RaDT
Rat für
Deutschsprachige
Terminologie

Terminologie trifft KI

Grundlagen für erfolgreiche KI-Projekte

RaDT
Rat für Deutschsprachige Terminologie

Inhaltsverzeichnis

1 Einführung 5

2 Grundlagen 7

2.1 Einleitung 7
2.2 Künstliche Intelligenz (KI) 7
2.3 Maschinelles Lernen (ML) 8
2.4 Generative KI 8
2.5 Große Sprachmodelle (Large Language Models, LLMs) 8
2.6 Agentische KI (Agentic AI) 9
2.7 Anbindung an externe Wissensquellen: Retrieval-Augmented Generation (RAG) 9
2.8 Terminologie 10
2.9 Wissen und Wissensrepräsentationen 11
2.9.1 Wissenspyramide 11
2.9.2 Wissenstripel als Basis 11
2.9.3 Knowledge Graphs 12
2.9.4 Ontologien 12
2.9.5 Semantic Web 13
2.9.6 Begriffssysteme 14
2.9.7 Gemeinsamkeiten und Unterschiede verschiedener Wissensrepräsentationen 14

3 Beziehungen zwischen KI und Terminologie 15
3.1 Die Evolution der Kommunikation: von Mensch zu Mensch hin zu von Maschine zu Maschine 15
3.2 LLMs und Terminologie: vom Neben- zum Hauptdarsteller 15
3.3 Agentische KI und Terminologie 16
3.4 RAG mit Terminologie 17

4 Terminologie als Erfolgsfaktor für KI 18
4.1 Ausgangslage 18
4.2 LLMs lernen aus Text – ohne Wissenserwerb 18
4.3 Terminologiedatenbanken: explizites Wissen in strukturierter Form 18
4.4 Terminologie als Rückgrat vernetzter Wissenssysteme 19
4.5 Die besonderen Stärken von Terminologie als RAG-Quelle 20
4.5.1 Sparsame Nutzung des Kontextraums (Token-Wirtschaftlichkeit) 20
4.5.2 Optimierte Antwortzeiten (Performanz) 20
4.5.3 Aktuelle und gepflegte Daten 20
4.5.4 Transparenz und Nachvollziehbarkeit 21
4.5.5 Datenstrukturen statt Wortlisten 21

4.6 Die Rolle von Terminologiefachleuten in KI-Projekten 21
4.6.1 Qualifikationsprofil von Terminologiefachleuten 21
4.6.2 Einsatzfelder in KI-Projekten 22
4.7 Zusammenfassung 23

5 Best Practices: KI gezielt mit Fachwissen anreichern 24
5.1 Terminologie dynamisch in LLMs bereitstellen 24
5.2 Umsetzung in der Praxis 25
5.2.1 Terminologie-Prompts 25
5.2.2 Terminologiedatenbanken 25
5.3 Warum Terminologiemanagement hilfreich ist 26

6 Fazit: gute Daten für gute Prozesse 27
6.1 Terminologie als semantisches Fundament der digitalen Wertschöpfung 27
6.2 Terminologische Kompetenz als Voraussetzung für Automation 27

7 Quellenangaben und weiterführende Literatur 28

8 Publikationen des RaDT 29

Anhang: Glossar 30

Impressum 36

1 Einführung

Die Einführung von Künstlicher Intelligenz (KI) zur Optimierung von Geschäfts- und Verwaltungsprozessen schreitet in Unternehmen wie auch in öffentlichen Institutionen mit großer Dynamik voran. Besonders im Fokus stehen dabei Anwendungen, die auf großen Sprachmodellen (Large Language Models, LLMs) basieren. Diese Modelle sind darauf spezialisiert, Sprache zu verarbeiten und zu generieren, und sie tun dies mit zunehmender Präzision und Tiefe.

Im Zentrum der Leistungsfähigkeit solcher Sprachmodelle steht ihre Fähigkeit, mit Sprache als Trägermedium von Wissen umzugehen. Ein zentrales Element hierbei ist die eindeutige, fachgerechte und konsistente Verwendung von Terminologie. Terminologie ist die strukturierende Kraft hinter erfolgreicher Kommunikation, präziser Wissensvermittlung und funktionierender Interaktion zwischen Menschen und zunehmend auch zwischen Mensch und Maschine und zwischen Maschinen.

In der Praxis verfügen viele Fachkräfte, die mit der Implementierung KI-gestützter Prozesse betraut sind, über fundiertes Wissen im Bereich des maschinellen Lernens sowie über detaillierte Kenntnisse der jeweiligen Anwendungsdomänen und Fachgebiete. Was jedoch häufig unterschätzt wird, ist die tiefgreifende Wirkung, die Sprache – insbesondere terminologische Präzision – auf die Effizienz und Qualität von KI-Anwendungen hat. Terminologie ist kein nebensächlicher Aspekt, sondern ein zentraler Hebel zur Qualitäts- und Leistungssteigerung, insbesondere bei Anwendungen mit großen Sprachmodellen.

Wer Terminologie systematisch berücksichtigt, kann an mehreren Stellschrauben gleichzeitig drehen: Prozesse beschleunigen (**Zeit**), Fehler und Nacharbeit reduzieren (**Kosten**) und die erzielte Leistung deutlich steigern (**Qualität**). Die Verbindung von Terminologiemanagement und KI ist damit ein zentrales Instrument, um das klassische Spannungsfeld zwischen Zeit, Kosten und Qualität nachhaltig zu entschärfen und Implementierungen erfolgreicher zu machen.

Diese RaDT-Broschüre richtet sich an Personen im Organisations-, Strategie- und Projektmanagement, die KI-Projekte verantworten oder begleiten, sei es in der Planungsphase oder bereits im operativen Betrieb.

Ziel ist es, das Bewusstsein für die zentrale Rolle von Sprache und Terminologie im Kontext von KI-Implementierungen zu schärfen und bislang ungenutzte Potenziale aufzuzeigen. Gleichzeitig soll die Bedeutung von Sprache und Terminologie als kritischer Erfolgsfaktor für KI-gestützte Prozesse ins Blickfeld gerückt werden, denn nicht immer ist den Beraterinnen und Beratern und Projektteams die strategische Relevanz der ihnen zur Verfügung stehenden Mittel bewusst.

Diese RaDT-Broschüre richtet sich auch an Terminologie- und weitere Sprachfachleute. Eine praxisorientierte Handreichung – etwa zur konkreten Anwendung von Sprachmodellen im Terminologiemanagement – wird Gegenstand einer weiteren Broschüre des RaDT sein.

Für alle Anwendungsszenarien sollte allerdings beachtet werden: Die Entwicklung in diesem Bereich schreitet sehr schnell voran. Modelle werden leistungsfähiger, kosteneffizienter und besser integrierbar. Was heute als „State of the Art“ gilt, kann in wenigen Monaten überholt sein.

Organisationen sollten sich daher nicht nur mit der aktuellen Technologie befassen, sondern auch Strukturen schaffen, um flexibel auf neue Entwicklungen reagieren zu können.

2 Grundlagen

2.1 Einleitung

Um ein gemeinsames Verständnis zu schaffen und die Lesbarkeit dieser RaDT-Broschüre zu fördern, werden im Folgenden zentrale Begriffe erläutert (siehe auch Glossar im Anhang). Diese Begriffe bilden die Grundlage für die nachfolgenden Ausführungen. Wer gute Vorkenntnisse in KI besitzt, kann diese Kapitel überspringen.

2.2 Künstliche Intelligenz (KI)

Künstliche Intelligenz ist eine Technologie, die menschliche kognitive Fähigkeiten imitiert, indem sie Informationen aus Eingabedaten erkennt, sortiert und so Computern und Maschinen ermöglicht, menschliches Lernen, Verständnis, Problemlösung, Entscheidungsfindung, Kreativität und Autonomie zu simulieren (vgl. IBM o. J. a und Fraunhofer IKS o. J.).

Reasoning bezeichnet die Fähigkeit eines KI-Systems, über die Simulation logischer Schlussfolgerungen zu einer Lösung zu gelangen, anstatt lediglich statistische Wahrscheinlichkeiten für das nächste Wort zu berechnen. Dabei zerlegt das Modell komplexe Problemstellungen in geordnete Zwischenschritte (**Chain-of-Thought**), überprüft diese auf Konsistenz und korrigiert sich bei Bedarf selbst. Während Standardmodelle oft musterbasierte „Bauchantworten“ geben, simuliert Reasoning einen Denkprozess, der besonders bei mathematischen, logischen oder strategischen Aufgaben die Genauigkeit und Nachvollziehbarkeit erhöht.

Planning bezeichnet die Fähigkeit eines KI-Systems, zielgerichtete Handlungsabläufe zu entwerfen, zu strukturieren und schrittweise umzusetzen. Dabei wird ein übergeordnetes Ziel in Teilziele zerlegt und in eine sinnvolle Reihenfolge von Aktionen überführt. Das System berücksichtigt verfügbare Ressourcen, zeitliche Abhängigkeiten sowie mögliche Einschränkungen und passt den Plan bei neuen Informationen dynamisch an. Planning ermöglicht eine vorausschauende Problemlösung, wie sie etwa bei mehrstufigen Aufgaben, komplexen Workflows oder strategischen Entscheidungen erforderlich ist.

