das finale Ergebnis bekannt, aber es ist nicht erkennbar wie die KI dieses Ergebnis erreicht hat [57]. Ziel von erklärbarer KI soll also sein, dass ein Mensch nachvollziehen kann, wie ein Algorithmus ein Ergebnis erreicht [57].

Die meisten Systeme generativer KI arbeiten mit einer sogenannten **Prompt** (z.Dt. Eingabeaufforderung) als Eingabe. Im Rahmen der Verarbeitung natürlicher Sprache und des maschinellen Lernens bezeichnet ein Prompt ein Stück Text, das als eine Anweisung an ein System verstanden werden kann und welches den Kontext und alle dazugehörigen und notwendigen Informationen enthält, damit ein Modell relevante Ausgaben erzeugen kann [58, 59]. Anwendungsbeispiele für Prompts können

Erklärungen, Fragen oder anderweitige Texteingaben sein. Das Modell eines Systems verarbeitet die Eingabe und produziert in Abhängigkeit der Eingabe und der Fähigkeiten des jeweiligen Modells entsprechende Ausgaben, wie z.B. Vervollständigungen von Text, Erklärungen, Übersetzungen, Zusammenfassungen oder Bilder. Beim Dirigieren des Verhaltens eines Systems spielen Prompts daher eine entscheidende Rolle [60]. Grundlegende Elemente oder Inhalte einer Prompt können Anweisungen, Kontext, Eingabedaten oder Angaben zur gewünschten Ausgabe umfassen [61]. Bereits durch einfache Prompts können viele Aufgaben ausgeführt werden, jedoch hängt die Qualität der Ausgabe von der Qualität der Eingabe ab [61]. Durch das Bereitstellen von mehr Informationen in einer Prompt können die Wahrscheinlichkeiten zum Generieren gewünschter Ausgaben erhöht werden. Dieser Ansatz, Prompts anzupassen, um optimale Ergebnisse zu erzielen, heißt **Prompt Engineering** (z.Dt. Entwicklung von Eingabeaufforderungen) [62]. Prompt Engineering ist, in Essenz, ein iterativer, experimenteller Prozess. Begonnen wird zumeist mit einfachen Prompts, die iterativ und inkrementell weiterentwickelt werden. Einfache Anpassungsmöglichkeiten sind Prompts schlicht, kurz und knapp zu halten und eine möglichst präzise Formulierung zu verwenden. Diese Anpassungsmöglichkeiten können durch bekannte Techniken des Prompt Engineering, wie Zero-Shot [63], Few-Shot [64] und Chain-of-Thought (CoT) [65] erweitert werden.

**Zero-Shot** ist eine Technik, in der Modelle wie GPT-3.5 aufgrund der großen Datenmengen, mit welchen sie trainiert wurden, in der Lage sind Prompts ohne vorher spezifizierte Informationen oder Beispiele wirksam auszuführen [63]. Ein Anwendungsbeispiel für Zero-Shot Prompting ist die folgende Instruktion einen Text in eine andere Sprache zu übersetzen.

**Eingabe:** Übersetze „Das Auto hat einen Platten“ in Englisch.

**Ausgabe:** „The car has a flat tire“.

**Few-Shot** beschreibt eine Technik, die sich dadurch auszeichnet, dass der Prompt Instruktionen in Form von ergänzenden Informationen zum Ausführen der Prompt beigefügt werden [64]. Ein Anwendungsbeispiel für Few-Shot Prompting ist die folgende Reihe an Prompts, die als Beispiele dienen.

**Eingabe:** „Hauptstadt der Niederlande ist ...“.

**Ausgabe:** „Amsterdam“.

**Eingabe:** „Hauptstadt von Norwegen ist ...“.

**Ausgabe:** „Oslo“.

**Eingabe:** „Hauptstadt von Deutschland ist ...“.

**Ausgabe:** „Berlin“.

In diesem Beispiel wird dem Modell eine Reihe an Beispielen für das gewünschte Verhalten bereitgestellt. Anhand dieser Beispiele soll das Modell das Muster erkennen und erlernen, um ähnliche Fragen beantworten zu können.

**Chain-of-Thought** hingegen zeichnet sich dadurch aus, dass die Anweisung, die mit einer Prompt vermittelt werden soll, in eine Reihe an Schritten heruntergebrochen wird [66]. Für jeden Schritt in dieser Reihe wird wiederum ein Prompt definiert. Ein Anwendungsbeispiel für eine leicht modifizierte Technik von CoT ist die folgende Prompt.

**Aussage:** Die ungeraden Zahlen in folgender Zahlenreihe ergeben eine gerade Zahl: 6, 4, 7, 12, 19, 1, 2. Gehe schrittweise vor.

**Antwort:**

1. Alle ungeraden Zahlen in der Reihe identifizieren: 7, 19, 1.
2. Alle ungeraden Zahlen addieren:  $7 + 19 + 1 = 27$ .
3. Feststellen, dass 27 eine ungerade Zahl ist.
4. Erkennen, dass die ursprüngliche Aussage „Die ungeraden Zahlen in folgender Zahlenreihe ergeben eine gerade Zahl: [...]“ eine falsche Aussage ist.

In einem ersten Schritt werden die gegebenen Informationen betrachtet und aus der gegebenen Zahlenreihe alle ungeraden Zahlen identifiziert. In einem zweiten Schritt formuliert das Modell eine Gleichung zum Summieren der ungeraden Zahlen. Subsequent wird die aufgestellte Gleichung gelöst und erkannt, dass die Summe der ungeraden Zahlen wiederum eine ungerade Zahl ist. Im letzten Schritt wird die ursprüngliche Aussage mit den neu gewonnenen Fakten verglichen und es wird erkannt, dass die ursprüngliche Aussage falsch ist.

Das Beispiel endet mit dem Satz „Gehe schrittweise vor“ (engl. Lets think step by step). Durch diesen Satz wird die Anweisung in der Prompt in eine Reihe an Schritten zerteilt. Das schrittweise Vorgehen hilft dem Modell das Problem besser zu verstehen und die richtige Antwort zu liefern. Diese Technik wurde zuerst in [67] beschrieben und wird auch Zero-Shot CoT genannt und kombiniert Zero-Shot und CoT miteinander. Es ist eine leistungsstarke Technik, die eine nicht zu unterschätzende Wirkung auf Schlussfolgerungen des Modells hat und die zur Verbesserung der Leistung von Modellen bei einer Vielzahl von Aufgaben eingesetzt werden kann.

Prompt Engineering ist ein sich rasant entwickelndes Feld im Bereich der generativen KI und stellt mittlerweile auch einen eigenen Beruf dar. Da Prompt Engineering ein iterativer und inkrementeller Prozess ist, werden stetig neue Techniken entwickelt, oder alte Techniken in neuer Weise miteinander kombiniert.

## *Systeme generativer KI*

Systeme generativer KI haben einen enormen Einfluss auf die Weltwirtschaft [68]. So kann, laut Schätzungen, generative KI die Produktivität der weltweiten Wirtschaften erhöhen und, unter Berücksichtigung vieler Anwendungsfälle, pro Jahr einen Mehrwert in Höhe von \$2.6 bis \$4.4 Billionen schaffen [68]. Dieser steigende wirtschaftliche Anreiz führt zu einer explosionsartigen und rasanten Entwicklung im Bereich der generativen KI und weltweit zur Entwicklung, Integration und Anwendung generativer KI [68, 69]. Aus diesem Grund ist eine Aufführung, oder gar eine ausführlichere Liste solcher Systeme im Rahmen dieser Arbeit nicht machbar. Daher werden in diesem Abschnitt nur ausgewählte Systeme generativer KI vorgestellt und kurz beschrieben. Die Auflistung dieser Systeme erfolgt nach den Kategorien Image und Text, sowie Audio und Video.

Die Kategorie **Image** zählt eine Auswahl an Programmen auf, die auf KI basieren und in der Lage sind, aus Ausdrücken natürlicher Sprache, Bilder und Kunst zu erzeugen. Diese Prozesse der Bildzeugung nennt man Text2Image (z.Dt. Text zu Bild). Ein paar der bekanntesten Programme sind:

- **DALL · E 2** [17] ist eine von OpenAI [70] entwickelte KI, die mittels Text To Image realistische Bilder und Kunst schaffen kann.
- **Stable Diffusion** [71] ist eine von StabilityAI [72] entwickelte KI, die ebenfalls auf Grundlage von Ausdrücken natürlicher Sprache Bilder erzeugen kann. Im Vergleich zu DALL · E 2 ist Stable Diffusion eine Open Source Software.
- **Midjourney** [73] ist eine KI, die auf Grundlage einer Textbeschreibung Bilder und Kunst erstellen kann.

Die Kategorie **Text** zählt eine Auswahl an Programmen auf, die auf KI basieren und anhand von Eingaben eines Nutzers, meist einer sogenannten „Prompt“, die auf Ausdrücken natürlicher Sprache basieren, wiederum Text zu generieren.

- **OpenAI API** bietet Zugang zu OpenAIs Sprachmodellen, wie GPT-3.5 [74] und GPT-4 [74]. Diese Sprachmodelle sind in der Lage anhand Instruktionen in Form von Ausdrücken natürlicher Sprache eine Reihe an Aufgaben durchzuführen.
- **Google Vertex AI** ist eine Plattform für maschinelles Lernen. Die **Vertex AI API** bietet Zugang zu Googles PaLM 2 [75], einer Familie an Sprachmodellen [76]. Spezialisierungen von PaLM 2 umfassen Med-PaLM für medizinische Anwendungsfälle und Sec-PaLM 2 für sicherheitsbezogene Anwendungsfälle [76]. PaLM 2 treibt generative KI-Funktionen, wie die PaLM API, -Tools, sowie Bard [77], an [76].

Im Rahmen dieser Arbeit sind vor allem Systeme zur Text- sowie Bilderzeugung von vorrangigem Interesse. Generative KI ist allerdings nicht auf Text- oder Bilderzeugung beschränkt, sondern kann auch in anderen Bereichen, wie der Audio- oder Videoerzeugung Verwendung finden. Diese Systeme können, wie bereits in Abschnitt 1.1 beschrieben, einen echten Mehrwert für Benutzer darstellen. Daher wird im Folgenden noch ein kurzer Überblick über die Kategorien Audio und Video geben.

In der Kategorie der Programme zur Audioerzeugung gibt es bekannte Systeme, wie die von Eleven Labs [78], Resemble AI [79] oder Bark [80]. Anwendungsfälle für diese Systeme können multilinguale Synchronisation, das Klonen von Stimmen, oder die automatische Erzeugung von Untertiteln in Videos sein. In der Kategorie der Programme zur Videoerzeugung sind bekannte Systeme, wie die von RunwayML [81], Synthesia [82] oder Rephrase AI [83]. Anwendungsfälle für diese Art von Systemen können schnelle Videoerzeugung, Videobearbeitung, vor allem im Bereich Marketing oder Nachrichtensendungen sein.

Die nachfolgende Abb. 2.6 zeigt eine Zusammenfassung und kurze Übersicht über ausgewählte aktuelle Systeme generativer KI zusammen mit ihren entsprechenden Anwendungsmöglichkeiten. Es ist wichtig zu betonen, dass diese Übersicht keinen Anspruch auf Vollständigkeit erhebt, da sich diese Landschaft rapide verändert und ständig neue Techniken, Technologien und Systeme entwickelt werden.

Anwendungsschicht	Design		Synchronisation	
	Schreiben	Bild erzeugung	Klonen von Stimmen	
	Notizen			
	Support (Chat / E-Mail)	Medien / Werbung	Video- bearbeitung / erzeugung	Automatische Erzeugung von Untertiteln
	Text	Image	Video	Audio
Modell ebene	OpenAI GPT-3.5 / GPT-4	OpenAI Dall-E 2	RunwayML	Eleven Labs
	DeepMind Gopher	Stable Diffusion	Synthesia	Resemble AI
	Anthropic Claude	Midjourney	Rephrase AI	Bark

Abb. 2.6: Übersicht ausgewählter aktueller Systeme generativer KI nach [84].

## **Einschränkungen**

Systeme generativer KI, wie z.B. das große Sprachmodell GPT-3 von OpenAI, sind bereits in der Lage realistische und höchst überzeugende Inhalte zu erzeugen. Trotzdem unterliegen Systeme generativer KI einer Reihe an Einschränkungen, die sich potentiell auf ihre Fähigkeiten und Anwendungsfälle auswirken können. Einige der wichtigsten Einschränkungen von Systemen generativer KI sind:

1. **Mangelnde Transparenz und Erklärbarkeit:** Modelle generativer KI sind sog. Black-Boxen und bieten keinen Einblick in ihren Entscheidungsprozess [57]. Aufgrund dieser mangelnder Transparenz ist es nicht ersichtlich, wie ein Modell zu einer Entscheidung und zu einem Ergebnis kommt [57]. Resultat ist eine eingeschränkte Erklärbarkeit von Modellen.
2. **Abhängigkeit von Trainingsdaten und ethische Bedenken:** Modelle generativer KI sind Produkte ihrer Trainingsdaten und daher stark von ihnen abhängig [85]. Aufgrund dieser Abhängigkeit können sich Ungenauigkeiten aus den Trainingsdaten in den generierten Ergebnissen wiederspiegeln [85]. Des Weiteren sind diese Modelle zumeist speziell auf Bereiche angepasst und lassen sich nur schwer auf andere Bereiche anwenden, die nicht von den Trainingsdaten vorgesehen sind [85]. Ein weiteres Problem generativer KI sind ethische Bedenken [85]. Durch Voreingenommenheiten in den Trainingsdaten können Modelle selbst voreingenommene oder gar schädliche Inhalte erzeugen [85].
3. **Mangelndes Verständnis des Kontextes:** Die Länge des Kontextfensters eines Modells begrenzt die Menge an vorausgehendem Text, welches ein Modell bei der Generierung einer Antwort auf eine Eingabeaufforderung berücksichtigt [86]. Informationen, die außerhalb dieses Kontextfensters liegen, werden unter Umständen nicht mehr berücksichtigt und können zu einem unvollständigen Verständnis des Textes oder zu kontextuell verfälschten Antworten führen [86]. In Kombination mit den ausgewählten Trainingsdaten beeinflusst die Länge des Kontextfensters somit maßgeblich die Fähigkeiten des Modells den Kontext zu verstehen, langfristige Abhängigkeiten zu erfassen und kohärente Texte zu generieren [86].

Trotz der bemerkenswerten Fähigkeit ansprechende und überzeugende Inhalte zu generieren, ist es wichtig die Einschränkungen und Grenzen generativer KI im Blick zu behalten, um so einen effektiven, verantwortungsbewussten und informierten Einsatz von generativer KI garantieren zu können. Im Bereich der Forschung gibt es laufende Bemühungen und Weiterentwicklungen, um diese Einschränkungen und Grenzen zu adressieren, abzustellen, die Zuverlässigkeit von generativer KI zu verbessern, sowie Fähigkeiten von Systemen generativer KI weiter auszubauen. Diese Bestrebungen können allgemein unter dem Begriff „Responsible AI“ zusammenge-

fasst werden, welche zum Ziel haben KI auf verantwortungsvolle Art und Weise zu entwickeln und die oben genannten Einschränkungen zu berücksichtigen [87].

### 2.1.5 Diskussion

In den ersten Abschnitten 2.1.1 bis 2.1.3 wurden zuerst grundlegende Technologien in Bezug auf das GMAF, welches das Bezugssystem dieser Arbeit darstellt, vorgestellt und detailliert erläutert. Weiter wurden in Abschnitt 2.1.4 grundlegende Techniken und Konzepte im Rahmen der generativen KI genannt und erklärt, sowie aktuelle Systeme generativer KI nach Kategorien gelistet.

Im Rahmen der Diskussion dieses Forschungsziels werden die Anwendungsmöglichkeiten der in Abschnitt 2.1.4.2 vorgestellten Systeme in Bezug auf das GMAF diskutiert. Anhand dieser Diskussion wird dann eine fundierte Entscheidung getroffen, welche Systeme generativer KI für die Umgebung „Erklärbarkeit von MMIR mittels generativer KI“ am besten geeignet sind.

Damit eine fundierte und informierte Entscheidung für die Auswahl von Systemen generativer KI getroffen werden kann, gilt es einige Kriterien zur Auswahl von Systemen zu beachten. Wichtige Kriterien zur Auswahl eines Systems sind:

- **Anpassungsmöglichkeiten:** Wie flexibel ist ein System? Ermöglicht es die Anpassung anhand Anforderungen besonderer Anwendungsfälle?
- **Kompatibilität und Integration:** Ist ein System mit der Programmiersprache oder der Umgebung des GMAF kompatibel? Bietet das System eine Schnittstelle, um eine einfache Integration des Systems in das GMAF zu ermöglichen? Bietet das System eine klare Dokumentation, um die Integration der Schnittstelle zu unterstützen?
- **Skalierbarkeit und Leistung:** Ist das System in der Lage größere Datens Mengen zu verarbeiten? Kann es auf Anfrage zeitnah Ergebnisse generieren?
- **Kosten und Lizenzierung:** Einige Systeme bieten kostenfreie Nutzung oder berechnen die Kosten abhängig von der Nutzung, wiederum erfordern andere Systeme eine kommerzielle Lizenz zur Nutzung. Welche Lizenzbedingung oder Preise sind mit der Nutzung eines Systems verbunden?

Das Berücksichtigen dieser Kriterien in der Auswahl eines Systems generativer KI ermöglicht eine wohl informierte und fundierte Entscheidung über die Integration eines Systems in das GMAF und legt das Fundament für das Erreichen von Anforderungen und Zielen. Daher wird zuerst die Auswahl eines Systems aus der Kategorie Image zur Bildgenerierung diskutiert. Danach folgt eine Diskussion zur Auswahl eines Systems aus der Kategorie Text zur Textgenerierung. Die Diskussion orientiert sich dabei jeweils nach den Systemen und nicht nach den oben genannten Kriterien.

### **Ausgewählte Systeme zur Bildgenerierung**

Zur Auswahl eines Systems zur Bildgenerierung stehen OpenAIs DALL·E 2, StabilityAIs Stable Diffusion und Midjourney. Um einen besseren Eindruck zu den jeweiligen Modellen bzw. Systemen zu ermöglichen, wird für jedes System ein Bild anhand der Textbeschreibung bzw. Prompt „*Beautiful Venice canals with gondolas and bridges, charming*“ generiert.

**Midjourney** ist eine generative KI, die auf Grundlage einer Textbeschreibung überzeugende und ästhetisch ansprechende Bilder erzeugen kann. Abb. 2.7 zeigt ein Beispiel für ein von Midjourney generiertes Bild. Um auf Midjourney zugreifen zu



Abb. 2.7: Beispiel für ein von Midjourney generiertes Bild anhand oben genannter Textbeschreibung [88].

können, wird eine Anmeldung über den Onlinedienst Discord vorausgesetzt. Weitere Voraussetzungen zum Nutzen Midjourneys sind das Hinzufügen und Beitreten des Discord-Servers von Midjourney und das Akzeptieren der allgemeinen Nutzungsbedingungen. Sind diese Voraussetzungen erfüllt, kann Midjourney über Befehle im Discord-Chat gesteuert werden. Bilder können über den Befehl /imagine und einer

Textbeschreibung **prompt** erzeugt werden. Midjourney bietet eine Reihe an Anpassungsmöglichkeiten seiner Prozesse, die gezielt durch Befehle, die in [89] dokumentiert sind, vorgenommen werden können. Midjourney bietet in einem begrenztem Testversuch ein initiales Guthaben von 25 Credits zum Erstellen von Bildern. Sind diese Credits aufgebraucht, kann aus einer Reihe an unterschiedlichen, gestaffelten Mitgliedschaften (siehe [90]), welche jeweils eigene Vorteile bieten, gewählt werden. Eine Integration von Midjourney in eine andere Software ist aufgrund einer fehlenden Schnittstelle oder Anbindung bzw. Bibliothek nicht möglich. Zwar gibt es Drittanbieter, die eine Schnittstelle bereitzustellen scheinen, allerdings verweist Midjourney in seinen allgemeinen Nutzungsbedingungen ausdrücklich, dass es keine Schnittstelle bereitstellt und Automatisierungen nicht duldet [91].

**Stable Diffusion** von StabilityAI ist eine quelloffene generative KI, die ebenfalls auf Grundlage einer Textbeschreibung Bilder erzeugen kann. Abb. 2.8 zeigt ein Beispiel für ein von Stable Diffusion generiertes Bild. Zugriff auf Stable Diffusion kann

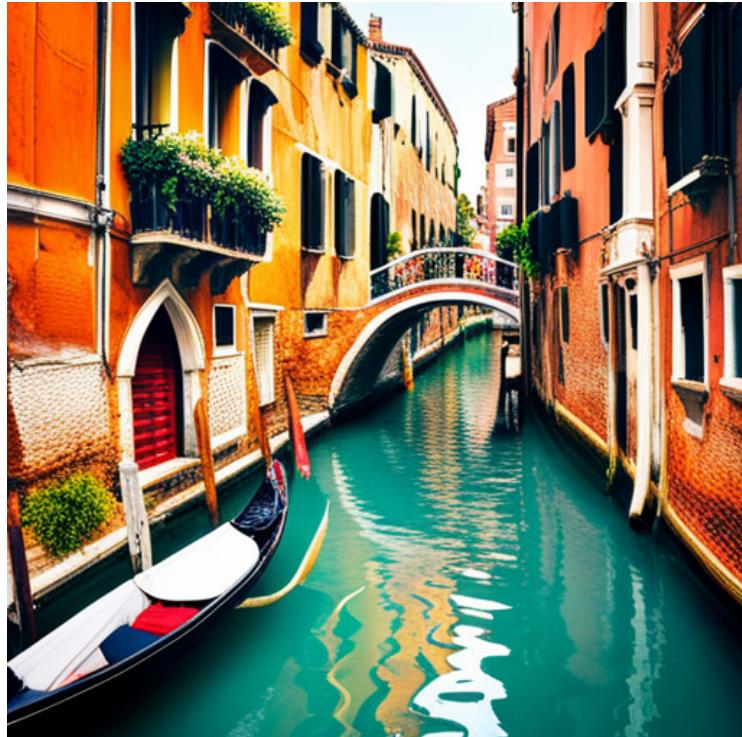


Abb. 2.8: Beispiel für ein von Stable Diffusion generiertes Bild anhand oben genannter Textbeschreibung.

durch eine Installation auf einer lokalen Maschine oder durch das WebInterface DreamStudio erfolgen. In beiden Fällen erfolgt der Zugriff durch ein WebInterface, in welchem eine Reihe an Anpassungsmöglichkeiten durch Einstellungen vorgenom-

men werden können. Aufgrund der hohen Rechenanforderungen benötigt eine lokale Installation von Stable Diffusion eine hohe Mindestanforderung an VRAM. Auch Stable Diffusion bietet in einem begrenztem Testversuch ein initiales Guthaben von 25 Credits. Sind diese Credits aufgebraucht, können für einen Mindestwert von 10 Euro Credits hinzugekauft werden. Die Preise für die Generierung von Bildern sind abhängig von den gewählten Eigenschaften der Bilder und können in [92] eingesehen werden. Eigenschaften können die Bildgröße, aber auch die Anzahl der Iterationen der Bildgenerierung sein. Diese Einstellungen haben einen Einfluss auf die Qualität der generierten Bilder und folglich auch auf die Rechenintensität und beeinflussen dementsprechend die Preise zur Generierung von Bildern. Eine Integration von Stable Diffusion ist durch eine offizielle REST (Representational State Transfer) [93] Schnittstelle möglich.

**DALL · E 2** ist eine generative KI, die ebenfalls auf Grundlage einer Textbeschreibung Bilder erzeugen kann. Abb. 2.9 zeigt ein Beispiel für ein von DALL · E 2 generiertes Bild. Zugriff auf DALL · E 2 erfolgt über ein WebInterface. OpenAI bietet

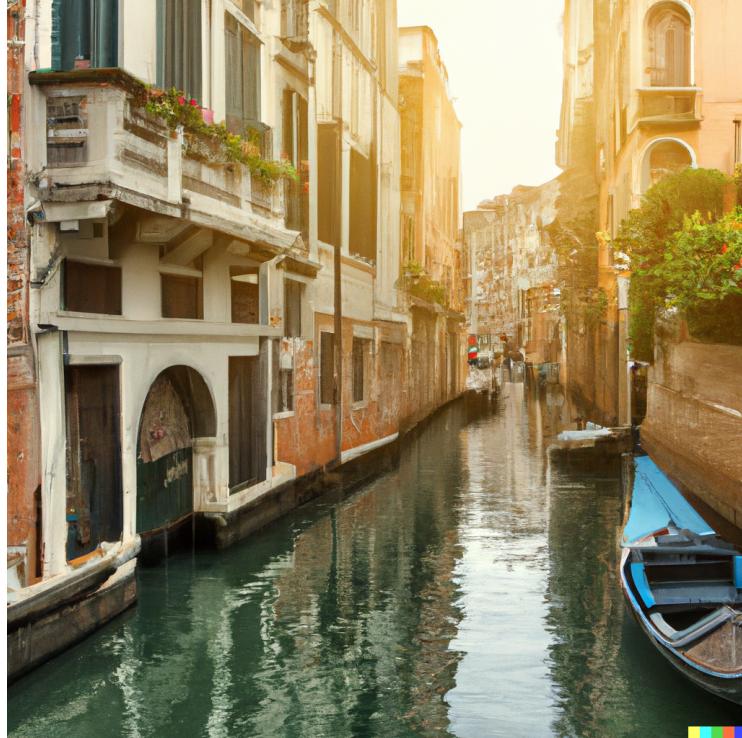


Abb. 2.9: Beispiel für ein von DALL · E 2 generiertes Bild anhand oben genannter Textbeschreibung.

monatlich eine freie Anzahl an Credits, mit welchen Bilder erzeugt werden können. Sind diese Credits aufgebraucht, kann für einen Mindestwert von 15 Dollar 115 Cre-

dits hinzugekauft werden. Im Vergleich zu Stable Diffusion sind die Preise zur Bildgenerierung von DALL · E 2 nur von der Bildgröße abhängig (siehe [94]). OpenAI bietet über eine Schnittstelle Zugriff auf DALL · E 2, sowie andere Modelle. Diese Schnittstelle verfügt über eine wohl definierte Dokumentation, die in [95] eingesehen werden kann. Des Weiteren unterstützt OpenAI zur Integration seiner Modelle eine Reihe an Community-Bibliotheken für unterschiedliche Sprachen, darunter eine Anbindung der Schnittstelle für Java (siehe [96]). Auch die Dokumentation zur Anbindung der Schnittstelle in Java unterstützt eine klare Dokumentation (siehe [96]).

