

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Automatische Termextraktion (ATE) der deutschsprachigen Termini zum Thema
Energiekrise mithilfe diverser Ansätze der ATE

verfasst von | submitted by

Miriama Smolková BA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Arts (MA)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt |
Degree programme code as it appears on the
student record sheet:

UA 070 369 331

Studienrichtung lt. Studienblatt | Degree pro-
gramme as it appears on the student record
sheet:

Masterstudium Translation Tschechisch Deutsch

Betreut von | Supervisor:

Assoz. Prof. Mag. Dr. Dagmar Gromann BSc

Danksagung

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich während des Schreibens dieser Masterarbeit unterstützt und motiviert haben. Zuerst möchte ich mich bei meiner Betreuerin, Frau Assoz. Prof. Mag. Dr. Dagmar Gromann, BSc, bedanken, die mich mit ihrer fachlichen Expertise und wertvollen Ratschlägen durch diesen langen und stellenweise nicht so einfachen Prozess begleitet hat. Ihre Geduld und Unterstützung, vor allem bei der Auswertung der Ergebnisse, haben maßgeblich zum Gelingen dieser Arbeit beigetragen. Ein besonderer Dank gilt auch meinen Studienkolleginnen, die mir stets mit Rat und Tat zur Seite standen und mich in schwierigen Momenten ermutigt haben. Insbesondere möchte ich meine Annotatorin hervorheben, die mich bei der wochenlangen, nie enden wollenden Aufgabe der Erstellung des Goldstandards unterstützt hat. Nicht zuletzt möchte ich meiner Familie danken, die mir stets den Rücken gestärkt und mir die nötige Kraft und Zeit gegeben hat, dieses Projekt abzuschließen. Diese Masterarbeit ist das Ergebnis vieler Stunden harter Arbeit, aber auch der Unterstützung und Ermutigung vieler wunderbarer Menschen.

Abstract

Ziel dieser Masterarbeit ist es, die Leistung verschiedener Systeme zur automatischen Termextraktion zu bewerten, die nicht nur Wissenschaftler*innen, sondern auch allgemeinen Nutzer*innen zugänglich sind. Ein Blackbox-Experiment wird mit Artikeln österreichischer Zeitungen aus dem Bereich der Energiewirtschaft zum Thema Energiekrise durchgeführt. Das Experiment ist so konzipiert, dass es auf andere Themen und Sprachen anwendbar ist, vorausgesetzt, das jeweilige System unterstützt die entsprechende Sprache und die benötigten Textformate. Im theoretischen Teil dieser Arbeit werden die Grundlagen der Terminologielehre sowie der Terminologearbeit mit Fokus auf den Begriff Terminologie sowie auf die automatische Termextraktion unter Berücksichtigung der neuesten Entwicklungen, insbesondere im Bereich des Deep Learning, behandelt. Ein wesentlicher Bestandteil dieser Arbeit ist die Erstellung eines Goldstandards und der Richtlinien für die Annotation, die für den Vergleich der Systeme benötigt werden. Sowohl neuronale als auch nicht-neuronale Systeme werden anhand desselben annotierten Korpus ausgewertet, um Genauigkeit, Trefferquote und die F1-Metrik berechnen zu können. Anhand dieser Ergebnisse wird die Leistungsfähigkeit der Systeme ermittelt. Zusätzlich wird neben dieser quantitativen Analyse auch eine qualitative Analyse durchgeführt, um die Schwächen und Stärken der verfügbaren Systeme genauer aufzuzeigen.

Inhaltsverzeichnis

Danksagung	1
Abstract	3
Inhaltsverzeichnis	4
Abbildungs- und Tabellenverzeichnis	7
Abkürzungsverzeichnis	8
Formelverzeichnis	9
1 Einleitung	10
1.1 Forschungsfrage und Annahmen	12
1.2 Motivation und Zielsetzung	13
2 Theoretische Grundlagen.....	14
2.1 Terminologie als Grundlage der Fachkommunikation	14
2.1.1 Gegenüberstellung von Terminologie und Fachwortschatz	15
2.1.2 Bedeutung von Terminologie	16
2.1.3 Gliederung von Termini	18
2.1.4 Gliederung von Termini in der ATE	20
2.1.5 Unterscheidung zwischen Fachgebiet und Fachbereich.....	21
2.1.6 Fachtext	22
2.1.6.1 Merkmale von Fachtexten	22
2.1.6.2 Vielfalt der Fachtexte	23
2.1.7 Fachsprache	24
2.1.8 Wirtschaftsfachsprache	26
2.1.8.1 Definition der Wirtschaftssprache.....	27
2.1.8.2 Wirtschaftliche Texte im journalistischen Kontext.....	29
2.1.8.3 Terminologische Besonderheiten in der Wirtschaftssprache	30
2.2 Grundlagen der Terminologielehre	32
2.2.1 Gegenstand	32
2.2.2 Begriff	33
2.2.3 Benennung.....	34
2.2.4 Beziehungen zwischen den Grundelementen der Terminologielehre	36
2.2.4.1 Semiotisches Dreieck	36
2.2.4.2 Beziehungen zwischen Begriffen	37
2.2.4.3 Beziehungen zwischen Begriffen und Benennungen	39
2.2.4.3.1 Idealfall Eineindeutigkeit	39
2.2.4.3.2 Synonymie und Teilsynonymie.....	40
2.2.4.3.3 Mehrdeutigkeit	41

2.2.4.3.4	Äquivalenz	42
2.3	Grundlagen der Terminologearbeit.....	43
2.3.1	Arbeitsformen der Terminologearbeit.....	44
2.3.1.1	Deskriptive und präskriptive Terminologearbeit	44
2.3.1.2	Punktuelle, textbezogene und fachgebietsbezogene Terminologearbeit.....	45
2.3.1.3	Einsprachige und mehrsprachige Terminologearbeit.....	46
2.3.2	Praktische Terminologearbeit	46
2.3.2.1	Terminologiegewinnung	47
2.3.2.2	Menschliche Terminologierecherche	48
2.3.2.3	Menschliche Terminologierecherche mit einfacher technischer Unterstützung... ..	49
2.4	Maschinelle bzw. (halb-)automatische Termextraktion (ATE)	50
2.4.1	Arten und Ansätze der ATE	55
2.4.1.1	Konventionelle Ansätze	56
2.4.1.2	Neue Ansätze in der ATE.....	59
2.4.2	Zeitgemäße Unterteilung der ATE-Methoden	61
2.4.2.1	Methoden basierend auf Statistiken über das Auftreten der Termini.....	61
2.4.2.2	Methoden basierend auf Kontextinformationen über das Auftreten der Termini.....	62
2.4.2.3	Methoden basierend auf Themenmodellierung.....	63
2.4.2.4	Methoden basierend auf externen Korpora	63
2.4.2.5	Methoden basierend auf Suchmaschinen	65
2.4.2.6	Methoden basierend auf Ontologien	65
2.4.2.7	Methoden basierend auf Wikipedia.....	66
2.4.2.8	Methoden basierend auf nicht neuronalem maschinellern Lernen.....	67
2.4.2.9	Methoden basierend auf neuronalen Netzen bzw. Deep Learning.....	68
2.4.3	Vor- und Nachteile der ATE	71
2.4.4	Anwendungsbereiche	72
2.5	Evaluation in der ATE	73
2.5.1	Evaluationsmethoden	74
2.5.2	Überblick der gängigen Metriken in der ATE.....	76
3	Aktueller Stand der Forschung	78
4	Methode	82
4.1	Korpuserstellung	83
4.2	Erstellung eines Goldstandards.....	88
4.2.1	Generelle Richtlinien für Annotation von Termini	89
4.2.2	Kategorien der Termini	90
4.3	Systemwahl und Vorgehensweise der Automatischen Termextraktion	92

4.4 Evaluierung	95
5 Ergebnisse.....	97
5.1 Quantitative Analyse der ATE-Ergebnisse	97
5.2 Qualitative Analyse der ATE-Ergebnisse	99
5.2.1 Qualitative Analyse der Ergebnislisten von D-Terminen	99
5.2.2 Qualitative Analyse der Ergebnislisten von OneClick Terms.....	101
5.2.3 Qualitative Analyse der nicht im Goldstandard vorkommenden Termini	103
5.2.4 Zusammenfassung der Ergebnisse	105
6 Diskussion.....	108
7 Fazit.....	110
Literaturverzeichnis.....	112
Anhang A	116
Anhang B.....	118
Anhang C.....	119
Anhang D	149
Anhang E.....	174

Abbildungs- und Tabellenverzeichnis

Abbildung 1: Gliederung von Fachsprach- und Fachtextwortschatz nach fachlicher Zugehörigkeit der betreffenden Fachsprach- und Fachtextwörter (Roelcke 2020: 71)	19
Abbildung 2: Vorkommen eines Ausdrucks in Nachschlagwerken (Adamzik 2018: 299)	19
Abbildung 3: Textsortengliederung: Übersicht und Beispiele (Roelcke 2020: 64)	24
Abbildung 4: Fachsprachliche Gesamttypologie (Roelcke 2020: 66)	25
Abbildung 5: Horizontale Gliederung von Fachsprachen in Wissenschafts-, Technik- und Institutionensprachen (mittel unterlegt), sowie stark vereinfachte Zuordnung einiger ausgewählten Ergänzungen (hell unterlegt) (Roelcke 2020: 45)	28
Abbildung 6: Modell von Wirtschaftssprache (Crestani 2010: 33)	29
Abbildung 7: Semiotisches Dreieck nach Ogden und Richards (1956: 11)	37
Abbildung 8: Anpassung des semiotischen Dreiecks (Drewer & Schmitz 2017: 7)	37
Abbildung 9: Beispiel für Polysemie (Drewer & Schmitz 2017: 18)	41
Abbildung 10: Begriffssystem „Terminologiewerk“ (DIN 2342 2022)	43
Abbildung 11: Übersicht der Arten des maschinellen Lernens (Mahesh 2020: 381)	60
Abbildung 12: Unterteilung der neuesten Methoden nach Art und Weise, wie die Systeme mit Sprachen im Trainingskorpus umgehen (Tran et al. 2023: 18)	70
Abbildung 13: Überblick der Evaluationsmethoden in der ATE (Tran et al. 2023: 16)	75
Abbildung 14: Medien-Analyse 2022 (VMA 2022)	84
Abbildung 15: Annotationsschema (Rigouts Terryn 2021a: 4)	90
Abbildung 16: Einstellung des Schwellenwertes von System OneClick Terms	94
Abbildung 17: Auszug der ATE-Ergebnisse von OneClick Terms ohne Schwellenwert	98
 Tabelle 1: Übersicht der verwendeten Metriken in ausgewählten Studien	 76
Tabelle 2: Suchkriterien der Artikelsuche und Gesamtergebnisse	85
Tabelle 3: Auswahl der Artikel aus Der Standard	86
Tabelle 4: Auswahl der Artikel aus Kurier	87
Tabelle 5: Quantitative Ergebnisse der Evaluation von Termini in Zahlen	97
Tabelle 6: Quantitative Ergebnisse berechnet mittels der Metriken Genauigkeit, Trefferquote und F1-Metrik	98
Tabelle 7: Anzahl der Termini je Kategorie, die nicht gefunden worden sind	99
Tabelle 8: Termini, die von beiden Systemen erkannt worden sind, sich aber nicht im Goldstandard befinden.	104