Während **Reasoning** primär auf das logische Durchdringen und Lösen einzelner Problemstellungen abzielt, fokussiert **Planning** auf die strukturierte Organisation von Handlungen über mehrere Schritte hinweg, um ein Ziel effizient zu erreichen.

2.3 Maschinelles Lernen (ML)

Maschinelles Lernen ist ein Teilgebiet der KI, bei dem Systeme aus Daten lernen, Muster erkennen und auf dieser Grundlage Vorhersagen treffen, ohne explizit dafür programmiert zu sein.

LLMs (siehe 2.5) basieren auf speziellen ML-Architekturen (z. B. Transformer), die auf Sprachverarbeitung spezialisiert sind.

Deep Learning ist der Teil des maschinellen Lernens, der künstliche neuronale Netze (KNN) verwendet, um komplexe Muster in großen Datenmengen automatisch zu erkennen (vgl. TH Köln 2022).

2.4 Generative KI

Generative KI ist eine Technologie, die ausgehend von Anweisungen (Prompts) und erlernten Mustern aus Trainingsdaten neue Inhalte wie Texte, Bilder, Videos, Audio oder Code erzeugt.

2.5 Große Sprachmodelle (Large Language Models, LLMs)

Große Sprachmodelle sind KI-Modelle, die auf künstlichen neuronalen Netzen basieren und in der Lage sind, natürliche Sprache überzeugend zu verarbeiten, beispielsweise indem sie Texte generieren, übersetzen oder zusammenfassen. Derzeit bekannte Vertreter sind ChatGPT, Claude, Gemini oder Mistral. LLMs bilden das Rückgrat vieler moderner KI-Anwendungen mit sprachlichem Fokus.

Ein modernes Sprachmodell basiert auf einer Technologie, die es ermöglicht, Wörter nicht isoliert, sondern in ihrem Kontext zu erfassen. Trainiert wird das Modell mit Milliarden von sprachlichen Daten, wobei es im Kern lernt, das jeweils nächste Wort vorherzusagen (sog. *Next Token Prediction*). Anschließend wird das Sprachmodell durch gezieltes menschliches Feedback verfeinert, damit es präziser, hilfreicher und angemessener antwortet. Das Zusammenspiel dieser drei Schritte – Kontextverständnis, Training und Feintuning – bildet die Grundlage heutiger Sprachmodelle.

Nichtsdestoweniger haben LLMs keine echte Intelligenz oder ein Weltverständnis. Sie simulieren Sprache auf Basis von Wahrscheinlichkeiten, was in vielen Fällen zu erstaunlich menschlich wirkenden Antworten führt, aber auch zu Fehlern oder irreführenden Ergebnissen (Halluzinationen).

2.6 Agentische KI (Agentic AI)

Agentische KI bezeichnet eine Generation von KI, die über das reine Generieren von Inhalten hinausgeht und selbstständig Aufgaben plant, Entscheidungen trifft und Prozesse ausführt, ähnlich wie ein digitaler Assistent mit Zielorientierung und Autonomie. Agentische KI wird häufig in Form von KI-Agenten umgesetzt.

Diese Agenten sind in der Lage, komplexe Abläufe zu steuern, sich Informationen aktiv zu beschaffen (z. B. über Schnittstellen zu anderen Systemen/Daten/Wissenssammlungen) und sogar mit anderen Systemen zu kooperieren.

Ein Beispiel: Ein KI-Agent könnte in einem Unternehmen automatisch Kundenanfragen analysieren, relevante Vertragsinformationen abrufen, rechtlich korrekte Antworten generieren und diese über verschiedene Kanäle versenden – je nach Anwendungsfall sogar ohne menschliches Zutun.

2.7 Anbindung an externe Wissensquellen: Retrieval-Augmented Generation (RAG)

Für die Einbindung unternehmens- oder projektspezifischen Wissens in Large Language Models hat sich in der Praxis RAG etabliert.

RAG ist ein Architekturkonzept, das die Antwortqualität von Sprachmodellen verbessert, indem die Ergebnisse nicht nur aus den Trainingsdaten, sondern auch unter Verwendung externer Daten erzeugt werden. Der Prozess erfolgt in drei Stufen:

1. **Retrieval (Datenabruf)**: Abrufen von Informationen aus einer oder mehreren externen Datenquellen
2. **Augmentation (Augmentierung)**: Anreicherung des Prompts mit relevanten Informationsbestandteilen aus den verknüpften Datenquellen
3. **Generation (Generierung)**: Erstellung einer Antwort unter Zuhilfenahme des zusätzlichen Wissens

Somit werden externe Wissensquellen (z. B. Dokumente, Datenbanken oder weitere strukturierte Wissensbestände) in Echtzeit abgefragt und die abgerufenen Inhalte als Kontext für die Generierung von Antworten durch das Sprachmodell genutzt.

RAG wird im Folgenden als Oberbegriff für unterschiedliche Ausprägungen der Wissensanbindung verwendet.

Die konkrete Ausgestaltung der Wissensanbindung ist abhängig von fachlichen und technischen Rahmenbedingungen und wird in der Praxis an den jeweiligen Anwendungsfall angepasst. Je nach Art der zugrundeliegenden Daten sowie den Anforderungen (Aktualität, Nachvollziehbarkeit, Performanz, Echtzeitzugriff, Granularität der Information, Governance oder Sicherheits- und Zugriffsmodelle) kommen spezialisierte Varianten zum Einsatz.

2.8 Terminologie

Terminologie ist die Gesamtheit aller Benennungen, d. h. aller Fachwörter in einem Fachgebiet. Terminologie spielt in der Privatwirtschaft, in öffentlichen Einrichtungen sowie in Bildung und Wissenschaft eine wichtige Rolle. Sie fördert Effizienz und Eindeutigkeit in der Fachkommunikation. Eine saubere Verwendung von Terminologie optimiert die Unternehmenssprache (Corporate Language), vermeidet Verständnisprobleme, erhöht die Lesefreundlichkeit, ermöglicht die optimale Nutzung von Wissenspotenzialen, führt zu größerer Rechtssicherheit und spart Kosten (vgl. RaDT 2013).

Leider wird Terminologie nicht immer so eindeutig und konsistent eingesetzt, wie es wünschenswert wäre. Fachwörter werden unreflektiert verwendet. So bezeichnen z. B. „Gesteinskunde“, „Felsenkunde“, „Petrografie“ und „Petrographie“ allesamt dasselbe: die Wissenschaft von den Gesteinen. „Verkleidung“ hingegen bezeichnet sowohl das Anbringen oder Montieren einer Verkleidung, also einen Vorgang, als auch das Ergebnis des Verkleidens (vgl. RaDT 2013).

Terminologiemanagement, auch Terminologiearbeit genannt, besteht darin, Fachwortschätze nach einschlägigen Regeln zusammenzustellen, zugänglich zu machen, bei Bedarf zu vereinheitlichen und das dahinterstehende Fachwissen zu systematisieren, indem die Beziehungen zwischen den Begriffen erarbeitet und bereitgestellt werden. Die so erarbeiteten terminologischen Daten (wie Benennungen, Definitionen oder Kontextbeispiele) werden meist in Terminologiedatenbanken verwaltet und zur Verfügung gestellt (vgl. RaDT 2013).

Im Kontext von KI – insbesondere beim Einsatz großer Sprachmodelle – spielt Terminologie eine Schlüsselrolle: LLMs verarbeiten Sprache statistisch und probabilistisch. Ohne eindeutige und konsistent verwendete Benennungen steigt die Wahrscheinlichkeit für Fehlinterpretationen, ungenaue Ausgaben oder Inkonsistenzen. Terminologie ist daher nicht nur eine Frage der sprachlichen Korrektheit, sondern auch der inhaltlichen Korrektheit und damit ein strategischer Qualitätsfaktor.

Durch den Einsatz großer Sprachmodelle kommt es zu einer Erweiterung der Terminologiearbeit: Während die bisherige Zielgruppe vorwiegend Menschen waren, werden Maschinen als Zielgruppe immer wichtiger. Terminologische Daten müssen maschinenlesbar sein.

2.9 Wissen und Wissensrepräsentationen

2.9.1 Wissenspyramide

Daten, Informationen und Wissen bilden eine hierarchische Wissenspyramide:

- **Daten** sind rohe Zeichenfolgen oder Fakten ohne Kontext (z. B. die Zahl „39,5“).
- **Informationen** sind Daten, die interpretiert und in einen Kontext gesetzt werden (z. B. „39,5 Grad Celsius Körpertemperatur“).
- **Wissen** entsteht durch die Verknüpfung von Informationen durch Erfahrungen und Regeln, um Handlungen abzuleiten (z. B. „39,5 Grad Celsius Körpertemperatur beim Menschen bedeutet Fieber, ich brauche Medikamente“).