Alle drei vorgestellten Systeme ermöglichen das Erzeugen von qualitativ hochwertigen Bildern. Hinter allen drei Systemen stehen in Bezug auf die Bepreisung der Generation von Bildern sogenannte Credits. Dabei bieten alle drei Systeme am Anfang eine gewissen Zahl an freien Credits. Midjourney und Stable Diffusion vergeben beim Erstellen eines Kontos eine endliche Anzahl an Credits. OpenAI hingegen vergibt ein anfängliches Guthaben von 5 Dollar und monatlich eine freie Anzahl an Credits, die jeden Monat zurückgesetzt werden. Credits gelten allerdings nicht für die Schnittstelle von OpenAI und werden separat berechnet [97]. Die Anzahl der Credits sind im Rahmen von OpenAI somit nicht von Belang. Die Berechnung der Preise geschehen im Fall der Schnittstelle von OpenAI nutzungsbasiert und werden am Ende des Monats abgerechnet, anstatt einer monatlichen Zahlung wie im Fall von Midjourney. Die Tatsache, dass Midjourney keine Schnittstelle anbietet, disqualifiziert Midjourney für die Auswahl eines Systems zur Bildgenerierung im Rahmen dieser Arbeit. Offen für die Auswahl bleiben StabilityAIs Stable Diffusion und OpenAIs DALL · E 2. Ein Argument gegen Stable Diffusion ist, dass sobald Credits aufgebraucht sind, auch zum Anwenden der Schnittstelle neue Credits für einen Mindestwert von 10 Euro hinzugekauft werden müssen, was umgerechnet 1000 standardmäßig generierten Bildern entspricht. Stable Diffusion setzt die Credits, im Gegensatz zu OpenAIs DALL · E 2, voraus, um die REST Schnittstelle nutzen zu können. Ein Argument für OpenAIs DALL · E 2 ist die bestehende Java Anbindung für die Schnittstelle, die eine einfache Integration ermöglicht. Im Rahmen dieser Arbeit wird für die folgende Modellierung und Implementierung Dall-E 2 als Vertreter zur Bildgenerierung ausgewählt.

### ***Ausgewählte Systeme zur Textgenerierung***

Zur Auswahl eines Systems zur Textgenerierung stehen OpenAIs Sprachmodelle, wie GPT-3.5 und GPT-4, sowie Googles PaLM 2. Die Anwendung der Systeme zur Textgenerierung werden an folgendem Beispiel demonstriert: „Auf meiner Reise durch die Städte Frankreichs ist mir in einem kleinen Dorf in einem Dialog mit einer Bewohnerin ein Satz aufgefallen, dessen Sinn sich mir nicht erschließen will. Sie sagte: "Tomber dans les pommes". Wollte sie mir Pommes zum Essen anbieten?“

**Pathway Language Model 2 (PaLM 2)** [76] ist ein von Google entwickeltes und auf enormen Datenmengen trainiertes großes Sprachmodell [75]. PaLM 2 zeichnet sich durch Fähigkeiten, wie das Verstehen, Erzeugen und Übersetzen von natürlicher Sprache, das Erzeugen von Code, sowie aber auch durch logisches Denken aus [76, 75]. Einige der wichtigsten Merkmale von PaLM 2 ist das verbesserte Verarbeiten von Daten, das Unterstützen vieler Sprachen und verbessertes logisches Denken im Vergleich zum Vorgänger PaLM [75]. Ungeachtet dessen kann PaLM 2 von Zeit zu Zeit unvollständige oder gar fälschliche Erklärungen bereitstellen [98]. Die PaLM 2 Familie besteht aus vier verschiedenen Modellen, die sich jeweils in aufsteigender Größe voneinander unterscheiden: *Gecko*, *Otter*, *Bison*, *Unicorn* [76]. Die unterschiedlichen Größen ermöglichen einen großen Bandbreite an Anwendungsmöglichkeiten [76]. Kleinere Modelle, wie Gecko könnten z.B. in mobilen Endgeräten Anwendung finden [76]. PaLM 2 befindet sich noch in der Entwicklung und Google bietet zum Zeitpunkt des Verfassens dieser Arbeit keine offizielle öffentliche Schnittstelle für PaLM 2 an [99]. Eine Möglichkeit dennoch Zugriff auf PaLM 2 zu erhalten, ist über Vertex AI oder über eine Warteliste [100]. Eine weitere Möglichkeit PaLM 2 zu nutzen, ist Bard [77], Googles Pendant zu ChatGPT [101]. Die Bepreisung für das Nutzen der Vertex AI Schnittstelle ist abhängig von der mit den Anfragen verbundenen Rechennutzung [102]. Da sich die PaLM Schnittstelle noch in einer Entwicklungsphase bzw. Vorschauphase befindet, ist der Zugriff für Entwickler, sofern freigeschaltet, kostenfrei [103]. Andernfalls steht die Ankündigung der Bepreisung der PaLM API durch Google noch aus [103]. Folgender Textausschnitt zeigt die Antwort auf das oben genannte Beispiel.



Bard (PaLM 2)

**Bard:** Nein, die Bewohnerin wollte dir keine Pommes anbieten. Die französische Redewendung „tomber dans les pommes“ bedeutet „sich übergeben“. Es ist eine gängige Redensart in Frankreich, die verwendet wird, um auszudrücken, dass jemand sich sehr schlecht fühlt und sich übergeben muss. Die Redensart hat ihren Ursprung im 18. Jahrhundert und leitet sich von dem französischen Wort „pomme“ ab, das „Apfel“ bedeutet. Im 18. Jahrhundert glaubte man, dass Äpfel eine heilende Wirkung haben und bei Übelkeit und Erbrechen helfen können. Daher wurde die Redewendung „tomber dans les pommes“ verwendet, um auszudrücken, dass jemand sich so schlecht fühlt, dass er Äpfel essen muss. Heute wird die Redensart „tomber dans les pommes“ immer noch verwendet, um auszudrücken, dass jemand sich sehr schlecht fühlt und sich übergeben muss. Es ist eine gängige Redensart in Frankreich und wird auch in anderen französischsprachigen Ländern verwendet.

**GPT-3.5** [74] und **GPT-4** [74] sind von OpenAI entwickelte und ebenfalls auf enormen Datenmengen trainierte große Sprachmodelle. Ähnlich wie PaLM 2 sind GPT-3.5 und GPT-4 in der Lage Text zu verstehen, Text zu erzeugen, Fragen zu beantworten und Übersetzungen oder Zusammenfassungen anzufertigen [104]. GPT-4 stellt die neuste Version der GPT Modellreihe dar und soll laut OpenAI besseres Textverständnis und eine höhere Leistung als die Vorgängermodelle liefern [105]. Da OpenAI keine genauen Spezifikationen zu GPT-3.5 und GPT-4 veröffentlicht hat, ist die genaue Anzahl der Parameter in GPT-3.5 bzw. GPT-4 nicht bekannt, aber laut Gerüchten soll sich die Anzahl auf 175 Milliarden bzw. 1 Billion Parameter belaufen [106]. Im Vergleich dazu umfasst das Vorgängermodell GPT-3, von welchem GPT-3.5 nur eine optimierte Version ist, nur 175 Milliarden Parameter [104]. Eine der neusten Funktionen von GPT-4 ist die Eingabe und Verarbeitung von Bildern, anhand welcher es Erklärungen generieren können soll [105]. Dies macht GPT-4 zu einem echten multimodalen Modell, welches mehrere Arten von Eingaben akzeptieren kann [105]. Bildeingaben sind aber nach wie vor in einer begrenzten Alpha Testphase und nicht unmittelbar verfügbar [105]. OpenAI bietet über eine REST Schnittstelle einfachen Zugriff auf die Sprachmodelle, wie GPT-3.5 und GPT-4, sowie Derivate, wie z.B. GPT-3.5-Turbo-(4k/16k). Unterschiede in diesen Derivaten liegen in der Größe des Kontextes, also der Länge an Textzeichen, die ein Modell bei der Generierung berücksichtigt. Die für die Sprachmodelle angebotene Schnittstelle ist die gleiche Schnittstelle wie die für DALL·E 2 und bietet ebenfalls eine umfassende Dokumentation (siehe [95]). Diese Dokumentation bietet nützliche Information, Leitfaden zur Nutzung, Beispiele für Quellcode und Beschreibungen für effektives Nutzen der Schnittstelle. Seit dem 6. Juli 2023 erlaubt OpenAI den Zugriff auf das Sprachmodell GPT-4-8k für alle Anwender der Schnittstelle, die einen Verlauf erfolgreich abgeschlossener Zahlungen aufweisen [107]. Die Bepreisung hängt dabei, sehr ähnlich zu PaLM 2, von der Anzahl an der in der Anfrage gesendeten Zeichen, sowie der in der Antwort enthaltenen Zeichen ab, und kann in [94] eingesehen werden. Folgender Textausschnitt zeigt die Antwort auf das oben genannte Beispiel.



### ChatGPT (GPT-3.5)

**ChatGPT:** Nein, der Satz "Tomber dans les pommes" bedeutet nicht, dass die Bewohnerin dir Pommes zum Essen anbieten wollte. Tatsächlich ist es eine französische Redewendung, die einen Zustand beschreibt, in dem jemand ohnmächtig wird oder in Ohnmacht fällt. Die wörtliche Übersetzung ist "in die Äpfel fallen", aber im übertragenen Sinne bedeutet es, dass jemand kollabiert oder bewusstlos wird.

Diese Redewendung wird umgangssprachlich verwendet und ist im Französi-

schen sehr geläufig, um einen plötzlichen Zustand der Bewusstlosigkeit oder des Zusammenbruchs zu beschreiben. Also, die Bewohnerin wollte dir keine Pommes anbieten, sondern sprach darüber, dass jemand ohnmächtig wird.

**Benutzer:** Sehr interessant, vielen Dank!

**ChatGPT:** Gern geschehen!

Verglichen mit GPT-3.5 bietet GPT-4 grundlegend verbesserte Funktionen und Fähigkeiten [105]. Eine weiterer Vergleich verbleibt somit zwischen OpenAIs Sprachmodellen GPT-3.5 und GPT-4, sowie Googles PaLM 2.

Ein grundlegendes Problem von Sprachmodellen ist, dass sie zu Halluzinationen neigen. Es ist wichtig zu verstehen, dass wenn große Sprachmodelle auf enormen Datensätzen trainiert werden, sie die in diesen Daten enthaltenen statistische Muster und Strukturen erlernen [108]. Als solches sind sie in der Lage Sätze oder Texte zu generieren, die kohärent wirken, ohne dabei wirklich ein wahres Verständnis des Textes oder ein Bewusstsein zu besitzen [108]. Dies kann zu faktisch falschen oder unsinnigen Textzeugnissen, auch Halluzinationen genannt, führen und sind Resultat von mangelndem Kontext und Textverständnis. [109]. Dieses Problem wird verstärkt durch veraltetes Wissen. So gilt für die meisten Sprachmodelle von OpenAI, auch GPT-4, dass ihre Trainingsdaten respektive ihr Wissen nach September 2021 abrupt enden [74]. Dies ist für viele Anwendungsfälle dieser Sprachmodelle eine einschneidende Einschränkung, besonders für den Aspekt der Erklärbarkeit. Ungeachtet dessen sind alle vorgestellten Sprachmodelle nach wie vor von Natur aus Black-Boxes und weisen somit nur bedingt Erklärbarkeit auf. Ein Argument für GPT-4 oder PaLM 2 ist, dass sie im Vergleich zu ihren Vorgängern jeweils eine wesentliche Verbesserung darstellen. Eine besonders interessante Eigenschaft von PaLM 2 ist der Fokus auf Multilingualität, d.h. PaLM 2 wurde mit Daten in vielen Sprachen trainiert [75]. Jedoch ist ihr Zugang entweder zur Zeit nicht vorhanden oder kann nur über eine Warteliste angefragt werden. Seit dem 6. Juli 2023 erlaubt OpenAI den Zugriff auf das Sprachmodell GPT-4-8k für alle Anwender der Schnittstelle, die einen Verlauf erfolgreich abgeschlossener Zahlungen aufweisen [107]. Ein Argument für GPT-3.5 ist hingegen die Schnelligkeit im Vergleich zu GPT-4 und ein weitaus einfacher Zugriff über die Schnittstelle, die bereits über eine Anbindung für Java verfügt. Im Rahmen dieser Arbeit wird für die folgende Modellierung und Implementierung GPT-3.5 / 4 als Vertreter zur Textgenerierung ausgewählt.

## **Schlussfolgerung**

Anhand der Erkenntnisse der beiden Diskussionen zur Wahl von Systemen zur Bild-, sowie Textgenerierung, fällt die Wahl eines Systems zur Bildgenerierung auf OpenAIs DALL · E 2, sowie die Wahl eines Systems zur Textgenerierung auf das Sprachmodelle GPT-3.5 und seiner Derivate, sowie eventuell auf andere von OpenAI angebotenen Sprachmodelle. Diese Entscheidungen liegen in den folgenden Eigenschaften begründet:

- Einfache Integration durch die offiziell von OpenAI unterstützte Community Bibliothek „OpenAI-Java“ [96] für Java.
- Beide Systeme, DALL · E 2 und die GPT-Sprachmodelle, werden bereitgestellt durch eine einzige Plattform: OpenAI.
- Zahlung einfach und flexibel an einen einzigen Anbieter. Zudem muss nur für das gezahlt werden, was wirklich genutzt wird, *Only pay for what you use*, anstatt einer monatlichen Fixzahlung.

Die Wahl der primären Systeme liegt somit bei den oben genannten Systemen, deren Einsatz im weiteren Verlauf dieser Arbeit verfolgt wird. Unabhängig von der Wahl der primären Systeme verbleiben die Systeme PaLM 2 oder Stable Diffusion weiterhin als optionale Kandidaten, die ebenfalls integriert werden könnten.

Die vorgestellten und ausgewählten Systeme bieten umfassende Möglichkeiten zur Bild- und Textgenerierung. Jedoch existiert keine Benutzerschnittstelle im GMAF, die geeignete Interaktionsmöglichkeiten für Systeme generativer KI ermöglicht. Dies führt zu einer weiteren, zweiten Herausforderung im Rahmen dieses Forschungsziels.

### **Offene Herausforderung OH 1.2 - Benutzerschnittstelle**

Sollen diese Systeme im Rahmen des GMAF sinnvoll eingesetzt werden, um Erklärungen generieren zu können, so müssen Anwendern des GMAF durch eine Benutzerschnittstelle geeignete Interaktionsmöglichkeiten mit diesen Systemen geboten werden.

Die offene Herausforderung **OH 1.2** wird im Rahmen einer Modellierung, sowie Implementierung ebenfalls in jeweils FZ 1.2/TB und FZ 1.3/I angesprochen und untersucht.

Im nächsten Abschnitt werden die im Rahmen dieses Forschungsziels identifizierten offenen Herausforderungen zusammengefasst.

## ***Offene Herausforderungen***

In Abschnitte 2.1.1 bis 2.1.3 wurden Technologien in Bezug auf das GMAF vorgestellt. Es konnte festgestellt werden, dass der Ansatz MMFGs bzw. Graph Codes erklärbar zu machen, sehr statistisch und rein mathematischer Natur ist. Eine Verbesserung bzw. einen alternativen Ansatz, um MMFGs bzw. Graph Codes erklärbar machen zu können, verspricht der Einsatz von Systemen generativer KI. Dies führt zur ersten offenen Herausforderung **OH 1.1** und der Frage „Wie kann generative KI eingesetzt werden, um Erklärbarkeit zu erreichen?“. Es gilt daher im Rahmen der Umgebung „Erklärbarkeit von MMIR mittels generativer KI“ zu ermitteln, wie Systeme generativer KI genutzt werden können, um Erklärbarkeit zu erreichen bzw. Erklärungen zu generieren.

Im weiteren Verlauf wurden im Abschnitt 2.1.4 grundlegende Technologien, Techniken und Eigenschaften beschrieben und eine Reihe an Systemen generativer KI vorgestellt. Aus dieser Reihe an Systemen wurde in Abschnitt 2.1.5 eine Auswahl an Systemen getroffen, deren Einsatz im weiteren Verlauf dieser Arbeit verfolgt werden. Die in diesem Abschnitt vorgestellten und ausgewählten Systeme bieten umfassende Möglichkeiten zur Bild- und Textgenerierung. Sollen diese Systeme im Rahmen des GMAF sinnvoll eingesetzt werden, um Graph Codes erklärbar zu machen, so müssen Anwendern des GMAF durch eine Benutzerschnittstelle geeignete Interaktionsmöglichkeiten mit diesen Systemen geboten werden. Es konnte festgestellt werden, dass das GMAF für seine Anwender keine Interaktionsmöglichkeiten mit Systemen generativer KI bietet. Dies führt zur zweiten offenen Herausforderung **OH 1.2** und der Frage „Wie muss eine Benutzerschnittstelle beschaffen sein, um geeignete Interaktionsmöglichkeiten mit Systemen generativer KI zu ermöglichen?“. Es gilt daher anhand von Anwendungsfällen geeignete Schnittstellen zur Interaktion mit den in diesem Abschnitt vorgestellten Systemen generativer KI für die Umgebung „Erklärbarkeit von MMIR mittels generativer KI“ im Rahmen des GMAF in einem eigens dafür bestimmten Programm oder als Erweiterung in das GMAF umzusetzen.

Die beiden offenen Herausforderungen **OH 1.1** und **OH 1.2** werden im Rahmen einer Modellierung durch das Forschungsziel FZ 1.2/TB angesprochen und durch das Forschungsziel 1.3/I im Rahmen einer prototypischen Implementierung untersucht. Forschungsziel 1.4/E wird sich der Evaluierung der in diesen Kapiteln erlangten Forschungsergebnissen widmen.

## 2.2 FZ 2.1/O Integration generativer KI in das GMAF

Dieser Abschnitt gibt einen Überblick über ausgewählte Ergebnisse und Erkenntnisse aus der Forschung im Bereich der Integration von Systemen generativer KI und deckt technische Aspekte ab, die sich auf die zweite Problembeschreibung beziehen. Abschnitt 2.2.1 widmet sich den Integrationsmöglichkeiten von Graph Codes. Weiter befasst sich Abschnitt 2.2.2 mit den Integrationsmöglichkeiten von Systemen generativer KI.

### 2.2.1 Integrationsmöglichkeiten von Graph Codes

Dieser Abschnitt befasst sich mit den Integrationsmöglichkeiten von Graph Codes. Wie bereits in Abschnitt 2.1.3 beschrieben, sind Graph Codes eine Technologie zur effizienten und effektiven Darstellung und Verarbeitung von Merkmalen im GMAF. Als solches ist eine Integration von Graph Codes im Rahmen des GMAF bereits gegeben. Diese Integration bietet im Rahmen des GMAF folgende Komponenten:

- *GraphCode* bietet eine Implementierung der zweidimensionalen Darstellung von Merkmalen und ermöglicht Zugriff auf das Wörterbuch eines Graph Codes.
- *GraphCodeCollection* bietet eine Reihe an Operationen, wie z.B. die Vereinigung oder Subtraktion von Graph Codes an.
- *GraphCodeGenerator* ist eine Implementierung, die das Erstellen eines Graph Codes aus einem MMFG ermöglicht.
- *GraphCodeIO* bietet eine Reihe an nützlichen Operationen zum Import oder Export von Graph Codes aus oder in Dateien.
- *GraphCodeMetric* bietet eine Implementierung der in Abschnitt 2.1.3.3 beschriebenen Metrik zur Ähnlichkeit von Graph Codes.

Die grundsätzliche Darstellungsform eines Graph Codes erfolgt im textbasierten Dateiübertragungsformat JSON (JavaScript Object Notation) und kann somit leicht von Menschen, aber auch Maschinen gelesen und verarbeitet werden. Listing 2.1 zeigt das Beispiel des Graph Codes  $GC_{ex}$  in JSON-Format.

---

```
1 {  
2   "dictionary": [  
3     "root-image",  
4     "person",  
5     "body",  
6     "arm",  
7     "watch"  
8   ],
```

```

9   "collectionElements": [],
10  "matrix": [
11    [1,1,0,0, 0],
12    [0,1,1,0, 0],
13    [0,0,1,1, 0],
14    [0,0,0,1,17],
15    [0,0,0,0, 2]
16  ]
17 }

```

Listing 2.1: Beispiel des Graph Codes  $GC_{ex}$  im JSON-Format.

Dieses Beispiel zeigt, wie die Informationen eines Graph Codes, wie z.B. das Wörterbuch oder die Wertungsmatrix, als ein Objektgraph bestehend aus mehreren Paaren von Schlüssel und Werten im Syntax **Schlüssel : Wert** dargestellt werden. Im Falle der Paare aus Schlüsseln und Werten stellen Schlüssel immer eine Zeichenfolge, wie „dictionary“, „collectionElements“ oder „matrix“ dar, während Werte Zahlen, Strings, Arrays oder wiederum Objekte sein können. Die Elemente „dictionary“ und „collectionElements“ haben jeweils Arrays als Werte, in welchen eine Menge an Bezeichnern bzw. Graph Codes enthalten sein können. Da die gesamte Darstellung eines Graph Codes im JSON-Format wiederum ein Objekt ist, kann „collectionElements“ eine Menge bzw. Sammlung an Darstellungen von Graph Codes enthalten. Auf diese Weise kann eine rekursive Darstellung von Graph Codes erreicht werden. Das Element „matrix“ hat ebenfalls ein Array als Wert. Da allerdings die Adjazenzmatrix eine zweidimensional Matrix ist, sind die Werte im Array wiederum Arrays aus ganzen Zahlen mit einer Länge, die äquivalent zu Größe des Wörterbuchs ist. Das GMAF integriert bereits für die Verarbeitung und Operationen mit JSON die GSON [110] Bibliothek von Google ein.

Auch wenn Systeme wie GPT-3.5 oder GPT-4 und PaLM 2 in der Lage sind Daten im JSON-Format in Textform zu verarbeiten, so bleibt jedoch offen, wie die Informationen aus einem Graph Code und die damit assoziierten Informationen in die Eingabe für ein System generativer KI überführt werden kann, und welche Eigenheiten berücksichtigt werden müssen. Dies führt zu einer ersten offenen Herausforderung im Rahmen dieses Forschungsziels.

### Offene Herausforderung OH 2.1 - Transformieren von Graph Codes

Umwandlung von Graph Codes in eine passende Eingabeform für Systeme generativer KI.

Die offene Herausforderung **OH 2.1** wird im Rahmen einer Modellierung, sowie Implementierung in jeweils FZ 2.2/TB und FZ 2.3/I angesprochen und untersucht.

## 2.2.2 Integrationsmöglichkeiten von Systemen generativer KI

Dieser Abschnitt befasst sich mit den Möglichkeiten der Integration der in Abschnitt 2.1.5 ausgewählten Systeme generativer KI. Dies sind in erster Linie die Systeme, GPT-3.5/4 und Dall-E, die von OpenAI entwickelt wurden. OpenAI bietet zur Übersicht eine Plattform für seine Produkte an. Diese Plattform umfasst neben den Produkten auch Anleitungen, Beispiele und Tipps zu bewährten Praktiken zu diesen Produkten und beschreibt eine Reihe an Möglichkeiten zur Integration seiner Produkte. Diese Integrationsmöglichkeiten umfassen:

- Eine Schnittstelle (API), die es Entwicklern ermöglicht Modelle in ihre eigene entwickelte Software zu integrieren und aufzurufen.
- Unterstützte Systeme zur Softwareentwicklung (SDKs) in vielen Programmiersprachen, die u.a von OpenAI bereitgestellt werden, oder von Mitgliedern der Community in anderen, nicht offiziell von OpenAI unterstützten Programmiersprachen entwickelt wurden [111].
- Online-Dienste, die es Entwicklern ermöglichen Modelle ohne eigens entwickelte Software zu verwenden. Beispiele für diese Online-Dienste sind ChatGPT, Dall-E, Tokenizer oder ein Textklassifikator für KI.
- Cloud-Dienste, die es Entwicklern ermöglichen Modelle ohne eigene Infrastruktur zu verwenden [112].

Im nächsten Abschnitt wird gesondert auf die Schnittstelle und die Systeme zur Softwareentwicklung eingegangen. Im darauffolgenden Abschnitt 2.2.2 werden dann weitere interessante Werkzeuge im Rahmen der Schnittstelle von OpenAI besprochen und Aspekte hervorgehoben.

### ***API***

Eine Schnittstelle (engl. **Application Programming Interface**) ermöglicht Entwicklern die Kommunikation zwischen Softwareanwendungen umzusetzen. Eine API definiert und stellt Funktionen bereit, die Entwickler verwenden können, um Softwareanwendungen miteinander zu verbinden. Funktionen können das Auslesen und Schreiben von Daten, oder Ausführen von Aktionen in einer anderen Anwendung sein. Auf diese Weise ermöglicht eine Schnittstelle den Zugriff auf Funktionen einer Anwendung, ohne dass diese in einer Anwendung selbst neu geschrieben bzw. programmiert werden müssen.

Die API von OpenAI kann praktisch auf jede Aufgabe angewendet werden, die das Verstehen oder Erzeugen von natürlicher Sprache betrifft und erfordert [113]. Des Weiteren kann die API auch verwendet werden, um Bilder zu erzeugen, zu bearbeiten oder Sprache in Text zu konvertieren [113]. Die API von OpenAI bietet

Zugriff auf eine Reihe von Modellen, die unterschiedliche Fähigkeiten und Preise aufweisen [113].

Die Schnittstelle von OpenAI wird durch eine RESTful (Representational State Transfer) Schnittstelle bereitgestellt. REST-Schnittstellen sind eine der prominentesten Arten von APIs [114], die eine Reihe an Konventionen verwenden, um die Verarbeitung und Übertragung von Informationen zu gewährleisten. Diese RESTful Schnittstelle wird spezifiziert über eine OpenAPI Spezifikationen [115], die die Funktionen der OpenAI Schnittstelle genaustens dokumentiert. OpenAPI ist eine auf dem Architekturstil REST basierende SpezifikationsSprache für die Beschreibung der Struktur, der Reihenfolge und des Verhaltens von APIs. Diese Spezifikation wird für gewöhnlich in YAML oder JSON geschrieben und kann, neben der Dokumentation einer API, auch dazu verwendet werden, um mit spezialisierten Programmen, wie z.B. Swagger Codegen, automatisch Quellcode in einer spezifischen Programmiersprache zu generieren. Für die Programmiersprache Java wird diese Spezifikation bereits durch die SDK bzw. Bibliothek OpenAI-Java [96] abgedeckt.