Abkürzungsverzeichnis

ACTER.....	Annotated Corpora for Term Extraction Research
ATE	automatische Termextraktion
ATR.....	Automated Term Recognition
BERT.....	Bidirectional Encoder Representations from Transformers
CT.....	Common Term
ELG	European Language Grid
GSÖD.....	Gesellschaft für österreichisches Deutsch
IDF	Inverse Document Frequency
JATE.....	Java Automatic Term Extraction
KÜDES.....	Konferenz der Übersetzungsdienste europäischer Staaten
KWIC-Index.....	Keyword in Context Index
LREC.....	Language Resources and Evaluation
NE.....	Named Entity
NPs	Nominalphrasen
OOD	Out-of-Domain
POS.....	Part-of-speech
RIDF.....	Residual Inverse Document Frequency
ST	Specific Term
TCS.....	Terminological Concept System
TERM21	Terminology in the 21st century
TF	Term Frequency
VMA.....	Verein Arbeitsgemeinschaft Media
X-LMR	Cross-lingual Modeling-RoBERTa
ZTW	Zentrum für Translationswissenschaft

Formelverzeichnis

(1) TF-IDF Formel	61
(2) C-Value Formel	62
(3) LinkProbability Formel	67
(4) Genauigkeit Formel	76
(5) Trefferquote Formel	76
(6) F1-Score Formel	76

1 Einleitung

Um sich innerhalb eines Fachgebietes eindeutig und effektiv verständigen zu können und somit den Wissenstransfer zu ermöglichen, braucht es einen Wortschatz der Fachtermini des jeweiligen Fachgebietes, der sogenannten Terminologie (siehe Abschnitt 2.1.1). Sie bezieht sich auf die konkrete Fachsprache eines bestimmten Wissensgebietes und trägt maßgeblich dazu bei, dass sich auch hochkomplexere Konzepte eindeutig und daher verständlich beschreiben lassen. Dadurch werden Erkenntnisgewinn, Wissenstransfer und Kommunikation innerhalb eines Fachgebietes gewährleistet. Mit der immer steigenden Menge an Informationen, deren schnelle Verarbeitung und dem Streben nach höherer Qualität der Wissensgewinnung und Wissenstransfers tritt die automatische Termextraktion (ATE) in den Vordergrund. Diese ermöglicht es, große Mengen an Textdaten effizient und rasch zu analysieren und relevante Termini hervorzuheben. In diesem Kontext gewinnt die ATE zunehmend an Bedeutung, weil sie Forschung, Bildung und industrielle Anwendungen durch die Extraktion präziser und relevanter Termini unterstützt.

Trotzt breitgefächerten Anwendungsmöglichkeiten (siehe Abschnitt 2.4.4) stellt die ATE nicht nur für die Praxis, sondern auch für die Forschung selbst eine Herausforderung dar. Einige Forscher*innen weisen explizit auf diese Tatsache in ihren wissenschaftlichen Arbeiten hin, indem sie die ATE als „challenging task“ (Rigouts Terryn et al. 2020: 85) oder „an unsolved problem with considerable room of improvement“ (Hazem et al. 2020: 95) bezeichnen. Dieser herausfordernde Charakter der ATE hängt von verschiedenen Faktoren und Entwicklungen im Bereich der Computerlinguistik, wie z.B. der Einsatz der künstlichen Intelligenz, im Bereich der Technik und Sprachtechnologien, wie z.B. die Entwicklung neuer Tools und Methoden, und im Bereich der Linguistik, wie z.B. die Erweiterung der Definition von Terminus und ihre diversen Auffassungen, ab. Auf die einzelnen Faktoren und Entwicklungen wird in der Masterarbeit vor allem im theoretischen Teil näher eingegangen. Was die ATE noch zusätzlich herausfordernder macht, ist die Breite an Anwendungsmöglichkeiten der Tools in verschiedensten Anwendungsbereichen (siehe Abschnitt 2.4.4), die sich nicht ausschließlich auf reine Terminologielistenausgabe beschränken, sondern tief in das Gebiet des Data Mining reichen. Es ist deutlich, dass auf die Termextraktion-Systeme heutzutage in vielen Wissensbereichen nicht mehr verzichtet werden kann. Diese Unverzichtbarkeit resp. Nachfrage gekoppelt mit dem Einzug des maschinellen Lernens in die ATE öffnet diverse Möglichkeiten für die Entwicklung neuer Tools, die immer komplexer werden. Zurzeit gibt es Bemühungen, Tools resp. Konzepte zu entwickeln, die mit den neusten Entwicklungen und Anforderungen mithalten können. Als

Beweise für diese Entwicklung fungieren zum Beispiel die aktuellen Tools wie D-Terminer (Rigouts Terryn et al. 2022) und Text2TCS (Gromann et al. 2021). Bei der Geschwindigkeit, mit der die neuen Tools entwickelt und präsentiert werden, ist es nicht nur für die Forscher*innen, sondern auch für die potenziellen Anwender*innen fundamental zu wissen, was das neue System leisten kann, wo die Unterschiede zu den Systemen, die bis dato verwendet worden sind, bestehen, und welches System die besten Ergebnisse liefert.

Diese Überlegungen führten auch zum Thema dieser Masterarbeit, die darauf abzielt, neueste Methoden der automatischen Termextraktion mit unterschiedlichen Architekturen miteinander zu vergleichen. Allerdings geht es hier nicht um einen umfangreichen, funktionalen und methodologischen Vergleich von ATE-Systemen per se, sondern um einen Vergleich der generierten Ergebnisse, d. h. Termkandidatenlisten, da vor allem diese für den Einsatz als Ausgangsressource in der Praxis, ungeachtet dessen, ob es sich um eine Terminologiedatenbankerstellung oder um eine Termextraktion für Übersetzungszwecke handelt, von höchster Relevanz sind. Der Vergleich erfolgt unter Anwendung der bei der Evaluation der automatischen Termextraktion üblichen Metriken Genauigkeit, Trefferquote und F1-Score, wobei man in den aktuellen wissenschaftlichen Arbeiten die Meinung beobachten kann, dass die Anwendung dieser Metriken für die weitere Forschung nicht mehr ausreichend ist, da viele Ansätze des maschinellen Lernens unvorhersehbar sind und die Ergebnisse deshalb variieren können (Rigouts Terryn et al. 2020). Für die Zwecke und Rahmenbedingungen dieser Masterarbeit sind diese Metriken allerdings ausreichend und werden durch manuelle Analysen der Ergebnisse ergänzt, um detailliert konkrete Herausforderungen einzelner Methoden festzustellen, wie z. B. Länge von Termini oder Umgang mit Akronymen.

Diese Masterarbeit ist wie folgt aufgeteilt. Im ersten Unterkapitel der Einleitung wird die Forschungsfrage präsentiert, anschließend wird im zweiten Unterkapitel auf die Motivation und Zielsetzung detailliert eingegangen. Im darauffolgenden Kapitel werden die Konzepte für die theoretischen Grundlagen näher erläutert. Im Kontext der Fachkommunikation und Terminologie werden verschiedene Definitionen und Auslegungen zu den Begriffen Fachkommunikation, Terminologie und Terminus verwendet und diskutiert. Anschließend werden Konzepte des eigentlichen Themas ATE, ihre Einbettung bzw. Kontext, Anwendungsbereiche, Klassifikationen der Tools und Methoden, u.v.m. vorgestellt. Das theoretische Kapitel wird mit der Problematik der Evaluationsmethodik abgeschlossen. Im darauffolgenden Kapitel wird der Stand der Forschung beschrieben. Hier werden vor allem die Beiträge aus den zwei Terminologie-Workshops im Rahmen der International Conference on Language Resources and Evaluation (LREC) (Daille et al. 2020; Rute Costa et al. 2022) präsentiert. Im praktischen Teil der Masterarbeit

werden die Methodik der Textauswahl, Auswahl der Termextraktionssysteme, und Evaluationsmethode beschrieben. Hier werden die wichtigsten Komponenten wie das Blackbox-Experiment, Korpus, Auswahl an Tools, manuelle Termextraktion und Goldstandard u.v.m. nähergebracht. Es folgt eine Ergebnisauswertung, die zur Diskussion und dem Fazit führt.

1.1 Forschungsfrage und Annahmen

Die Forschungsfrage der vorliegenden Masterarbeit lautet: Welcher der ausgewählten, aktuellen Termextraktionsansätze, erzielt in der Textsorte Zeitungsartikel und dem Fachgebiet Energiekrise auf Deutsch das beste Ergebnis nach Genauigkeit, Trefferquote, F1-Score und manueller Beurteilung? Die Forschungsfrage liefert den Hinweis auf die wichtigsten Themenschwerpunkte, die man grob unter folgenden Stichpunkten zusammenfassen kann: ATE, Entwicklungstrends, neue Ansätze und Evaluationsmetriken bzw. -methoden.

Die gewählte Textsorte der Zeitungsartikel impliziert, dass die Texte nicht nur fachspezifische Termini beinhalten werden, sondern auch allgemeinsprachliche Ausdrücke. Es ist anzumerken, dass in dieser Arbeit kein Vergleich der Ergebnisse unter Verwendung verschiedener Korpora beabsichtigt ist, allerdings ist die Tatsache, dass der Korpus aus Zeitungsartikeln besteht, für einige theoretischen und empirischen Grundlagen bzw. Entscheidungen ausschlaggebend, z.B. was wird als Terminus in diesem Vergleich klassifiziert und was nicht. Dass man diesen Punkt nicht außer Acht lassen darf, zeigt auch die wissenschaftliche Arbeit von Drouin (2020:1), der die Textsorte Zeitungsartikel als „atypical textual resources“ für die Terminologie bezeichnet, da die Zeitungsartikel sich nicht an Fachexpert*innen sondern an die allgemeine Öffentlichkeit wenden.

Die zentrale Annahme ist, dass die neuronalen Ansätze in allen gemessenen Kategorien bessere Ergebnisse erzielen, worauf auch die wissenschaftliche Arbeit von Hazem (2020) hinweist, die als Ergebnis hervorbrachte, dass ein Ansatz mit einem Deep Neural Network, die merkmalsbasierten Ansätze übertrifft. Daraus folgend wird in diesem Vergleich angenommen, dass neuronale Ansätze die gewählten klassischen ATE-Methoden übertreffen.

Weiters wird angenommen, dass trotz der Überlegenheit der Systeme mit neuronalen Ansätzen auch die auf nicht-neuronalen Ansätzen basierten Systeme ihre Stärken zeigen werden und für ihre Anwender*innen durchaus Vorteile bringen können.