Während Daten und Informationen in der Regel maschinell verarbeitet werden können, erfordert Wissen menschliche Kontextualisierung.

2.9.2 Wissenstripel als Basis

Ein Wissenstripel stellt eine elementare Einheit strukturierten Wissens dar. Es setzt sich aus einem Subjekt, einem Prädikat und einem Objekt zusammen. In sprachlicher Form können Wissenstripel beispielsweise wie folgt ausgedrückt werden:

- Paris ist die Hauptstadt von Frankreich.
- Ein Motor ist ein Bestandteil eines Kraftfahrzeugs.

Solche Aussagen lassen sich nicht nur in natürlicher Sprache formulieren, sondern auch grafisch darstellen (siehe Abb. 1). Üblicherweise werden dafür gerichtete Kanten in Form von Pfeilen verwendet. Der Ausgangsknoten des Pfeils repräsentiert das Subjekt, der Zielknoten das Objekt. Die Beschriftung der Kante bildet das Prädikat und beschreibt die Art der Relation zwischen den beiden Begriffen.

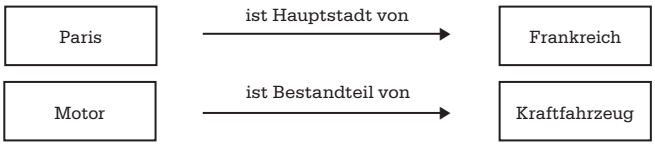


Abbildung 1: Grafische Darstellung von Wissenstripeln

Wissenstripel bilden die grundlegenden Bausteine formaler Wissensrepräsentationen wie Knowledge Graphs oder Begriffssysteme. Sie ermöglichen es, Wissen in strukturierter und maschinenlesbarer Form zu erfassen. Durch die systematische Verknüpfung vieler Wissenstripel entstehen komplexe Wissensnetze, die Zusammenhänge und Abhängigkeiten zwischen Begriffen sichtbar machen. Dies bildet die Grundlage für semantische Analysen und fördert die Interoperabilität zwischen unterschiedlichen Wissenssystemen (vgl. Drewer 2025).

2.9.3 Knowledge Graphs

Knowledge Graphs („Wissensgraphen“) sind strukturierte Formen der Wissensrepräsentation, in denen Entitäten – Begriffe oder konkrete Objekte wie Personen, Orte oder Gegenstände – sowie ihre Relationen in Form eines Netzwerks modelliert werden. Ein Knowledge Graph besteht aus Knoten sowie aus gerichteten Kanten, die die Relationen zwischen diesen Entitäten ausdrücken. Die Kanten werden in der Regel beschriftet, um die jeweilige Relationsart zu spezifizieren (vgl. IBM o. J. b). Auf diese Weise lassen sich Informationen in einer vernetzten Struktur organisieren.

Ein bekanntes Beispiel ist der Knowledge Graph von Google, der bei Suchanfragen Informationen aus verschiedenen Quellen aggregiert und zusätzliche Informationen zu den gesuchten Begriffen bereitstellt, um relevantere oder vollständigere Antworten zu liefern.

Knowledge Graphs eignen sich insbesondere für die Integration komplexer Datenbestände und die Identifikation semantischer Zusammenhänge. Sie kommen daher nicht nur bei Suchmaschinen und KI-Anwendungen zur Verbesserung der Ergebnisqualität zum Einsatz, sondern auch in Empfehlungssystemen (*recommender systems*), die kontextbezogene und personalisierte Vorschläge generieren (vgl. Drewer 2025).

2.9.4 Ontologien

Ontologien sind formale Modelle, die eine Menge von Begriffen sowie deren Beziehungen innerhalb eines definierten Wissensbereichs darstellen. Typischerweise umfassen sie Konzepte (*concepts*), die meist als Klassen (*classes*) realisiert werden, Instanzen (*instances* bzw. *individuals*), die konkrete Ausprägungen dieser Klassen darstellen, sowie Relationen (*properties*). Relationen dienen sowohl zur Beschreibung von Eigenschaften von Konzepten als auch zur Modellierung der Beziehungen zwischen den Konzepten (vgl. Studer o. J.).

Im Vergleich zu anderen Formen der Wissensrepräsentation sind Ontologien komplexer aufgebaut. Sie enthalten nicht nur eine Strukturierung von Begriffen, sondern legen zusätzlich Regeln und Einschränkungen für deren Interaktion fest. Dadurch ermöglichen sie logische Schlussfolgerungen innerhalb des Modells. Allerdings sind Ontologien in der Regel einsprachig ausgelegt. Bei der Überführung in mehrsprachige Anwendungen kann dies zu Problemen führen, z. B. bei Teiläquivalenzen oder terminologischen Lücken.

Ontologien werden insbesondere im Bereich des Semantic Web, in der Informatik und in KI-Anwendungen eingesetzt. Hier dienen sie dazu, Daten so zu modellieren, dass sie maschinell interpretierbar, austauschbar und automatisiert weiterverarbeitbar werden. Im Gegensatz zu informelleren Repräsentationen sind Ontologien in der Regel strenger definiert und basieren auf einem formalen Logiksystem (vgl. Studer o. J.).

2.9.5 Semantic Web

Das Semantic Web, auch Web 3.0 genannt, ist eine Erweiterung des World Wide Web. Im Gegensatz zum traditionellen Web, das vor allem Dokumente verlinkt, ermöglicht das Semantic Web die maschinelle Verarbeitung und Verknüpfung von Informationen nicht nur über verschiedene Quellen hinweg, sondern auch auf semantischer Ebene. Ziel ist die Schaffung eines globalen Netzes aus verknüpften, austauschbaren und vielseitig verwertbaren Daten. Dazu werden verschiedene Datenmodelle eingesetzt, die je nach Anwendungsfall ausgewählt oder kombiniert werden können, um Flexibilität, Interoperabilität und logische Konsistenz zu gewährleisten. Bei der Verwendung von Datenmodellen wird darauf geachtet, dass auch die Beschreibungen selbst eindeutig dem jeweiligen Datenmodell zugewiesen werden können und es keine Namenskonflikte gibt. Hierfür werden Namespaces eingesetzt. Ein Namespace im Semantic Web ist ein eindeutiger **Uniform Resource Identifier (URI)**, der anhand eines vorangestellten Präfixes das Datenmodell für die jeweiligen Klassen, Eigenschaften oder Ressourcen identifiziert, z. B.: @prefix owl: <http://www.w3.org/2002/07/owl#>.

In Kontext von KI und Terminologie gewinnt das Semantic Web zunehmend an Bedeutung, weil es dazu beitragen kann, die Genauigkeit von Such- und Analyseprozessen, die Erkennung von Zusammenhängen und die Vermeidung von falschen oder missverständlichen Antworten in KI-Anwendungen erheblich zu verbessern.

Zu den bekanntesten Datenmodellen des Semantic Web gehören:

- **Resource Description Framework (RDF)**
  Resource Description Framework ist das grundlegende Modell zur Beschreibung von Daten im Semantic Web. Es nutzt die in Abschnitt 2.9.2 erwähnte Tripel-Struktur (Subjekt-Prädikat-Objekt), um Zusammenhänge als Graph darzustellen, wobei die Tripel-Elemente eine eindeutige Kennung erhalten, z. B. einen URI. Während ein Prädikat immer durch einen vordefinierten URI dargestellt wird (z. B. rdf: type), um die Relation zwischen Subjekt und Objekt anzugeben, können die anderen Elemente auch durch Literalwerte oder Blank Nodes repräsentiert werden. Diese Flexibilität ist ein Hauptmerkmal von RDF, da sowohl strukturierte, global identifizierte Daten als auch lokale, individuelle Informationen einheitlich und maschinenlesbar modelliert werden können.

- **Web Ontology Language (OWL)**
  Web Ontology Language ist ein logisch fundiertes, standardisiertes Datenmodell zur Umsetzung von Ontologien im Semantic Web. OWL ist darauf spezialisiert, Wissen über die Form von Klassen, Eigenschaften und logischen Regeln (Axiomen) präzise zu beschreiben, damit Computer diese Informationen lesen („verstehen"), verknüpfen und neue Ergebnisse ableiten können. OWL eignet sich als Grundlage für verschiedenste Knowledge-Graph-Anwendungen und kann KI dabei unterstützen, die Korrektheit bzw. Wahrheit von Informationen zu verifizieren, weil durch logische Schlussfolgerungen (Inferenz) Widersprüche aufgedeckt werden können. Dies zeigt das nachfolgende Beispiel:

- Ein medizinisches LLM sagt: „Ein Patient mit Diabetes hat immer eine Nierenschädigung".

- Ein OWL-Modell sagt: „Diabetes kann eine Nierenschädigung verursachen, aber nicht immer."

OWL korrigiert die falsche, zu allgemeine Aussage des LLMs.