Die API von OpenAI bietet Zugriff auf folgende Endpunkte:

1. /v1/completions
2. /v1/chat/completions
3. /v1/edits
4. /v1/images/generations
5. /v1/images/edits
6. /v1/images/variations
7. /v1/embeddings
8. /v1/audio/transcriptions
9. /v1/audio/translations
10. /v1/files
11. /v1/fine-tunes
12. /v1/moderations

Im Rahmen dieser Arbeit sind nur die Endpunkte 2 und 4 zur Generierung von Text bzw. Bildern von Interesse. In [95] wird eine genaue Dokumentation, mitsamt Beispielen für diese Endpunkte beschrieben.

Eine Anfrage an den Endpunkt **/chat/completions/** nimmt als Eingabe eine Liste an Nachrichten in Form einer Unterhaltung zwischen verschiedenen Parteien. Eine

Nachricht besteht dabei aus einem Wertepaar **Rolle : Inhalt**. Beide Werte sind notwendige Parameter einer Anfrage an diesen Endpunkt. Der Parameter *Rolle* kann einen von drei Werten annehmen: *System, Assistant, User* [116]. Über den Endpunkt liefert das Modell eine Antwort auf diese Unterhaltung.

Eine Anfrage an den Endpunkt **/images/generations/** nimmt als Eingabe eine Eingabeaufforderung bzw. Prompt. Das Modell generiert anhand dieser Prompt ein oder mehrere Bilder und liefert diese über den Endpunkt als Antwort zurück.

Die in diesem Abschnitt vorgestellte Schnittstelle mit all ihren Endpunkten bietet bereits umfassende Möglichkeiten zur Generierung von Text und Bildern. Allerdings gibt es keine bestehende Integration dieser Systeme generativer KI im GMAF. Dies führt zu zweiten, offenen Herausforderung im Rahmen dieses Forschungsziels.

### Offene Herausforderung OH 2.2 - Integration generativer KI

Zum Zeitpunkt des Verfassens dieser Arbeit existiert keine Integration der von OpenAI angebotenen Systeme generativer KI im GMAF.

Die offene Herausforderung **OH 2.1** wird im Rahmen einer Modellierung, sowie Implementierung ebenfalls in jeweils FZ 2.2/TB und FZ 2.3/I angesprochen und untersucht.

### Tokenizer

Neben der Schnittstelle bietet OpenAI eine Reihe an Werkzeugen bzw. Online-Diensten an. Eines dieser Werkzeuge ist der Tokenizer. Die GPT-Modellreihe benutzt zum Verarbeiten von Text sogenannte Tokens [117]. Tokens stellen häufige im Text vorkommende Zeichenfolgen dar [117]. Die GPT-Modelle verstehen die statistischen Beziehungen zwischen diesen Token und sind anhand dieser Beziehungen und der bereits bestehenden Folge an Token im Text in der Lage das nächste Token im Text zu erzeugen [117]. Ein abgewandeltes Beispiel für eine Tokenaufteilung ist im folgenden Text einsehbar:

„Multiple models, each with different capabilities and price points. Prices are per 1,000 tokens. You can think of tokens as pieces of words, where 1,000 tokens is about 750 words. This paragraph is 49 tokens.“ Encoding: (*cl100k\_base*) [118]

Die Aufteilung der Zeichenfolgen in Tokens ist abhängig vom jeweiligen Modell und kann in [118] eingesehen werden. Die Aufteilung eines Textes in Tokens ist deshalb interessant, da durch dieses Werkzeug die Anzahl der Tokens bestimmt werden kann. In Kombination dieser Information und der Information zur Bepreisung kann ein potentielles Optimierungspotenzial in Bezug auf die Anzahl der Tokens und somit auch der Bepreisung ausgeschöpft werden.

Der Online-Dienst Tokenizer von OpenAI ist offiziell durch die für Python konzipierte Bibliothek „Tiktoken“ [119] oder inoffiziell durch die für Java konzipierte Bibliothek „JTokkit“ [120] als Schnittstelle für die Tokenisierung von Text verfügbar.

### 2.2.3 Diskussion

In diesem Abschnitt wurden die Möglichkeiten der Integration von Graph Codes, sowie den von OpenAI angebotenen Systemen generativer KI behandelt.

In Bezug auf die Integrationsmöglichkeiten von Graph Codes konnte eine Übersicht über die vom GMAF gebotenen Komponenten, die eine entscheidende Rolle in der Technik der Graph Codes spielen, zusammengestellt werden. Es wurde weiterhin bestimmt, dass die Darstellungsform der in Graph Codes enthaltenen Informationen im JSON-Format erfolgt. Diese Darstellungsform kann insbesondere von aktuellen Systemen generativer KI effizient und effektiv eingelesen werden.

In Bezug auf die Integrationsmöglichkeiten der von OpenAI entwickelten Systemen generativer KI konnte eine Übersicht über die von OpenAI bereitgestellte Schnittstelle gegeben werden. Diese Übersicht gibt auch Einblick über die verfügbaren Endpunkte und Einbindungsmöglichkeiten dieser Endpunkte über eine speziell für Java entwickelte Bibliothek bzw. SDK für eine einfache Integration in das GMAF. Des Weiteren wurde das Tool Tokenizer, welches im Umgang mit den Modellen eine wichtige Rolle spielen kann, genauer erklärt und in eine potentielle Nutzung eingeordnet.

### *Offene Herausforderungen*

In Abschnitt 2.2.1 wurde die Erkenntnis erlangt, dass Graph Codes zur Darstellung im bekannten Dateiübertragungsformat JSON dargestellt werden. Auch wenn die von OpenAI angebotenen Modelle in der Lage sind JSON zu verstehen und analysieren, so bleibt offen, wie eine Umwandlung von Graph Codes in eine geeignete Eingabeform für Systeme generativer KI erfolgen kann. Dies führt zur ersten, offenen Herausforderung **OH 2.1** im Rahmen dieses Forschungsziels. Es gilt daher eine geeignete Überführung der Graph Codes in Systeme generativer KI für die Umgebung „Integration generativer KI in das GMAF“ zu bestimmen.

Im weiteren Verlauf wurde in Abschnitt 2.2.2 eine Übersicht über die von OpenAI bereitgestellte Schnittstelle, mitsamt ihrer verfügbaren Endpunkte, und ihrer Einbindungsmöglichkeiten geschaffen. Es konnte festgestellt werden, dass das GMAF bisher keine Einbindung dieser Endpunkte über die Schnittstelle und folglich auch keinen Systeme generativer KI umsetzt. Dies führt zur zweiten, offenen Herausforderung **OH 2.2** im Rahmen dieses Forschungsziels. Es gilt daher die Funktionen der Schnittstelle zu den Endpunkten der in Abschnitt 2.1.5 ausgewählten Systeme

generativer KI für die Umgebung „Integration generativer KI in das GMAF“ zu integrieren.

Die beiden offenen Herausforderungen **OH 2.1** und **OH 2.2** werden im Rahmen einer Modellierung durch das Forschungsziel FZ 2.2/TB angesprochen und durch das Forschungsziel 2.3/I im Rahmen einer prototypischen Implementierung untersucht. Forschungsziel 2.4/E wird sich der Evaluierung der in diesen Kapiteln erlangten Forschungsergebnissen widmen.

## 2.3 Evaluierungsmethodiken

Die Evaluierung ist ein entscheidender Teil im Zyklus der Softwareentwicklung und dient der Beurteilung der Qualität, Korrektheit und Einhaltung von Standards und Vorschriften. In diesem Abschnitt werden eine Reihe bekannter Evaluierungsmethodiken zur systematischen Evaluierung in der Softwareentwicklung vorgestellt. Mit Hilfe dieser Evaluierungsmethodiken lassen sich besagte Bewertungen der im Rahmen dieser Arbeit implementierten prototypischen Software vornehmen.

Das IEEE [121] bietet mit der Norm IEEE 1028-2008 [122] Richtlinien zur Durchführung von Softwareüberprüfungen. Diese Softwareüberprüfungen sind, formal betrachtet, Evaluierungen von in der Entwicklungen entstandenen Software-Artifakten, um Defekte festzustellen, die Qualität zu verbessern und sicherzustellen, dass alle Vorschriften eingehalten werden. Die Norm 1028-2008 definiert mehrere Arten von Peer-Review-Verfahren, die im Zyklus der Softwareentwicklung Verwendung finden können. Ein Peer-Review-Verfahren ist ein Prozess, in welchem die Ergebnisse der Softwareentwicklung manuell von einem oder mehreren im selben Feld tätigen Experten bzw. Kollegen überprüft werden. Folgende Peer-Review-Verfahren sind in der Norm IEEE 1028-2008 definiert:

- Informal Review
- Technical Review
- Walkthrough
- Inspection
- Audit

Im Folgenden werden diese einzelnen Verfahren kurz erläutert und ihre Verwendung, der Ablauf, sowie die Vor- und Nachteile näher beschrieben.

**Informal Review** (z.Dt. informeller Review) ist ein einfaches und flexibles Verfahren, um in erster Instanz nach Defekten zu suchen und Rückmeldungen zu Software-Artifakten zu erhalten, ohne dabei aktiv Programmcode auszuführen. Es werden im Rahmen von informellen Reviews keine Dokumentationen angefertigt. Der Ablauf eines informellen Reviews sieht es vor, dass Prüfer die Dokumentation der Software oder ihren Quellcode individuell untersuchen und anhand ihrer Erfahrungen und Expertise Rückmeldungen liefern. Der Vorteil an informellen Reviews ist, dass sie einfach, schnell und kostengünstig umsetzbar sind. Des Weiteren kann dieses Verfahren bereits in den allerersten Schritten der Entwicklung durchgeführt werden, um frühzeitig Defekten vorzubeugen und Einblicke zu erlangen. Nachteilig an einem informellen Review ist der Mangel an einem strukturierten Vorgehen und mangelnder Gründlichkeit der Untersuchungen.

**Technical Review** (z.Dt. technischer Review) ist, verglichen mit einem informellen Review, ein detailliertes Verfahren zur Evaluierung von Software-Artifakten, um die Eignung des zu testenden Artifakts für die beabsichtigte Nutzung festzustellen. Schwerpunkte eines technischen Reviews liegen dabei im Einhalten von Standards und Normen, Korrektheit und Vollständigkeit der überprüften Artifakte. Der Ablauf eines technischen Reviews sieht es vor, dass Prüfer die Artifakte gründlich gegen die in den Vorgaben festgelegten Kriterien, wie Spezifikationen und Standards, prüfen und Abweichungen genaustens dokumentieren. Der Vorteil an technischen Reviews ist, dass sie durch eine eingehende Analyse und Bewertung der Software-Artifakte die Einhaltung von Spezifikationen, Normen und Standards garantieren und aufgrund ihrer gründlichen Natur technische Defekte identifizieren können.

**Walkthrough** (z.Dt. Durchlauf) ist ein Verfahren, in welchem die Vorgehensweise von einem Autor in einer Durchführung durch die Software beschrieben wird. Ziel dieser Durchführung ist es, die Beteiligten mit den Software-Artifakten vertraut zu machen, um Rückmeldungen und Klarstellungen von ihnen zu erhalten. Der oder die Autor(en) der Artifakte demonstrieren diese den Teilnehmern des Durchlaufs, um Rückmeldungen über die vorgestellten Artifakte zu erhalten. Teilnehmer dieses Durchlaufs sind für gewöhnlich eine andere Partei, wie z.B. Arbeitskollegen. Rückmeldungen über die vorgestellten Artifakte enthalten zumeist Fragen, Kommentare oder Verbesserungsvorschläge zu den wahrgenommenen Abweichungen oder Mängeln in der Software im Vergleich zum beschriebenen Ablauf. Der Vorteil an einem Durchlauf ist, dass er die Zusammenarbeit und den Austausch von Wissen zwischen den Beteiligten fördert. Mit Hilfe eines Durchlaufs können Fehler identifiziert, das Verständnis verbessert und unterschiedliche Perspektiven gesammelt werden. Es wird somit die Gültigkeit des beschriebenen Ablaufs bzw. Lösungsvorschlags für die Software geprüft. Eine weit verbreitete Variante des Walkthroughs ist der **Cognitive Walkthrough**. Im Vergleich zu einem simplen Durchlauf ist ein Cognitive Walkthrough speziell auf die Bewertung der Benutzerfreundlichkeit und Benutzererfahrung ausgelegt. Ein simpler Durchlauf kann eher als allgemeines Verfahren zur Prüfung von Software verstanden werden, welches verschiedene Aspekte der Softwareentwicklung umfasst. An einem Cognitive Walkthrough sind für gewöhnlich sogenannte **Usability-Experten** beteiligt. Aufgabe der Usability-Experten ist es, die Perspektive eines potentiellen Nutzers einzunehmen, während in einem simplen Durchlauf in der Regel nur Teammitglieder oder andere Interessensgruppen beteiligt sind, die anhand ihres Fachwissens Rückmeldungen und Erkenntnisse liefern.

**Inspection** ist ein streng formales Verfahren, welches die Eignung bzw. Beschaffenheit eines Arbeitsergebnisses für die beabsichtigte Verwendung untersucht. Das Verfahren hat zum Ziel Defekte in Software-Artifakten zu identifizieren und die Qualität der Software-Artifakte zu garantieren, sowie zu verbessern [122]. Der Ablauf einer Inspection sieht vor, dass ein Team, welches von einem Moderator geleitet wird, die

Software-Artifakte gründlich und systematisch gegen ihre vorgegebenen Standards und Ansprüche prüft [122]. Defekte und Widersprüche werden festgestellt und klar dokumentiert [122]. Der Vorteil an einer Inspection ist, dass sie einen klaren strukturierten und disziplinierten Ansatz zur Fehlererkennung verfolgen. Weiterhin werden im Rahmen einer Inspection die Untersuchungen von ausgebildeten Inspektoren vorgenommen [122]. Diese Eigenschaften machen eine Inspection zu einem wirksamen Ansatz zur Fehlerentdeckung und Verbesserung der Gesamtqualität von Software-Artifakten.

**Audit** ist ein Verfahren, das umfassende Bewertungen von Software-Artifakten zum Ziel hat, um Übereinstimmungen bzw. Diskrepanzen mit Anforderungen, Vorschriften oder Standards zu beurteilen. Der Ablauf eines Audits sieht vor, dass Auditoren Software-Artifakte und Dokumentationen prüfen und diese mit den vorgegebenen Anforderungen vergleichen. Abweichungen werden festgestellt und klar dokumentiert. Der Vorteil an einem Audit ist, dass durch umfassende Bewertungen von Software-Artifakten die Einhaltung von Vorschriften garantiert wird und zudem verbesslungsfähige Bereiche aufzeigt werden.

Nachteile der genannten Peer-Review-Verfahren sind vor allem zumeist die Zuweisung von Ressourcen, wie z.B. von Zeit, der Definition des genauen Vorgehens, potentielle Voreingenommenheiten bzw. Dynamiken zwischen Prüfern und Teilnehmern, sofern sie aus einer Organisation kommen und nicht unabhängig voneinander arbeiten und begrenzter Fachkenntnisse und Expertise von Prüfern.

Die in diesem Abschnitt vorgestellten Verfahren bzw. Methodiken bieten einen Überblick über eine Bandbreite an Möglichkeiten zur Evaluierung von Software. Eine Auswahl aus diesen Evaluierungsmethodiken wird am Anfang in Kapitel 5 vorgenommen.

## 2.4 Zusammenfassung

In diesem Kapitel wurde die Recherche zum bisherigen Stand der Wissenschaft und Technik zur Erklärbarkeit und Visualisierung von Graph Codes mittels generativer KI behandelt. In Abschnitt 2.4.1 werden die noch offenen und verbleibenden Herausforderungen zusammengefasst. Weiter werden in Abschnitt 2.4.2 die wesentlichen Ergebnisse und Erkenntnisse dieser Recherche nach den Forschungszielen der Beobachtungsphase, wie nach der Methodik von Nunamaker vorgesehen, aufgeschlüsselt und kurz zusammengefasst. In Abschnitt 2.4.3 werden die Entscheidungen der Diskussionen der behandelten Forschungsziele aufgezählt. Schlussendlich wird in Abschnitt 2.4.4 die Behandlung der noch offenen Herausforderungen angesprochen. Schlussendlich gibt Tab. 2.4 einen Gesamtübersicht über den aktuellen Stand der Arbeit.

### 2.4.1 Welche offenen Herausforderungen (OH) verbleiben?

In diesem Abschnitt werden die noch offenen und verbleibenden Herausforderungen, die sich im Rahmen der Recherche zum Stand der Wissenschaft und Technik ergeben haben, aufgezählt, kurz beschrieben und jeweils Verortungen der Lösungsansätze genannt.

Tabelle 2.3: Offene Herausforderungen

Nummer der offenen Herausforderungen	Inhalt der offenen Herausforderungen	Ansatz für Lösung
<b>OH 1.1</b>	Offen bleibt, wie durch den Einsatz von generativer KI Erklärbarkeit erreicht werden kann.	Modellierung, Implementierung
<b>OH 1.2</b>	Offen bleibt, wie Benutzerschnittstellen beschaffen sein müssen, um geeignete Interaktionsmöglichkeiten mit Systemen generativer KI für Anwender des GMAF zu bieten.	Modellierung, Implementierung
<b>OH 2.1</b>	Offen bleibt, wie die Darstellungsform von Graph Codes in eine geeignete Eingabeform für Systeme generativer KI überführt werden kann.	Modellierung, Implementierung

Nummer der offenen Herausforde- rungen	Inhalt der offenen Herausforderungen	Ansatz für Lösung
<b>OH 2.2</b>	Offen bleibt, wie die Funktionen der Schnittstellen und Endpunkte der vorgestellten und ausgewählten Systeme generativer KI in das GMAF integriert werden können.	Modellierung, Implementie- rung

#### 2.4.2 Gewonnene Erkenntnisse

In diesem Abschnitt werden die Erkenntnisse aus den Forschungszielen „FZ 1.1/O Erklärbarkeit von MMIR mittels generativer KI“ und „FZ 2.1/O Integration generativer KI in das GMAF“ zusammengefasst.

Im ersten Forschungsziel wurde die Erkenntnis erzielt, dass die bisherige Methodik der Erklärbarkeit von MMFGs bzw. Graph Codes sehr statisch, statistisch und rein mathematisch ist. Dies führt zur ersten offenen Herausforderung **OH 1.1**. Ein alternativer Ansatz zur Erklärbarkeit wird im Einsatz generativer KI gesehen. Nach der Auswahl aus vorgestellten Systemen wurde festgestellt, dass es keine Interaktionsmöglichkeiten für Anwender des GMAF mit diesen Systemen gibt. Dies führt zur zweiten offenen Herausforderung **OH 1.2**.

Im Rahmen des zweiten Forschungsziels wurden die Komponenten, die für die Technologie der Graph Codes eine entscheidende Rolle spielen, identifiziert. Weitere Erkenntnisse umfassen die Dateiübertragungsform der Graph Codes. Es wurde festgestellt, dass obwohl aktuelle Modelle diese Dateiübertragungsform einlesen und analysieren können, es offen bleibt, wie Graph Codes in eine geeignete Eingabeform für diese Systeme überführt werden können. Dies führt zur ersten offenen Herausforderung **OH 2.1**. Eine genaue Beschreibung der angebotenen Schnittstelle Durch eine genaue Beschreibung der von OpenAI angebotenen Schnittstelle konnte Erkenntnis über die von dieser Schnittstelle bereitgestellten Endpunkte für entsprechende Modelle bzw. Systeme erlangt werden. Nach der Beschreibung der Funktionen dieser Endpunkte konnte festgestellt werden, dass das GMAF keine Funktionen dieser Systeme einbindet. Dies führt zur zweiten offenen Herausforderung **OH 2.2**. Des Weiteren konnte festgestellt werden, dass OpenAI eine Reihe an Tools, darunter auch das Tool „Tokenizer“ anbietet, welches eine wichtige Rolle in der Verarbeitung spielen könnte. Erkenntnis hier ist, dass dieses Tool einen besseren Einblick bzw. mehr Informationen zur Verarbeitung von Tokens bietet.

### 2.4.3 Getroffene Entscheidungen

In diesem Abschnitt werden die Entscheidungen die in den Forschungszielen „FZ 1.1/O Erklärbarkeit von MMIR mittels generativer KI“ und „FZ 2.1/O Integration generativer KI in das GMAF“ getroffen wurden, rekapituliert.

Im Rahmen des Forschungsziels FZ 1.1/O wurde eine Auswahl aus den vorgestellten Systemen generativer KI getroffen. Die Auswahl fällt auf die von OpenAI angebotenen Modelle bzw. Systeme generativer KI. Diese Entscheidung liegt in der Tatsache begründet, dass diese Modelle beide Aspekte der Text- und Bildgenerierung abdecken und diese durch einen einzigen Anbieter vertreten werden und welcher eine einfache und flexible Zahlung ermöglicht.

Da im Rahmen des Forschungsziels nur die Integrationsmöglichkeiten von Graph Codes, sowie von Systemen generativer KI dargestellt wurden, wurde keine spezielle Entscheidung getroffen. Erwähnenswert bleibt jedoch das in Abschnitt 2.2.2.2 vorgestellte Tool Tokenizer, welches im Rahmen der Verarbeitung von Prompts eine wichtige Rolle spielen kann.

### 2.4.4 Behandlung offener Herausforderungen

Dieser Abschnitt benennt die Möglichkeiten zur Behandlung der in Tab. 2.3 genannten offenen Herausforderungen. Der Methodik nach Nunamaker folgend, wird in Kapitel 3 die Modellierung und in Kapitel 4 die Implementierung folgen. Hierbei werden die offenen Herausforderungen **OH 1.1** und **OH 1.2** im Rahmen des Forschungsziels „FZ 1.2/TB Erklärbarkeit von MMIR mittels generativer KI“ adressiert.

Die erste offene Herausforderung **OH 2.1** wird im Rahmen des Forschungsziels „FZ 2.2/TB Integration generativer KI in das GMAF“ adressiert und zuletzt wird die zweite offene Herausforderung **OH 2.2** im Rahmen des Forschungsziels „FZ 2.3/I Integration generativer KI in das GMAF“ adressiert.

Schlussendlich werden in der nachfolgenden Tabelle die Erkenntnisse dieses Kapitels in den aktuellen Stand der Arbeit eingeordnet. Diese Tabelle ist als eine Checkliste zu verstehen, die sich im weiteren Verlauf dieser Arbeit, schrittweise, Kapitel für Kapitel, füllt und den Fortschritt der Arbeit wiederspiegelt.

Tabelle 2.4: Tabelle zur Übersicht des aktuellen Arbeitstands.

<b>Forschungsziele</b>		
<b>FZ / OH</b>	<b>Beschreibung</b>	<b>Referenz</b>
<b>Erklärbarkeit von MMIR mittels generativer KI</b>		
FZ 1.1/O	Recherche zur Erklärbarkeit von MMIR mittels generativer KI Grundlegende Technologien: <ul style="list-style-type: none"> <li>• GMAF</li> <li>• MMFG</li> <li>• Graph Code</li> </ul>	Abschnitt 2.1.1 Abschnitt 2.1.2 Abschnitt 2.1.3
OH 1.1	Erste offene Herausforderung Systeme generativer KI und ein Überblick über aktuelle Systeme Diskussion und Auswahl von Systemen	<b>OH 1.1</b> Abschnitt 2.1.4
OH 1.2	Zweite offene Herausforderung	<b>OH 2.1</b> Abschnitt 2.1.5
FZ 1.2/TB	Modellierung der Erklärbarkeit von MMIR mittels generativer KI	
FZ 1.3/I	Implementierung der Erklärbarkeit von MMIR mittels generativer KI	
FZ 1.4/E	Evaluierung der Erklärbarkeit von MMIR mittels generativer KI	
<b>Integration generativer KI in das GMAF</b>		
FZ 2.1/O	Recherche zur Integration generativer KI in das GMAF Aufzeigen der Integrationsmöglichkeiten von: <ul style="list-style-type: none"> <li>• Graph Codes</li> </ul>	Abschnitt 2.2.1
	Erste offene Herausforderung <ul style="list-style-type: none"> <li>• Systemen generativer KI</li> </ul>	<b>OH 2.1</b>
	Zweite offene Herausforderung	Abschnitt 2.2.2 <b>OH 2.2</b>
FZ 2.2/TB	Modellierung der Integration generativer KI in das GMAF	
FZ 2.3/I	Implementierung der Integration generativer KI in das GMAF	

<b>Forschungsziele</b>		
<b>FZ / OH</b>	<b>Beschreibung</b>	<b>Referenz</b>
FZ 2.4/E	Evaluierung der Integration generativer KI in das GMAF	

### 3 Modellierung

Dieses Kapitel widmet sich der konzeptuellen Modellierung und es werden, entsprechend der Methodik nach Nunamaker, alle Forschungsziele der Theoriebildungsphase behandelt. Die Struktur dieses Kapitels folgt der logischen Abfolge der Forschungsziele. Am Anfang wird in Abschnitt 3.1 die Konzeptions- und Modellierungsmeethodik vorgestellt und beschrieben. In Abschnitt 3.2 wird das Forschungsziel „FZ 1.2/TB Erklärbarkeit von MMIR mittels generativer KI“ behandelt und es werden Möglichkeiten der Erklärbarkeit durch generative KI angesprochen, sowie Anwendungsfälle von Benutzern als Basis für weitere Modellierungsvorgehen identifiziert. In Abschnitt 3.3 wird das Forschungsziel „FZ 2.2/TB Integration generativer KI in das GMAF“ behandelt und es werden die technischen Aspekte der Integration von Systemen generativer KI besprochen. Dies umfasst die Überführung von Graph Codes in eine passende Form der Eingabe für Systeme generativer KI und verwandte Aspekte, sowie die Integration der verfügbaren Endpunkte der von den ausgewählten Systemen angebotenen Schnittstelle. Am Ende jedes Forschungsziels werden die Erkenntnisse aus diesen Bereichen diskutiert und das weitere Vorgehen festgehalten. Abschließend wird in Abschnitt 3.4 eine Zusammenfassung der in diesem Kapitel erarbeiteten Ergebnisse gegeben.