1.2 Motivation und Zielsetzung

Die Motivation dieses Themas liegt in der raschen Entwicklung der ATE-Ansätze und dem Bestreben nach einer vollautomatischen Termextraktion, die bis dato zwar noch nicht erreicht wurde. Derzeit erfordert das ausgegebene Ergebnis noch manuelle Nachbearbeitung bzw. -bereinigung. Allerdings zeigen aktuelle wissenschaftliche Arbeiten, wie z.B. von Rokas et al. (2020), Jerdhaf et al. (2022), Wachowiak et al. (2021) u.v.a., dass man mit den aktuellen neuronalen Ansätzen Ergebnisse erzielen kann, die für die jeweiligen Anwender*innen von Bedeutung sein könnten. Es werden immer neue Methoden bzw. Umgebungen entworfen, die wohl eine entscheidende Rolle bei der Wahl des geeigneten Tools spielen. In dieser Arbeit geht es nicht um einen Versuch eine neue Methode zu entwickeln, sondern um einen Vergleich der Leistung der ausgewählten Ansätze unter den gleichen experimentellen Bedingungen. Erstens soll dadurch der Fortschritt in der Entwicklung der ATE-Tools am praktischen Beispiel sichtbar gemacht werden und zweitens soll potenziellen Anwender*innen, eine Übersicht der Unterschiede der jeweiligen Ansätze präsentiert werden, um sich dann je nach Zweck der Nutzung der Terminologie für den richtigen Ansatz entscheiden zu können. Anhand der Ergebnisse könnte man die jeweiligen Probleme spezifisch in weiteren Forschungsansätzen angehen und zur Verbesserung der Ansätze zur Termextraktion beitragen.

Die Ergebnisse dieser Arbeit könnten auch für die Übersetzer*innen von Bedeutung sein, vor allem in ihrer Ausbildung und in der mit der Übersetzung einhergehenden Terminologearbeit. Durch die detaillierte Analyse der Vorteile und Problemstellungen der jeweiligen Ansätze kann das Verständnis für die geeignete Wahl eines konkreten Ansatzes je nach Situation geschärft werden und die Wichtigkeit von Terminologie in der Übersetzung nähergebracht werden.

Was für die Forschung relevant sein könnte, ist die Durchführung der ATE aus einem Korpus bestehend aus Zeitungsartikeln. Weiters könnte die manuelle Annotation von Termini im Korpus auch aus Sicht der Terminologiewissenschaft von Interesse sein. Es werden Zeitungsartikel aus einem bestimmten Zeitraum, konkret von 24.02.2022 bis 24.02.2023, als Korpus verwendet. Eine Analyse der Terminologie aus anderen Zeiträumen könnte die diachrone Entwicklung der Terminologie zum gleichen Thema der Energiekrise aufzeigen.

2 Theoretische Grundlagen

In diesem Kapitel werden die theoretischen Konzepte und Begriffe erläutert, die als Grundlage für die vorliegende Masterarbeit von Bedeutung sind. Ziel ist es, einen Überblick über die wissenschaftlichen und fachlichen Aspekte zu geben, die im Rahmen der Arbeit untersucht und angewandt werden. Dazu gehören die Grundlagen der Terminologie und ihre Bedeutung in der Fachkommunikation, die automatische Termextraktion (ATE) samt ihren unterschiedlichen Ansätzen, sowie die Methoden der Evaluation. Zunächst wird auf die Terminologie selbst eingegangen, um die Bedeutung von Termini im Kontext der Fachkommunikation und Wissensvermittlung darzustellen. Ein weiterer Schwerpunkt liegt auf der Terminologiarbeit und ihrer Rolle in der Praxis, insbesondere in der automatisierten Verarbeitung großer Textmengen. Dabei wird die ATE als zentrales Werkzeug zur Extraktion und Analyse von relevanten Termini in Texten eingeführt und die verschiedenen Ansätze zur Evaluierung der Extraktionsergebnisse vorgestellt. Im weiteren Verlauf des Kapitels werden die Terminologiekonzepte im Zusammenhang mit ATE detaillierter beschrieben, einschließlich der Kategorisierung von Termini. Abschließend wird auf die Herausforderungen und Probleme der Evaluationsmethodik bei der ATE eingegangen, insbesondere in Bezug auf die Verwendung klassischer Metriken wie Genauigkeit, Trefferquote und F1-Score.

2.1 Terminologie als Grundlage der Fachkommunikation

Bevor man sich mit der Rolle der Terminologie in der Fachkommunikation befasst, ist es wichtig, sich näher mit dem Begriff Terminologie auseinanderzusetzen, da der Begriff je nach Kontext unterschiedlich interpretiert werden kann (siehe Abschnitt 2.1.1). Der Schwerpunkt dieser Arbeit liegt auf der automatischen Termextraktion und daher wird im Verlauf dieses Unterkapitels näher auf die Termini und den Umgang mit derer Kategorisierung anhand der ausgewählten Forschungsarbeiten mit dem Fokus auf der ATE eingegangen. Anschließend werden weitere wichtige Bestandteile der Fachkommunikation erläutert, wie Fachbereich und Fachgebiet, Fachsprache und Fachtext. Abschließend wird die Wirtschaftssprache mit ihren typischen terminologischen Besonderheiten beschrieben.

2.1.1 Gegenüberstellung von Terminologie und Fachwortschatz

Drewer und Schmitz (2017: 5) definieren Terminologie als „[d]ie Gesamtheit der Fachwörter eines Fachgebietes“. Diese Definition lässt keine andere Interpretation des Begriffes Terminologie zu, und daher kann geschlussfolgert werden, dass Terminologie und Fachwortschatz eines bestimmten Fachgebietes gleichbedeutend sind. Allerdings betonen sie auch (ibid.: 6), dass im Laufe der Zeit durch das zunehmende Interesse an Fachwörtern fälschlicherweise auch die wissenschaftliche Disziplin Terminologiewissenschaft, die sich mit Grundsätzen der Terminologearbeit befasst, als Terminologie bezeichnet wird. Dies kann unter anderem auch daran liegen, dass der englische Begriff „terminology“ als Terminologie ohne Rücksicht auf den Kontext ins Deutsche übersetzt wird. In einigen wissenschaftlichen Publikationen (Kageura 2015; Shuttleworth & Cowie 1997) wird betont, dass der Begriff „terminology“ je nach Kontext unterschiedliche Bedeutungen haben kann. Shuttleworth und Cowie (1997) verstehen unter dem Begriff je nach Kontext erstens eine umfangreiche Sammlung spezifischen Fachwortschatzes und zweitens eine Bezeichnung einer bestimmten Disziplin, wie z.B. in der Translationswissenschaft das Terminologiemanagement (ibid.: 166). Kageura (2015) hebt hervor, dass dem Begriff sogar drei Bedeutungen zuzuschreiben sind: erstens handelt es sich hier um eine Bandbreite an verschiedenen Praktiken- und Methodensammlungen, die zur Erfassung der Termini dienen, zweitens um eine Theorie, die verschiedene Konzepte ihren jeweiligen Begriffen zuordnet und drittens um eine Wortschatzsammlung eines spezifischen Fachgebietes (ibid.:45). In dieser Arbeit wird unter Terminologie die Auffassung aus der Norm DIN 2342 (2022) übernommen, da die Definition keine verschiedenen Begriffe zulässt, wie es der Fall bei Kageura (2015) ist. Erstens wird Terminologie in der Norm (DIN 2342 2022) dem Begriff Fachwortschatz gleichgestellt, zweitens als ein Produkt der Terminologearbeit angesehen und drittens als „Gesamtbestand der Begriffe und ihrer Bezeichnungen in einem Fachgebiet“ definiert, wobei sie ein- oder mehrsprachig sein kann. Diese Definition deckt sich lediglich mit der dritten Auffassung der „terminology“ von Kageura und zeigt, dass der Begriff Terminologie im Deutschen nicht mit der Terminologearbeit (siehe Abschnitt 2.3) und -wissenschaft (siehe Abschnitt 2.2) zu verwechseln ist, obwohl man im Englischen dieselbe Bezeichnung für diese drei unterschiedliche Konzepte verwendet.

Wenn Terminologie mit Fachwortschatz gleichgesetzt wird, muss in weiterer Folge der Begriff Fachwortschatz näher erläutert werden. In der Sprachwissenschaft definiert Roelcke (2020: 70f.) den Fachwortschatz als eine Menge der Fachwörter, wobei er das Fachwort bzw. den Terminus als „die kleinste bedeutungstragende und zugleich frei verwendbare sprachliche

Einheit eines fachlichen Sprachsystems, die innerhalb der Kommunikation eines bestimmten menschlichen Tätigkeitsbereichs im Rahmen geäußerter Texte gebraucht wird“ betrachtet. An dieser Stelle wird von einer nicht näher charakterisierten kleinsten Einheit ausgegangen. Wenn diese sprachliche Einheit nicht nur auf ein einzelnes Wort reduziert ist, spricht man von einem Terminus (siehe Abschnitt 2.2.3). Allerdings kann nicht jede sprachliche Einheit als ein Terminus bestimmt werden. Diese Einheit benötigt ein bestimmtes Umfeld resp. fachliches Sprachsystem. Der Terminologiewissenschaftler Kageura (2015: 47) bezeichnet dieses Umfeld als „functional class“, d. h. dass eine sprachliche Einheit erst innerhalb einer Domäne, eines Gegenstandes oder eines Berufes zum Terminus wird, wenn sie durch ein Kollektiv anerkannt wird. Weiters betont er, dass die Berücksichtigung der extralinguistischen Symbole, wie z.B. chemische Formeln oder fremdsprachige Nomenklaturen, der individuellen Wahl obliegt, je nach der Beschaffenheit einer Terminologie (ibid.: 47f). Terminologie weist außer ihrer Fachbezogenheit auch weitere typische Eigenschaften auf. Erstens kommt eine typische Wortart Nomen vor, wobei andere Wortarten, wie vor allem Verben, Adjektive und Adverbien, nicht automatisch ausgeschlossen werden dürfen, zweitens sind häufig komplexe Termini vertreten, drittens zeichnet sie sich durch Eindeutigkeit der Termini aus, d. h. der dem Begriff zugewiesene Ausdruck hat eine bestimmte Bedeutung und bietet daher keinen Raum für Fehlinterpretationen, wobei man Synonyme (siehe Abschnitt 2.2.4.3.2) und Polysynonyme (siehe Abschnitt 2.2.4.3.3) nicht völlig ausschließen kann, und viertens beinhaltet sie fremdsprachige Termini, vor allem aus dem Englischem entlehnt (ibid.:47ff).

Abschließend ist kurz anzumerken, dass eine bestimmte Terminologie keine geschlossene Ansammlung von Fachwörtern und Sprachzeichen ist, sondern ständig erweitert werden kann. Dies geschieht nicht willkürlich, sondern anhand bestimmter Verfahren, an die näher im Abschnitt 2.1.8.3 eingegangen wird.

2.1.2 Bedeutung von Terminologie

Wenn man sich die theoretischen Publikationen (Arntz et al. 2014; Drewer & Schmitz 2017) zum Thema Terminologearbeit und Terminologiemanagement betrachtet, weisen die Autor*innen auf die Tatsache hin, dass Terminologie die Basis der mündlichen und schriftlichen Fachkommunikation bildet. Unter Fachkommunikation versteht Roelcke (2020) eine Art Kommunikation, die voraussetzt, dass Produzent*innen und Rezipient*innen über ein gemeinsames sprachliches Zeichensystem und gemeinsamen Text- und Sachkontext verfügen (ibid.:13), d. h. das gemeinsame Wissen bzw. Vorwissen ist ausschlaggebend, um erfolgreich an der

beabsichtigten Kommunikation teilnehmen zu können. Bei Drewer und Schmitz (2017) wird auf die zwangsläufige Verwendung eines sprachlichen Zeichensystems in der Fachkommunikation hingewiesen, und zwar der Fachsprache. Einer breiten Definition, bedienen sich auch Heine und Schubert (2013), die die Akteur*innen der Fachkommunikation in einem beruflichen Umfeld, das durch einen bestimmten fachlichen Inhalt und Handeln gekennzeichnet ist, ansiedeln. Dabei ist es erforderlich, dass die verwendeten Kommunikationsmittel zielgerichtet, informativ und optimiert sind (ibid.: 103). Daraus folgt, dass die Fachkommunikation eine wesentliche Rolle in all den Bereichen spielt, in denen das Fachwissen bzw. die Fachwissens- und Terminologiebestände von Belang sind, und dass die Fachkommunikation bestimmte Kriterien erfüllen muss: bestimmtes (Berufs-)Umfeld bzw. bestimmter Fachbereich, gemeinsame Sprache, (Vor-)Wissen und Interesse an fachlichen Inhalten der Kommunikationsteilnehmer*innen.