- **Simple Knowledge Organization System (SKOS)**
  Simple Knowledge Organization System baut auf RDF auf und ermöglicht die maschinenlesbare Darstellung von Wissensorganisationssystemen wie Thesauri, Taxonomien und Klassifikationen. Mit SKOS lassen sich Begriffe (Konzepte) und deren hierarchische (z. B. „broader", „narrower") und assoziative Beziehungen (z. B. „related") standardisiert definieren und systemübergreifend nutzbar machen. Insbesondere durch Mappings (z. B. über skos:exactMatch, skos:closeMatch) können Konzepte aus verschiedenen Vokabularen verknüpft und gemeinsam nutzbar gemacht werden. SKOS findet bei der Beschreibung und Verknüpfung vorhandener Wissensorganisationssysteme Anwendung, eignet sich aber wie OWL auch zum Aufbau neuer Systeme.

2.9.6 Begriffssysteme

Begriffssysteme sind ein zentraler Bestandteil der professionellen Terminologiearbeit, aber auch von Informationsarchitekturen und Wissensmanagement. Sie bieten eine strukturierte Darstellung von Begriffen und ihren Beziehungen innerhalb eines spezifischen Fach- oder Wissensgebiets. Dabei können Begriffssysteme sowohl hierarchisch als auch nichthierarchisch aufgebaut sein.

Für hierarchische Ober- und Unterbegriffe, die in einer Abstraktionsbeziehung stehen, lässt sich beispielsweise ein Wissenstripel der Form {X is_a Y} formulieren. Bei Bestandsbeziehungen entstehen Aussagen wie {X is_part_of Y}. In beiden Fällen bezeichnet die Variable X den untergeordneten und Y den übergeordneten Begriff.

2.9.7 Gemeinsamkeiten und Unterschiede verschiedener Wissensrepräsentationen

Trotz unterschiedlicher Ausprägungen und Einsatzkontexte verfolgen die genannten Formen der Wissensrepräsentation ein gemeinsames Ziel: Sie strukturieren und organisieren Wissen, um Verstehensprozesse und Kommunikation zu unterstützen. Dies gilt sowohl für die Interaktion zwischen Menschen als auch für den Austausch zwischen Mensch und Maschine oder zwischen Maschinen. Gemein ist ihnen die Annahme, dass Wissen wesentlich über Beziehungen zwischen Begriffen erfassbar ist.

Die zentralen Unterschiede liegen in Formalitätsgrad und Anwendungsschwerpunkt. Knowledge Graphs und Ontologien sind in der Regel datenorientiert und so gestaltet, dass sie automatisiert verarbeitet werden können. Begriffssysteme hingegen dienen primär der Unterstützung menschlicher Erkenntnis.

3 Beziehungen zwischen KI und Terminologie

3.1 Die Evolution der Kommunikation: von Mensch zu Mensch hin zu von Maschine zu Maschine

Traditionell dient Terminologie der präzisen Kommunikation zwischen Fachleuten, beispielsweise im Bereich Ingenieurwesen, Recht oder Medizin. Einheitliche Benennungen und Begriffsdefinitionen sorgen für Konsistenz, Klarheit, Nachvollziehbarkeit und Effizienz. In Dokumentation, Übersetzung oder Qualitätsmanagement ist Terminologie ein Mittel zur Fehlervermeidung und zur Einhaltung von Standards.

Mit der Digitalisierung und Automatisierung hat sich der Anwendungsbereich von Terminologie erweitert. Heute erstellen zunehmend Maschinen fachliche Inhalte. Zunächst unterstützten Maschinen menschliche Prozesse (z. B. in der Übersetzung). Aktuell sind wir in der Phase, in der Mensch-Maschine-Kommunikation durch dialogfähige LLMs beim Übersetzen oder in anderen Kommunikationssituationen alltäglich geworden ist.

Die nächste Stufe ist bereits absehbar: Maschine-zu-Maschine-Kommunikation, bei der Systeme eigenständig mit anderen Systemen kommunizieren, Entscheidungen vorbereiten oder Handlungsempfehlungen generieren. Damit steigt die Notwendigkeit für standardisierte, maschinenlesbare terminologische Datenbestände.

3.2 LLMs und Terminologie: vom Neben- zum Hauptdarsteller

Wie in Abschnitt 2.5 beschrieben, basieren LLMs auf riesigen Mengen unstrukturierter Textdaten. Doch obwohl diese Modelle beeindruckend „sprachmächtig“ erscheinen, fehlt ihnen oft eines: Fach- bzw. Organisationswissen. Hier kommt die Terminologie ins Spiel – nicht mehr nur als unterstützendes Element im Redaktions- oder Übersetzungsworkflow, sondern als essenzielle Wissensquelle für korrekte Antworten.

Ein besonders wirksamer Mechanismus ist die in Abschnitt 2.7 beschriebene RAG-Technologie. Dabei werden LLMs gezielt mit externem Wissen angereichert, etwa mit Glossaren, Definitionen oder Domänenwissen aus Terminologiedatenbanken. Anstatt sich nur auf ihr internes „Trainingsgedächtnis“ zu verlassen, greifen die Modelle auf kuratierte Fachinformationen zurück, was die Genauigkeit und Konsistenz der generierten Inhalte deutlich erhöht.

So wird Terminologie zu einem strategischen Asset: Unternehmen und öffentliche Einrichtungen, die hochwertige strukturierte terminologische Daten bereitstellen können, schaffen die Grundlage für zuverlässige KI-Anwendungen, sei es in Übersetzung, Wissensmanagement oder Kommunikation.

Heute wandelt sich die Nutzung von Terminologie grundlegend:

- Von statisch zu dynamisch: Terminologien werden in Wissensgraphen und Ontologien eingebettet, mit Kontextwissen verknüpft und sind automatisiert abrufbar.

- Von generell zu spezifisch: Über Schnittstellen können LLMs in Echtzeit kontextsensitiv terminologische Daten auswählen.

- Vom Compliance-Werkzeug zum Innovationsfaktor: Wer Terminologie strategisch pflegt, verbessert nicht nur die Qualität, sondern auch die Effizienz von KI-Anwendungen.

3.3 Agentische KI und Terminologie

Mit agentischer KI werden LLMs in die Lage versetzt, Aufgaben aufgrund von vorgegebenen Zielen mit Reasoning und Planning autonom zu lösen. Damit dies zuverlässig funktioniert, kommt Terminologie via RAG (vgl. Abschnitt 2.7) ins Spiel.

Terminologie via RAG ist aus folgenden Gründen wichtig für agentische KI:

- Fachliche Präzision: Agenten müssen in spezialisierten Kontexten arbeiten (z. B. Recht, Medizin, Technik). Terminologische Daten stellen sicher, dass Begriffe und Benennungen korrekt verwendet werden.

- Wissensverankerung: Über RAG können Agenten auf externe Terminologiequellen zugreifen, was Halluzinationen reduzieren kann und die Verlässlichkeit erhöht.

- Interoperabilität: In mehrsprachigen oder systemübergreifenden Anwendungen hilft Terminologie dabei, Begriffe und Benennungen konsistent und eindeutig zu verwenden – ein Muss für Maschine-zu-Maschine-Kommunikation.

- Compliance: Terminologie unterstützt die Nachvollziehbarkeit von KI-Ausgaben (generierte Texte, Chats oder Textanalysen).

Agentische KI braucht strukturiertes, domänenspezifisches Wissen als Grundlage. Genau das liefert das Terminologiemanagement: Terminologie, eingebunden über RAG-Mechanismen, wird so zum semantischen Treibstoff intelligenter, autonomer Systeme. Wer terminologische Daten bereitstellt und pflegt, gestaltet aktiv die Qualität und Vertrauenswürdigkeit von agentischer KI.

3.4 RAG mit Terminologie

Beim Zugriff auf bestehende terminologische Daten sollte beachtet werden, dass es unterschiedliche Reifegrade von terminologischen Daten gibt.

- Einfache Wortlisten oder Glossare: Diese Listen werden immer noch in Systemen der maschinellen Übersetzung (MÜ-Systeme) verwendet, um die Verwendung bestimmter Terminologie zu forcieren, d. h., um eine bestimmte Benennung anstelle eines probabilistisch vorgeschlagenen Äquivalents in eine Übersetzung einzufügen.

- Einfache Terminologiedatenbanken: Hier werden Terminologieeinträge begriffsorientiert organisiert, so dass für jeden Begriff die passenden Benennungen sowie die üblichen Zusatzinformationen wie Definitionen oder Anmerkungen – in der Regel mehrsprachig – dargestellt werden.

- Wissensbasierte Terminologieverwaltungssysteme: Diese Terminologieverwaltungssysteme können zusätzlich zu den begriffsorientierten Daten auch Relationen verwalten und (grafisch) darstellen.

Um die terminologischen Daten in KI-Anwendungen zu verwenden, können die Inhalte im einfachsten Fall exportiert und dem Zielsystem zur Verfügung gestellt werden.