Tabelle 3.1: Struktur dieses Kapitels.

FZ	Kurze Beschreibung	Abschnitt
Konzeptions- und Modellierungsmeethodik	Konzeptuelle Gestaltung Modellierung	Abschnitt 3.1
FZ 1.2/TB	Modellierung: Erklärbarkeit & gen. KI	Abschnitt 3.2
FZ 2.2/TB	Modellierung: Integration generativer KI	Abschnitt 3.3
Zusammenfassung	Ergebnisse / Erkenntnisse	Abschnitt 3.4

#### 3.1 Konzeptions- und Modellierungsmethodik

Die konzeptuelle Gestaltung und Modellierung in diesem Kapitel folgt dem Paradigma „User Centered System Design“ [26] von Norman & Draper. Ziel dieses Paradigmas ist eine Modellierung aus der Perspektive eines Benutzers oder eines Anwendungsfalls. Die konzeptuelle Gestaltung folgt dem Vorgehen des Rational Unified Process (RUP) [123] und wird mittels der Unified Modeling Language beschrieben [124].

### 3.1.1 Rational Unified Process

Der Rational Unified Process (RUP) ist ein für die Nutzung der Modellierungssprache Unified Modeling Language (UML) beschriebenes Vorgehensmodell bzw. Phasenmodell für die Anwendungsentwicklung [124]. Das Vorgehensmodell des RUP wird dabei in UML beschrieben [124]. Der Arbeitsablauf im RUP-Vorgehensmodell leitet sich aus dem evolutionären bzw. inkrementellen Vorgehen ab und lässt sich in vier zeitlich geordnete Phasen und Iterationen jener aufteilen: *Planung*, *Ausarbeitung*, *Errichtung*, *Übergang* [124]. Die Struktur der Phasen und der Iterationen im RUP-Vorgehensmodell werden durch sechs Kernarbeitsprozesse abgedeckt [124]:

- Geschäfts- und Modellierungskonzept
- Anforderungen
- Analyse & Design
- Implementierung
- Test
- Inbetriebnahme

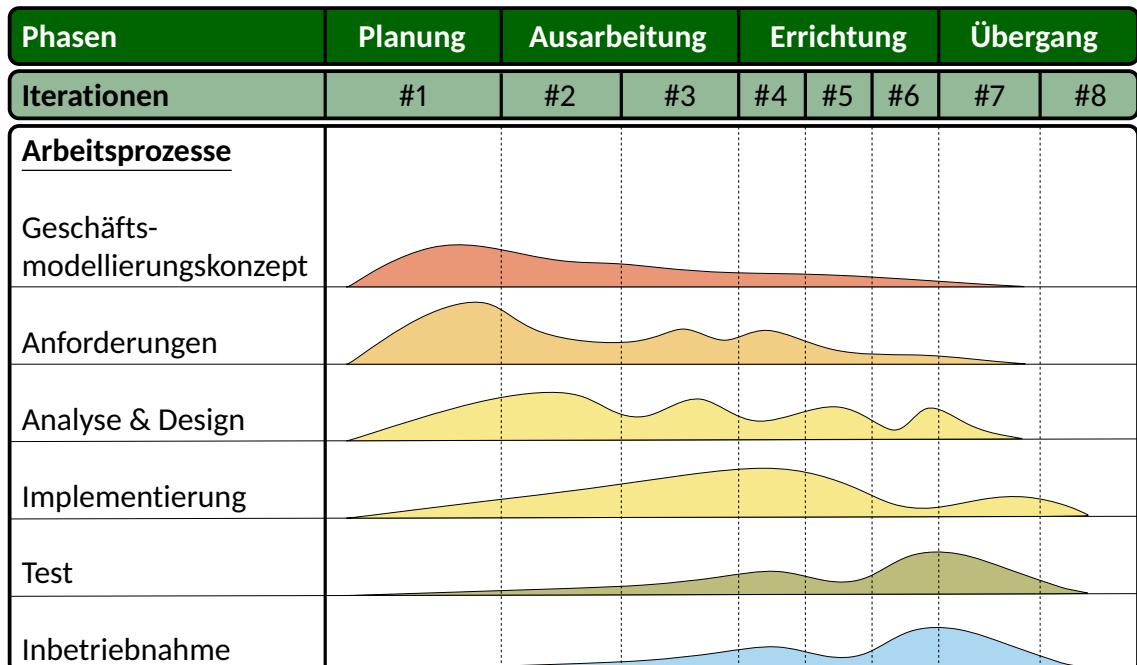


Abb. 3.1: Modellierung zum Vorgehen des Rational Unified Process.

Die Kern-Arbeitsprozesse „Geschäfts-Modellierungskonzept“, „Anforderungen“ und „Analyse & Gestaltung“ sind Teil dieses Kapitels und setzen die Methodik für die

Bearbeitung der Forschungsziele der Theoriebildungsphase nach Nunamaker um. Der Arbeitsprozess „Implementierung“ des RUP-Vorgehensmodells entspricht den Forschungszielen der Implementierungsphase nach Nunamaker und wird in Kapitel 4 behandelt. Der Arbeitsprozess „Test“ des RUP-Vorgehensmodells entspricht den Forschungszielen der Experimentphase nach Nunamaker und wird in Kapitel 5 behandelt. Der Arbeitsprozess „Inbetriebnahme“ wird in dieser Arbeit nicht behandelt. Somit wird jeder konzeptionelle Entwurf und jede Modellierung aus der Perspektive eines Benutzers oder eines Anwendungsfalls und auf Grundlage der Arbeitsprozesse des RUP-Vorgehensmodells entwickelt und im weiteren Verlauf dieser Arbeit unter Verwendung der UML formalisiert. Dabei ist es Ziel der Modellierung wiederverwendbare Softwarekomponenten zu erstellen, die erweiterbar, konfigurierbar sowie wartbar sind. Die Auswahl der Konzeptions- und Modellierungsmethodiken erfordert eine spezielle Anordnung der Forschungsziele. Diese Anordnung sieht wie folgt aus:

- FZ 1.2/TB Erklärbarkeit von MMIR mittels generativer KI
- FZ 2.2/TB Integration generativer KI in das GMAF

Die Notwendigkeit dieser Anordnung ergibt sich nach der Methodik von Norman & Draper, in welcher die Modellierung aus der Perspektive eines Benutzers oder eines Anwendungsfalls beginnt. Des Weiteren bauen die ersten Kern-Arbeitsprozesse „Geschäfts-Modellierungskonzept“ und „Anforderungen“ des RUP-Vorgehensmodells auf Anwendungsfällen auf. Somit ist logischer Startpunkt die Erklärbarkeit von MMIR mittels generativer KI.

### 3.1.2 Unified Modeling Language

Die Unified Modeling Language (UML) [125] ist eine standardisierte, visuelle Modellierungssprache, die ihre Verwendung in der Softwareentwicklung findet [126]. UML kann allerdings auch in anderen technischen Bereichen, oder der Geschäftsmodellierung verwendet werden [127]. Die UML hat zum Ziel, komplexe Systeme und deren Prozesse grafisch darzustellen, zu analysieren, zu spezifizieren und zu dokumentieren [126]. Die UML wurde dazu entwickelt, um die Kommunikation zwischen Softwareentwicklern, Analysten, Designern und anderen involvierten Parteien bzw. Interessensgruppen zu erleichtern [126]. Um diese Kommunikation zu ermöglichen, bietet die UML eine Vielzahl von Diagrammtypen zur einheitlichen, verständlichen und flexiblen Darstellung von Systemstrukturen- und verhalten. Folgende Aufzählung umfasst die bekanntesten und am häufigsten verwendeten Diagrammtypen, sowie ihre jeweiligen Aufgaben.

- Klassendiagramm: Stellt die simple, statische Struktur eines Systems durch seine Klassen, Attribute, Methoden und deren Beziehungen untereinander dar.

- Anwendungsfalldiagramm: Stellt die Interaktion zwischen Akteuren, die Benutzer oder wiederum andere Systeme sein können, und dem System dar. Zweck ist die Darstellung der für die Akteure verfügbaren Funktionalitäten des Systems.
- Sequenzdiagramm: Stellt die Interaktion zwischen Objekten eines Systems in einem zeitlichen Verlauf dar. Interaktion geschieht durch das Senden von Nachrichten zwischen den Objekten.
- Aktivitätsdiagramm: Stellt den Ablauf von Aktivitäten und Prozessen dar, um den Kontrollfluss innerhalb eines Systems zu visualisieren.
- Zustandsdiagramm: Stellt den Lebenszyklus eines Objekts dar und zeigt auf, wie es auf Ereignisse und Zuständsänderungen reagiert.

Die wichtigste Eigenschaft der UML ist seine Unabhängigkeit zu spezifischen Technologien, wie z.B. Programmiersprachen [128]. Diese Eigenschaft ermöglicht die UML zur Modellierung verschiedenster Systeme in vielen unterschiedlichen Bereichen. Bekannte Beispiele sind Softwareanwendungen, Datenbanken und Geschäftsprozesse. Mittels eines wohl definierten, sowie gepflegten Standards, der durch die OMG [129] erfolgt und eine einheitliche Sprache für Kommunikation schafft, ermöglicht die UML eine nahtlose Zusammenarbeit zwischen verschiedenen involvierten Gruppen. Angesichts dieser Eigenschaft ist die Unified Modeling Language ein wertvolles Werkzeug für die Softwaremodellierung- und Entwicklung komplexer Systeme.

## 3.2 FZ 1.2/TB Erklärbarkeit von MMIR mittels generativer KI

Dieser Abschnitt befasst sich mit der Modellierung der „Erklärbarkeit von MMIR mittels generativer KI“. Dabei werden die in FZ 1.1/O identifizierten, offenen Herausforderungen **OH 1.1** und **OH 1.2** adressiert. Ziel dieses Abschnitts ist es Konzepte für den Einsatz von Systemen generativer KI für Erklärbarkeit hervorzu bringen und Anwendungsfälle zur Interaktion mit selbigen zu identifizieren. Hierzu wird in Abschnitt 3.2.1 die offene Herausforderung **OH 1.1** und in Abschnitt 3.2.2 die offene Herausforderung **OH 1.2** adressiert.

### 3.2.1 Erklärbarkeit durch generative KI

In diesem Abschnitt wird die erste offene Herausforderung **OH 1.1** adressiert und es werden mögliche Techniken zum Erzeugen von Erklärungen durch generative KI angesprochen. In Abschnitt 2.1.4.1 wurde das Konzept des Prompt Engineering erklärt und eine Reihe an Techniken, sowie dazugehörige Beispiele präsentiert. Um mit Hilfe von Systemen generativer KI Erklärungen generieren zu können, müssen die Eingaben in ein System möglichst präzise formuliert sein. Ziel muss es daher sein, mit gezielten und präzise formulierten Eingabeaufforderungen das System so zu instruieren, dass es möglichst genaue Erklärungen generieren kann. In [67] wurde bereits gezeigt, dass das recht simple Textanhängsel „*Lets think step by step.*“ einen großen Einfluss auf die Schlussfolgerungen und somit auch die Ausgabe einer generativen KI haben kann. Es ist daher nicht abwegig anzunehmen, dass eine simple Instruktion, wie z.B. „*Create a coherent textual explanation.*“ oder „*Create a visual explanation.*“ einen positiven Einfluss auf erzeugte Texte bzw. Erklärungen haben könnte. Eine weitere interessante Idee Erklärungen generieren zu lassen, wäre die Methode des Geschichtenerzählens. Eine beispielhafte Instruktion könnte sein: „*Make up a story.*“. Weitere „*elementare Bausteine*“ einer Instruktion können dann beispielhafte Erklärungen umfassen, um dem System zu zeigen, welche Art von Erklärung letztlich gewünscht wird, oder die Vorgabe eines unvollständigen Satzes, welcher dann vom System komplementiert wird.

### 3.2.2 Anwendungsfälle

Ein sinnvoller Einsatz von Systemen generativer KI im GMAF zur Erklärung von Graph Codes setzt eine Benutzerschnittstelle mit geeigneten Interaktionsmöglichkeiten voraus. In diesem Abschnitt wird die zweite offene Herausforderung **OH 1.2** adressiert und es werden eine Reihe an Anwendungsfällen aus der Perspektive von Benutzern bzw. Anwenden des GMAF identifiziert und beschrieben. Abb. 3.2 zeigt ein UML-Diagramm zur Übersicht der identifizierten Anwendungsfälle. Diese identifizierten Anwendungsfälle stellen die Basis für weitere Modellierungsvorgehen dar

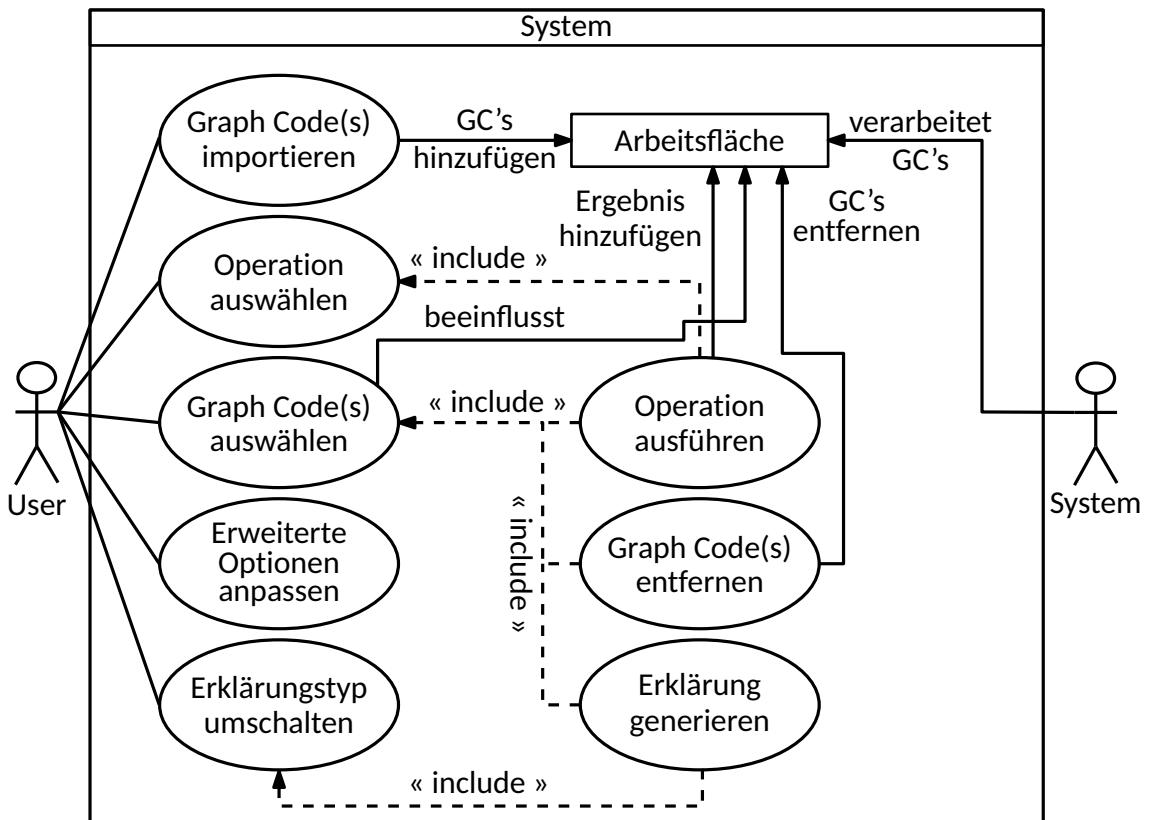


Abb. 3.2: Übersicht über alle Anwendungsfälle.

und werden im Laufe dieses Abschnitts weiter ausgeführt. Weitere Ausführungen umfassen textuelle Beschreibungen, erste Wireframes, die das allgemeine Skelett der Anwendung darstellen sollen, Mechanismen, die die an Anwendungsfällen beteiligten Komponenten identifizieren, sowie Sequenzdiagramme um die Interaktion zwischen den Objekten bzw. Komponenten eiens Systems darzustellen. Anhand dieser Modellierungsvorgehen werden die Wireframes weiter ausgeführt bzw. ausgebaut.

### ***Textuelle Beschreibungen***

In diesem Abschnitt werden die in Abb. 3.2 dargestellten Anwendungsfälle detailliert beschrieben. Diese detaillierte Beschreibung umfasst jeweils eine Beschreibung der Aufgabe des Anwendungsfalles, die Akteure, die an diesem Anwendungsfall beteiligt sind, die Vorbedingungen, die vor dem Anwendungsfall gelten, einen Ablauf an Schritten zum Durchführen des Anwendungsfalls und schlussendlich die Nachbedingungen, die jeweils nach dem Anwendungsfall gelten.

### UC-1.1 Graph Code(s) importieren

**Beschreibung:** Benutzer klicken in der Arbeitsfläche einen Knopf „Select Graph Code(s)“. Daraufhin öffnet sich ein Filechooser, in welchem Benutzer eine Auswahl von einem oder mehreren Graph Code Datei(en) treffen können. Die ausgewählten Graph Code Datei(en) werden einer Liste in der Arbeitsfläche hinzugefügt und angezeigt.

**Akteure:** Benutzer, System

**Vorbedingung:** Keine.

**Ablauf:**

1. System zeigt eine Liste in der Arbeitsfläche an.
2. Benutzer klickt Knopf „Select Graph Code(s).“
3. System zeigt Filechooser an.
4. Benutzer wählt ein oder mehrere Graph Code Datei(en) aus.
5. System fügt Graph Code Datei(en) der Liste in der Arbeitsfläche zu.

**Nachbedingung:** Graph Code Datei(en) sind der Liste in der Arbeitsfläche hinzugefügt worden.

### UC-1.2 Graph Code(s) entfernen

**Beschreibung:** Benutzer wählen aus der Liste in der Arbeitsfläche ein oder mehrere Graph Code Datei(en) aus und können über einen Knopf „Remove selected Graph Code(s)“ diese Graph Codes aus der Liste und somit der Arbeitsfläche entfernen.

**Akteure:** Benutzer, System

**Vorbedingung:**

- Liste enthält ein oder mehrere Graph Code(s).
- Benutzer haben ein oder mehrere Graph Code(s) ausgewählt.

**Ablauf:**

1. System zeigt eine Liste an Graph Codes in der Arbeitsfläche an.
2. Benutzer wählt ein oder mehrere Graph Code(s) aus.
3. Benutzer klickt auf Knopf „Remove selected Graph Code(s).“

4. System entfernt ausgewählte Graph Codes aus der Liste in der Arbeitsfläche.

**Nachbedingung:** Ausgewählte Graph Code Datei(en) sind aus der Liste in der Arbeitsfläche entfernt worden.

### UC-1.3 Graph Code(s) auswählen

**Beschreibung:** Benutzer wählen aus der Liste in der Arbeitsfläche ein oder mehrere Graph Code Datei(en) aus. Die Auswahl von Graph Code Dateien dient als Grundlage für Anwendungsfälle, wie UC-1.2, UC-1.5 oder UC-1.8.

**Akteure:** Benutzer, System

**Vorbedingung:** Liste enthält ein oder mehrere Graph Code(s) zum Auswählen.

**Ablauf:**

1. System zeigt eine Liste an Graph Code(s) in der Arbeitsfläche an.
2. Benutzer wählen ein oder mehrere Graph Code(s) an.

**Nachbedingung:** Keine.

### UC-1.4 Operation auswählen

**Beschreibung:** Benutzer wählen aus einem Feld eine auszuführende Operation aus. Verfügbare Optionen sind: Vereinigung, Subtraktion, Gemeinsamkeiten, Unterschiede. Das Auswählen einer Operation ist die Vorbedingung für den Anwendungsfall UC-1.5, dem Ausführen einer Operation.

**Akteure:** Benutzer, System

**Vorbedingung:** Keine.

**Ablauf:**

1. System bietet in einem Feld eine Reihe an auszuwählenden Operationen.
2. Benutzer wählt eine Operation aus.

**Nachbedingung:** Keine.

### UC-1.5 Operation ausführen

**Beschreibung:** Benutzer klicken auf den Knopf „Execute“, um die vorher ausgewählte Operation auf den in der Arbeitsfläche ausgewählten Graph Code Datei(en) auszuführen. Die vorher ausgewählte Operation wird daraufhin auf den ausgewählten Graph Code Datei(en) ausgeführt.

**Akteure:** Benutzer, System

**Vorbedingung:**

- Operation in der Arbeitsfläche ausgewählt.
- In der Arbeitsfläche wurden Graph Code Datei(en) ausgewählt.

**Ablauf:**

1. Benutzer klickt auf Knopf „Execute“.
2. System führt vorher ausgewählte Operation auf den ausgewählten Graph Code Datei(en) aus.

**Nachbedingung:** Erfolgreich ausgeführte Operation auf ausgewählten Graph Code(s).

### UC-1.6 Erklärungstyp umschalten

**Beschreibung:** Benutzer wählen in der Arbeitsfläche über Knöpfe „Image“ oder „Text“ den Typ der Erklärung aus. Anhand der ausgewählten Erklärung wird die Benutzerschnittstelle in der Arbeitsfläche für den jeweiligen Erklärungstyp umgeschaltet.

**Akteure:** Benutzer, System

**Vorbedingung:** Keine.

**Ablauf:**

1. Benutzer wählen über die Knöpfe „Image“ oder „Text“ den Typ der Erklärung aus.
2. System schaltet auf für Typ spezifische Benutzerschnittstelle um.

**Nachbedingung:** Benutzerschnittstelle für spezifischen Erklärungstyp umgeschaltet.

### UC-1.7 Erweiterte Optionen anpassen

**Beschreibung:** Benutzer passen in einem dafür vorgesehenen Feld erweiterte Optionen für den Endpunkt, der für die zu generierende Erklärung zuständig ist, an. Die anpassbaren Optionen sind abhängig vom jeweiligen Endpunkt bzw. Erklärungstypen.

**Akteure:** Benutzer, System

**Vorbedingung:** Keine.

**Ablauf:**

1. System bietet, abhängig vom Endpunkt bzw. Erklärungstypen, eine Reihe an anpassbaren Optionen an.
2. Benutzer passen Optionen nach eigenen Bedürfnissen an.

**Nachbedingung:** Keine.

### UC-1.8 Erklärung generieren

**Beschreibung:** Benutzer klicken auf einen Knopf „Generate ...“, um eine Erklärung für eine zuvor ausgewählte Graph Code Datei zu generieren. Die generierte Erklärung ist abhängig vom zuvor gewählten Endpunkt bzw. Erklärungstypen.

**Akteure:** Benutzer, System

**Vorbedingung:** Graph Code Datei wurde ausgewählt.

**Ablauf:**

1. Benutzer klickt auf Knopf „Generate ...“
2. System lässt durch Endpunkt Erklärung generieren.
3. System zeigt Erklärung in einem dafür vorgesehenen Bereich der Benutzerschnittstelle an.

**Nachbedingung:** Keine.

**Zweiglauf:**

- 2.a Endpunkt für „Image“ generiert eine visuelle Erklärung bzw. ein Bild.
- 2.b Endpunkt für „Text“ generiert eine textuelle Erklärung.

### **Wireframe für die Interaktion mit Graph Codes**

Auf Basis der Anwendungsfälle und den textuellen Beschreibungen dieser, wird in diesem Abschnitt ein Wireframe, sprich ein Konzept für eine Benutzerschnittstelle, vorgestellt. Abb. 3.3 zeigt eine erste Ausführung eines Wireframes, welches die Grundbereiche und somit das Skelett der Benutzerschnittstelle darstellen soll.

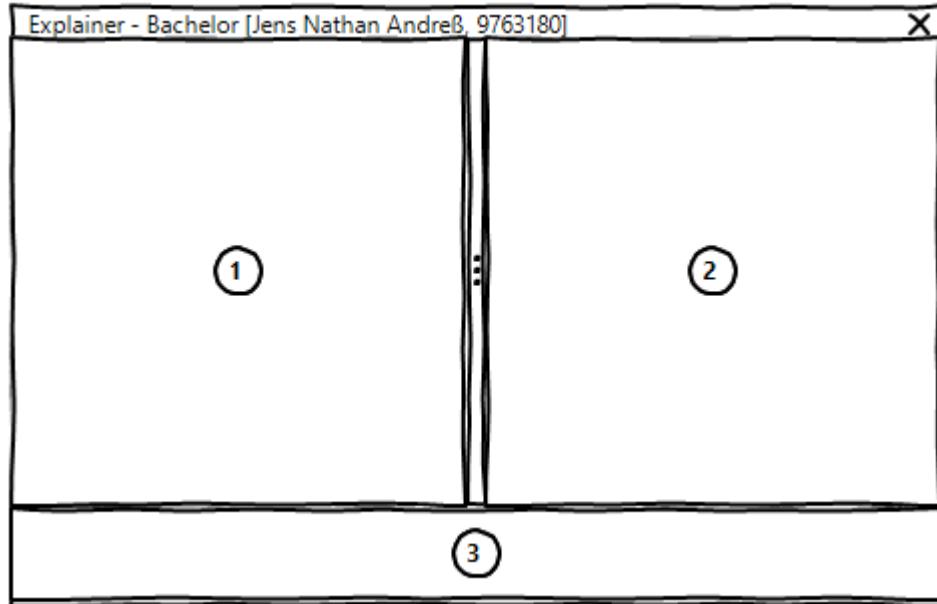


Abb. 3.3: Wireframe für die Grundbereiche der Benutzerschnittstelle.

In dieser Abbildung sind drei Grundbereiche zu erkennen, die im Folgenden genauer beschrieben werden:

- ① ist der Bereich der Arbeitsfläche, in der die Bearbeitung von Graph Code Dateien stattfindet. Bearbeitungen von Graph Codes umfassen das Importieren, Auswählen und Entfernen von Graph Code Dateien, sowie das Auswählen und Ausführen von Operationen auf Graph Code Dateien.
- ② ist der Bereich der Arbeitsfläche, in der Benutzer Erklärungen generieren können. Benutzer können hier zwischen unterschiedlichen Typen an Erklärungen hin und her schalten.
- ③ ist eine Konsole, in der alle wichtigen Informationen über Aktionen und Prozesse festgehalten werden.