Der fortlaufende Zuwachs an Wissen hat einen enormen Einfluss auf zahlreiche Bereiche. Dies betrifft nicht nur hochspezialisierte Fachgebiete wie Informatik, Technik und Wirtschaft (siehe Abschnitt 2.1.5 für eine Differenzierung zwischen Fachbereich und Fachgebiet) und die Forschung und die Kommunikation zwischen Expertinnen, sondern auch die Kommunikation zwischen Expertinnen und Lai*innen ist betroffen. Dabei findet das Fachwissen, insbesondere die Terminologie, oft Einzug in das alltägliche Leben. Es kommt zur Popularisierung des Fachwissens (Roelcke 2020). Dies zeigt sich zum Beispiel, wenn man beim Durchblättern einer Zeitung auf Begriffe wie *Brownout* oder *Blackout* stößt. Auch beim Kauf eines Elektrogeräts wird man zwangsläufig mit einem Fachtext oder einer Betriebsanleitung konfrontiert, die zahlreiche Termini enthalten.

Ein weiteres Beispiel, warum die Terminologie immer mehr an Bedeutung gewinnt, liefern Drewer und Schmitz (2017), die Terminologie als einen maßgeblichen Qualitätsfaktor für Unternehmen bezeichnen. Dieser Faktor sorgt nicht nur für eine konsistente Vermittlung des Fachwissens innerhalb des Unternehmens, sondern stärkt dessen Image außerhalb des Unternehmens bei den Verbraucher*innen, und zwar mittels einer sogenannten Corporate Language, d. h. einer Unternehmenssprache, die auch durch die firmenspezifische Terminologie geprägt ist (ibid: 1f.).

Ein weiteres Beispiel erwähnen Arntz et al. (2014), die die Vorteile der Terminologie und Terminologearbeit für die Tätigkeit der Fachübersetzer*innen hervorheben und betonen, dass bei einem heutzutage derart rasanten Zuwachs an immer spezialisierteren Fachtexten und dementsprechend auch an Termini, bloße Fachwörterbücher beim Übersetzen nicht mehr ausreichend sind. Außerdem erwähnen die Autor*innen etliche andere Berufsgruppen, für die die Terminologie und Terminologearbeit eine wichtige Rolle spielen wie technische

Redakteur*innen, Sprachmittler*innen, Terminolog*innen, Forscher*innen, Ingenieur*innen u.v.a. (ibid: 1f.). Eine eindeutige Antwort auf die oben gestellte Frage liefert unter anderen auch Witschel (2004: 1), indem er die Terminologie als „Schlüssel zum Verständnis vieler Fachgebiete“ bezeichnet.

Ausgehend von den angeführten Beispielen kommt man zur Schlussfolgerung, dass Terminologie aus vielen Bereichen nicht mehr wegzudenken ist. Daher ist es notwendig, sich mit ihr sowohl auf der praktischen als auch auf der theoretischen Ebene näher auseinanderzusetzen.

2.1.3 Gliederung von Termini

Die Struktur des Fachwortschatzes spielt für die Terminologiewissenschaft und folglich auch für die ATE eine wesentliche Rolle, indem man anhand der Kategorien der Termini z.B. präzisere Text- und Tools-Analysen durchführen kann. Allerdings gibt es keine allgemeingültige Struktur für den Fachwortschatz, da dieser erstens von Fachgebiet zu Fachgebiet unterschiedlich ist und zweitens, wie Arntz et al. (2014) betonen, es keine eindeutige Grenze zwischen der Gemeinsprache und Fachsprache gibt. Diese Tatsache verdeutlichen Arntz et al. (2014: 26f.) nachdem Sie einige Modelle mit unterschiedlicher Komplexität untersucht haben, wie z.B. das Modell der Dreiteilung in allgemeinen, allgemeinwissenschaftlichen und speziellen Wortschatz oder das Modell von Filipec (1969: 57), indem er sich mit der Spezifik des spezialsprachlichen Wortschatzes bzw. „Sonderwortschatzes“ gegenüber dem allgemeinen Wortschatz beschäftigt. Vereinfacht zusammengefasst unterteilt er Termini in eigentliche Termini und Quasitermini, d. h. ungenaue, unbefriedigende Termini, und stellt sie den Nicht-Termini gegenüber (ibid.). Arntz et al. (2014: 27) betonen, dass die sprachliche Realität so vielfältig ist, dass solche Einteilungen des Wortschatzes nur als ein bloßer Versuch einer Kategorisierung bleiben. Ein weiterer Versuch einer solchen Einteilung kommt aus der Sprachwissenschaft, und zwar von Roelcke (2020). In seinem Gliederungsmodell unterscheidet er zwischen vier Gruppen der Fachwörter (Abb. 1). Laut seiner Auffassung gibt es erstens den intrafachlichen Fachwortschatz, d. h. Fachwörter, die ausschließlich einer bestimmten Fachsprache zugeordnet werden können, zweitens den interfachlichen Fachwortschatz, d. h. Fachwörter, die einer bestimmten Fachsprache zugeordnet werden können, aber nicht ausschließlich, drittens den extrafachlichen Fachwortschatz, d. h. Fachwörter, die anderer Fachsprache angehören, aber in der untersuchten Fachsprachen vorkommen und viertens den nichtfachlichen Fachwortschatz, d. h. Fachwörter, die dem allgemeinen Sprachwortschatz angehören (Roelcke 2020: 71).

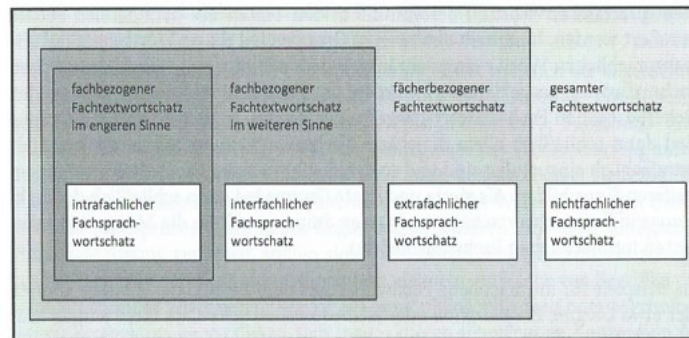


Abbildung 1: Gliederung von Fachsprach- und Fachtextwortschatz nach fachlicher Zugehörigkeit der betreffenden Fachsprach- und Fachtextwörter (Roelcke 2020: 71)

Auf den ersten Blick scheint das Modell unproblematisch zu sein, allerdings, wenn man sich in der Praxis mit der Zuordnung der Fachwörter zu jeweiliger Gruppe befasst, stößt man auf ein wesentliches Problem, das darin besteht, dass nicht ersichtlich ist, wer und nach welchen Kriterien es entschieden wird, welches Fachwort zu welcher Gruppe gehört. Adamzik (2018) geht sogar so weit, dass sie diese Gliederung zwar in der Theorie sinnvoll betrachtet, in der Praxis allerdings nicht unbedingt (ibid.: 297). Sie schlägt eine grobe Unterteilung der Fachwörter je nach ihrer Erkennbarkeit (Abb. 2), d. h. ob sie in ihrer Form bekannt oder unbekannt sind, wobei man die Subjektivität dieser Erkennung durch Nachschlagwerke ausgleichen soll.

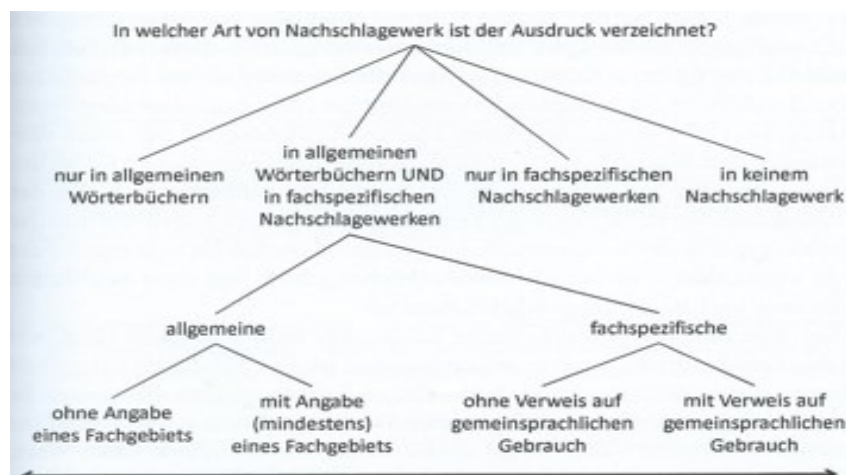


Abbildung 2: Vorkommen eines Ausdrucks in Nachschlagwerken (Adamzik 2018: 299)

Adamzik (2018) selbst, sieht diese Gliederung nicht als optimal an, da es auch hier einige Fragen gibt, die vor allem mit der Tatsache zu tun haben, dass der Wortschatz des jeweiligen Fachgebiets sich ständig weiterentwickelt und eine einheitliche und vor allem vollständige Liste an Fachwörtern in keinem Nachschlagwerk zu finden ist (ibid.: 300).

Wie die oben erwähnten Gliederungsmodelle aufzeigen, ist an eine universale und vor allem absolut gültige Gliederung nicht zu denken, allerdings helfen solche Gliederungen die Terminologien zu ordnen und daher auch leichter verwalten resp. aktualisieren.

2.1.4 Gliederung von Termini in der ATE

Die Kategorisierung von Termini spielt eine entscheidende Rolle bei der ATE. Bereits in der Grundidee der Termini-Erkennung ist sie unverzichtbar. Vereinfacht ausgedrückt, für die Extraktion angewendetes Tool muss in erster Linie Termini von Nicht-Termini unterscheiden können. Schon diese einfache binäre Kategorisierung erweist sich als eine schwierige Aufgabe, da es in der ATE-Forschung keinen Konsensus darüber gibt, wie spezifisch ein Termkandidat, also ein potenzieller Terminus, sein soll, um tatsächlich als „true term“ (siehe unten) klassifiziert zu werden (Rigouts Terryn et al. 2020: 87). Um diese Aufgabe zu lösen, bedienen sich die Forscher*innen in der ATE unterschiedlicher Methoden bzw. Konzepte, unter denen „termhood“ und „specificity“ öfters ihre Anwendung finden, wie bei Kageura und Umino (Kageura & Umino 1996) und Hazem et al. (2020). Unter dem linguistischen Konzept „termhood“ verstehen Kageura und Umino (Kageura & Umino 1996: 274) die Verknüpfung einer sprachlichen Einheit mit einem bestimmten Fachgebiet. Somit ist Termhood für komplexe und auch einfache sprachliche Einheiten relevant. Das Konzept „specificity“ wird angewandt, indem man die Häufigkeit des Vorkommens eines Terminus im bestimmten Kontext bzw. Fachgebiet in Relation zu seiner Häufigkeit des Vorkommens außerhalb dieses Fachgebietes setzt (Hazem et al. 2020: 96), d. h. dass auch ein Terminus, der nur selten vorkommt, identifiziert werden kann.