Viele Systeme verfügen über Schnittstellen, damit externe Anwendungen direkt auf die terminologischen Daten zugreifen können. Dabei gibt es unterschiedliche Tiefen der Integration: Übernahme von Benennungen, Einbezug von Fachgebietsklassifikationen zur Begriffsunterscheidung, Informationen zu Relationen usw.

4 Terminologie als Erfolgsfaktor für KI

4.1 Ausgangslage

LLMs sind in der Lage, beeindruckend menschlich klingende Texte zu erzeugen, Fragen zu beantworten und komplexe sprachbezogene Aufgaben zu bearbeiten. Grundlage dafür ist das Training mit riesigen Mengen bestehender Texte, von Fachartikeln bis zu Webseiten. Doch was bedeutet das konkret für den Umgang mit Wissen aus solchen Systemen?

4.2 LLMs lernen aus Text – ohne Wissenserwerb

Beim Training erkennen LLMs statistische Muster in Sprache. Sie lernen, wie Wörter in bestimmten Kontexten vorkommen und verknüpft sind. Auf diese Weise entsteht eine Art „gefühltes" Sprachverständnis, allerdings ohne Weltwissen. Was ein LLM also z. B. über medizinische Diagnosen, rechtliche Begriffe oder technische Normen ausgibt, basiert auf Mustern aus unstrukturierten Texten. Eine logische Herleitung oder zuverlässige Verifizierung des enthaltenen Wissens findet dabei nicht statt, so dass auch die fachliche Präzision nicht gewährleistet werden kann.

4.3 Terminologiedatenbanken: explizites Wissen in strukturierter Form

Im Gegensatz zu unstrukturierten Texten ist eine Terminologiedatenbank ein strukturierter und expliziter Wissensspeicher. Sie enthält Benennungen, Definitionen, Begriffsbeziehungen sowie Kontextinformationen, die oft durch Fachleute validiert worden sind. Terminologische Daten bilden das präzise Fundament fachlicher Kommunikation und fachlichen Wissens.

Durch das Zusammenspiel von KI-Anwendungen und Terminologie ergeben sich vielfältige Vorteile:

- KI-Anwendungen können über RAG gezielt auf terminologische Daten statt nur auf zufällig gefundene Textmuster zugreifen.
- Terminologische Daten liefern eine semantische Basis – also Bedeutung – in einer Form, die von KI-Systemen verarbeitet werden kann.
- Terminologische Daten helfen, fachsprachliche Vorgaben einzuhalten und interne Wissenssysteme konsistent zu halten.
- Terminologische Daten schaffen Eindeutigkeit: Ein Beispiel für fehlende Eindeutigkeit ist die Benennung „Flügel", die in den Fachgebieten Mechanik, Musik, Fleischverarbeitung, Zoologie, Architektur usw. vorkommt und jeweils andere zielsprachliche Äquivalente erfordern kann. Terminologische Daten sind durch vielfältige Metadaten angereichert und damit in der Lage, Mehrdeutigkeiten aufzulösen, die sonst zu großen Problemen führen können (speziell im mehrsprachigen Bereich).

- Die terminologischen Daten sind in der Regel durch kontinuierliche Pflege stets aktuell und qualitätsgesichert.

- Terminologiemanagement beschränkt sich nicht auf Begriffe; es zeigt auch, wie Begriffe zueinander in Beziehung stehen. Die Darstellung erfolgt klar modelliert, z. B. in Form von Knowledge Graphs, also Wissensgraphen, in denen Konzepte und ihre Relationen logisch verknüpft sind. Solche Verknüpfungen machen Wissen für KI-Anwendungen explizit zugänglich und interpretierbar.

4.4 Terminologie als Rückgrat vernetzter Wissenssysteme

Als semantische Infrastruktur wird Terminologie zum gemeinsamen Bezugspunkt zwischen Datenquellen. Sie verknüpft technische, sprachliche und betriebliche Informationen über eine einheitliche Begriffsbasis.

Beispiele:

- Digitale Zwillinge (*digital twins*) sind virtuelle Repräsentationen von realen Objekten und Prozessen. Sie benötigen einheitliche semantische Beschreibungen ihrer Komponenten (z. B. „Weiche“, „Signal“, „Antriebseinheit“). Terminologie liefert dafür die kontrollierte Begriffsbasis, die Datenmodelle und Sensorinformationen verknüpft.

- In domänenübergreifenden Projekten fungieren terminologische Daten als *Linked Data Layer*, der Datenflüsse zwischen Organisationen oder einzelnen Organisationseinheiten semantisch absichert.

- KI-Agenten können über terminologisch modellierte Beziehungen eigenständig Informationen einordnen und validieren.

So bildet Terminologie den semantischen Kitt zwischen Datenmanagement, KI und Prozessautomation.

4.5 Die besonderen Stärken von Terminologie als RAG-Quelle

4.5.1 Sparsame Nutzung des Kontextraums (Token-Wirtschaftlichkeit)

Kommerziell verfügbare LLMs bieten nur einen begrenzten Kontextraum. Das bedeutet, dass sie nur eine bestimmte Menge an Informationen für einen Chat oder einen Abfragezyklus verarbeiten können. Übersteigt die übergebene Menge an Informationen diesen Kontextraum, werden überschüssige Informationen nicht für die Generierung der Ausgabe berücksichtigt. Damit steigt das Risiko unzutreffender Ausgaben stark an. Selbst bei steigendem Kontextraum muss diese Einschränkung im Blick behalten werden.

Die Anbindung von Terminologiedatenbanken ermöglicht eine sparsame und damit kostengünstige Nutzung des Kontextraums, da nicht die komplette Terminologiedatenbank, sondern nur die relevanten Informationen in den Kontextraum geladen werden müssen.

4.5.2 Optimierte Antwortzeiten (Performanz)

Beim Einsatz von RAG sind neben den Bearbeitungszeiten für die reine Generierung der KI-Ausgaben zusätzlich noch die Antwortzeiten, beispielsweise bei der Abfrage von Datenbanken, zu berücksichtigen. Diese hängen von vielen Parametern ab und sind nicht generell prognostizierbar.

Im Vergleich zu anderen Quellen sind Terminologiedatenbanken tendenziell auf kurze Antwortzeiten ausgelegt. Dies kann ein erheblicher Vorteil sein, insbesondere, wenn bei den anderen Quellen neben der reinen Abfrage auch noch möglicherweise rechenintensive Schritte zur Extraktion der benötigten Informationen erforderlich sind.

Dieser Vorteil zeigt sich insbesondere bei allen Echtzeit-Anwendungen wie Chatbots.

4.5.3 Aktuelle und gepflegte Daten

Terminologieverwaltungssysteme verfügen über ausgereifte Mechanismen und Funktionalitäten zum Terminologiemanagement.

Sie sind sehr gut geeignet, um terminologische Bestände effizient aufzubauen, kontinuierlich zu pflegen und bereitzustellen. So stellen Terminologiedatenbanken eine zentrale, redundanzarme und aktuelle Quelle terminologischer Daten bereit. Dies ist ein erheblicher Vorteil gegenüber dem Ansatz, Terminologie aus möglicherweise umfangreichen und widersprüchlichen Textkorpora jeweils neu zu extrahieren.

4.5.4 Transparenz und Nachvollziehbarkeit

Gut gepflegte Terminologiedatenbanken verfügen über eine große Menge an wertvollen Metainformationen zu den terminologischen Einträgen. Nutzen KI-Systeme diese Daten, stellt das einen erheblichen Vorteil für die Transparenz und Nachvollziehbarkeit der generierten Ausgaben dar.

Gerade im Hinblick auf regulatorische Anforderungen in diesem Bereich, beispielsweise in Bezug auf das KI-Gesetz (*AI Act*) der Europäischen Union, kann so der Einsatz von entsprechend hochwertigen Terminologiedatenbanken zu einer Notwendigkeit werden, um die entsprechenden Anforderungen mit vertretbarem Aufwand überhaupt erfüllen zu können.

4.5.5 Datenstrukturen statt Wortlisten

Eine Terminologiedatenbank ist deutlich mehr als nur eine lose Sammlung von Benennungen oder Excel-Glossaren mit unklarem Status oder unklarer Herkunft. Für eine eindeutige, effiziente und konsistente Verwaltung von mehrsprachigen Daten sind solche Listen nicht hilfreich; eine Terminologiedatenbank ist unerlässlich.

Terminologiedatenbanken enthalten neben Benennungen und Definitionen auch Metadaten zu Herkunft, Fachgebieten, Verwendung, Beziehungen zwischen den einzelnen terminologischen Einträgen oder genutzten Quellen.

Diese Daten werden strukturiert abgelegt und verwaltet.

Genau diese Struktur ermöglicht es KI-Systemen, die terminologischen Daten maschinell so aufzubereiten, dass sie mit den aus dem Training der KI-Systeme entstandenen Strukturen interagieren können. Dies ist eine elementare Voraussetzung für hochwertige Ausgaben der KI-Systeme.