Anhand der Anwendungsfälle UC-1.1 bis UC-1.5 können in der linken Arbeitsfläche, die der Bearbeitung von Graph Codes dienen soll, wichtige Komponenten in der Benutzerschnittstelle identifiziert werden. Analog können anhand der Anwendungsfälle

UC-1.6 bis UC-1.8 in der rechten Arbeitsfläche, die sich der Erklärung von Graph Code Dateien widmet, weitere wichtige Komponenten in der Benutzerschnittstelle identifiziert werden. Abb. 3.4 zeigt die linke und rechte Arbeitsfläche mit diese neuen Komponenten in jeweils einem partiellen Wireframe.

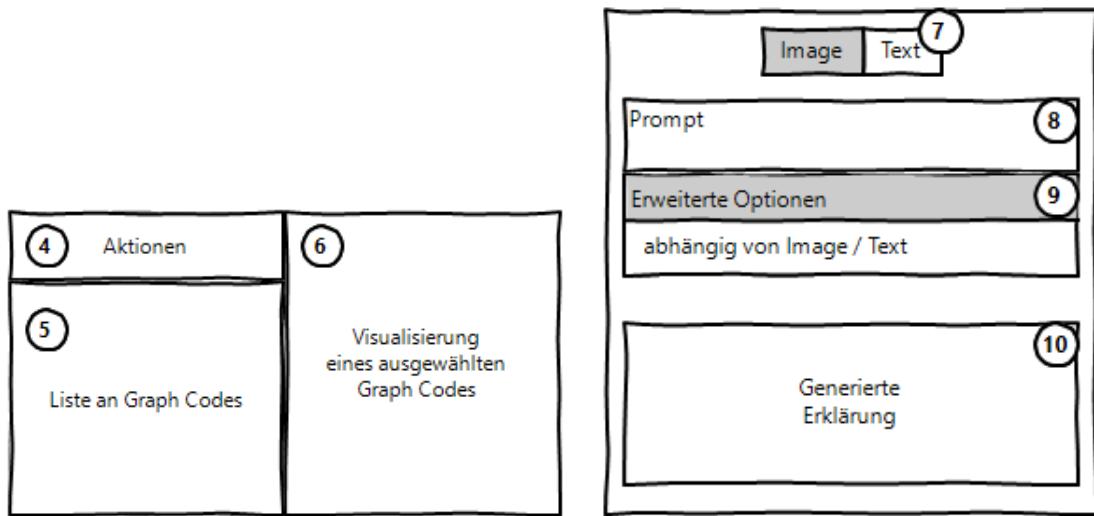


Abb. 3.4: Wireframe für den linken Arbeitsbereich (links) und den rechten Arbeitsbereich (rechts) der Benutzerschnittstelle.

Diese neuen Komponenten werden im Folgenden genauer beschrieben:

- (4) ist der Aktionsbereich, in welchem Interaktionsmöglichkeiten, wie das Importieren und Entfernen von Graph Codes in die Arbeitsfläche eingebunden werden. Des Weiteren werden in diesem Aktionsbereich auch Interaktionsmöglichkeiten für die Auswahl und das Ausführen von Operationen auf den zuvor in der Arbeitsfläche ausgewählten Graph Code Dateien eingebunden.
- (5) ist der Bereich in der Benutzerschnittstelle, in der die Graph Code Dateien in einer Liste dargestellt werden. Diese Liste ist die Voraussetzung für Anwendungsfälle UC-1.2, UC-1.3, UC-1.5 und UC-1.8.
- (6) ist der Bereich der Benutzerschnittstelle, in der ausgewählte Graph Code Datei(en) in einer geeigneten Form visualisiert werden. Da Graph Codes quadratische Matrizen sind, eignet sich für die visuelle Form der Darstellung eine Tabelle.
- (7) sind die Knöpfe, mit welchen Benutzer zwischen den Benutzerschnittstellen für die jeweiligen Erklärungstypen umschalten können.
- (8) ist die Prompt, die später über den entsprechenden Endpunkt in das System generativer KI übermittelt werden soll.

- (9) zeigt erweiterte anpassbare Optionen für den jeweiligen Endpunkt.
- (10) ist der Bereich, in welchem das Endergebnis der Anfrage an das System generativer KI dargestellt wird.

Zusammengeführt ergeben diese partiellen Wireframes ein ganzes Wireframe, welches in Abb. 3.5 abgebildet ist und die komplette Benutzerschnittstelle darstellt.

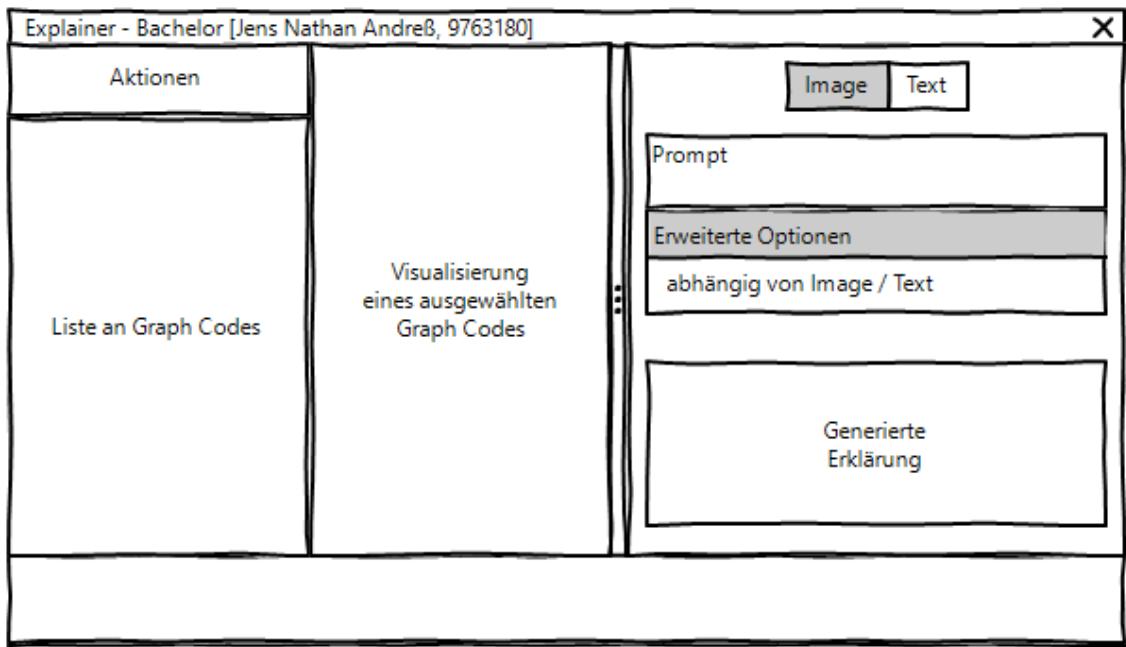


Abb. 3.5: Komplettes Wireframe für die Benutzerschnittstelle.

Die in diesem Abschnitt dargestellten Wireframes bilden die Struktur der Benutzerschnittstelle und werden im weiteren Verlauf dieser Arbeit in FZ 1.3/I im Rahmen einer prototypischen Implementierung untersucht. Im weiteren Verlauf dieses Forschungsziels werden nun mittels Mechanismen die an Anwendungsfällen beteiligten Komponenten identifiziert, sowie deren Zusammenspiel in einer kurzen Erklärung angerissen. Aufbauend auf diesen Erkenntnissen wird weiter mittels Sequenzdiagrammen das Verhalten zwischen diesen Komponenten untersucht und festgehalten.

### ***Mechanismen***

Ein Mechanismus ist ein partielles Klassendiagramm, das die Menge aller Klassen, von denen Instanzen in irgendeinem Objektdiagramm für einen Anwendungsfall vorkommen, darstellt. Mechanismen haben somit zum Ziel, die an einem Anwendungsfall beteiligten Objekte bzw. Klassen zu identifizieren und darzustellen. In einem

Mechanismus wird der Anwendungsfall selbst gestrichelt dargestellt und die mit diesem Anwendungsfall assoziierten Klassen werden ebenfalls durch gestrichelte Linien mit dem Anwendungsfall verbunden. Im weiteren Verlauf dieses Kapitels werden die in Mechanismen identifizierten Klassen zusammen mit den in den Anwendungsfällen beschriebenen Abläufen die Basis für Sequenz- und Klassendiagramme bilden. Im weiteren Verlauf dieses Abschnitts werden nun für ausgewählte Anwendungsfälle Mechanismen erstellt.

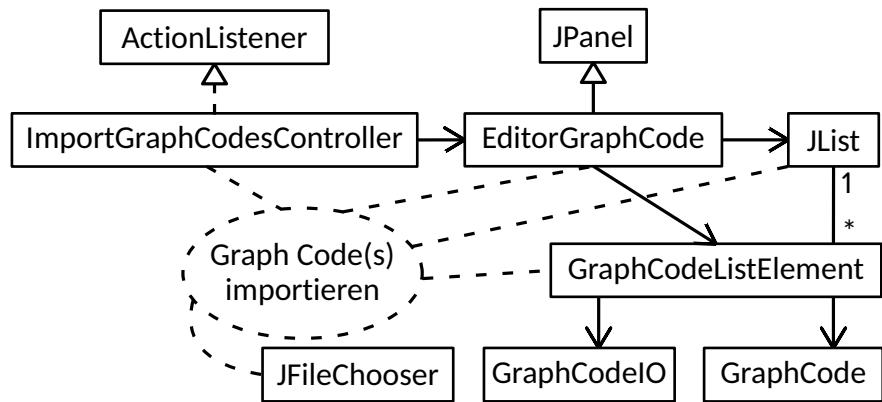


Abb. 3.6: Mechanismus für den Anwendungsfall UC-1.1.

Abb. 3.6 zeigt den Mechanismus für den Anwendungsfall Anwendungsfall UC-1.1. Sobald der Anwendungsfall UC-1.1 beginnt, benötigt es eine Komponente, die die in diesem Anwendungsfall enthaltenen Subaktionen steuert. Diese Aufgabe übernimmt die Komponente *ImportGraphCodesController*. Da das Importieren von Graph Codes das Auswählen der jeweiligen Graph Code Datei(en) erfordert, benötigt es einen Auswahldialog *JFileChooser*. Die nun durch den Benutzer ausgewählten Dateien werden einer Liste in der Arbeitsfläche hinzugefügt. Genauer ist diese Arbeitsfläche die linke Arbeitsfläche (siehe Abb. 3.3 ①) und wird in diesem Diagramm durch *EditorGraphCode* dargestellt. Diese Arbeitsfläche zeigt eine (1) Liste *JList* (siehe Abb. 3.4 ⑤), welcher beliebig viele (\*) Graph Codes *GraphCodeListElement* hinzugefügt werden können.

Abb. 3.7 zeigt den Mechanismus für den Anwendungsfall UC-1.2. Der Mechanismus für den Anwendungsfall UC-1.2 ist sehr ähnlich im Vergleich zum Mechanismus für den Anwendungsfall UC-1.1. Einzig erwähnenswert ist die Komponente *RemoveSelectedGraphCodesController*, die die notwendigen Subaktionen zum Entfernen von Graph Codes steuert.

Abb. 3.8 zeigt den Mechanismus für den Anwendungsfall UC-1.3. Der Anwendungsfall UC-1.3 beginnt mit der Auswahl von Graph Code(s). Das Auswählen von Graph Codes beeinflusst mehrere Komponenten in der gesamten Benutzerschnittstelle. Die

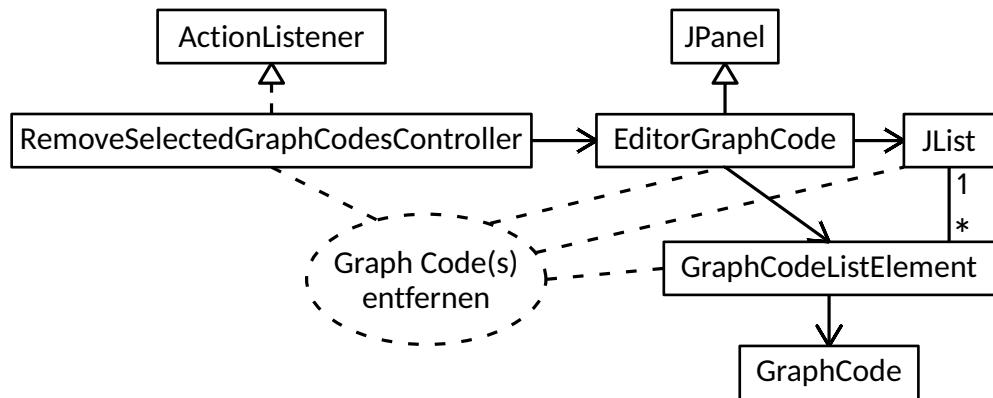


Abb. 3.7: Mechanismus für den Anwendungsfall UC-1.2.

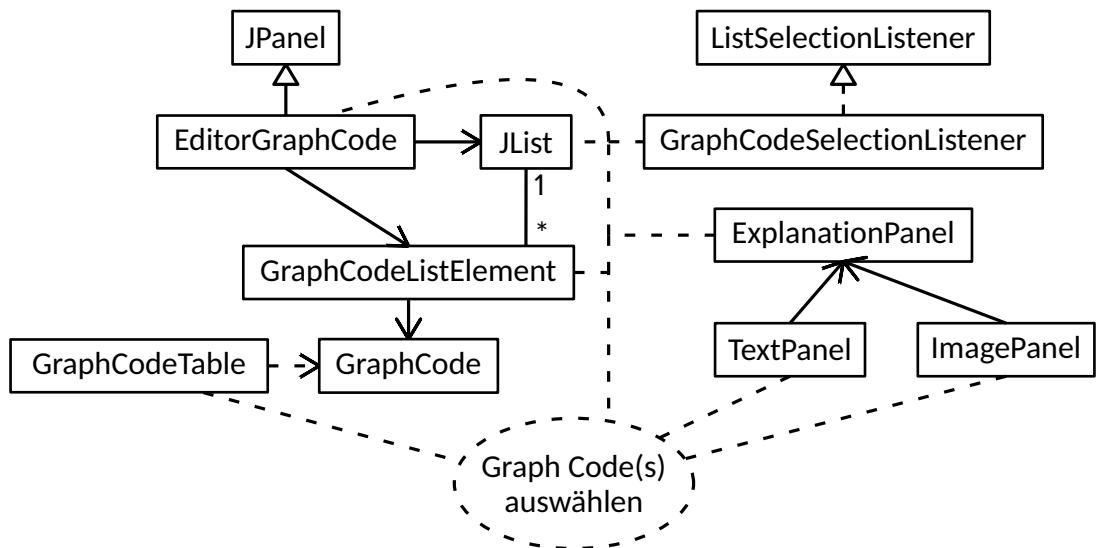


Abb. 3.8: Mechanismus für den Anwendungsfall UC-1.3.

Aktionen zur Einflussnahme auf diese Komponenten werden durch die Komponente *GraphCodeSelectionListener* gesteuert. Beeinflusste Komponenten umfassen die Komponente *GraphCodeTable* (siehe Abb. 3.4 (6)), sowie die Komponente *ExplanationPanel* (siehe Abb. 3.4 (2)), in der Benutzer zwischen Erklärungstypen umschalten können. Die Komponenten *ImagePanel* und *TextPanel* sind die für die Erklärungstypen spezifischen Benutzerschnittstellen, und können durch Abb. 3.4 (7) umhergeschaltet werden.

Abb. 3.9 zeigt den Mechanismus für den Anwendungsfall UC-1.5. Der Anwendungsfall UC-1.5 „Operation ausführen“ beginnt durch das Klicken des Knopfes „Execute“. Die an diesem Anwendungsfall beteiligten Aktionen werden durch die Komponente *GraphCodeCalculationController* gesteuert und geleitet. Dies beinhaltet auch das

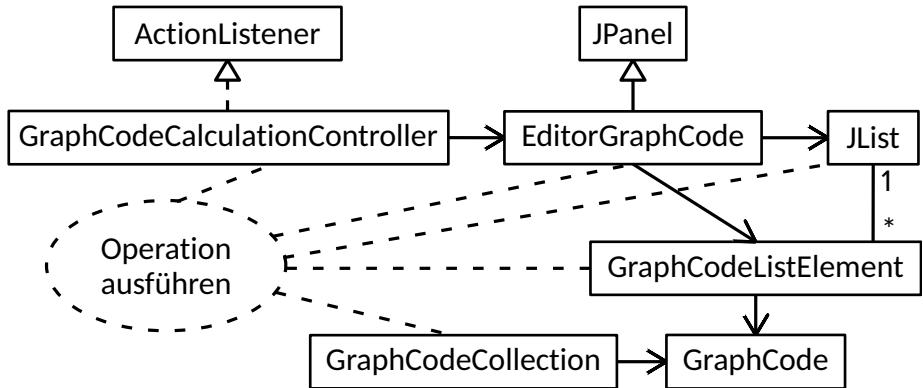


Abb. 3.9: Mechanismus für den Anwendungsfall UC-1.5.

Differenzieren der in Anwendungsfall UC-1.4 zuvor ausgewählten Operation. Diese Operation kann die *Vereinigung (Union)*, *relative Differenz / Unterschied bzw. Subtraktion*, *Gemeinsamkeiten (Similarities)* oder *Unterschiede (Differences)* zwischen mehreren Graph Codes sein. Besonders hierbei ist die *Subtraktion*, die nur auf zwei Graph Codes angewandt werden kann. Die Algorithmen zur Berechnung der Operationen der *Vereinigung* von mehreren Graph Codes, sowie der *Subtraktion* zweier Graph Codes werden bereits durch die Komponente *GraphCodeCollection* bereitgestellt. Im Folgenden werden Konzepte für die weiteren Operationen *Gemeinsamkeiten* und *Unterschiede* in Form von Pseudoalgorithmen vorgestellt, und die an diesen Operationen beteiligten Komponenten, ihre Aufgaben, sowie durchzuführende Schritte beschrieben.

Algorithmus 3.1 zeigt einen Pseudoalgorithmus für die Operation *Gemeinsamkeiten*. Diese Algorithmus nimmt als Eingabe eine Liste an Graph Codes und gibt als Ergebnis einen Graph Code zurück. Der Algorithmus kann in drei Bereiche unterteilt werden: *Vorbereitung*, *Berechnung* und *Verwertung*. Der Bereich *Vorbereitung* umfasst die Zeilen 2 bis 4. In diesem Bereich wird das Ergebnis als Objekt zum ersten Mal initialisiert. Weiterhin wird die Vereinigung aller Graph Codes aus der Eingabe der Funktion mit der Funktion `getUnion`, welche durch die Komponente *GraphCodeCollection* bereitgestellt wird, berechnet. Ergebnis dieser Vereinigung ist wiederum ein Graph Code, dessen Vokabular zuletzt für die weitere Berechnung in den folgenden Bereichen aufbereitet wird. Der zweite Bereich *Berechnung* umfasst die Zeilen 5 bis 6. In diesem Bereich wird das Vokabular der Vereinigung (`unionDic`) mit dem Vokabular aller Graph Codes aus der Eingabe miteinander verglichen und es werden alle Merkmale aus `unionDic`, die nicht in irgendeinem Vokabular eines Graph Codes aus der Eingabe vorkommen, entfernt. Übrig bleiben die Merkmale, die in jedem Graph Code aus der Eingabe vorkommen. Der dritte Bereich *Verwertung* umfasst die übrigen Zeilen 7 bis 14. Dem Objekte `sim`, welches das Ergebnis repräsentiert, wird das Vokabular `unionDic` gesetzt. Im weiteren Verlauf wird für

---

**Algorithm 3.1** Berechne Gemeinsamkeiten

---

**Eingabe:** Liste an Graph Codes: gcs

**Ausgabe:** Graph Code

```
1: function GETSIMILARITIES(gcs)
2:   sim ← new GraphCode
3:   union ← getUnion(gcs)
4:   unionDic ← new Vector(union.getDictionary())
5:   for gc in gcs do
6:     Entferne alle Elemente aus unionDic,
7:     die nicht im Vokabular von gc enthalten sind
8:   dictionary ← new Vector(unionDic)
9:   sim.setDictionary(dictionary)
10:  for gci in gcs do
11:    for s in dictionary do
12:      for t in dictionary do
13:        i ← gci.getEdgeValueForTerms(s, t)
14:        sim.setValueForTerms(s, t, i)
15:  return sim
```

---

jeden Graph Code zeilen- und spaltenweise der Wert für ein Paar aus Merkmalen in der Adjanzenzmatrix bestimmt und dieser Wert in der Matrix des Ergebnisses eingetragen. Schlussendlich wird das Ergebnis als Ausgabe zurückgegeben.

Algorithmus 3.2 zeigt einen Pseudoalgorithmus für die Operation *Unterschiede*. Dieser Algorithmus ist sehr ähnlich zum Algorithmus 3.1 und nimmt ebenfalls eine Liste an Graph Codes und gibt als Ergebnis einen Graph Code zurück. Zudem kann dieser Algorithmus ebenfalls in die gleichen Bereiche unterteilt werden. Mehr noch sind die Bereiche *Vorbereitung* (Zeile 2 bis 4) und *Verwertung* (Zeile 10 bis 15) identisch und Unterschiede in den Algorithmen bestehen nur im Bereich *Berechnung* (Zeile 6 - 7). Genauer besteht der einzige Unterschied im Vergleich zum Algorithmus 3.1, dass unterschiedliche Merkmale in mehreren Graph Codes das Gegenteil der gemeinsamen Merkmale in mehreren Graph Codes ist.

Zusammengefasst und einfacher ausgedrückt wird in der Operation *Gemeinsamkeiten* die Schnittmenge aller Vokabulare von Graph Codes, welche wiederum nur Mengen von Merkmalen sind, gebildet. Analog wird in der Operation *Unterschiede* die symmetrische Differenz aller Vokabular von Graph Codes gebildet, welches dem Gegenteil der Schnittmenge aller Vokabulare von Graph Codes entspricht. Dies unterscheidet die Operation *Unterschiede* von der durch die Komponente bereitgestellten Operation *Subtraktion*, welche nur die relative Differenz zweier Graph Codes bildet.

---

**Algorithm 3.2** Berechne Unterschiede

---

```
1: function GETDIFFERENCES(gcs)
2:   diff ← new GraphCode
3:   union ← getUnion(gcs)
4:   unionDic ← new Vector(union.getDictionary())
5:   for gc in gcs do
6:     Entferne alle Elemente aus unionDic,
7:     die nicht im Vokabular von gc enthalten sind
8:     Alle Merkmale bestimmen, die im Vokabular der Vereinigung union,
9:     aber nicht im Vokabular unionDic enthalten sind.
10:    dictionary ← new Vector(unionDic)
11:    diff.setDictionary(dictionary)
12:    for gci in gcs do
13:      for s in dictionary do
14:        for t in dictionary do
15:          i ← gci.getEdgeValueForTerms(s, t)
16:          diff.setValueForTerms(s, t, i)
17: return diff
```

---

### Sequenzdiagramme

Sequenzdiagramme modellieren den konkreten zeitlichen Verlauf von mitunter komplexen Operationen in einem Klassenmodell unter Einbeziehung der in diesen Operationen beteiligten Objekten. Die in einer Operation beteiligten Objekte werden dabei am oberen Rand des Diagramms als ein Rechteck mit einer gestrichelten vertikalen Linie dargestellt. Diese gestrichelte vertikale Linie wird als *Lebenslinie* bezeichnet und symbolisiert die Lebenszeit eines Objektes, welche von oben nach unten voranschreitet. Interaktionen zwischen Objekten, wie z.B. dem Aufrufen einer Operation eines Objektes durch ein anderes Objekt, geschehen durch das Senden von Nachrichten zwischen diesen. Die gängisten Formen von Nachrichten sind:

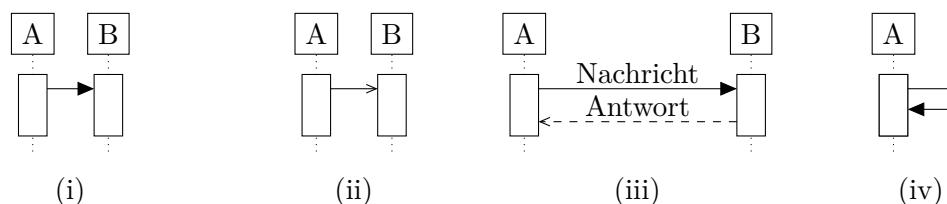


Abb. 3.10: (i) Synchrone Nachrichten, (ii) Asynchrone Nachrichten, (iii) Antwort auf Nachricht, (iv) Selbstdelegation / Selbstaufruf.

- (i) *Synchrone Nachrichten* werden durch einen geschlossenen, gefüllten Pfeil → dargestellt. Eine synchrone Nachricht repräsentiert einen direkten Aufruf einer

Operation vom Dienstnutzenden (Sender) zum Dienstleistenden (Empfänger). Eine wichtige Eigenschaft, die diesem Nachrichtentyp seinen Namen verleiht ist, dass der Sender auf die Antwort seiner Nachricht wartet, bevor er weitere Nachrichten sendet.

- (ii) *Asynchrone Nachricht* werden durch einen offenen, spitzen Pfeil → dargestellt. Im Gegensatz zu einer synchronen Nachricht wartet der Sender nach Senden der Nachricht nicht mehr auf eine Antwort des Empfängers, sondern kann unmittelbar weitere Nachrichten senden. Folglich sind bei dieser Art von Nachricht die Sender und Empfänger nicht mehr synchron zueinander.
- (iii) Eine Nachricht vom Typ *Antwort* ist eine Antwort auf eine asynchrone Nachricht und wird durch einen gestrichelten, entgegengesetzten Pfeil <--- dargestellt. Eine Antwort führt den Kontrollfluss zurück auf den Sender der Nachricht nachdem die Operation vom Empfänger abgeschlossen wurde.
- (iv) *Selbstdelegation / Selbstaufruf* ist eine Nachricht eines Objektes auf sich selbst und tritt meist bei rekursiven Methoden auf.

Im Folgenden werden für ausgewählte Anwendungsfälle Sequenzdiagramme vorgestellt und beschrieben.