Um die Leistung eines ATE-Tools überhaupt auswerten zu können, benötigt es eines Vergleichs der gelieferten Ergebnisse mit einem Standard, sog. Goldstandard (siehe Abschnitt 2.4.2.4), der nach vorher bestimmten Regeln annotiert wird. Wie Rigouts Terryn et al. (Rigouts Terryn et al. 2020: 87f.) hervorheben, tragen diese Regeln zur Transparenz des Experiments und zur Vereinfachung des Entscheidungsprozesses bei der Annotation und daher zur Konsistenz bei. Wenn man sich die Forschungsarbeiten in der ATE näher betrachtet, stellt man fest, dass auch hier keine einheitliche Kategorisierung vorliegt. Zum Beispiel wählen Jerdhaf et al. (2022: 30) für ihren Goldstandard die Kategorien „yes-terms (good candidates), no-terms (bad candidates) and u-terms (unsure and ambiguous candidates)“ aus. Ähnlich auch Drouin et al. (Drouin et al. 2020: 4), die bei der Evaluation der Termkandidaten eine Einteilung in „Term“, „Non-Term“ und „Uncertain“ angewendet haben. Spezifischer ist es bei Rigouts Terryn et al. (2020), die für ihre Forschung sogar vier Kategorien unter der Beachtung der Kriterien

„domain-specificity“, d. h. der Relevanz des Terminus im Fachgebiet, und der „lexicon-specificity“, d. h. der Bekanntheit des Terminus im Expert*innenkreis, identifizieren. Sie unterscheiden zwischen „Specific Terms“ (ST), die fachspezifisch und lexikonspezifisch sind, „Common Terms“ (CT), die fachspezifisch, aber nicht lexikonspezifisch sind, „Out-of-Domain (OOD) Terms“, die weder fachspezifisch noch lexikonspezifisch sind und „Named Entities“ (NE), d. h. Personen, Organisationen, geographische Benennungen usw., die zwar nicht zwangsläufig als Termini gelten, allerdings so oft in fachspezifischen Kontext vorkommen, dass es für Terminologie eines bestimmten Fachgebietes Sinn macht, diese einzubeziehen (ibid.: 88). Die ersten drei Kategorien werden in der Forschung von Rigouts Terryn et al. (2020: 88) unter der Bezeichnung true terms in eine Kategorie zusammengefasst.

2.1.5 Unterscheidung zwischen Fachgebiet und Fachbereich

Wie aus bisherigen Abschnitten dieses Unterkapitels ersichtlich ist, spielen außer dem zentralen Begriff Terminologie, auch die Begriffe Fachgebiet, Fachtext und Fachsprache eine wesentliche Rolle. Im allgemeinsprachlichen Gebrauch werden Begriffe Fachbereich und Fachgebiet häufig synonym verwendet. Wenn man allerdings die Definitionen näher betrachtet, stößt man auf Unterschiede. Im Falle des Fachbereiches handelt es sich um eine größere organisatorische Einheit, die mehrere verwandte Disziplinen oder Fachrichtungen abdeckt, wobei es sich im Falle von Fachgebiet um einen engeren Wissensbereich handelt. Die DIN 2342 Terminologiewissenschaft und Terminologiarbeit (2022) definiert das Fachgebiet als einen „spezialisierte(n) Bereich des Wissens“ (ibid.), wobei die Grenzen zwischen Fachgebieten je nach Zweck festgelegt werden. Das bedeutet, dass ein Wissens- bzw. Fachbereich mehrere (Fach-)gebiete beinhalten kann. Auch, wenn es bei Roelcke (2020) nicht explizit erwähnt ist, kann man bei ihm beobachten, dass er wohl zwischen Wissens- bzw. Fachbereichen und Fachgebieten bzw. einzelnen Fächern unterscheidet (ibid: 68), wobei Fach als „ein [...] spezialisierter menschlicher Tätigkeitsbereich“ (ibid.: 14) zu definieren ist. Adamzik (2018) führt diesen Begriff noch weiter aus, indem sie vorschlägt den Begriff um einen weiteren Aspekt der Gegenstandsperspektive zu erweitern und das aufgrund der Tatsache, dass wenn sich verschiedene Expert*innengruppen mit demselben Gegenstand beschäftigen, diese infolgedessen ihren eigenen Gegenstand innerhalb ihres Faches schaffen. Sie definiert Fach als „ein Gegenstandsreich, auf dessen (kommunikative) Bearbeitung unter einer bestimmten Perspektive sich in einer Gesellschaft bestimmte Menschen spezialisiert haben. Die Aufgliederung in verschiedene Fächer variiert historisch und kulturell sehr stark (ibid.: 95)“.

Zum Zweck der Übersichtlichkeit wird in dieser Masterarbeit der Begriff Fachgebiet bzw. Fach verwendet. Das gewählte Fachgebiet ist jenes der Energiekrise, welches dem größeren Gebiet der Wirtschaft angehört.

2.1.6 Fachtext

Die Fachkommunikation bringt Fachtexte hervor, die im Gegensatz zu allgemeinen Texten bestimmte Merkmale aufweisen, z.B. das häufige Vorkommen von Termini. Ziel der Fachtexte ist eine fehlerfreie und effiziente Kommunikation und deshalb gelten die Fachtexte als unverzichtbare Instrumente für den Austausch von Informationen.

2.1.6.1 Merkmale von Fachtexten

Crestani (2010) geht davon aus, dass der Fachtext ein Produkt ist, das durch die Tätigkeit in einer Gesellschaft entsteht, wobei gewisse Voraussetzungen gegeben werden sollten, und zwar die kommunikative zwischenmenschliche Interaktion und eine produktive Tätigkeit (ibid.: 34). Roelcke (2020:129), ausgehend von zahlreichen wissenschaftlichen Konzeptionen des Fachtexts wie von Lothar Hoffmann und Klaus-Dieter Baumann, versteht unter diesem Begriff komplexe und kohärente Äußerungen, die nicht nur einen sprachlichen sondern auch nichtsprachlichen Charakter haben, z.B. Formeln und Graphiken, die einem kommunikativen Zweck in einer bestimmten fachlichen Umgebung dienen. Die Gemeinsamkeit der Definitionen liegt in der kommunikativen Funktion im konkreten Fachbereich.

Fachtexte weisen bestimmte Eigenschaften und Merkmale auf, damit man sie von nicht-fachlichen Texten unterscheiden kann. Zu den Eigenschaften, die sie gemeinsam mit den nicht-fachlichen Texten aufweisen, wenn auch in anderen Graden der Ausprägung, gehören nach Roelcke (2020): Intentionalität, Akzeptabilität, Informativität, Situationalität, Intertextualität und Objektivität. Die Merkmale eines Fachtexts unterteilt Roelcke (ibid.) allgemein in lexikalische, grammatische und in textuelle Besonderheiten (ibid.: 159). Da für diese Arbeit vor allem die Merkmale betreffend Lexik von hoher Relevanz sind, wird auf diese vorzugsweise eingegangen, vor allem auf die lexikalischen/terminologischen Besonderheiten im bestimmten Fachbereich, am konkreten Beispiel des Fachbereichs Wirtschaft (siehe Abschnitt 2.1.8.3). Einen kurzen Überblick über die wichtigsten grammatischen morphosyntaktischen Merkmale haben Arntz et al. (2014: 26 f.) in sieben Punkten zusammengefasst: 1. Verben findet man meistens im Präsens vor, 2. Passiv wird häufig verwendet, 3. Verb weicht anderen Wortarten, 4.

Pronomen sind selten zu finden, 5. Substantiv ist sehr wichtig, 6. Singular dominiert über Plural, 7. Adjektive werden häufiger als in Nichtfachtexten verwendet. Unter den textuellen Besonderheiten versteht Roelcke (2020) die Elemente, die die Kohärenz und Kohäsion der Fachtexte gewährleisten. Diese sind auf der Makroebene zu finden: 1. Textbaupläne, 2. Textbausteine wie Zitate oder Fußnoten, 3. metasprachliche Elemente, 4. typographische Konventionen und 5. nonverbale Elemente wie Bilder. Auch auf Mikroebene gibt es diese Elemente, z.B. Frage/Antwort-Konstruktionen u.v.m. (ibid.: 154).

2.1.6.2 Vielfalt der Fachtexte

Wenn es auf die Frage der Typologisierung der Fachtexte bzw. -sorten kommt, stellt man fest, dass das Interesse nach einer universalen Typologie einem beschränkten wissenschaftlichen Kreis der Linguist*innen vorbehalten ist, da es kein einheitliches Kriterium geben kann, um die Vielfalt an Fachtexten einer einzelnen Fachsprache, soweit es diese überhaupt gibt (siehe Abschnitt 2.1.7), zuzuordnen. Vielmehr waren schon seit den 90er Jahren die Grobunterscheidungen von Interesse, die als Grundlage der bekannten Typologien von Gläser (1990) und Göpferich (1995) dienen (Adamzik 2018: 259ff.). Gläser (1990) schlägt ein Modell zur Typologie von Fachtexten vor, dem verschiedene Kategorisierungsstufen, wie z.B. fachinterne und fachexterne Kommunikation, zugrunde liegen und unterscheidet zwischen Primärtextsorten, abgeleiteten Textsorten (z.B. Rezension), Prä-Textsorten (z.B. Abstract) oder Quasi-Textsorten (z.B. Protokoll). Die von Göpferich (1995) vorgeschlagene Typologie beschränkt sich auf den Fachbereich Naturwissenschaften und Technik, wobei die kommunikative Funktion der Fachtexte ausschlaggebend ist, d. h. es werden nur Informationstexte einbezogen. Roelcke (2020) versucht eine Gliederung vorzuschlagen, indem er sich an einer Typologie fachlicher Textsorten orientiert. Diese Fachtextsorten werden von ihm „als Typen oder Klassen von Fachtexten angesehen, die im Rahmen bestimmter Verwendungsweisen innerhalb fachlicher Kommunikation jeweils bestimmte funktionale und formale Gemeinsamkeiten aufweisen“ (Roelcke 2020: 56). Unter der Beachtung erstens der symbolischen Dimension, in der Informationstexte erfasst werden, zweitens der appellativen Dimension, in der Anweisungstexte erfasst werden und drittens der konzeptionellen Dimension, in der fachliche Texte entweder in mündlicher oder schriftlicher Form erfasst werden, unterscheidet er insgesamt acht Typen von Fachtexten. (Abb. 3):

	APPELLATIV			
	regulierend		sanktionierend	
	KONZEPTIONELL			
SYMBOLISCH	schriftlich	mündlich	schriftlich	mündlich
aktualisierend	Typ 1: Aufsatz in einer wissenschaftlichen Zeitschrift	Typ 2: Gespräch über neue Befunde im laufenden Laborbetrieb	Typ 3: Urteil über eine Freiheitsstrafe nach einem Prozess	Typ 4: Arbeitsanweisung im wissenschaftlichen Laborbetrieb
konfirmierend	Typ 5: Aufsatz in einer populärwissenschaftlichen Zeitschrift	Typ 6: Chat unter Auszubildenden über Prüfungsgegenstände	Typ 7: Mahnung bei säumiger Rechnung im Zahlungsverkehr	Typ 8: Erinnerung an Maßnahmen zur Soforthilfe

Abbildung 3: Textsortengliederung: Übersicht und Beispiele (Roelcke 2020: 64)

Wie aus der Abb. 3 hervorgeht, geht es um vier mündliche und vier schriftliche Prototypen für Fachtexte, wobei für den Zweck dieser Arbeit die schriftlichen Fachtexte von Bedeutung sind. Diese sind Typ 1 aktualisierend-regulierende Texte, wie wissenschaftliche Zeitschriftenartikel, Typ 3 aktualisierend-sanktionierende Texte, wie z.B. Urteile, Typ 5 konfirmierend-regulierende Texte, wie z.B. populärwissenschaftliche Zeitschriftenartikel, und Typ 7 konfirmierend-sanktionierende Texte, wie z.B. Mahnung (ibid.: 63f.). Dieses Gliederungsmodell besteht aus prototypischen Textbeispielen, die als Mittel zur Einordnung neuer Fachtexte behilflich sein sollten und das auf Basis eines Vergleichs zu diesen Prototypen. Weiter ist sich Roelcke (2020: 65) bewusst, dass solche Gliederungen ihre Schwächen haben, indem sie nur einen gewissen Teil der Fachtexte abbilden und plädiert für eine fachsprachliche Gesamttypologie (siehe Abschnitt 2.1.7), die es erlaubt, alle Fachtexte unabhängig von Fach und Bereich zu klassifizieren.