4.6 Die Rolle von Terminologiefachleuten in KI-Projekten

4.6.1 Qualifikationsprofil von Terminologiefachleuten

Die erfolgreiche Entwicklung und Implementierung von KI-Systemen, insbesondere solcher, die auf LLMs basieren, erfordert weit mehr als rein technisches Know-how in der Softwareentwicklung oder im Data Engineering. Terminologiefachleute können hierbei eine Schlüsselrolle als KI-Enabler einnehmen: Sie schlagen die Brücke zwischen der unstrukturierten Welt der Sprache und der strukturierten Welt der Wissensrepräsentation.

Terminologiefachleute bringen spezifische Kompetenzen mit (siehe RaDT 2020), die auch für die Qualitätssicherung und Leistungssteigerung von KI-Anwendungen unerlässlich sind:

- Strukturierungskompetenz: Sie verfügen über Expertise darin, implizites Fachwissen zu explizieren und in maschinenlesbare Strukturen (z. B. Wissenstripel, Begriffssysteme) zu überführen.

- Semantische Präzision: Durch die Definition von Begriffen, die Zuweisung von eindeutigen Benennungen und die Auflösung von Mehrdeutigkeiten (Polysemie/Homonymie) reduzieren sie aktiv das Risiko von KI-Halluzinationen oder Fehlinterpretationen und unterstützen das kontextsensitive ***Reasoning***.

- Domänenübergreifendes Denken: Sie verstehen es, zwischen verschiedenen Fachdisziplinen zu vermitteln und konsistente Sachgebietsklassifikationen zu erstellen, die für das gezielte „Retrieval" in RAG-Systemen entscheidend sind.

- Data Governance: Sie sind erfahren im Management von Lebenszyklen von Datenbeständen. Diese Fähigkeit ist für die notwendige kontinuierliche Pflege der KI-Wissensbasis essenziell.

4.6.2 Einsatzfelder in KI-Projekten

Der Einsatz von Terminologiefachleuten sollte nicht erst am Ende eines Projekts, sondern als integraler Bestandteil des Entwicklungsprozesses erfolgen:

1. **Kuratierung der Wissensbasis:** Anstatt KI-Modelle mit ungeprüften Massen daten zu „füttern", stellen Terminologiefachleute hochwertige, validierte Datensätze bereit, die als „Golden Source" für RAG-Anwendungen dienen.

2. **Optimierung des Prompt-Engineerings:** Durch die Bereitstellung kontrollierter Terminologie unterstützen Terminologiefachleute die Erstellung präziserer System-Prompts, was die Relevanz der KI-Antworten deutlich erhöht.

3. **Entwicklung semantischer Netze:** Terminologiefachleute unterstützen den Aufbau von Knowledge Graphs und Ontologien, die es KI-Agenten ermöglichen, logische Schlussfolgerungen (*Reasoning*) auf Basis korrekter Fakten zu ziehen.

4. **Qualitätssicherung und Evaluation:** Terminologiefachleute definieren die sprachlichen Standards, an denen die Ausgaben der KI gemessen werden, und identifizieren systematische Fehler in der Benennungswelt der Modelle.

Durch die aktive Einbindung von Terminologie-Expertise wandelt sich das Terminologiemanagement von einer rein sprachlichen Unterstützungsfunktion zu einem **strategischen Instrument der KI-Wertschöpfung.**

4.7 Zusammenfassung

LLMs sind leistungsstarke Werkzeuge, die mit einer Vielzahl von unterschiedlichen Texten trainiert wurden. Dementsprechend liefern sie Ausgaben auf Basis einer hohen Variabilität des Trainingsmaterials. Terminologie stellt explizites, strukturiertes und domänenspezifisches Wissen bereit. Wer terminologische Daten in KI-Anwendungen integriert – insbesondere über RAG – erhöht Qualität, Zuverlässigkeit und übergreifende Nutzbarkeit enorm.

Gleichzeitig ebnet die Verbindung von KI und Terminologie den Weg zu semantischen Architekturen, die für die zukunftsfähige, nachvollziehbare und regulierungskonforme Nutzung der KI unverzichtbar sind. Terminologie ist damit kein linguistisches Detail, sondern ein strategischer Schlüssel zur wissensbasierten KI.

5 Best Practices: KI gezielt mit Fachwissen anreichern

5.1 Terminologie dynamisch in LLMs bereitstellen

LLMs sind keine Datenbanken, aus denen Wissen explizit und reproduzierbar abgerufen werden kann. Sie wurden darauf trainiert, auf Eingaben mit einer Ausgabe zu reagieren. Somit können sie spezifische Terminologie nicht dauerhaft bereithalten und konsistent verwenden.

Angesichts der enormen Größe des Trainingsmaterials bleibt der Versuch, Terminologie explizit in das Modell „hineinzutrainieren", wirkungslos. Um die gewünschte Terminologie zu forcieren, sind RAG-Mechanismen erforderlich.

Im Fall von Terminologie bedeutet RAG:

- Die KI erhält im Prompt strukturierte Benennungen samt Definitionen und Kontextinformationen, um präziser, konsistenter und robuster zu antworten. Diese Vorgehensweise erfordert aufmerksames und fachgerechtes Formulieren der Prompts. Dieses Vorgehen ist nur geeignet, wenn lediglich wenige terminologische Vorgaben benötigt werden (für Beispiele siehe Abschnitt 5.2.1).

- Die gesamte Terminologiedatenbank wird systematisch angebunden, und so werden die Informationen aus den einzelnen terminologischen Einträgen genutzt. Dadurch erübrigt sich die manuelle Eingabe von Terminologie, und die terminologische Konsistenz wird automatisch erzwungen. Vorteile dieses Vorgehens sind der unmittelbare Zugriff auf Live-Daten, die Nutzung eines vollständigen Terminologiebestands sowie die Bereitstellung einer Infrastruktur für skalierbare und nachhaltige Projekte (für Beispiele siehe Abschnitt 5.2.2). Gegebenenfalls werden zusätzlich auch Begriffsbeziehungen genutzt, die in der Terminologiearbeit oft zu Begriffssystemen ausgebaut werden. Die Nutzung von Begriffsbeziehungen führt aber auch zu einer deutlichen Vergrößerung der Daten(austausch)mengen zwischen den Systemen, was unter Umständen zu Performanzproblemen und erhöhten Kosten führt.

5.2 Umsetzung in der Praxis

5.2.1 Terminologie-Prompts

Bei diesem Ansatz werden strukturierte Terminologie-Prompts mit wenigen relevanten Begriffen verwendet (z. B. über Prompt-Templates, die wiederverwendet und angepasst werden können).

- Beispiel 1)
  Erstelle eine Meldung für die Website der Polizei über die Razzia am … in …; verwende dabei nicht *verhaften*, sondern *festnehmen*.

- Beispiel 2)
  Erkläre einem Fachpublikum, wie Energieeffizienz zur Dekarbonisierung beiträgt. Halte dich dabei an folgende Definitionen:
  - „Energieeffizienz“: Maß für den Energieverbrauch eines Systems im Verhältnis zur erbrachten Leistung
  - „Dekarbonisierung“: Prozess der Reduktion von CO_2-Emissionen durch strukturverändernde Maßnahmen

5.2.2 Terminologiedatenbanken

Für eine systematische Nutzung der Terminologiedaten werden vollständige Terminologiedatenbanken über RAG bereitgestellt, z. B. über standardisierte Schnittstellen wie Model Context Protocol (MCP) oder strukturierte Formate, um so eine semantische Infrastruktur aufzubauen. So können LLMs auf terminologische und ontologische Informationen (z. B. aus Begriffssystemen bzw. Wissensgraphen mit RDF- oder OWL-Strukturen) in Echtzeit zugreifen und kontextsensitiv „verstehen“, anstatt nur Texte zu zitieren.

- Beispiel:
  Prompt: „Erstelle eine Meldung für die Website der Polizei über die Razzia am … in …“.

Im Hintergrund wird automatisch auf die Terminologiedatenbank zugegriffen, ohne dass dies im Prompt gefordert werden muss. In der Folge können unter anderem die Vorzugsbenennungen aus der Terminologiedatenbank in die Ausgabe des LLMs übernommen werden, so dass terminologische Konsistenz, Eindeutigkeit und Verständlichkeit sichergestellt sind.