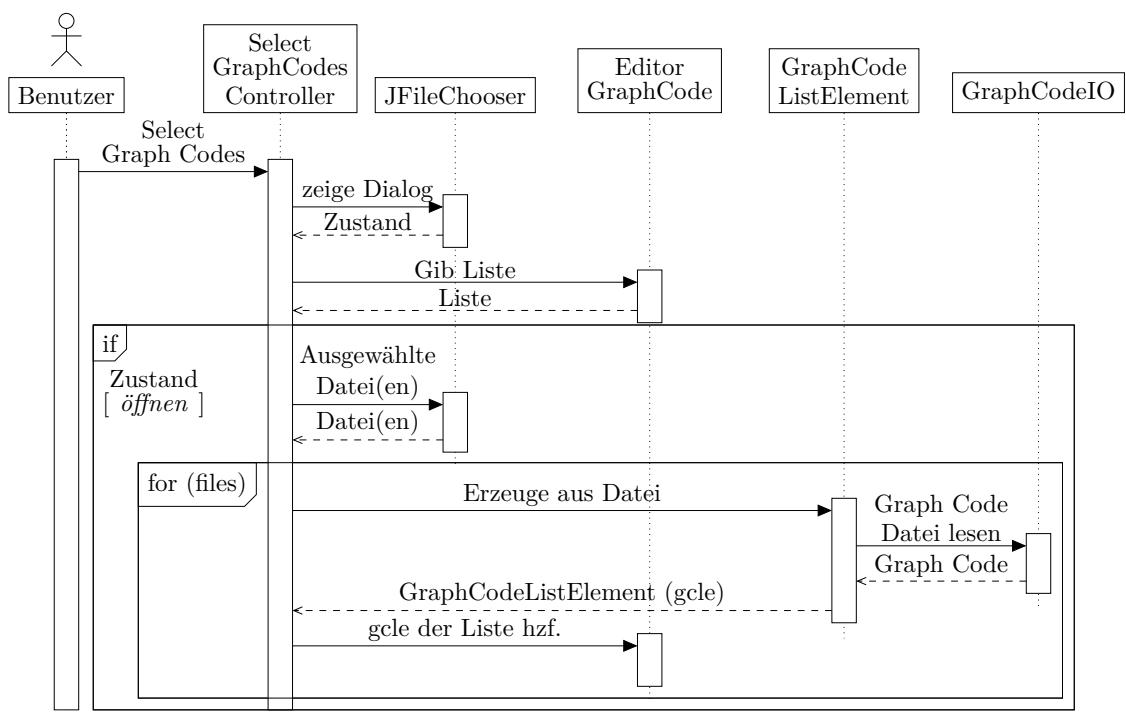


Abb. 3.11: Sequenzdiagramm für den Anwendungsfall UC-1.1.

Abb. 3.11 zeigt ein Sequenzdiagramm für den Anwendungsfall UC-1.1. Es zeigt, wie der Benutzer die Aktion „Graph Code(s) importieren“ durch einen Klick auf den Knopf „Select Graph Code(s)“ initiiert. Weitere Aktionen werden dann von der Komponente *SelectGraphCodesController* durchgeführt. Diese Aktionen umfassen das Anzeigen eines Dateiauswahldialogs und der in diesem Dialog ausgewählten Dateien. Für alle ausgewählten Dateien *for (files)* werden dann Listenelemente erzeugt, die der Liste hinzugefügt werden können. Hierfür müssen die Graph Code Dateien durch die Komponente *GraphCodeIO* eingelesen und in ein Objekt umgewandelt werden.

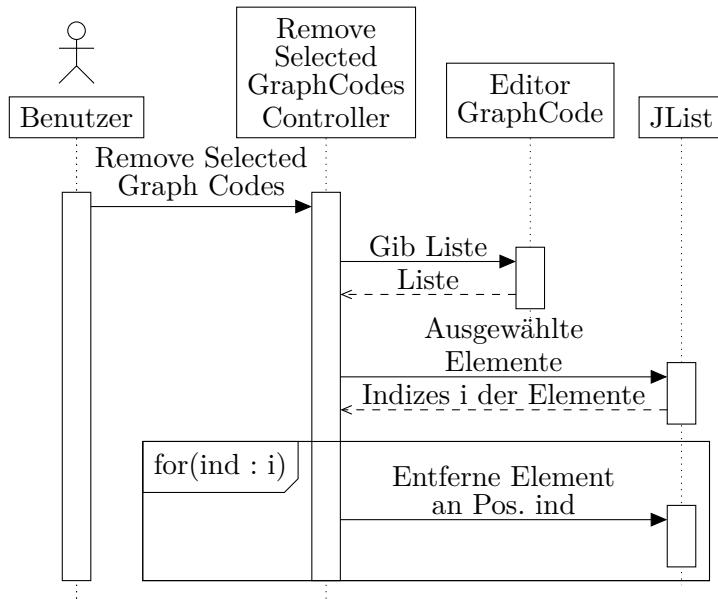


Abb. 3.12: Sequenzdiagramm für den Anwendungsfall UC-1.2.

Abb. 3.12 zeigt ein Sequenzdiagramm für den Anwendungsfall UC-1.2. Es zeigt, wie der Benutzer die Aktion „Graph Code(s) entfernen“ durch einen Klick auf den Knopf „Remove selected Graph Code(s)“ initiiert. Weitere Aktionen werden dann von der Komponente *RemoveSelectedGraphCodesController* durchgeführt. Diese Aktionen umfassen die Abfrage der ausgewählten Elemente in der Liste und die Rückführung der Indizes, die die Positionen dieser Elemente wiedergeben. Anhand dieser Indizes kann der Liste mitgeteilt werden das Element an der Position eines Indexes zu entfernen.

Abb. 3.13 zeigt ein Sequenzdiagramm für den Anwendungsfall UC-1.3. Es zeigt, wie der Benutzer die Aktion „Graph Code(s) auswählen“ durch eine Auswahl auf ein Element in der Liste initiiert. Die Ausführungen der weiteren Aktionen werden dann von der Komponente *GraphCodeSelectionListener* übernommen. Zusammengefasst wird hier das aktuell ausgewählte Element in der Liste abgefragt und von der

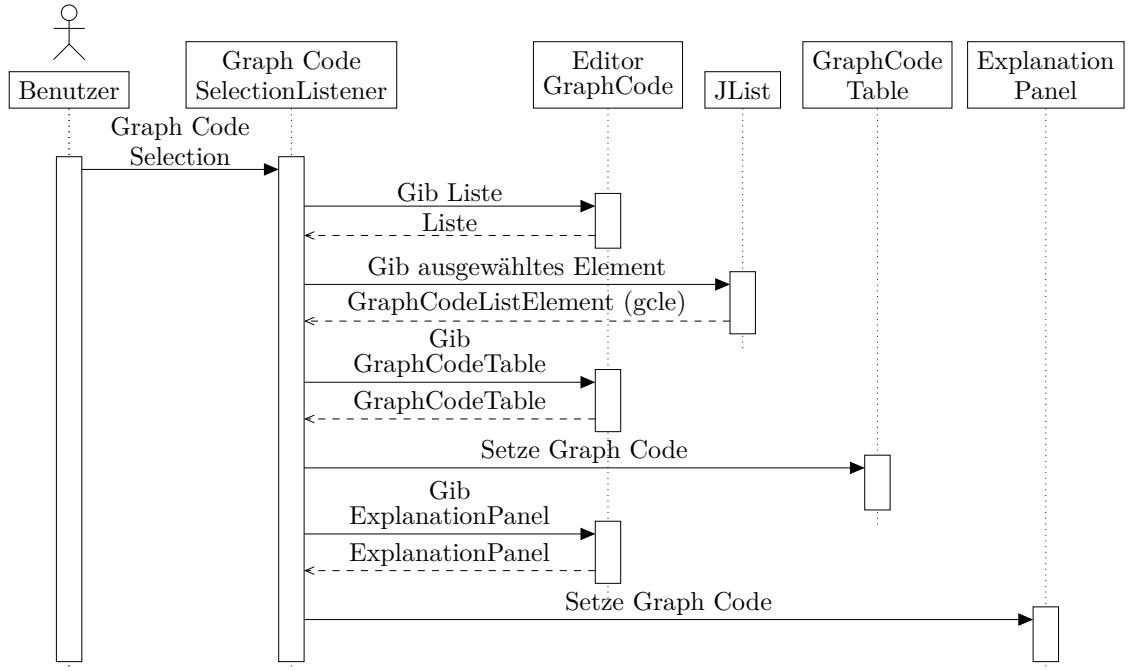


Abb. 3.13: Sequenzdiagramm für den Anwendungsfall UC-1.3.

Liste zurückgeführt. Dieses Komponente *GraphCodeListElement* enthält wiederum die Komponente *GraphCode*, welche dann anderen Schnittstellen, wie der Komponente *GraphCodeTable* oder *ExplanationPanel* zur weiteren Verarbeitung mitgeteilt werden kann. Die Komponente *ExplanationPanel* stellt wiederum die Komponenten *ImagePanel* und *TextPanel* zur Verfügung bzw. hält diese Komponenten inne. Die weitere Verarbeitung eines Graph Codes erfolgt dann analog durch diese Komponenten.

Abb. 3.14 zeigt ein Sequenzdiagramm für den Anwendungsfall UC-1.5. Es zeigt, wie der Benutzer die Aktion „Operation ausführen“ durch einen Klick auf den Knopf „Execute“ initiiert. Die weiteren Aktionen werden dann von der Komponenten *GraphCodeCalculationController* übernommen. Diese Aktionen umfassen die Abfrage der ausgewählten Elemente aus der Liste und die Rückführung dieser Elemente, damit die in diesen Elementen gespeicherten Graph Codes im weiteren Verlauf verarbeitet werden können. Die Verarbeitung ist dabei abhängig von der in Anwendungsfall UC-1.4 ausgewählten Operation. Die Differenzierung dieser Aktionen wird im Sequenzdiagramm durch die Box *switch* dargestellt. Schlussendlich wird mit dem Ergebnis ein neues Element für die Liste erzeugt und der Liste hinzugefügt.

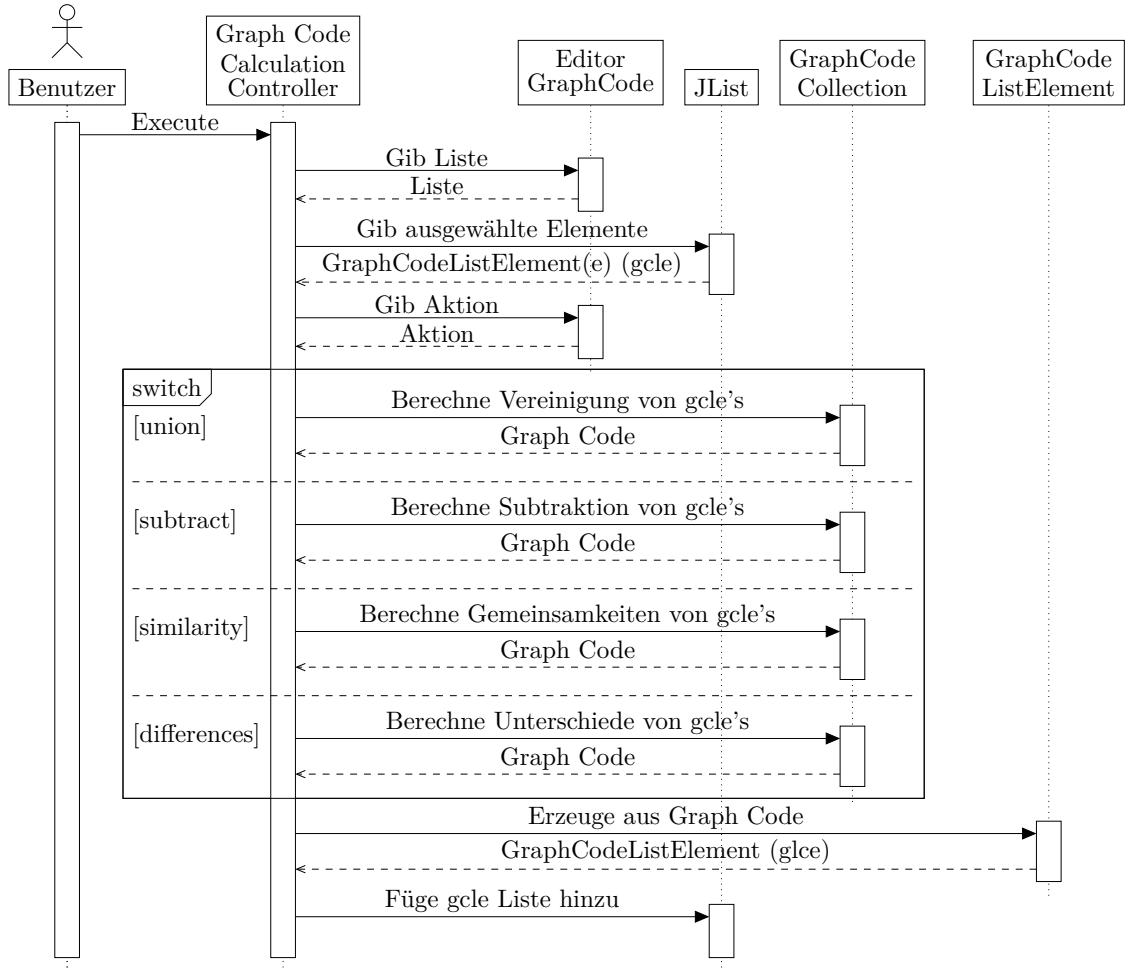


Abb. 3.14: Sequenzdiagramm für den Anwendungsfall UC-1.5.

### 3.2.3 Diskussion

In diesem Abschnitt wurden die offenen Herausforderungen **OH 1.1** und **OH 1.2** adressiert. In Abschnitt 3.2.1 wurde die offene Herausforderung **OH 1.1** behandelt und Möglichkeiten vorgestellt, wie durch generative KI Erklärbarkeit erreicht werden kann. In Abschnitt 3.2.2 wurde die offene Herausforderung **OH 1.2** behandelt und es wurden Anwendungsfälle aus der Perspektive von Benutzern des GMAF identifiziert und beschrieben. Darüber hinaus wurden Wireframes für eine Benutzerschnittstelle mit geeigneten Interaktionsmöglichkeiten für den sinnvollen Einsatz von Systemen generativer KI zur Erklärung von Graph Codes konzipiert und weiter ausgeführt. Weiterhin wurden anhand der identifizierten Anwendungsfälle und den dazu entsprechenden textuellen Beschreibungen Mechanismen erstellt, um die an den Anwendungsfällen beteiligten Komponenten zu identifizieren. Anhand der Anwen-

dungsfälle und der mit diesen Mechanismen identifizierten Komponenten konnten daraufhin Sequenzdiagramme zur genaueren Ablaufbeschreibung erstellt werden.

Die in diesem Abschnitt gewonnenen Erkenntnisse bieten die Grundlage zur Entwicklung einer geeignete Benutzerschnittstelle zur Erklärbarkeit von Graph Codes mittels generativer KI.

### 3.3 FZ 2.2/TB Integration generativer KI in das GMAF

Dieser Abschnitt befasst sich mit der Modellierung der „Integration generativer KI in das GMAF“. Dabei werden die in FZ 2.1/O identifizierten, offenen Herausforderungen **OH 2.1** und **OH 2.2** adressiert. Ziel dieses Abschnitts ist es Konzepte für die Überführung von Graph Codes in eine geeignete Form zur Eingabe in ein System generativer KI zu entwickeln, sowie Konzepte für die Integration generativer KI zu entwickeln. Hierzu wird in Abschnitt 3.3.1 die offene Herausforderung **OH 2.1** und in Abschnitt 3.3.2 die offene Herausforderung **OH 2.2** adressiert.

#### 3.3.1 Transformation von Graph Codes

Dieser Abschnitt widmet sich der Entwicklung von Konzepten zur Überführung der in Graph Codes gespeicherten Informationen, sowie der mit Graph Codes assoziierten Informationen in eine geeignete Form, die als Eingabe in ein System generativer KI verwendet werden kann. Wie bereits in Kapitel 2 recherchiert wurde, erfolgt die Eingabe in ein System generativer KI durch eine Eingabeaufforderung bzw. Prompt, die im Endeffekt eine beliebig lange Zeichenkette sein kann. Die geeignete Form, in die es also Graph Codes zu Überführen gilt, sind somit Prompts. Bevor allerdings mit dem eigentlichen Konzipieren der Transformation von Graph Codes in Prompts begonnen werden kann, ist es sinnvoll zu bestimmen, welche Informationen im Rahmen von Graph Codes überhaupt für eine Erklärung relevant sein können. Von besonderer Relevanz für eine Erklärung sind in erster Linie die Merkmale bzw. das Vokabular, gespeichert im Wörterbuch eines Graph Codes, sowie die Informationen über bestehende Beziehungen zwischen diesen Merkmalen, gespeichert in einer zweidimensionalen, quadratischen Matrix. Daher werden in diesem Abschnitt erst Konzepte für die Grundlagen von Graph Codes entwickelt. Im Anschluss werden dann Konzepte für die mit Graph Codes assoziierten Informationen besprochen.

##### *Transformation des Vokabulars*

Das Vokabular eines Graph Codes ist eine Menge aus Merkmalen, welche wiederum beliebig lange Zeichenketten sein können. Eine einfache Art das Vokabular eines Graph Codes in einen Prompt zu überführen, ist somit einfach die Merkmale in Textform aufzuzählen. Zusätzlich wird dem System durch eine in der Prompt vorangestellte Instruktion erklärt, welche Bedeutung die Elemente in der Aufzählung besitzen. Ein Beispiel hierfür wäre folgende Prompt: „The elements in the following list represent feature vocabulary terms and describe features extracted from multimedia files: root-image, cloud, water, sky, lake, dusk“. Eine solche Transformation des Vokabulars würde in der entsprechenden Komponente des Erklärungstyps *ImagePanel* oder *TextPanel* durch eine Funktion *setUpPrompt* umgesetzt werden. Abb. 3.15 zeigt ein einfaches Klassendiagramm mit den Komponenten *ImagePanel*,

*TextPanel* und der Funktion *setUpPrompt*.

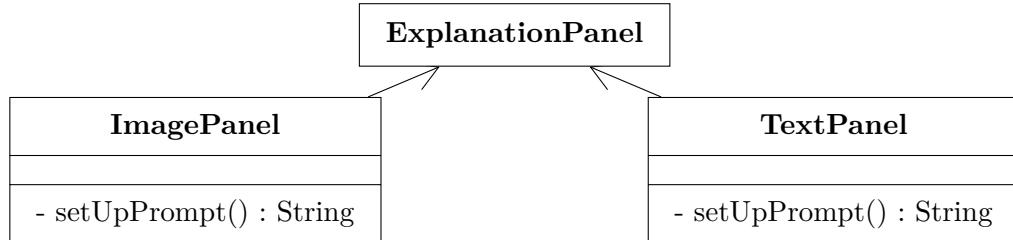


Abb. 3.15: Einfaches Klassendiagramm für die Erklärungspaneale.

### ***Transformation der Matrix***

Die Adjazenzmatrix eines Graph Codes ist eine zweidimensionale, quadratische Matrix, dessen Anzahl an Zeilen bzw. Spalten äquivalent zur Größe des Vokabulars ist. Die Werte der Einträge in der Matrix geben Aufschluss über das Vorhandensein von Beziehungen, sowie über die Bedeutung bzw. Art der Beziehung. Bei der Transformation der Matrix müssen daher zwei Aspekte berücksichtigt werden: Das allgemeine Überführen der Informationen aus der Matrix, sowie das Überführen der Werte bzw. der Kodierung der Beziehungen und ihren Bedeutungen.

Eine einfache und naive Transformation kann bereits durch die Funktion *toString* erfolgen, die durch die Komponente *GraphCode* bereitgestellt wird. Diese Funktion überführt die Informationen eines Graph Codes in das JSON-Format. Ein Beispiel für so ein JSON-Format kann in Listing 2.1 eingesehen werden. Die Tatsache, dass die GPT-Modelle in der Lage sind dieses Format zu analysieren und zu verarbeiten, begünstigt diese Herangehensweise. Ein Beispiel hierfür kann folgende Prompt sein: „The following JSON string represents a matrix. The matrix is a square twodimensional matrix where each element represents the relationship between two words: {"matrix":[[1,1,1,1,1],[0,2,0,0,0,0],[0,0,2,0,0,0],[0,0,0,2,0,0],[0,0,0,0,2,0],[0,0,0,0,0,2]]}“.

Die Umsetzung dieser Transformation findet in der Funktion *toString* in der Komponente *GraphCode* statt und kann in der entsprechenden Komponente des Erklärungstypes *ImagePanel* oder *TextPanel* in der Funktion *setUpPrompt* eingesetzt werden.

Ein Nachteil an dieser Herangehensweise ist das Verhältnis zum Platzverbrauch im Zeichenformat. Aufgrund der quadratischen Natur der Adjazenzmatrix wird überproportional viel Platz verbraucht, was insbesondere anhand des immer weiter steigenden Detaillierungsgrades (LOD) problematisch ist. Dies führt auch zu einem zunehmenden Tokenverbrauch, was in Bezug auf einige Modelle, wie z.B. Dall-E 2, die Zeichen- bzw. Tokenbeschränkungen besitzen, ebenfalls ein Problem darstellen könnte. Eine Möglichkeit diesem Problem zu begegnen, wäre durch das Prüfen der

Tokenlänge in der entstandenen Prompt. Das Werkzeug *Tokenizer*, vorgestellt in Abschnitt 2.2.2.2, kann hierfür beim Zusammensetzen der Prompt verwendet werden. Da hierdurch das Problem allerdings nicht gelöst wird, wird im Folgenden eine verbesserte Herangehensweise zum Transformieren der Informationen eines Graph Codes vorgestellt.

Die verbesserte Herangehensweise zur Überführung der Informationen eines Graph Codes reduziert die zu transformierenden Informationen auf ein Minimum. Anders ausgedrückt, werden in dieser Herangehensweise nur die Informationen betrachtet und transformiert, die von Interesse sind. Dies sind insbesondere die Einträge in einer Matrix, die überhaupt über eine Beziehung verfügen, ungeachtet der Kodierung bzw. des Typs der Beziehung. Da jedes Merkmal  $n$  Beziehungen zu anderen Merkmalen aufweisen kann, könnte sich das Format  $< i_t > - < i_{t_1}, \dots, i_{t_n} >$  gut zur Darstellung von Beziehungen zwischen Merkmalen eignen. Der Bezeichner  $i_t$  im Format bezeichnet den Index eines Merkmals  $t$  im Vokabular des Graph Codes. Effektiv sagt dieses Format aus, dass das Merkmal  $t$  mit den anderen Merkmalen  $t_1, \dots, t_n$  verbunden ist. Durch das Abbilden des Merkmals auf den entsprechenden Index im Vokabular kann ein Optimierungspotenzial in Bezug auf die Tokenlänge ausgeschöpft werden. Dies ist möglich, da die Merkmale in einem Vokabular in einer fest definierten Reihenfolge geordnet sind. Ein Beispiel für so ein Format kann sein: „Some of the dictionary terms are connected through a relationship. These relationships will be noted as  $< i_t > - < i_{t_1}, \dots, i_{t_n} >$ , where  $i_t$  denotes the index of a feature vocabulary term:  $< 1 > - < 1, 2, 3, 4, 5, 6 >$ .“

Dieses Format trifft allerdings nur Aussagen darüber, ob eine Beziehung zwischen diesen Merkmalen vorhanden ist, oder nicht. Aussagen über den Typ einer Beziehung erfolgen über die Kodierung bzw. den spezifischen Wert des Eintrages. Durch eine Erweiterung des Formats zu  $< i_t > k < i_{t_1}, \dots, i_{t_n} >$  kann die Information der Kodierung in das Format integriert werden.  $k$  steht hierbei für den spezifischen Wert bzw. Typ der Beziehung. Ein Beispiel für so ein erweitertes Format kann z.B. „ $< 2 > 3 < 4 > . < 3 > 3 < 5 >$ .“ sein. In diesem Beispiel symbolisiert der Wert 3 die Beziehung „ist Teil von“.

Auf diese Weise kann für jeden Beziehungstyp ein Format erstellt und in die Prompt eingefügt werden. Zusätzlich gilt es im Kontext eines wertespezifischen Formats dem System generativer KI klarzumachen, welche Bedeutung dem Wert  $k$  beigemessen wird. Dies kann durch eine Erweiterung der Instruktionen erfolgen. Ein Beispiel hierfür ist: „Some of the dictionary terms are connected through a relationship. These relationships will be noted as  $< i_t > - < i_{t_1}, \dots, i_{t_n} >$  or  $< i_t > k < i_{t_1}, \dots, i_{t_n} >$ , where  $i_t$  denotes the index of a feature vocabulary term. Additionally – represents an undefined relationship, while  $k$  represents a specific relationship. A value of 3 represents the relationship 'is part of'. 4 represents the relationship 'consists of'. The relationships are as follows:  $< 1 > - < 1, 2, 3, 4, 5, 6 >, < 2 > 3 <$

$4 >, < 3 > 3 < 5 >, < 5 > 4 < 3 >.$ “

Zusammengefasst sieht der gesamte Ablauf vor, dass zuerst die Werte für die Beziehungstypen aus der Adjanzenzmatrix gefiltert werden und für die Merkmale, die über diese Beziehung mit anderen Merkmalen verbunden sind, jeweils ein Format erzeugt und in die Prompt integriert wird. Mitsamt Instruktionen kann auf die Weise eine umfassende Prompt erzeugt werden, die die Essenz von Graph Codes erfasst und dem System generativer KI zum Zweck der Generierung einer Erklärung vermittelt.

### ***Anwendung der Graph Code Metriken***

Metriken von Graph Codes sind mit Graph Codes assoziierte Informationen, die in Abhängigkeit ihrer Aufgabe Aussagen über Graph Codes in Form von Gleitkomma-werten in einem Wertebereich, der meistens zwischen 0 und 1 liegt, ermöglichen. In Abschnitt 2.1.3.3 wurde bereits die Ähnlichkeitsmetrik vorgestellt, die Aussagen über die Ähnlichkeit zwischen zwei Graph Codes ermöglicht. Im Folgenden wird aufgezählt, welchen Nutzen diese Metrik beim Erreichen des Ziels eine Erklärung zu Graph Codes zu erzielen, beitragen könnte und wie so eine Nutzung bzw. Anwendung dieser Metrik aussehen könnte:

- **Kontrollfluss:** Metriken könnten im Kontrollfluss des Programms verwendet werden, um Fallunterscheidungen zu ermöglichen. In Bezug auf die in Anwendungsfall UC-1.4 genannten Operationen könnten die Metriken genutzt werden, um bei der Verarbeitung
- **Prompt / Instruktion:** Die Werte von Metriken könnten, erweitert um Instruktionen, in eine Prompt transformiert bzw. integriert werden, um die Aussagekraft von Graph Codes bzw. in ihnen enthaltenen Merkmalen zu verstärken oder zu gewichten.