2.1.7 Fachsprache

Wie Fachsprache definiert wird, hängt davon ab welches wissenschaftliche Modell appliziert wird und welche Konzeptionen berücksichtigt werden. Im Rahmen des systemlinguistischen Inventarmodells spricht Roelcke (2020) von Fachsprachen als Zeichensystemen. Unter Berücksichtigung des pragmalinguistischen Kontextmodells spricht Roelcke beispielsweise von Fachsprachen als Textäußerungen (Roelcke 2020: 40). Da es den Rahmen dieser Arbeit übersteigt, sich den einzelnen Modellen und allen verschiedenen zugrundeliegenden Theorien im Detail zu widmen, wird an dieser Stelle die Definition aus der Norm DIN 2343 (2022) als ausschlaggebend betrachtet, da diese aus terminologischer Sicht konzipiert und daher am besten für die Zwecke dieser Arbeit geeignet. In der Norm wird Fachsprache als der „Bereich der Sprache, der auf eindeutige und widerspruchsfreie Kommunikation in einem Fachgebiet gerichtet ist“ (ibid.).

Wie aus dieser Definition hervorgeht, und wie auch Arntz et.al (2014), Roelcke (2020) und Adamzik (2018) betonen, gibt es zahlreiche unterschiedlichen Fachgebiete und daher auch zahlreiche Fachsprachen, wie z.B. Rechtssprache(n), Verwaltungssprache(n), Wirtschaftssprache(n), Fachsprache der Medizin, Biologie u.v.m. Adamzik (2018) geht von einer Einzelsprache aus, die kein geschlossenes System darstellt, sondern heterogen ist, d. h. unter anderen aus mehreren Fachsprachen bzw. Varietäten besteht.

Da es die eine Fachsprache nicht gibt und die Grenzen zwischen Fachbereich, Fachgebiet und Fach nicht eindeutig sind, ist es schwierig, eine eindeutige Gliederung der Fachsprachen vorzunehmen. Roelcke (2020) widmet diesem Thema ein ganzes Kapitel und bietet etliche schematische Darstellungen, die die unterschiedlichen Kriterien und Gesichtspunkte berücksichtigen. Zusammenfassend bietet er drei Kriterien, mithilfe welcher sich Fachsprachen gliedern lassen, und zwar nach Fachbereichen und einzelnen Fächern, die horizontale Gliederung, nach Abstraktions- und Kommunikationsebenen, die vertikale Gliederung und schließlich die Gliederung nach Fachtexten, wobei er darauf hinweist, dass es einer fachsprachlichen Gesamttypologie (Abb. 4) bedarf, um die Bandbreite an Fachgebieten und Fächern, ihrer unterschiedlichsten fachkommunikativen Situationen und die Breite an unterschiedlichsten Textsorten zu erfassen (ibid.: 68).

Horizontal: Wirtschaftssektoren	Vertikal: Experten und Laien	Textsorten: Symbol, Appell, Diskurs
Primär: Rohstoffgewinnung	Experten ein und desselben Faches (Typ 1)	aktualisierend-regulierende Texte mit schriftlichem Charakter (Typ 1)
Sekundär: Rohstoffverarbeitung	Laien in einem bestimmten Sachbereich (Typ 2)	aktualisierend-regulierende Texte mit mündlichem Charakter (Typ 2)
Tertiär: Dienstleitung	Experten verschiedener Ebenen oder Bereiche eines Faches (Typ 3)	aktualisierend-sanktionierende Texte mit schriftlichem Charakter (Typ 3)
Quartär: Informationsverarbeitung	Experten eines Faches und Laien im entsprechenden Sachbereich (Typ 4)	aktualisierend-sanktionierende Texte mit mündlichem Charakter (Typ 4)
	Experten eines Faches und Experten eines anderen Faches (Typ 5)	konfirmierend-regulierende Texte mit schriftlichem Charakter (Typ 5)
		konfirmierend-regulierende Texte mit mündlichem Charakter (Typ 6)
		konfirmierend-sanktionierende Texte mit schriftlichem Charakter (Typ 7)
		konfirmierend-sanktionierende Texte mit mündlichem Charakter (Typ 8)

Abbildung 4: Fachsprachliche Gesamttypologie (Roelcke 2020:66)

Diese Typologie vereint alle drei oben erwähnten Typologien und ist daher auf alle fachkommunikativen, schriftlichen und mündlichen Texte anwendbar (Roelcke 2020).

Die Definitionen in den DIN-Normen werden häufig durch Zusatzanmerkungen ergänzt. So ist es auch im Falle der Fachsprache, wobei als Anmerkung angeführt ist, dass die Terminologie ein wichtiger Bestandteil der Fachsprache ist (DIN 2342:2022). Da stellt sich die Frage, was die anderen Bestandteile der Fachsprache sind.

Einen allgemeinen Überblick der Merkmale bietet Crestani (2010), die insgesamt vier Merkmale der Fachsprache beschreibt, und zwar lexikalische, d. h. bestimmte Terminologie, die die Fachsprache nicht nur von der Allgemeinsprache, sondern von anderen Varietäten der Fachsprache, morphologische und syntaktische, d. h. bestimmte Wort- und Syntaxbildungsmuster, und textuelle, d. h. bestimmte Texte mit spezifischen Textbauplänen (Crestani 2010: 22). Wiederum Roelcke (2019) widmet seine Aufmerksamkeit bestimmten Ebenen, auf denen die Merkmale auftreten können. Auf lexikalischer Ebene begegnet man unter anderem z.B. den Entlehnungen aus dem Englischen, auf grammatischer Ebene z.B. den diversen Wortbildungen wie Komposita und auf Textebene den typischen Textsorten und Textbauplänen. Seine Unterteilung beinhaltet noch eine semiotische Ebene, die Elemente wie Formeln, Graphiken, Bilder usw. einbezieht, und eine pragmatische Ebene, die die Organisation des Textes und der Gedanken sowie die sog. „Beachtung von Writer- und Reader-Responsibility“ (Roelcke 2019: 2), d. h. Verantwortung für das (Vor-)Wissen sowohl bei den Textproduzent*innen als auch bei Textrezipient*innen, einbezieht (ibid.: 2). Bei beiden Autor*innen spielt bei den Merkmalen besonders die Terminologie eine wichtige Rolle.

2.1.8 Wirtschaftsfachsprache

Die gescheiterte Suche nach einer einheitlichen Definition der Fachsprache und einer einheitlichen Unterteilung der Termini lassen erahnen, dass dies auch der Fall bei der Suche nach der Definition der Wirtschaftssprache sein wird. Zudem muss an dieser Stelle auch die Diskussion der Zugehörigkeit der Wirtschaftssprache zu den Fachsprachen angesprochen werden. Maghețiu (2015) erwähnt, dass einige Wissenschaftler*innen in der Fachsprachenforschung die Zugehörigkeit der Wirtschaftssprache zu den Fachsprachen hinterfragen, indem sie argumentieren, dass viele wirtschaftliche Termini durch ihre ständige Medienpräsenz mittlerweile zu alltagssprachlichen Mitteln geworden sind (ibid.: 87). Diese Argumentation ist für Maghețiu eine schwache Begründung und deshalb macht sie deutlich, dass für sie die Frage der Zugehörigkeit der Wirtschaftssprache zu den Fachsprachen durch den Aspekt der Verständlichkeit, d.

h. dass die Wirtschaftssprache über eigene Terminologie, die ohne Erläuterungen für Lai*innen nicht verständlich sind und über spezifischen Metaphern, Phraseologismen und Wendungen wie z.B. *Boom* verfügt, geklärt ist (2015: 87f.). Was aber zu beachten ist, ist die Tatsache, dass es die eine Wirtschaftssprache so nicht gibt und dass einige Merkmale für die Wirtschaftssprachen typisch sind, was in den folgenden Abschnitten erläutert wird.

2.1.8.1 Definition der Wirtschaftssprache

Crestani (2010:24) befasst sich in ihrer Publikation mit der Frage nach einer Definition der Wirtschaftssprache und bezieht sich auf Buhlmann und Fearn (2000: 306), die sie als „die Gesamtheit aller Fachsprachen, [...], die in einem fachlich begrenzten Kommunikationsbereich, [...] verwendet werden, um die Verständigung der in diesem Bereich tätigen Menschen zu gewährleisten“. So wie schon bei der Definition der Fachsprache, geht man also von keiner Einzelsprache aus, sondern von einem Gefüge aus mehreren Fachsprachen, d. h. einer Mehrzahl an Wirtschaftssprachen. Crestani (2010) spricht von verschiedenen „Subvarietäten“ (ibid.:26), die abhängig von unterschiedlichen Fachgebieten wie z.B. Finanzen, Recht, Technologie usw., von ihren kommunikativen Funktionen und Kommunikationsteilnehmer*innen, z.B. Lai*innen, Expert*innen der Wirtschaft oder Expert*innen anderes Fachbereiches, sind.

Der Problematik der Definition Fachsprache Wirtschaft bzw. Wirtschaftsdeutsch widmet sich auch Maghețiu (2015) und hebt die Vielschichtigkeit des Begriffes hervor. In ihrer Forschung betont sie insbesondere die Bedeutung der Zielgruppe in all ihren Facetten, wobei sie die Wirtschaftssprache als fachsprachliches Handeln im Bereich der wirtschaftlichen Zusammenarbeit, das in einem komplexen, interkulturell geprägten Umfeld stattfindet, betrachtet. (2015: 91). Im Gegensatz zu der oben erwähnten Definition wird hier der Aspekt der Interkulturalität und des (fach-)sprachlichen Handelns hervorgehoben, d. h. dass die Kommunikationsteilnehmer*innen; unabhängig davon, ob sich um Lai*innen oder Expert*innen handelt; sich im beabsichtigten Kommunikationsbereich, z.B. Finanzen zurechtfinden und dementsprechend auch handeln.