5.3 Warum Terminologiemanagement hilfreich ist

KI-Ausgaben werden durch sprachliche Gegebenheiten auf unterschiedliche Weise beeinflusst. Die folgende Tabelle stellt dar, wie Terminologiemanagement zu einer Verbesserung der KI-Ausgaben beitragen kann:

Sprachliche Gegebenheit	Beschreibung/ Beispiel	KI-Ausgabe ohne Terminologie-unterstützung	Mehrwert durch Einbindung von terminologischen Daten
Es gibt mehrere Benennungen für dieselbe Bedeutung (Synonymie).	„Kunde“ vs. „Klient“ „Bestellung“ vs. „Auftrag“ vs. „Order“ „Frontscheibe“ vs. „Windschutzscheibe“	Das Modell generiert inkonsistente Texte, in denen die verschiedenen Synonyme scheinbar willkürlich verwendet bzw. gemischt werden.	Das Terminologieverwaltungssystem stellt unter anderem Vorzugsbenennungen sowie abgelehnte Benennungen bereit. Abgelehnt werden Benennungen, die z. B. veraltet, negativ konnotiert oder verwirrend sind.
Es gibt Benennungen mit mehreren Bedeutungen (Mehrdeutigkeit).	„Bank“ (Sitzgelegenheit vs. Geldinstitut) „Flügel“ (Musikinstrument vs. Vogelschwinge vs. Gebäudeteil) „Einsatz“ (Operation vs. Engagement vs. Gebrauch)	Wenn nur wenig Kontext und keine Fachgebietseinordnung vorhanden sind, trifft das Modell u. U. falsche Annahmen und generiert Texte mit Benennungen aus falschen Fachgebieten.	In Terminologiedatenbanken gibt es eindeutige Fachgebietszuordnungen und Begriffsdefinitionen, welche die Auswahl der passenden Terminologie ermöglichen.
Mehrdeutige Benennungen erfordern unterschiedliche Übersetzungen (Äquivalente). Es gibt keine 1:1-Relation zwischen den Benennungen in verschiedenen Sprachen.	DE: „Bank“ = EN: „bench“ (Sitzgelegenheit) vs. DE: „Bank“ = EN: „bank“ (Geldinstitut) DE: „Flügel“ = EN: „grand piano“ (Musikinstrument) vs. DE: „Flügel“ = EN: „wing“ (Vogelschwinge und Gebäudeteil)	Wenn nur wenig Kontext und keine Fachgebietseinordnung vorhanden sind, trifft das Modell u. U. falsche Annahmen und liefert fehlerhafte Übersetzungen, indem es Äquivalente aus dem falschen Fachgebiet verwendet.	Das Terminologieverwaltungssystem verwaltet unterschiedliche Bedeutungen begriffsorientiert in getrennten Einträgen. Dadurch wird eine eindeutige Zuordnung der verschiedenen Äquivalente überhaupt erst ermöglicht.
Kurzformen und Abkürzungen werden ohne Nennung der Langform verwendet. Das Problem ist besonders relevant bei organisationsspezifischen Abkürzungen.	SPG = Sicherheitspolizeigesetz oder = strategische Planungsgruppe? AG = Aktiengesellschaft oder = Arbeitsgemeinschaft oder = Amtsgericht oder = Auftraggeber?	Kurzformen, die ohne Erklärung und klare Kontextangabe eingeführt und verwendet werden, sind für das Modell nicht klar zuzuordnen, so dass vage oder sachlich falsche Inhalte generiert werden.	Das Terminologieverwaltungssystem liefert zu jeder Kurzform die entsprechende Langform sowie die weiteren Informationen, so dass das KI-System die passende Benennung ermitteln kann.

Eine saubere, konsistente und dokumentierte Terminologie ist nicht nur hilfreich für menschliche Kommunikation, sondern ein **kritischer Erfolgsfaktor** für die Effektivität von Sprachmodellen in produktiven KI-Anwendungen. Besonders in regulierten, technischen, rechtlichen oder sicherheitskritischen Kontexten kann unpräzise Terminologie zu erheblichen inhaltlichen, rechtlichen und reputationsbezogenen Risiken führen.

6 Fazit: gute Daten für gute Prozesse

6.1 Terminologie als semantisches Fundament der digitalen Wertschöpfung

Wenn Terminologie als strukturierte, vernetzte Wissensbasis erkannt wird, entsteht eine Infrastruktur, die alle Informationssysteme verbindet, unabhängig von Sprache, Medium oder Technologie. Dadurch wird Wissen auffindbar, vergleichbar und interoperabel.

Damit wird Terminologie zum semantischen Fundament der digitalen Wertschöpfung und zu einem entscheidenden Erfolgsfaktor für die Integration von KI, Wissensmanagement und datengetriebenen Prozessen. Somit ist es unerlässlich, hochwertige und bereits verfügbare Quellen wie Terminologiedatenbanken zu nutzen.

Für Entscheidungsträgerinnen und Entscheidungsträger bedeutet dies, dass Terminologie mehr als ein optionales Asset ist: Vielmehr ist sie ein strategisches Instrument zur Steigerung der Präzision und Qualität KI-generierter Ergebnisse.

6.2 Terminologische Kompetenz als Voraussetzung für Automation

Die treibende Kraft hinter aktuellen KI-Projekten ist das Versprechen der konsequenten Prozessautomation: Abläufe sollen effizienter, schneller und idealerweise mit minimaler menschlicher Beteiligung realisiert werden. Dabei darf die Qualität nicht zu kurz kommen. Doch in dieser Transformation verbirgt sich eine oft übersehene Herausforderung. In klassischen, menschenzentrierten Prozessen fungieren Menschen als flexibles Sicherheitsnetz: Mit allgemeinem Weltwissen, beruflicher Erfahrung und fachlicher Kompetenz können sie beispielsweise Unschärfe in Prozessdefinitionen, lückenhafte Vorbereitungen oder fehlerbehaftete Daten ausgleichen. Menschen verstehen meist, was gemeint ist, auch wenn es nicht explizit dasteht.

KI verfügt nach heutigem Stand – und auf absehbare Zeit – nicht über dieses tiefgreifende Weltverständnis oder die Fähigkeit zur intuitiven Fehlerkorrektur. Während Menschen eine mehrdeutige Anweisung durch Kontextwissen ausgleichen können, kann eine KI scheitern oder halluzinieren, da sie lediglich statistische Muster verarbeitet. An dieser Stelle wird Terminologie zum entscheidenden Hebel für den Projekterfolg. Terminologiearbeit liefert das explizite, strukturierte Fachwissen, das der Maschine fehlt, um Mehrdeutigkeiten aufzulösen und fachlich konsistent zu agieren.

Terminologiefachleute agieren hierbei als unverzichtbare Brückenbauer. Ihre Kompetenz, implizites Wissen in maschinenlesbare Strukturen zu transformieren, führt dazu, dass Automatisierung nicht an sprachlicher Unschärfe oder unzureichender Wissensrepräsentation scheitert. Für Entscheidungsträgerinnen und Entscheidungsträger bedeutet das: Der Erfolg von KI-Projekten hängt nicht allein von der Wahl des Modells ab, sondern von der Qualität und Präzision der semantischen Grundlage, die diesem Modell zur Verfügung gestellt wird.

Für erfolgreiche KI-Projekte macht Terminologie den Unterschied.

7 Quellenangaben und weiterführende Literatur

Deutsches Institut für Terminologie e. V. (2017): Was haben Wissensmodellierung, Wissensstrukturierung, künstliche Intelligenz und Terminologie miteinander zu tun?. <https://downloads.ditev.org/publikationen/DITeV_org_Terminologie_und_KI_2017_03_22_v2.pdf>. Zugriff: 19.05.2026.

DIN 2342 (2022): Terminologiewissenschaft und Terminologiearbeit – Begriffe. Berlin: DIN-Media.

Drewer, Petra (2025): „Knowledge Graphs, Concept Maps, Ontologien, Begriffssysteme & Co.: Warum Terminologiemanagement (auch) Wissensmanagement ist". In: Drewer, Petra / Mayer, Felix / Pulitano, Donatella (Hrsg.) (2025): Terminologie in der KI – KI in der Terminologie. Akten des DTT-Symposions 2025. Karlsruhe, München, Bern: Deutscher Terminologie-Tag / SDK Systemdruck, S. 203-220.

Drewer, Petra / Mayer, Felix / Pulitano, Donatella (Hrsg.) (2025): Terminologie in der KI – KI in der Terminologie. Akten des DTT-Symposions 2025. Karlsruhe, München, Bern: Deutscher Terminologie-Tag / SDK Systemdruck.

Fraunhofer IKS – Institut für Kognitive Systeme (o. J.): „Künstliche Intelligenz (KI) und maschinelles Lernen". <https://www.iks.fraunhofer.de/de/themen/kuenstliche-intelligenz.html>. Zugriff: 19.05.2026.

IBM (o. J. a): „Was ist künstliche Intelligenz (KI)?". <https://www.ibm.com/de-de/think/topics/artificial-intelligence>. Zugriff: 19.05.2026.

IBM (o. J. b): „Was ist ein Wissensgraph?". <https://www.ibm.com/de-de/topics/knowledge-graph>. Zugriff: 11.11.2025.

IBM (o. J. c): „Was sind neuronale Netzwerke?". <https://www.ibm.com/de-de/think/topics/neural-networks>. Zugriff: 07.07.2026.

RaDT – Rat für Deutschsprachige Terminologie (2013): „Terminologisches Basiswissen für Fachleute". <https://publikationen.radt.org/RaDT_Basiswissen_2013_deutsch.pdf>. Zugriff: 13.07.2026.

RaDT – Rat für Deutschsprachige Terminologie (2020): „Berufsprofil 2.0 – Terminologin–Terminologe." <https://publikationen.radt.org/RaDT_Berufsprofil_2-0_2020_deutsch.pdf>. Zugriff: 13.07.2026.