Die in diesem Abschnitt konzipierten Herangehensweisen werden im Rahmen einer Implementierung durch das FZ 2.3/I genauer untersucht.

### ***Transformation der Graph Code Metriken***

Metrik zur Ähnlichkeit ist nur ein Tripel aus Metriken. Nichtsdestotrotz sind auch diese Metriken auf mehrere Graph Codes angewiesen...

Grundlegender Startpunkt ist nur ein Graph Code... Dieser Graph Code enthält das Element CollectionElements... Dieses Element ist ein Array aus Darstellungen von Graph Codes als Json Objekt bzw. Element (rekursive Struktur) Könnte man anhand dieser Graph Codes in CollectionElements Metriken berechnen?

Wie kann die Info zu einer Metrik mit in die Erklärung integriert werden? (Benötigt

offensichtlich Berechnung der Metrik)

Wie kann generative KI genutzt werden um Metriken zu (generieren?) erweitern?

Idee: Sentimentalitätsmetrik - Analyse... Idee: Individuelle Metriken (wie zur Sentimentalität) können angefragt und dann durch einen Funktionsaufruf ins Programm übernommen werden... (Um diesen Ansatz zu verfolgen wäre auch eine zweite, kleine Erweiterung des Kap.2 FZ2 eventuell sinnvoll? ansprechen der Funktionen und Aufrufe

Das Verwenden der bestehend(en) Metriken ist nur im Anwendungsfall Operation ausführen sinnvoll und dort auch nur für die Operation Subtraktion... Hier dann begründen? Anhand Implementierung begründen? Oder anhand Definition? Eher Definition... Andere Operationen berechnen anhand mehrerer Graph Codes, was hier für diese Metrik kein Sinn ergibt... Metrik der Ähnlichkeit nur zwischen zwei Graph Codes... Hier muss auch nochmal überprüft werden, wie denn überhaupt für diese Operation die Berechnung dieser Metriken zustande kommen würde... (Aber eher unwahrscheinlich... Hier in der Modellierung nur die Idee hervorheben... Umsetzbarkeit wird in der Implementierung untersucht...)

Allgemein Vorstellen wie (sofern berechnet) die Metrik transformiert werden könnte und integriert werden könnte... (Erklärung der Metrik + Instruktion)

Erweiterung der bestehenden Metriken...

Neuer Anwendungsfall... (Graph Codes vergleichen)

Allgemein nur kurz über die Möglichkeiten der Transformation von Metriken sprechen... Wo und wie könnten diese Metriken eingesetzt werden?

Wie könnten Metriken von Graph Codes allgemein durch generative KI erweitert werden? + Beispiel: Sentimentalität

### 3.3.2 Einbindung von Systemen generativer KI

Dieser Abschnitt widmet sich der Entwicklung von Konzepten zur Integration von Systemen generativer KI in das GMAF. Die Integrationsmöglichkeiten der in Kapitel 2 vorgestellten und ausgewählten Systeme generativer KI wurden bereits in Abschnitt 2.2.2 hervorgehoben. Genauer wurde in Abschnitt 2.2.2 eine Übersicht über die von OpenAI bereitgestellte Schnittstelle und ihre Endpunkte geschaffen. Bevor allerdings mit der Entwicklung von Konzepten zur Integration von Systemen generativer KI begonnen wird, ist es sinnvoll zu bestimmen, welche Komponenten an der Erstellung einer Anfrage an diesen Endpunkten beteiligt sind.

Abb. 3.16 zeigt ein partielles Klassendiagramm, welches die an der Erstellung von Anfragen an die Endpunkte der Schnittstelle beteiligten Komponenten zeigt. Die

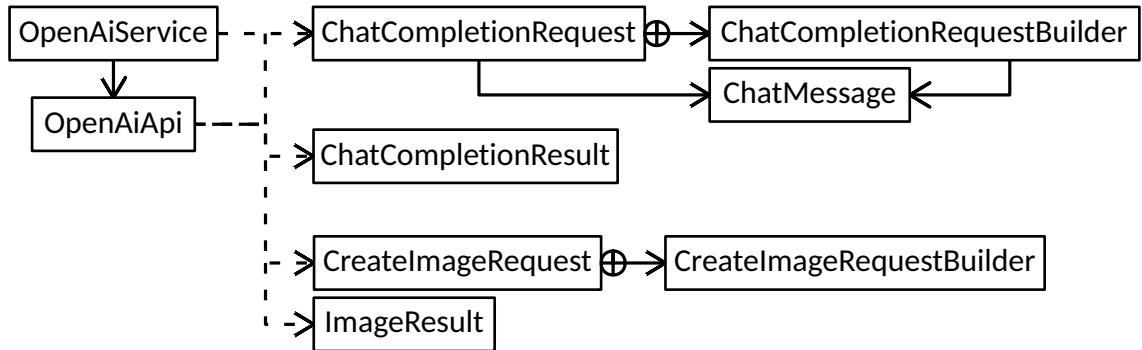


Abb. 3.16: Endpunkte der OpenAI Schnittstelle.

Komponenten `OpenAiService` und `OpenAiApi` bieten Zugriff auf die Funktionalitäten der Schnittstelle und behandeln zudem die Konfiguration der Schnittstelle. Weitere Komponenten sind `ChatCompletionRequest` und `CreateImageRequest`, die jeweils als Objekt eine Anfrage an einen Endpunkt darstellen. Die Anfragen müssen dann noch über die Komponente `OpenAiService` an die Schnittstelle gesendet werden. Analog zu den Anfragen sind die Komponenten `ChatCompletionResult` und `ImageResult` Container, die die Ergebnisse der Anfragen an die jeweiligen Endpunkte speichern. Folglich bestehen diese Komponenten wiederum aus kleinen weiteren Komponenten, die die jeweiligen Informationen speichern. Die Komponenten `ChatCompletionRequestBuilder` und `CreateImageRequestBuilder` sind jeweils die Komponenten, die die Parameter der Anfrage spezifizieren. Inbesondere ist in Bezug auf die Anfragen zur Generierung von Text die Komponente `ChatMessage` von Bedeutung. Während bei normalen Anfragen, wie z.B. der Anfrage zur Generierung von Bildern nur einfache Texteingaben benötigt werden, wird bei der Anfrage von Text eine Reihe an Textnachrichten (`ChatMessage`) benötigt, die einen Dialog zwischen mehreren Parteien darstellen. Eine Textnachricht erwartet als Eingabe eine Rolle im Dialog, sowie den Inhalt dieser Partei im Dialog.

Die genannten Komponenten stellen die Grundlage zum Erstellen einer Anfrage an die Endpunkte der Schnittstelle von OpenAI dar. Im Folgenden ist es nun von Interesse herauszufinden, welche Informationen für das Erstellen einer effektiven und sinnvollen Anfrage an einen Endpunkt der Schnittstelle notwendig sind, und welche Komponenten am Erzeugen dieser Informationen beteiligt sind. Speziell für die Textgenerierung bestehen die Informationen für den Dialog, der durch eine Reihe an Textnachrichten `ChatMessage` dargestellt wird, zumeist aus Instruktionen (siehe z.B. Abschnitt 3.2.1), sowie den in Textformat transformierten Informationen aus einem Graph Code.

Der Prozess der Verarbeitung der Informationen eines Graph Codes wird, wie bereits durch das Sequenzdiagramm für den Anwendungsfall UC-1.3 dargestellt, von

einem Benutzer durch die Auswahl eines Graph Codes initiiert. Schlussendlich wird dann der Komponente *ExplanationPanel* ein Graph Code übergeben, welche diesen Graph Code dann an die spezifischen Komponenten *ImagePanel* und *TextPanel* der Erklärungstypen *Image* oder *Text* delegiert. Die Verarbeitung der Informationen kann dann gesondert und abhängig vom Erklärungstyp durch die Komponenten *TextPanel* oder *ImagePanel* erfolgen.

Mit diesen neuen Informationen kann eine Erweiterung bzw. Ergänzung des Sequenzdiagramms für den Anwendungsfall UC-1.3 vorgenommen werden. Abb. 3.17 zeigt die Erweiterung des Sequenzdiagramms um die neuen Abläufe der Verarbeitung von Graph Codes in den Komponenten *ImagePanel* und *TextPanel*.

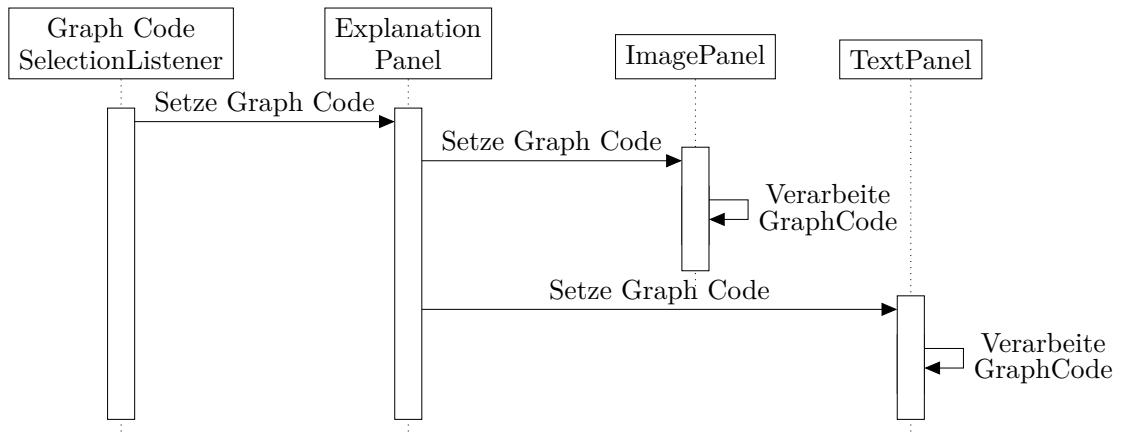


Abb. 3.17: Erweiterung zum Sequenzdiagramm für den Anwendungsfall UC-1.3.

Weitere Parameter zum Anpassen einer Anfrage an einen Endpunkt der Schnittstelle können durch die Benutzerschnittstellen, dargestellt in Abb. 3.4, generiert werden. Genauer würden die Spezifikationen der Anfragen an die Endpunkte der Schnittstelle durch die gewählten Einstellungen im Rahmen des Anwendungsfall UC-1.7 in den erweiterten Optionen (siehe Abb. 3.4 (9)) zustande kommen.

Anhand dieser Informationen kann nun auch ein Mechanismus für den Anwendungsfall UC-1.8 erstellt werden (siehe Abb. 3.18). Der Anwendungsfall UC-1.8 „Erklärung generieren“ beginnt durch das Klicken des Knopfes „Generate ...“. Im weiteren Verlauf erzeugt das System dann die notwendigen Komponenten für eine Anfrage an einen Endpunkt der Schnittstelle. Diese Objekte werden dann durch die in den erweiterten Optionen gewählten Einstellungen parametrisiert und die Anfrage finalisiert. Schlussendlich übergibt das System die Anfrage an die Komponente *OpenAiService*, welche die Anfrage dann an die Schnittstelle sendet und auf diese Weise eine Erklärung generieren lässt. Es ist wichtig anzumerken, dass diese Beschreibung für das Generieren von Erklärungen sehr generisch ist und für beide Erklärungstypen gilt.

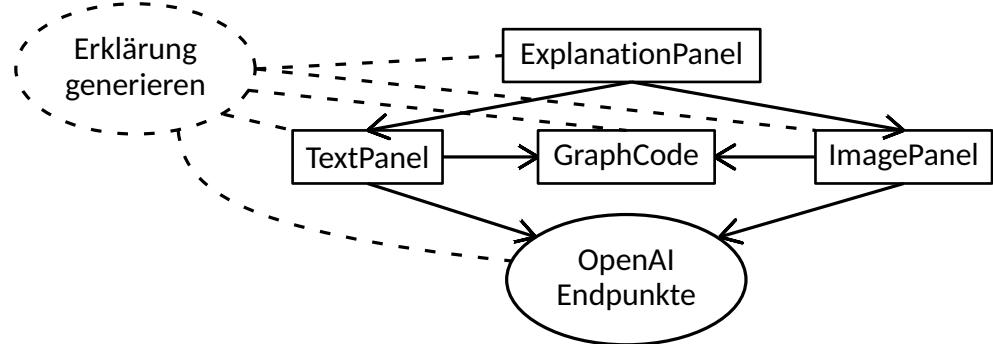


Abb. 3.18: Mechanismus für den Anwendungsfall UC-1.8.

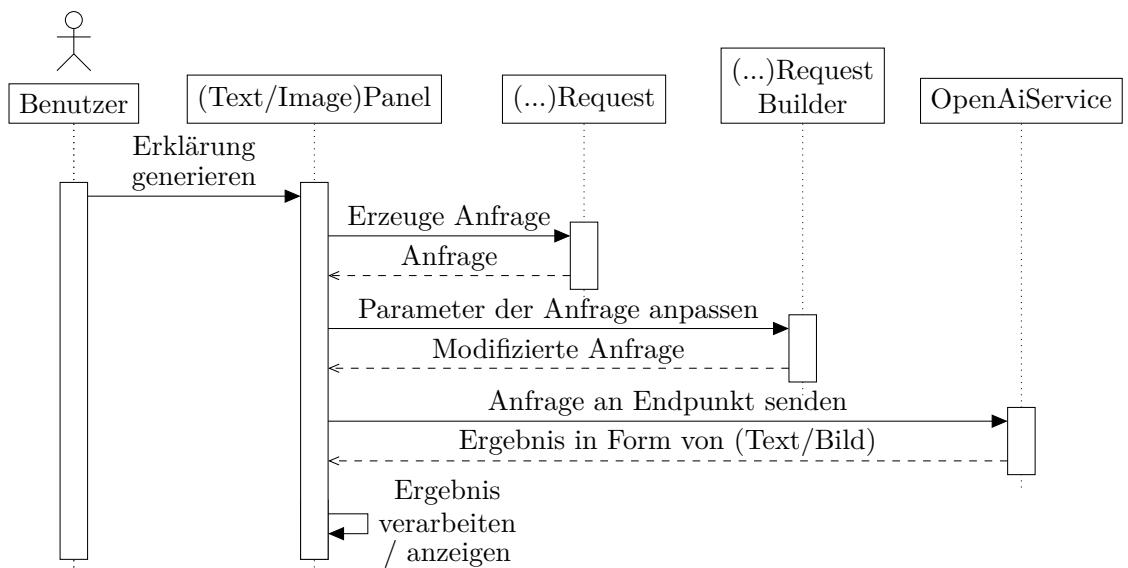


Abb. 3.19: Sequenzdiagramm für den Anwendungsfall UC-1.8.

Abb. 3.19 zeigt ein Sequenzdiagramm für den Anwendungsfall UC-1.8. Es zeigt, wie der Benutzer die Aktion „Erklärung generieren“ durch einen Klick auf den Knopf „Generate ...“ initiiert. Die Abbildung stellt dabei eine generische Darstellung eines Ablaufs zum Generieren einer Erklärung dar und gilt für beide Komponenten *Image-Panel* und *TextPanel*. Im weiteren Verlauf wird dann die notwendigen Komponenten, die für eine Anfrage benötigt werden, erzeugt und anhand von Benutzereingaben parametrisiert bzw. deren Parameter modifiziert. Hierzu gehören im Übrigen auch die Eingabeaufforderungen / Prompt, die die bereits verarbeiteten Informationen aus einem Graph Code enthalten. Diese angepasste Anfrage kann dann über den Endpunkt an die Schnittstelle gesendet werden. Nach Fertigstellung der Verarbeitung seitens des Systems generativer KI, wird über die Schnittstelle ein Ergebnis, abhängig vom Erklärungstyp, zurückgeliefert. Dieses Ergebnis kann dann schlussendlich

von der Komponente *ImagePanel* oder *TextPanel* weiterverarbeitet oder angezeigt werden.

### 3.3.3 Diskussion

In diesem Abschnitt wurden die offenen Herausforderungen **OH 2.1** und **OH 2.2** adressiert. In Abschnitt 3.3.1 wurde die offene Herausforderung **OH 2.1** behandelt und Konzepte zur Transformation von Graph Codes vorgestellt, mit welchen Graph Codes in eine geeignete Form zur Eingabe in Systeme generativer KI transformiert werden sollen. In Abschnitt Abschnitt 3.3.2 wurde die offene Herausforderung **OH 2.2** behandelt und Konzepte zur Integration von Systemen generativer KI vorgestellt.

## 3.4 Zusammenfassung

In diesem Kapitel wurde die Modellierung behandelt und es wurden in Abschnitt 3.2 Konzepte für die Erklärbarkeit von MMIR mittels generativer KI, sowie in Abschnitt 3.3 Konzepte für die Integration von generativer KI in das GMAF entwickelt. In Abschnitt 3.4.1 werden die in diesem Kapitel gewonnenen Erkenntnisse zusammengefasst, in Abschnitt 3.4.2 wird das weitere Vorgehen festgehalten und schlussendlich in Tab. 3.3 eine Übersicht des aktuellen Arbeitstands in einer Tabelle dargestellt.

### 3.4.1 Gewonnene Erkenntnisse

Forschungsziel vom Typ Modellierung	Wesentliche Arbeitsergebnisse der Modellierung
Forschungsziel 1.2: Erklärbarkeit von MMIR mittels generativer KI	Test12
Forschungsziel 2.2: Integration von generativer KI in das GMAF	Test22

Tabelle 3.2: Wesentliche Arbeitsergebnisse der Modellierung

Die in den Abschnitte 3.2 und 3.3 gewonnenen Erkenntnisse münden in folgendem Klassendiagramm, in welchem die identifizierten und für die Modellierung wichtigen Klassen, sowie die Beziehungen in denen sie zueinander stehen übersichtlich dargestellt werden.

### 3.4.2 Weiteres Vorgehen

Im folgenden Kapitel werden die in diesem Kapitel entwickelten Konzepte im Rahmen einer Implementierung durch entsprechende Forschungsziele untersucht. Genauer werden in Kapitel 4 in FZ 1.3/I Benutzerschnittstellen und Funktionen implementiert und in FZ 2.3/I die Integration der Systeme generativer KI, sowie die Überführung der Informationen von und um Graph Codes zur Generierung von Erklärungen umgesetzt.

Tabelle 3.3: Tabelle zur Übersicht des aktuellen Arbeitstands.

<b>Forschungsziele</b>		
<b>FZ / OH</b>	<b>Beschreibung</b>	<b>Referenz</b>
<b>Erklärbarkeit von MMIR mittels generativer KI</b>		
FZ 1.1/O	Recherche zur Erklärbarkeit von MMIR mittels generativer KI Grundlegende Technologien: <ul style="list-style-type: none"> <li>• GMAF</li> <li>• MMFG</li> <li>• Graph Code</li> </ul>	Abschnitt 2.1.1 Abschnitt 2.1.2 Abschnitt 2.1.3
OH 1.1	Erste offene Herausforderung Systeme generativer KI und ein Überblick über aktuelle Systeme Diskussion und Auswahl von Systemen	<b>OH 1.1</b> Abschnitt 2.1.4
OH 1.2	Zweite offene Herausforderung	<b>OH 2.1</b> Abschnitt 2.1.5
FZ 1.2/TB	Modellierung der Erklärbarkeit von MMIR mittels generativer KI	
FZ 1.3/I	Implementierung der Erklärbarkeit von MMIR mittels generativer KI	
FZ 1.4/E	Evaluierung der Erklärbarkeit von MMIR mittels generativer KI	
<b>Integration generativer KI in das GMAF</b>		
FZ 2.1/O	Recherche zur Integration generativer KI in das GMAF Aufzeigen der Integrationsmöglichkeiten von: <ul style="list-style-type: none"> <li>• Graph Codes</li> </ul>	Abschnitt 2.2.1
	Erste offene Herausforderung <ul style="list-style-type: none"> <li>• Systemen generativer KI</li> </ul>	<b>OH 2.1</b>
	Zweite offene Herausforderung	Abschnitt 2.2.2 <b>OH 2.2</b>
FZ 2.2/TB	Modellierung der Integration generativer KI in das GMAF	

<b>Forschungsziele</b>		
<b>FZ / OH</b>	<b>Beschreibung</b>	<b>Referenz</b>
FZ 2.3/I	Implementierung der Integration generativer KI in das GMAF	
FZ 2.4/E	Evaluierung der Integration generativer KI in das GMAF	

## 4 Implementierung

Dieses Kapitel widmet sich der Implementierung und es werden, entsprechend der Methodik von Nunamaker [25], alle Forschungsziele der Implementierungsphase behandelt. Ziel der Implementierung ist die Validierung der in Kapitel 3 beschriebenen Modellierung, und die Vorbereitung einer detaillierten Evaluierung in der Experimentphase, sowie Untersuchungen weiterer Beobachtungen, die während der Implementierung gemacht wurden. Die Struktur dieses Kapitels folgt der logischen Abfolge der Forschungsziele der Implementierungsphase. Doch zuvor werden in Abschnitt 4.1 die Grundlagen bzw. Vorbereitungen für die Implementierung getroffen. Dies umfasst die Beschreibung der geplanten Struktur und Architektur der Implementierung, sowie weitere technische Voraussetzungen und die in dieser prototypischen Implementierung genutzten Bibliotheken bzw. Abhängigkeiten. In Abschnitt 4.2 werden ausgewählte Implementierungsdetails der Erklärbarkeit von MMIR mittels generativer KI vorgestellt. Analog dazu werden in Abschnitt 4.3 ausgewählte Implementierungsdetails der Integration generativer KI in das GMAF vorgestellt.

### 4.1 Grundlagen der Implementierung

In diesem Abschnitt werden die Grundlagen für die in diesem Kapitel beschriebene prototypische Proof-of-Concept Implementierung festgehalten. Dies umfasst die allgemeine Architektur, ihr zugrunde liegendes Prinzip, ihr hierarchischer Aufbau, sowie die Abhängigkeiten zu Drittanbieter-Bibliotheken.

#### *Architekturmuster*

Die Implementierung der Software folgt dem Muster des *Model-View-Controller*-Prinzips und teilt die Architektur der Software in drei separate Bereiche: *Model* für die Datenstrukturen, *View* für das Frontend bzw. die Benutzerschnittstellen und *Controller* zum Handhaben von Ereignissen und Interaktionen zwischen den Bereichen *Model* und *View*. Ziel dieser Architektur ist durch das gezielte Trennen dieser Bereiche die Eigenschaften wie Wiederverwendbarkeit, Wartbarkeit und Robustheit der Software zu erhöhen und zu stärken.

Das erste Forschungsziel FZ 1.3/I widmet sich ... und deckt somit die Bereiche *View*, *Controller* ab. Das zweite Forschungsziel FZ 2.3/I widmet sich ... und deckt somit den Bereich *Model* ab.

Abb. 4.1 zeigt die Hierarchie des Quellcodes der in diesem Kapitel beschriebenen prototypischen Proof-of-Concept Implementierung.

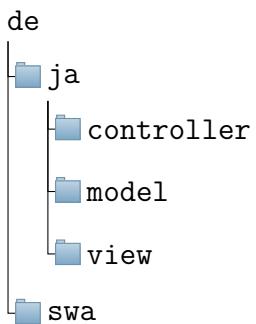


Abb. 4.1: Hierarchie des Quellcodes.

### ***Abhängigkeiten (Maven)***

Für die prototypische Proof-of-Concept Implementierung der in Kapitel 3 vorgestellten konzeptuellen Modellierung wird die Programmiersprache Java [130] ausgewählt. Der in diesem Kapitel vorgestellte Quellcode ist auf GitHub verfügbar. Die Implementierung in Java baut hierbei auf Maven [131], einem Build-Tool der Apache Software Foundation für die Projektverwaltung und Management von Abhängigkeiten, auf.

## **4.2 FZ 1.2/TB Erklärbarkeit von MMIR mittels generativer KI**

### **4.3 FZ 2.3/I Integration generativer KI in das GMAF**

## **4.4 Zusammenfassung**

## **5 Evaluierung**

## **A Anhang**

## Literatur

- [1] BENDEL, O. „Digitalisierung • Definition | Gabler Wirtschaftslexikon.“ Springer Nature. (2021), Verfügbar: <https://wirtschaftslexikon.gabler.de/definition/digitalisierung-54195/version-384620> (Zugriff am 29.08.2023).
- [2] VAILSHERY, L. S. „Number of IoT connected devices worldwide 2019-2021, with forecasts to 2030.“ (2022), Verfügbar: <https://www-statista-com.ub-proxy.fernuni-hagen.de/statistics/1183457/iot-connected-devices-worldwide/> (Zugriff am 29.08.2023).
- [3] RABE, L. „Anzahl der Internetnutzer weltweit bis 2022.“ ITU. (2022), Verfügbar: <https://de-statista-com.ub-proxy.fernuni-hagen.de/statistik/daten/studie/805920/umfrage/anzahl-der-internetnutzer-weltweit/> (Zugriff am 29.08.2023).
- [4] McKinsey & Company. „How COVID-19 has pushed companies over the technology tipping point—and transformed business forever.“ SEILER, D., Hrsg. (2020), Verfügbar: <https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/how-covid-19-has-pushed-companies-over-the-technology-tipping-point-and-transformed-business-forever> (Zugriff am 29.08.2023).
- [5] RABE, L. „Wie wirkt sich das Coronavirus auf die Nutzung digitaler Medien aus?“ (2022), Verfügbar: <https://de.statista.com/themen/6289/auswirkungen-des-coronavirus-covid-19-auf-digitale-medien/#topicOverview> (Zugriff am 29.08.2023).
- [6] Instagram.com. „Instagram.“ (2023), Verfügbar: <https://www.instagram.com> (Zugriff am 29.08.2023).
- [7] Facebook.com. „Facebook.“ (2023), Verfügbar: <https://www.facebook.com> (Zugriff am 29.08.2023).
- [8] Snapchat.com. „Snapchat.“ (2023), Verfügbar: <https://www.snapchat.com> (Zugriff am 29.08.2023).
- [9] TAYLOR, P. „Mobile data usage worldwide, by device 2020-2027.“ (2023), Verfügbar: <https://www-statista-com.ub-proxy.fernuni-hagen.de/statistics/1370201/global-mobile-data-usage/> (Zugriff am 29.08.2023).
- [10] KWAŚNICKA, H. und JAIN, L. C., *Bridging the Semantic Gap in Image and Video Analysis*. Berlin, Germany: Springer Cham, 2018, ISBN: 978-3-319-73891-8. DOI: <https://doi.org/10.1007/978-3-319-73891-8>. (Zugriff am 29.08.2023).