An dieser Stelle ist anzumerken, dass auch wenn es im Bereich Wirtschaft mehrere Fachsprachen gibt, in dieser Arbeit von einer Wirtschaftssprache ohne der Unterteilung in andere spezialisiertere Sprachen, wie z.B. Sprache des Bankwesens oder Sprache der Energiekrise, ausgegangen wird. Das schließt aber solche Unterscheidungen bzw. Unterteilungen in anderen wissenschaftlichen Arbeiten nicht aus.

Da es sich bei der Wirtschaftssprache um keine eindeutig definierbare Einzelsprache handelt, ist die Gliederung der etlichen Wirtschaftssprachen eine Herausforderung. Bei Roelcke (2020) wird die bekannteste horizontale Gliederung aus der Sprachlinguistik, bei der man zwischen drei fachlichen und sprachlichen Bereichen unterscheidet, vorgestellt (Abb. 5). Wie man schon bei der Gliederung der Fachsprachen feststellen konnte, sind solche Gliederungen problematisch, da nicht alle Fächer bzw. Bereiche erfasst und die sprachlichen Verwandtschaften zwischen den Fachsprachen nicht ersichtlich werden. Aus diesem Grund wurde diese klassische Gliederung um weitere Bereiche, Wirtschaftssprache und Konsumtionssprache, durch Kalverkämper erweitert (Roelcke 2020: 68). Allerdings erhebt dieses Gliederungsmodell auch nach dieser Erweiterung keinen Anspruch auf Vollständigkeit.

Fachsprachen							
Theoriesprache				Praxisprache			
Wissenschafts- sprache		Techniksprache		Institutio- nensprache	Wirtschafts- sprache	Konsumtions- sprache	
Spra- che der <u>Natur-</u> <u>wiss.</u>	Sprache der <u>Geistes-</u> <u>wiss.</u>	Sprache der Produk- tion	Sprache der <u>Ferti-</u> <u>gung</u>	Sprache des Dienstleistungssektors		[...]	

Abbildung 5: Horizontale Gliederung von Fachsprachen in Wissenschafts-, Technik- und Institutionensprachen (mittel unterlegt), sowie stark vereinfachte Zuordnung einiger ausgewählten Ergänzungen (hell unterlegt) (Roelcke 2020: 45)

Diese einfache horizontale Gliederung erlaubt allerdings nur eine Zuordnung einer begrenzten Anzahl an Fächern und man kann auf den ersten Blick sehen, dass hier von einer einzelnen Wirtschaftssprache ausgegangen wird.

Einen anderen Zugang hat Crestani (2010) gewählt, die von mehreren Wirtschaftssprachen ausgeht. Anhand der Auffassung der Wirtschaftssprachen als Varietäten der Sprache im Allgemeinen und einer Zusammenfassung ausgewählter Typologien schlägt Crestani ein Gliederungsmodell vor, der sowohl eine vertikale, die die drei kommunikativen Funktionen der Wirtschaftssprache berücksichtigt, wie auch eine horizontale Gliederung in einem Modell vereint (Abb. 6).

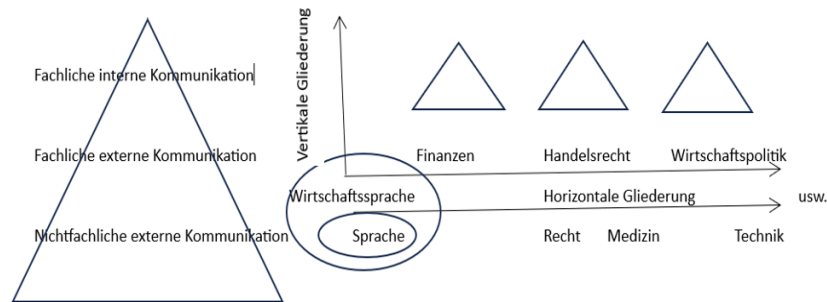


Abbildung 6: Modell von Wirtschaftssprache (Crestani 2010: 33)

Aus dem Modell geht hervor, dass die vertikale Gliederung graphisch als eine Pyramide dargestellt wird, was insofern Sinn macht, da sich in dem engsten Bereich der Pyramide die fachliche interne Kommunikation befindet, d. h. z.B. eine sehr spezifische Terminologie oder typische syntaktische Muster u.v.a. Auf der horizontalen Ebene kann man die Subvarietäten der Wirtschaftssprache beobachten, die man auch als Sammelbereiche auffassen kann und ihnen weitere Unterbereiche bzw. Subkategorien zuordnen kann (Crestani 2010). Ein großer Vorteil dieses Modells ist, dass es beliebig erweiterbar und flexibel ist, was bei der Schnelligkeit, mit der sich die Wirtschaftswelt entwickelt, ausschlaggebend ist.

Auch Maghețiu (2015) plädiert für ein Modell, das aus mehreren Ebenen besteht, da erstens das Definieren des Begriffs Wirtschaft an sich schon problematisch ist und es daher mehrere Auffassungen des Begriffes gibt, zweitens die Gliederung der Wirtschaftswissenschaften je nach Betrachtungsweise voran stattfindet, drittens eine exakte Zahl an Fachsprachen nicht genau feststellbar ist, da kein Konsens darüber herrscht, was eine Fachsprache überhaupt ist und viertens immer wieder neue Fachgebiete entstehen (ibid.: 82f.). Weiters betont Maghețiu (2015), dass zur Beschreibung und Gliederung der Fachsprachen ein Schichtungsmodell, das die horizontale und vertikale Ebene vereint, das geeignetste ist, da man auf der horizontalen Ebene die (Sub-)Fachsprachen innerhalb verschiedener Bereiche ansiedeln kann und auf der vertikalen die verschiedenen Textsorten und die Kommunikationssituationen in Betracht gezogen werden können, d. h. dass der Grad der Spezialisierung der Texte und das (Vor-)Wissen der Kommunikationsteilnehmer*innen entscheidend sind (ibid.: 84f.).

2.1.8.2 Wirtschaftliche Texte im journalistischen Kontext

Was man unter wirtschaftlichen Texten versteht, definiert Crestani (2010:34) als „Resultate und Instrumente einer wirtschaftlichen Tätigkeit [...], die sich in einer sprachlichen Tätigkeit [...]

kommunikativ widerspiegelt“. Diese sprachliche Tätigkeit erfasst bestimmte Ereignisse, die z.B. von der Industrie oder Banken verursacht werden und die z.B. von Fachjournalist*innen oder Akademiker*innen in der Form einer Abhandlung oder eines Fachzeitschriftartikels verarbeitet und weitervermittelt werden (ibid.).

Zeitungsartikeln haben unter den Fachtexten eine besondere Stellung, da es sich hier um eine Textsorte handelt, die sich einer Verflechtung aus zwei Sprachbereichen bedient: der Wirtschaftssprache und der Journalistenfachsprache. Die wirtschaftlichen Texte werden somit durch journalistische Elemente für die breitere Öffentlichkeit lesbar gemacht, wobei man davon ausgehen muss, dass die breite Öffentlichkeit über ein bestimmtes Wissen verfügt. Crestani (2010:33) betont, dass es sich im Falle der journalistischen Artikel bzw. Zeitungsartikel um wirtschaftliche, popularisierende Texte handelt.

2.1.8.3 Terminologische Besonderheiten in der Wirtschaftssprache

Wie schon in Abschnitt 2.1.7 kurz erwähnt, zeichnet sich die Fachsprache durch bestimmte lexikalische, morphologisch-syntaktische und textuelle Merkmale aus, die sie von den nicht-fachlichen Sprachen unterscheiden. Zu ergänzen ist, dass in diesem Abschnitt und generell in dieser Arbeit von der deutschen Sprache ausgegangen wird und daher sind diese Merkmale nicht auf jede beliebige Sprache anwendbar.

Im Abschnitt 2.1.1 wurde bereits näher auf die Eigenschaften der Termini, die einen großen Teil der Lexik der Wirtschaftssprache ausmachen, näher eingegangen. Es bleibt nur zu ergänzen, dass im Idealfall folgende Eigenschaften der Termini beachtet werden sollen: erstens die Exaktheit, d. h. einem Gegenstand wird ein präziser Ausdruck zugewiesen, zweitens die Eindeutigkeit (siehe Abschnitt 2.1.1) und drittens die Kontextabhängigkeit bzw. -unabhängigkeit, d. h. die Bedeutung des Fachwortes bleibt dieselbe unabhängig vom Kontext (Maghețiu 2015: 42ff.; Roelcke 2020: 93). Allerdings werden diese Idealfallvoraussetzungen in der Fachliteratur, z.B. von Maghețiu (2015: 42ff.), hinterfragt, da sie in der Praxis nicht unbedingt anzutreffen sind.

Gyáfrás (1999) und Maghețiu (2015) bieten eine umfassende Übersicht der morphologisch-syntaktischen Besonderheiten der Wirtschaftssprache, wobei sie betonen, dass die Wirtschaftssprache im Vergleich zur Allgemeinsprache keine spezifische Grammatik aufweist, jedoch bestimmte sprachliche Mittel bevorzugt verwendet. Im Bereich der Morphologie fällt auf, dass die Substantive und ihre verschiedenen Kombinationen mit anderen Wortarten eine besondere Rolle spielen. Insbesondere fällt der häufige Gebrauch von Funktionsverbgefügen auf, wie

z.B. *den Vertrag auflösen*, wo durch laut Maghețiu (2015: 47) die Kriterien der Exaktheit, Eindeutigkeit und Kontextunabhängigkeit erfüllt werden. Überhaupt werden die Wirtschaftstexte durch den Nominalstil, insbesondere durch die Nominalisierungen von Verben und die Genitivkonstruktionen, geprägt (Gyáfrás 1999: 81). Weitere morphologische Besonderheiten sind erstens die Verwendung der 3. Person statt dem Singular, um die Objektivität der Texte zu fördern, zweitens die Nutzung von Modalverben und Infinitivkonstruktion, drittens die Verwendung der Gegenwartsform, viertens die häufige Verwendung des Passivs. Gyáfrás (1999: 82) bezeichnet die Verwendung des Passivs als das wichtigste Merkmal der Wirtschaftssprache.

Im Bereich der Syntax zeichnet sich die Wirtschaftssprache vor allem durch hohe Satzkomplexität aus. Wirtschaftstexte enthalten viele Nebensätze vor allem Konjunktionalsätze, Relativsätze, Infinitivkonstruktionen und Temporalsätze, die allerdings zu Verständigungsschwierigkeiten auch unter Fachleuten führen können (Gyáfrás 1999: 80f.; Maghețiu 2015: 49ff.).