Studer, Rudi (o. J.): „Ontologien". <https://wi-lex.de/index.php/lexikon/informations-daten-und-wissensmanagement/wissensmanagement/wissensmodellierung/wissensrepraesentation/ontologien/>. Zugriff: 11.11.2025.

TH Köln (2022): „Kurz erklärt: Über KI, maschinelles Lernen und Deep Learning". <https://www.th-koeln.de/hochschule/kurz-erklaert-ueber-ki-maschinelles-lernen-und-deep-learning_95656.php>. Zugriff: 07.07.2026.

8 Publikationen des RaDT

Drewer P./Pulitano D. (Hrsg.) (2019): Terminologie: Epochen – Schwerpunkte Umsetzungen. Zum 25-jährigen Bestehen des Rats für Deutschsprachige Terminologie. Berlin/Heidelberg: Springer Vieweg.

Unter anderem können die folgenden Broschüren des RaDT unter www.radt.org > Publikationen kostenlos heruntergeladen werden:

Rat für Deutschsprachige Terminologie (2010): „Wissen, Marken, Kundenbindung – Kritischer Erfolgsfaktor Terminologie.“ <https://publikationen.radt.org/RaDT_Wirtschaftspapier_2010_deutsch.pdf>. Zugriff: 13.07.2026.

Rat für Deutschsprachige Terminologie (2013): „Terminologisches Basiswissen für Fachleute“. <https://publikationen.radt.org/RaDT_Basiswissen_2013_deutsch.pdf>. Zugriff: 13.07.2026.

Rat für Deutschsprachige Terminologie (2015): „Terminologie – Grundlage für Fachkommunikation, Fachinformation, Fachwissen“. <https://publikationen.radt.org/RaDT_Terminologie_2015_deutsch.pdf>. Zugriff: 13.07.2026.

Rat für Deutschsprachige Terminologie (2017): „Terminologiewissenschaft – Eine Standortbestimmung.“ <https://publikationen.radt.org/RaDT_Terminologiewissenschaft_2017_deutsch.pdf>. Zugriff: 13.07.2026.

Rat für Deutschsprachige Terminologie (2020): „Berufsprofil 2.0 – Terminologin – Terminologe.“ <https://publikationen.radt.org/RaDT_Berufsprofil_2-0_2020_deutsch.pdf>. Zugriff: 13.07.2026.

Anhang: Glossar

Grundlegende terminologische Begriffe, die in diesem Glossar nicht aufgeführt sind, finden sich mit Definitionen in DIN 2342 (2022): Terminologiewissenschaft und Terminologiearbeit – Begriffe.

agentische KI
agentic AI
Form der KI, bei der KI-Systeme über das reine Generieren von Inhalten hinausgehen und Ziele autonom planen, Entscheidungen treffen und Handlungen ausführen

Augmentation
Augmentierung
Anreicherung des Frage-Prompts mit relevanten Informationsbestandteilen aus den verknüpften Datenquellen

Begriff
Denkeinheit, die aus einer Menge von Gegenständen unter Ermittlung der diesen Gegenständen gemeinsamen Eigenschaften mittels Abstraktion gebildet wird
Anmerkung: In einigen Disziplinen werden Begriffe oft auch als Konzepte bezeichnet.

Begriffssystem
Menge von Begriffen eines Fachgebiets, die nach bestimmten Begriffsbeziehungen geordnet sind

Benennung
sprachliche Bezeichnung eines Allgemeinbegriffs aus einem Fachgebiet

Datenmodell
standardisierte Struktur zur Repräsentation von Daten und deren Relationen

Definition
Begriffsdefinition
Begriffsbestimmung mit sprachlichen Mitteln

Entität
<Fachgebiet: Wissensgraphen> eindeutig definiertes Objekt oder Konzept in einem Wissensnetz

generative KI
Gen KI
Form der KI, die neue Inhalte auf Basis von Anweisungen und erlernten Mustern erzeugt

Halluzination
falsche oder erfundene Informationen, die von einem KI-System generiert werden

Interoperabilität
Fähigkeit verschiedener Systeme, Daten auszutauschen und zusammenzuarbeiten

KI-Modell
mathematisches System, das aus großen Datenmengen Muster gelernt hat und dieses Wissen nutzt, um für neue Eingaben Vorhersagen, Entscheidungen oder Inhalte wie Text oder Bilder zu erzeugen
Anmerkung: Large Language Models (LLMs) sind eine spezielle Unterart davon.

Knowledge Graph
Wissensgraph
strukturierte Wissensrepräsentation in Form eines Netzwerks aus Entitäten und ihren Relationen

Kontextraum
begrenzter Speicherbereich für Informationen, die ein Modell in einem Chat oder Abfragezyklus verarbeiten kann

Konzept → Begriff

künstliche Intelligenz
KI
artificial intelligence
AI
Technologie, die darauf abzielt, menschliche kognitive Fähigkeiten wie Lernen, Problemlösen und Entscheidungsfindung durch Maschinen nachzuahmen, indem sie Informationen aus Eingabedaten erkennt und sortiert

künstliches neuronales Netz
neuronales Netz
KNN
maschinelles Lernmodell mit mehrschichtiger Struktur, das aus in Schichten angeordneten, gewichtet verbundenen Recheneinheiten besteht und durch Anpassung dieser Gewichtungen anhand von Trainingsdaten Muster erlernt

Large Language Model
LLM
großes Sprachmodell
großes KI-Modell, das auf die Verarbeitung und Generierung von Sprache spezialisiert ist

maschinelles Lernen
machine learning
Teilbereich der KI, bei dem Systeme Algorithmen nutzen, um aus Daten zu lernen, Muster zu erkennen und Vorhersagen zu treffen

maschinelles Übersetzungssystem
MÜ-System
Softwareanwendung, die linguistische Regeln und Algorithmen oder künstliche Intelligenz und maschinelles Lernen nutzt, um Texte automatisch aus einer Ausgangssprache in eine Zielsprache zu übersetzen

Metadaten
strukturierte Informationen, die andere Daten beschreiben, kontextualisieren und deren Verwaltung erleichtern

Ontologie
<Fachgebiet: Informatik> formale Wissensrepräsentation mit expliziten Spezifikationen von Konzepten und deren Relationen

32

Prompt
Eingabeaufforderung, Frage oder Anweisung an ein KI-System, um dieses zu einer spezifischen Reaktion oder Antwort zu veranlassen

Prompt-Engineering
Gestaltung und Optimierung von Eingabeaufforderungen für KI-Systeme

RAG
Retrieval-Augmented Generation
Architekturkonzept zur Verbesserung der Antwortqualität von Sprachmodellen durch Verwendung externer Daten

Reasoning
Fähigkeit eines KI-Systems, über logische Schlussfolgerungen zu einer Lösung zu gelangen

Terminologie
Fachwortschatz
Gesamtbestand der Begriffe und ihrer Bezeichnungen in einem Fachgebiet

Terminologiedatenbank
Termbank
Datenbank zur Verwaltung terminologischer Daten

Terminologiemanagement
Planung, Organisation und Leitung der terminologischen Aktivitäten eines Unternehmens oder einer Institution

Terminologieverwaltungssystem
Softwaresystem zur systematischen Verwaltung terminologischer Datenbestände

terminologischer Eintrag
Terminologieeintrag
eigenständige, strukturierte Einheit, die terminologische Daten zu einem Begriff enthält

Token
kleinste verarbeitbare sprachliche Einheit, in die ein Text zerlegt wird

Trainingsdaten
Daten, die zum Training eines KI-Modells verwendet werden

Wissensmanagement
systematische Verwaltung und Organisation von Wissen in Organisationen

Wissensrepräsentation
formale Darstellung von Wissen in strukturierter Form

Wissenstripel
grundlegende Struktureinheit der Wissensrepräsentation bestehend aus Subjekt, Prädikat und Objekt

RaDT
Rat für
Deutschsprachige
Terminologie

Der Rat für Deutschsprachige Terminologie (www.radt.org) ist 1994 auf Initiative der UNESCO-Kommissionen Deutschlands, Österreichs und der Schweiz gegründet worden mit dem Ziel, die Terminologie im deutschsprachigen Raum zu fördern. Es handelt sich um ein Gremium aus Fachleuten in dem Organisationen, Verbände, Behörden, die Wirtschaft sowie die Bildungseinrichtungen auf dem Gebiet der Terminologie aus Deutschland, Österreich, Luxemburg, Belgien, Südtirol und der Schweiz vertreten sind.

36

Impressum:

Herausgeber:
Rat für Deutschsprachige Terminologie (RaDT)
eine Initiative der UNESCO-Kommissionen
Deutschlands, Luxemburgs, Österreichs und der Schweiz
www.radt.org

Produktion:
Deutscher Terminologie-Tag e.V. (DTT)
www.dttev.org

Gestaltung: M. Reisner, Brühl

Stand: Oktober 2026

© RaDT 2026