- [11] ANYOHA, R. „The History of Artificial Intelligence.“ (2017), Verfügbar: <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/> (Zugriff am 29.08.2023).
- [12] HEINRICH, L. „Anwendungsgebiete von künstlicher Intelligenz.“ (2022), Verfügbar: <https://datasolut.com/anwendungsgebiete-von-kuenstlicher-intelligenz/> (Zugriff am 29.08.2023).
- [13] Edley. „Artificial Intelligence vs. Machine Learning vs. Deep Learning.“ (2023), Verfügbar: <https://edley.de/insights/unterschied-artificial-intelligence-machine-deep-learning/> (Zugriff am 29.08.2023).
- [14] BENDEL, O. „Generative KI • Definition | Gabler Wirtschaftslexikon.“ (2023), Verfügbar: <https://wirtschaftslexikon.gabler.de/definition/generative-ki-124952/version-388556> (Zugriff am 29.08.2023).
- [15] McKinsey & Company, „What is generative AI?“ 2023. Verfügbar: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai> (Zugriff am 29.08.2023).
- [16] OpenAI. „Introducing ChatGPT.“ (2022), Verfügbar: <https://openai.com/blog/chatgpt> (Zugriff am 29.08.2023).
- [17] OpenAI. „DALL · E 2.“ (2022), Verfügbar: <https://openai.com/product/dall-e-2> (Zugriff am 29.08.2023).
- [18] WAGENPFEIL, S.: „Smart Multimedia Information Retrieval“. Fernuniversität in Hagen, Hagen, 2022. Fernuniversität in Hagen, Fakultät für Mathematik und Informatik, Dissertation, 2022. DOI: 10.18445/20220718-214959-0. Verfügbar: <https://doi.org/10.18445/20220718-214959-0> (Zugriff am 29.08.2023).
- [19] MAZZONI, D. „Using AI to give people who are blind the ‚full picture‘.“ (2019), Verfügbar: <https://blog.google/outreach-initiatives/accessibility/get-image-descriptions/> (Zugriff am 29.08.2023).
- [20] Lego Group. „Easter Chick.“ (2019), Verfügbar: <https://zusammengebaut.com/anleitung-jetzt-dieses-knuffige-lego-kueken-nachbauen-68584/> (Zugriff am 29.08.2023).
- [21] HEUWERS, I. „Schmetterling.“ (2019), Verfügbar: <https://www.gedichte-oase.de/gedicht/kleiner-schmetterling> (Zugriff am 29.08.2023).
- [22] WAGENPFEIL, S., P. MCKEVITT und M. HEMMJE, „Smart Multimedia Information Retrieval,“ *Analytics*, Jg. 2, Nr. 1, S. 198–224, 2023, ISSN: 2813-2203. DOI: 10.3390/analytics2010011. Verfügbar: <https://www.mdpi.com/2813-2203/2/1/11> (Zugriff am 29.08.2023).

- [23] KFZ Serviceportal. „Sicherheitsrelevante Bauteile an Kraftfahrzeugen.“ (2023), Verfügbar: <https://kfz-serviceportal.de/ratgeber/sicherheitsrelevante-bauteile-am-kfz/> (Zugriff am 29.08.2023).
- [24] Kia Austria GmbH. „Kia-Stinger-Preisliste.pdf.“ (2018), Verfügbar: <https://www.kia.com/content/dam/kwcmss/kme/at/de/assets/contents/utility/Pricelist/Kia-Stinger-Preisliste.pdf> (Zugriff am 29.08.2023).
- [25] J. NUNAMAKER JR., M. CHEN und T. PURDIN, „Systems Development in Information Systems Research.,“ *Journal of Management Information Systems*, Jg. 7, Nr. 3, S. 89–106, 1991, ISSN: 0742-1222. DOI: 10.1080/07421222.1990.11517898. (Zugriff am 29.08.2023).
- [26] NORMAN, D. A. und DRAPER, S. W., *User Centered System Design - New Perspectives on Human-Computer Interaction*. USA: L. Erlbaum Associates Inc., 1986, ISBN: 978-0-898-59872-8. (Zugriff am 29.08.2023).
- [27] M. Fowler, *UML Distilled - A Brief Guide to the Standard Object Modeling Language*. Addison-Wesley Professional, Boston, ISBN: 978-0-321-19368-1. (Zugriff am 29.08.2023).
- [28] WAGENPFEIL, S. „Github Repository of GMAF and MMFVG.“ (2023), Verfügbar: <https://github.com/stefanwagenpfeil/GMAF> (Zugriff am 29.08.2023).
- [29] WAGENPFEIL, S., F. ENGEL, P. MCKEVITT und M. HEMMJE, „AI-Based Semantic Multimedia Indexing and Retrieval for Social Media on Smartphones,“ *Information*, Jg. 12, Nr. 1, 2021, ISSN: 2078-2489. DOI: 10.3390/info12010043. Verfügbar: <https://www.mdpi.com/2078-2489/12/1/43> (Zugriff am 29.08.2023).
- [30] WAGENPFEIL, S., P. MCKEVITT, A. CHEDDAD und M. HEMMJE, „Explainable Multimedia Feature Fusion for Medical Applications,“ *Journal of Imaging*, Jg. 8, Nr. 4, 2022, ISSN: 2313-433X. DOI: 10.3390/jimaging8040104. Verfügbar: <https://www.mdpi.com/2313-433X/8/4/104> (Zugriff am 29.08.2023).
- [31] WAGENPFEIL, S., B. VU, P. MCKEVITT und M. HEMMJE, „Fast and Effective Retrieval for Large Multimedia Collections,“ *Big Data and Cognitive Computing*, Jg. 5, Nr. 3, 2021, ISSN: 2504-2289. DOI: 10.3390/bdcc5030033. Verfügbar: <https://www.mdpi.com/2504-2289/5/3/33> (Zugriff am 29.08.2023).
- [32] WAGENPFEIL, S., P. MCKEVITT und M. HEMMJE, „Towards Automated Semantic Explainability of Multimedia Feature Graphs,“ *Information*, Jg. 12, Nr. 12, 2021, ISSN: 2078-2489. DOI: 10.3390/info12120502. Verfügbar: <https://www.mdpi.com/2078-2489/12/12/502> (Zugriff am 29.08.2023).

- [33] W3C.org. „W3C Semantic Web Activity.“ (2001), Verfügbar: <https://www.w3.org/2001/sw/> (Zugriff am 29.08.2023).
- [34] R. R. HAUSSER, *Grundlagen der Computerlinguistik: Mensch-Maschine-Kommunikation in natürlicher Sprache*. Springer Berlin Heidelberg, 2013, ISBN: 978-3-64257-306-4. Verfügbar: [https://books.google.de/books?id=FW\\_zBQAAQBAJ](https://books.google.de/books?id=FW_zBQAAQBAJ) (Zugriff am 29.08.2023).
- [35] WAGENPFEIL, S., F. ENGEL, P. MCKEVITT und M. HEMMJE, „Graph Codes-2D Projections of Multimedia Feature Graphs for Fast and Effective Retrieval,“ *International Journal of Computer and Information Engineering*, Jg. 15, Nr. 12, S. 636–648, 2021, ISSN: eISSN: 1307-6892. Verfügbar: <https://publications.waset.org/vol/180> (Zugriff am 29.08.2023).
- [36] G. LAWTON. „What is generative AI? Everything you need to know.“ (2023), Verfügbar: <https://www.techtarget.com/searchenterpriseai/definition/generative-AI> (Zugriff am 29.08.2023).
- [37] K. MARTINEAU. „What is synthetic data?“ (2023), Verfügbar: <https://research.ibm.com/blog/what-is-synthetic-data> (Zugriff am 29.08.2023).
- [38] R. RIEMANN. „Synthetic Data.“ (2023), Verfügbar: [https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data\\_en](https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en) (Zugriff am 29.08.2023).
- [39] MOSTLY AI. „What is synthetic data?“ (2023), Verfügbar: <https://mostly.ai/synthetic-data/what-is-synthetic-data> (Zugriff am 29.08.2023).
- [40] J. FRUHLINGER. „Was ist Generative AI?“ (2023), Verfügbar: <https://www.computerwoche.de/a/was-ist-generative-ai,3614061> (Zugriff am 29.08.2023).
- [41] D. FOSTER, *Generatives Deep Learning*. O'Reilly, 2020, [https://www.assets.dpunkt.de/leseproben/13535/2\\_GenerativeModellierung.pdf](https://www.assets.dpunkt.de/leseproben/13535/2_GenerativeModellierung.pdf), ISBN: 978-3-96009-128-8. Verfügbar: <https://dpunkt.de/produkt/generatives-deep-learning/> (Zugriff am 29.08.2023).
- [42] Turing. „Generative Models vs Discriminative Models: Which One to Choose?“ (2023), Verfügbar: <https://www.turing.com/kb/generative-models-vs-discriminative-models-for-deep-learning> (Zugriff am 29.08.2023).
- [43] K. RONKOWITZ. „Eliza, a chatbot therapist.“ (2017), Verfügbar: <https://web.njit.edu/~ronkowit/eliza.html> (Zugriff am 29.08.2023).

- [44] M. DWHITE. „A Brief History of Generative AI.“ (2023), Verfügbar: [http://matthewdwhite.medium.com/a-brief-history-of-generative-ai-cb1837e67106](https://matthewdwhite.medium.com/a-brief-history-of-generative-ai-cb1837e67106) (Zugriff am 29.08.2023).
- [45] L. LUBER, S. „Was ist Machine Learning?“ (2016), Verfügbar: <https://www.bigdata-insider.de/was-ist-machine-learning-a-592092/> (Zugriff am 29.08.2023).
- [46] „Machine Learning: Definition, Algorithmen, Methoden und Beispiele.“ (2022), Verfügbar: <https://datasolut.com/was-ist-machine-learning/> (Zugriff am 29.08.2023).
- [47] S. Ray, „A Quick Review of Machine Learning Algorithms,“ in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 2019, S. 35–39. DOI: [10.1109/COMITCon.2019.8862451](https://doi.org/10.1109/COMITCon.2019.8862451).
- [48] IBM. „What is unsupervised learning?“ (2023), Verfügbar: <https://www.ibm.com/topics/unsupervised-learning> (Zugriff am 29.08.2023).
- [49] IBM. „Was ist Deep Learning?“ (2023), Verfügbar: <https://www.ibm.com/de-de/topics/deep-learning> (Zugriff am 29.08.2023).
- [50] IBM Data and AI Team. „AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What’s the difference?“ (2023), Verfügbar: <https://www.ibm.com/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks/> (Zugriff am 29.08.2023).
- [51] WUTTKE, L. „Deep Learning: Definition, Beispiele & Frameworks.“ (2023), Verfügbar: <https://datasolut.com/was-ist-deep-learning/> (Zugriff am 29.08.2023).
- [52] I. J. Goodfellow u. a., *Generative Adversarial Networks*, 2014. Verfügbar: <https://arxiv.org/abs/1406.2661>.
- [53] D. P. Kingma und M. Welling, „An Introduction to Variational Autoencoders,“ *Foundations and Trends® in Machine Learning*, Jg. 12, Nr. 4, S. 307–392, 2019. DOI: [10.1561/2200000056](https://doi.org/10.1561/2200000056). Verfügbar: <https://doi.org/10.1561%2F2200000056> (Zugriff am 29.08.2023).
- [54] Nvidia. „Large Language Models Explained.“ (2023), Verfügbar: <https://www.nvidia.com/en-us/glossary/data-science/large-language-models/> (Zugriff am 29.08.2023).
- [55] A. Vaswani u. a., *Attention Is All You Need*, 2017. Verfügbar: <https://arxiv.org/abs/1706.03762>.
- [56] NABI, J. „In-Context Learning Approaches in Large Language Models.“ (2023), Verfügbar: <https://towardsdatascience.com/in-context->

learning - approaches - in - large - language - models - 9c0c53b116a1 (Zugriff am 29.08.2023).

- [57] Kobold AI. „Was ist Explainable AI (XAI)? Alles was man wissen muss.“ (2023), Verfügbar: <https://www.kobold.ai/explainable-ai/> (Zugriff am 29.08.2023).
- [58] GAO, T., „Prompting: Better Ways of Using Language Models for NLP Tasks,“ *The Gradient*, 2021. Verfügbar: <https://thegradient.pub/prompting/> (Zugriff am 29.08.2023).
- [59] Hopsworks. „Prompt Engineering.“ (2023), Verfügbar: <https://www.hopsworks.ai/dictionary/prompt-engineering> (Zugriff am 29.08.2023).
- [60] SkimAI. „Understanding the Role of Prompts in ChatGPT and other Language Models.“ (2023), Verfügbar: <https://skimai.com/role-of-prompts-in-gpt-and-ai-language-models/> (Zugriff am 29.08.2023).
- [61] DAIR.AI. „Basics of Prompting.“ (2023), Verfügbar: <https://www.promptingguide.ai/introduction/basics> (Zugriff am 29.08.2023).
- [62] DAIR.AI. „Prompt Engineering Guide.“ (2023), Verfügbar: <https://www.promptingguide.ai/> (Zugriff am 29.08.2023).
- [63] DAIR.AI. „Zero-Shot Prompting.“ (2023), Verfügbar: <https://www.promptingguide.ai/techniques/zeroshot> (Zugriff am 29.08.2023).
- [64] DAIR.AI. „Few-Shot Prompting.“ (2023), Verfügbar: <https://www.promptingguide.ai/techniques/fewshot> (Zugriff am 29.08.2023).
- [65] DAIR.AI. „Chain-of-Thought Prompting.“ (2023), Verfügbar: <https://www.promptingguide.ai/techniques/cot> (Zugriff am 29.08.2023).
- [66] J. Wei u. a., *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models*, 2023. Verfügbar: <https://arxiv.org/abs/2201.11903>.
- [67] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo und Y. Iwasawa, *Large Language Models are Zero-Shot Reasoners*, 2023. Verfügbar: <https://arxiv.org/abs/2205.11916>.
- [68] McKinsey & Company. „The economic potential of generative AI: The next productivity frontier.“ (2023), Verfügbar: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#key-insights> (Zugriff am 29.08.2023).
- [69] THORMUNDSSON, B. „Artificial intelligence (AI) market size worldwide in 2021 with a forecast until 2030.“ (2023), Verfügbar: <https://www.statista.com/statistics/1365145/artificial-intelligence-market-size/> (Zugriff am 29.08.2023).

- [70] OpenAI. „OpenAI.“ (2023), Verfügbar: <https://openai.com/> (Zugriff am 29.08.2023).
- [71] StabilityAI. „Stable Diffusion — Stability AI.“ (2023), Verfügbar: <https://stability.ai/stablediffusion> (Zugriff am 29.08.2023).
- [72] StabilityAI. „Stability AI.“ (2023), Verfügbar: <https://stability.ai/> (Zugriff am 29.08.2023).
- [73] Midjourney. „Midjourney Documentation and User Guide.“ (2023), Verfügbar: <https://docs.midjourney.com/> (Zugriff am 29.08.2023).
- [74] OpenAI. „Models - OpenAI API.“ (2023), Verfügbar: <https://platform.openai.com/docs/models> (Zugriff am 29.08.2023).
- [75] Google. „PaLM 2 Technical Repor.“ (2023), Verfügbar: <https://ai.google/static/documents/palm2techreport.pdf> (Zugriff am 29.08.2023).
- [76] GHAHRAMANI, Z. „Introducing PaLM 2.“ (2023), Verfügbar: <https://blog.google/technology/ai/google-palm-2-ai-large-language-model/> (Zugriff am 29.08.2023).
- [77] Google. „Bard.“ (2023), Verfügbar: <https://bard.google.com/> (Zugriff am 29.08.2023).
- [78] ElevenLabs Inc. „ElevenLabs - Generative AI Text to Speech & Voice Cloning.“ (2023), Verfügbar: <https://elevenlabs.io/> (Zugriff am 29.08.2023).
- [79] Resemble.AI. „AI Voice Generator with Text-to-Speech and Speech-to-Speech.“ (2023), Verfügbar: <https://www.resemble.ai/> (Zugriff am 29.08.2023).
- [80] Suno.AI. „Text-Prompted Generative Audio Model.“ (2023), Verfügbar: <https://github.com/suno-ai/bark> (Zugriff am 29.08.2023).
- [81] Runway AI, Inc. „Runway - Advancing creativity with artificial intelligence.“ (2023), Verfügbar: <https://runwayml.com/> (Zugriff am 29.08.2023).
- [82] Synthesia. „Synthesia | #1 AI Video Generator.“ (2023), Verfügbar: <https://www.synthesia.io/> (Zugriff am 29.08.2023).
- [83] Rephrase.ai. „Rephrase.ai: Convert Text into Engaging AI Videos in Minutes.“ (2023), Verfügbar: <https://www.rephrase.ai/> (Zugriff am 29.08.2023).
- [84] Sequoia. „Generative AI: A Creative New World.“ (2023), Verfügbar: <https://www.sequoiacap.com/article/generative-ai-a-creative-new-world/> (Zugriff am 29.08.2023).

- [85] KLINKHAMMER, D., *Künstliche Intelligenz in der Radikalisierungsforschung: Synopse von Grundlagenwissen, Weiterbildungsangeboten und Limitationen* (MOTRA-Spotlight). Wiesbaden: MOTRA-Verbund Monitoringsystem und Transferplattform Radikalisierung, 2023, Bd. 06/23, S. 13. DOI: <https://doi.org/10.57671/motra-2023006>. (Zugriff am 29.08.2023).
- [86] KUNDU, R. „Large Language Models (LLMs): Challenges, Predictions, Tutorial.“ (2023), Verfügbar: <https://www.v7labs.com/blog/large-language-models-llms#limitations-of-large-language-models> (Zugriff am 29.08.2023).
- [87] BENDEL, O. „Responsible AI • Definition | Gabler Wirtschaftslexikon.“ (2021), Verfügbar: <https://wirtschaftslexikon.gabler.de/definition/responsible-ai-123232/version-384809> (Zugriff am 29.08.2023).
- [88] Maintenance\_Mental. „Famous Cities as a Venice : r/midjourney.“ (2023), Verfügbar: [https://www.reddit.com/r/midjourney/comments/13dzenp/famous\\_cities\\_as\\_a\\_venice/](https://www.reddit.com/r/midjourney/comments/13dzenp/famous_cities_as_a_venice/) (Zugriff am 29.08.2023).
- [89] Midjourney. „Midjourney Quick Start Guide.“ (2023), Verfügbar: <https://docs.midjourney.com/docs> (Zugriff am 29.08.2023).
- [90] Midjourney. „Midjourney Subscription Plans.“ (2023), Verfügbar: <https://docs.midjourney.com/docs/plans> (Zugriff am 29.08.2023).
- [91] Midjourney. „Midjourney Terms of Service.“ (2023), Verfügbar: <https://docs.midjourney.com/docs/en/community-guidelines> (Zugriff am 29.08.2023).
- [92] StabilityAI. „Credits + Billing.“ (2023), Verfügbar: <https://platform.stability.ai/docs/getting-started/credits-and-billing> (Zugriff am 29.08.2023).
- [93] W3C. „REST.“ (2011), Verfügbar: <https://www.w3.org/2001/sw/wiki/REST> (Zugriff am 29.08.2023).
- [94] OpenAI. „Pricing.“ (2023), Verfügbar: <https://openai.com/pricing> (Zugriff am 29.08.2023).
- [95] OpenAI. „API Reference - OpenAI API.“ (2023), Verfügbar: <https://platform.openai.com/docs/api-reference> (Zugriff am 29.08.2023).
- [96] KANNING, T. „OpenAI-Java.“ (2023), Verfügbar: <https://github.com/TheoKanning/openai-java> (Zugriff am 29.08.2023).
- [97] JANG, J. „How DALL·E Credits Work.“ (2023), Verfügbar: <https://help.openai.com/en/articles/6399305-how-dall-e-credits-work> (Zugriff am 29.08.2023).

- [98] Google. „Häufig gestellte Fragen zu Bard.“ (2023), Verfügbar: <https://bard.google.com/faq> (Zugriff am 29.08.2023).
- [99] W. HUFFMAN, S. „PaLM API & MakerSuite: an approachable way to start prototyping and building generative AI applications.“ (2023), Verfügbar: <https://developers.googleblog.com/2023/03/announcing-palm-api-and-makersuite.html> (Zugriff am 29.08.2023).
- [100] Google. „Meet the PaLM API.“ (2023), Verfügbar: <https://developers.generativeai.google/products/palm> (Zugriff am 29.08.2023).
- [101] OpenAI. „Introducing ChatGPT.“ (2022), Verfügbar: <https://openai.com/blog/chatgpt> (Zugriff am 29.08.2023).
- [102] Google. „Vertex AI - Preise.“ (2023), Verfügbar: <https://cloud.google.com/vertex-ai/pricing?hl=de#europe> (Zugriff am 29.08.2023).
- [103] Google. „Public preview FAQ.“ (2023), Verfügbar: [https://developers.generativeai.google/guide/preview\\_faq](https://developers.generativeai.google/guide/preview_faq) (Zugriff am 29.08.2023).
- [104] T. B. Brown u.a., *Language Models are Few-Shot Learners*, 2020. Verfügbar: <https://arxiv.org/abs/2005.14165>.
- [105] OpenAI. „GPT-4.“ (2023), Verfügbar: <https://openai.com/research/gpt-4> (Zugriff am 29.08.2023).
- [106] BASTIAN, M. „GPT-4 hat mehr als eine Billion Parameter - Bericht Mid-journey prompted by GPT-4.“ (2023), Verfügbar: <https://the-decoder.de/gpt-4-hat-eine-billion-parameter/> (Zugriff am 29.08.2023).
- [107] OpenAI. „GPT-4 API general availability and deprecation of older models in the Completions API.“ (2023), Verfügbar: <https://openai.com/blog/gpt-4-api-general-availability> (Zugriff am 29.08.2023).
- [108] IOANNIDOU, X. „Generative AI & Large Language Models for Life Sciences.“ (2023), Verfügbar: [https://ml-architects.ch/blog\\_posts/llm\\_for\\_life\\_sciences.html](https://ml-architects.ch/blog_posts/llm_for_life_sciences.html) (Zugriff am 29.08.2023).
- [109] SabrePC. „What Causes Large Language Models to Hallucinate.“ (2023), Verfügbar: <https://www.sabrepcc.com/blog/Deep-Learning-and-AI/what-causes-large-language-models-to-hallucinate> (Zugriff am 29.08.2023).
- [110] Google. „Gson.“ (2023), Verfügbar: <https://github.com/google/gson> (Zugriff am 29.08.2023).
- [111] OpenAI. „Libraries - OpenAI API.“ (2023), Verfügbar: <https://platform.openai.com/docs/libraries/community-libraries> (Zugriff am 29.08.2023).

- [112] OpenAI. „OpenAI and Microsoft extend partnership.“ (2023), Verfügbar: <https://openai.com/blog/openai-and-microsoft-extend-partnership> (Zugriff am 29.08.2023).
- [113] OpenAI. „Introduction.“ (2023), Verfügbar: <https://platform.openai.com/docs/introduction> (Zugriff am 29.08.2023).
- [114] Postman. „2023 State of the API Report | API Technologies.“ (2023), Verfügbar: <https://www.postman.com/state-of-api/api-technologies/#api-technologies> (Zugriff am 29.08.2023).
- [115] OpenAI. „OpenAPI.yaml.“ (2023), Verfügbar: <https://github.com/openai/openai-openapi/blob/master/openapi.yaml> (Zugriff am 29.08.2023).
- [116] J., J. „ChatGPT API Transition Guide.“ (2023), Verfügbar: <https://help.openai.com/en/articles/7042661-chatgpt-api-transition-guide> (Zugriff am 29.08.2023).
- [117] OpenAI. „Tokenizer.“ (2023), Verfügbar: <https://platform.openai.com/tokenizer> (Zugriff am 29.08.2023).
- [118] OpenAI. „Model.py.“ (2023), Verfügbar: <https://github.com/openai/tiktoken/blob/main/tiktoken/model.py> (Zugriff am 29.08.2023).
- [119] OpenAI. „tiktoken.“ (2023), Verfügbar: <https://github.com/openai/tiktoken> (Zugriff am 29.08.2023).
- [120] Knuddels GmbH & Co. KG. „JTokkit.“ (2023), Verfügbar: <https://jtokkit.knuddels.de/> (Zugriff am 29.08.2023).
- [121] IEEE. „IEEE - The world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.“ (2023), Verfügbar: <https://www.ieee.org/> (Zugriff am 29.08.2023).
- [122] IEEE, „IEEE Standard for Software Reviews and Audits,“ *IEEE Std 1028-2008*, S. 1–53, 2008. DOI: 10.1109/IEEESTD.2008.4601584. (Zugriff am 29.08.2023).
- [123] Q. S. bibinitperiod W. GmbH. „RUP - Rational Unified Process | QUALITY.DE.“ (2023), Verfügbar: <https://www.quality.de/lexikon/rup-rational-unified-process/> (Zugriff am 29.08.2023).
- [124] I. F. Simon. „Rational Unified Process (RUP).“ (2023), Verfügbar: <http://www.infforum.de/themen/anwendungsentwicklung/se-rup.htm> (Zugriff am 29.08.2023).
- [125] O. M. Group. „UML.“ (2015), Verfügbar: <http://www.omg.org/spec/UML/> (Zugriff am 29.08.2023).

- [126] V. Paradigm. „What is Unified Modeling Language (UML)?“ (2023), Verfügbar: <https://www.visual-paradigm.com/guide/uml-unified-modeling-language/what-is-uml/> (Zugriff am 29.08.2023).
- [127] KREISCHE, D., „Geschäftsprozessmodellierung mit der Unified Modeling Language (UML)“, Diss., Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2004.
- [128] BELL, D. (2003), Verfügbar: <https://developer.ibm.com/articles/introduction-to-uml/> (Zugriff am 29.08.2023).
- [129] O. M. Group. „OMG | Object Management Group.“ (2023), Verfügbar: <https://www.omg.org/> (Zugriff am 29.08.2023).
- [130] Oracle. „Java | Oracle.“ (2023), Verfügbar: <https://www.java.com/de/> (Zugriff am 29.08.2023).
- [131] A. M. Project. „Maven - Welcome to Apache Maven.“ (2023), Verfügbar: <https://maven.apache.org/> (Zugriff am 29.08.2023).