Abschließend ist anzumerken, dass der Fachwortschatz der Wirtschaftssprache dynamisch und daher keine geschlossene Einheit bildet. Er wird von aktuellen gesellschaftlichen Ereignissen geprägt, wobei neue Termini fortlaufend durch verschiedene Verfahren, wie durch die substantivische Wortbildung, der sich sehr detailreich Crestani (2010) widmet, entstehen. Die Autorin betont, dass einige dieser entstandenen Termini als Neologismen ins Lexikon aufgenommen werden, während andere nur für einen bestimmten Zeitraum existieren, die sogenannten Okkasionalismen oder Ad-hoc-Bildungen (ibid.: 64f.). Maghețiu (2015) identifiziert in Bezug auf die Wortbildung vier Verfahren: Komposition, Derivation, Konversion und Wortkürzungen. Vor allem die Substantivkomposita sind in der Wirtschaftssprache häufig vertreten und sind laut Crestani (2010: 67) die Prototypen der Komposita in der deutschen Sprache, z.B. *Volkswirtschaft*. Auch Komposita mit Abkürzungen und Zahlen, wie *EG-Richtlinie*, sind häufig in der Wirtschaftssprache vertreten (Maghețiu 2015: 51f.). Bei den Derivaten handelt es sich um Wörter, die durch Präfigierung und Suffigierung einer Basis entstehen, wobei Suffixe, wie z.B. -ung, -er, -keit u.v.a., Suffixiende wie -los, -frei u.v.a. und Präfixe wie in-, un- u.v.a. verwendet werden (Maghețiu 2015: 52f, Crestani 2010: 106). Bei den Konvertaten handelt es sich um Wörter, die von einer in die andere Wortart wechseln, wobei laut Maghețiu (2015) die Eponyme, d. h. von einem Namen abgeleitete Bezeichnung, für den fachsprachlichen Bereich eine wichtige Rolle spielen, wie z.B. *Parkinson* (ibid.: 53). Zuletzt müssen auch die Wortkürzungen oder Kurzwortbildungen erwähnt werden, da diese in ihren beiden Formen im Fachwortschatz der Wirtschaftssprache vorkommen, z.B. das Akronym *StGB*, in Langform *Strafgesetzbuch* (Crestani 2010: 134).

Alle diese Besonderheiten zeigen auf, was man bei der Termextraktion aus dem Fachbereich Wirtschaft erwarten kann.

2.2 Grundlagen der Terminologielehre

Die Terminologielehre ist eine eigene wissenschaftliche Disziplin und ist von der Terminologearbeit (siehe Abschnitt 2.3) abzugrenzen, da es sich um zwei unterschiedliche terminologische Disziplinen handelt (Roelcke 2020: 157). Die Norm DIN 2342 (2022) definiert die Terminologielehre als „die Wissenschaft von den Begriffen und ihren Bezeichnungen in den Fachsprachen“, wobei man zwischen der allgemeinen Terminologielehre, die sich fachübergreifend den theoretischen Grundlagen von Begriffen und Benennungen und ihrer Beziehungen widmet, und der speziellen Terminologielehre, die sich mit den Regeln und Richtlinien und daher den Gesetzmäßigkeiten konkreter Terminologien beschäftigt, d. h. die einzelnen Fächer und Sprachen im Fokus hat (Roelcke 2020: 158), unterscheidet. Die Terminologielehre und die Terminologiewissenschaft sind laut der DIN-Norm 2342 als synonym zueinander zu betrachten, wobei an dieser Stelle der Norm drauf hingewiesen wird, dass die Bezeichnung Terminologie im Sinne von Fachwortschatz als Synonym für Terminologielehre abzulehnen ist (2022), d. h. synonyme Verwendung dieser zwei Begriffe als Fehler zu betrachten ist. Das Ziel der Terminologielehre ist, eigene theoretische und methodische Grundlagen unter der Berücksichtigung anderer Bereiche, wie z.B. der Fachsprachenforschung zu generieren und zu etablieren (ibid.: 157). Der zu betrachtende Gegenstand der Terminologielehre ist in erster Reihe die Terminologie selbst, die im Abschnitt 2.1.1. ausführlich erläutert wurde. Wie man weiters der oben genannten DIN-Definition entnehmen kann, gehören weitere Grundelemente wie Begriff und die Bezeichnung zur Terminologielehre, was das genau bedeutet und in welcher Beziehung diese Elemente zueinanderstehen, wird unter anderem in diesem Unterkapitel erläutert.

2.2.1 Gegenstand

Die Norm DIN 2342 (2022) versteht unter Gegenstand ein Objekt, das einem „beliebige(n) Abschnitt aus der wahrnehmbaren oder vorstellbaren Welt“ (ibid.) entspricht, wobei dieses Objekt konkret oder abstrakt sein kann und weder Geschehnisse noch Sachverhalte als Objekte ausgeschlossen sind. Bei Arntz et al. (2014) wird der Gegenstand als die Grundlage der Begriffsbildung hervorgehoben. Nach einer genauen Betrachtung der früheren Überlegungen zu

diesem Thema, wie z.B. der Auffassung des Gegenstandes nach einem der Pioniere der Terminologiewissenschaft Eugene Wüster, der den Gegenstand mit allem von Menschen Denkbaren assoziiert (Wüster 1959), unterscheiden Arntz et al. (2014: 46f.) anhand des Kriteriums „materiell-immateriell“ unter folgenden acht Gegenstandsarten: 1. den materiellen Gegenständen, deren Existenz nicht von Menschen bedingt ist, 2. den „Objecta“, d. h. den materiellen Gegenständen, die zwar durch die Sinne zugänglich sind, aber nicht physisch anwesend sein müssen, 3. den „Subjecta“, d. h. physisch anwesenden Gegenständen, 4. den immateriellen Gegenständen, die zwar keine physische Form besitzen, aber vorstellbar, räumlich und zeitlich verortbar sind, 5. den materialisierbaren Gegenständen, die zwar in einer Vorstellung existieren, allerdings in materielle Gegenstände umwandelbar sind, 6. den gedachten Gegenständen, die rein in der Vorstellung existieren, 7. den real existierenden, gedachten Gegenständen, die in einer Vorstellung existieren, aber auch real sind und 8. den imaginären Gegenständen, die rein in der Welt der Vorstellung und des Glaubens existieren. Für konkrete Beispiele und eine weitere Unterteilung nach einem anderen Kriterium siehe Arntz et.al (ibid.: 46f.).

2.2.2 Begriff

Unter Begriff versteht Roelcke (2020:74) eine „Einheit aus Ausdruck und Bedeutung unter Betonung der Bedeutungsseite“. Diese knappe Definition wirft zwei Fragen auf, und zwar 1. was unter der genannten „Einheit“ genau verstanden wird und 2. wer und wie Entscheidungen über diese Einheit getroffen werden. Die Norm DIN 2342 (2022) definiert den Begriff präziser als Roelcke und versteht den Begriff als eine „Denkeinheit, die aus einer Menge von Gegenständen unter Ermittlung der diesen Gegenständen gemeinsamen Eigenschaften mittels Abstraktion gebildet wird“. Als Denkeinheit ist der Begriff also von den „Denker*innen“ abhängig, d. h. Menschen, die in einer Gesellschaft mit jeweiligem kulturellem Hintergrund einer Sprachgemeinschaft leben, was impliziert, dass nicht die einzelne Sprache an sich den Begriff bestimmt, sondern eine Gemeinschaft (ibid.). Diese Aspekte des Begriffes sind keinesfalls neu. Sie führen auf Wüster (1959) zurück, der sich seit den späten 50er Jahren mit terminologischen Konzepten eingehend befasst. Seine Auffassungen werden seitdem in der Forschung angewandt, wie bei Hohnhold (1990), der den Begriff als eine „sprachunabhängige Vorstellungs-, Denk- oder Wissensseinheit“ (ibid.: 44) definiert. Diese Sichtweise greifen auch Arntz et al. (2014) auf, die den Begriff als eine Wissensseinheit, in der das Wissen einer bestimmten Zeit gesammelt ist, ansehen und fügen noch den Aspekt der Erkenntnis hinzu, wodurch Sie den Begriff als eine Erkenntnis- und Wissensseinheit definieren (ibid.: 50).

Die Norm DIN 2342 (2022) unterscheidet je nach Kriterium folgende Arten von Begriffen: nach dem Kriterium Anzahl der vertretenen Gegenstände gibt es Individual- und Allgemeinbegriffe und nach dem Kriterium Begriffssystemhierarchie gibt es übergeordnete, untergeordnete und nebengeordnete Begriffe (ibid., siehe Abschnitt 2.2.4.2). Um Unterschiede zwischen Allgemein- und Individualbegriff zu veranschaulichen, führen Drewer und Schmitz (2017: 8) unter anderen das Beispiel *Planet* vs. *Saturn* an, wobei der Allgemeinbegriff *Planet* eine ganze Gruppe von verschiedenen Planeten bzw. Gegenständen, wie *Saturn* u.v.a., zusammenfasst. Im Gegenteil zu Allgemeinbegriff repräsentiert der Individualbegriff nur einen konkreten Gegenstand (Arntz et al. 2014: 51; DIN 2342 2022). Bei Drewer und Schmitz (2017) wird der Unterschied zwischen Allgemein- und Individualbegriff vereinfacht veranschaulicht, indem hervorgehoben wird, dass die Benennungen immer Bezeichnungen von Allgemeinbegriffen und die Eigennamen bzw. Namen die Bezeichnungen von Individualbegriffen sind (ibid.: 15). Aus der terminologischen Sicht spricht man daher im Falle von Allgemeinbegriffen von Benennungen und im Falle von Individualbegriffen von (Eigen-)Namen (siehe Abschnitt 2.2.3). Obwohl die Anzahl der vertretenen Gegenstände als Unterscheidungsmerkmal auf den ersten Blick schlüssig erscheint, ist fraglich, ob dies der Realität entspricht. Arntz et al. (2014: 51) betonen, dass der repräsentierte Einzelgegenstand, der in einem bestimmten Zeit-Raum-Verhältnis steht, durch die Abstraktion von Eigenschaften mehrerer Gegenstände zu einem Individualbegriff wird.

2.2.3 Benennung

Bevor erläutert wird, was man unter Benennung in der Terminologielehre versteht, muss noch ein terminologisches Gewirr angesprochen werden, das die vermeintliche Synonymie der Begriffe Fachwort, Terminus und Benennung betrifft. Die gängige Unterscheidung aus terminologischer Sicht ist in der Norm DIN 2342 (2022) verankert, in der man unter dem Stichwort Benennung im Abschnitt 3 nachlesen kann, dass sowohl Terminus als auch Fachausdruck synonym zu Benennung verwendet werden können. Im Falle von Fachwort und Terminus handelt es sich jedoch keinesfalls um 100% Synonyme. Es gibt in der Wissenschaft unterschiedliche Auffassungen von Terminus. Wendt (1997: 63) führt aus, dass mit dem Begriff Terminus unter anderem eine Funktion beschrieben wird, die unter bestimmten Bedingungen, wie fachlicher Kommunikation und der Repräsentation eines Fachbegriffes, von jeder Einheit der Sprache übernommen werden kann. Hohnhold (1990: 30f.) betont, dass der Terminus das Kriterium der Eindeutigkeit erfüllen muss, d. h. dass eine Benennung nur einen Begriff zu bezeichnen hat.

Wogegen beim Fachwort erst die funktionale Bedeutung, Kontext, und die Art der Kommunikation, fachliche resp. nicht-fachliche, darüber entscheiden, wann es sich um ein Fachwort und wann nicht handelt (ibid.: 31). Allerdings wird dieses Kriterium von einigen Wissenschaftler*innen in Frage gestellt, beispielsweise von Roelcke (2020), da das Kriterium der Eindeutigkeit in der Wirklichkeit nicht zutrifft. Es gibt Termini, die vage und mehrdeutig sind. Roelcke betont, dass es in der Fachsprachenforschung und der Terminologiewissenschaft notwendig ist, zwischen Wort und Begriff wie auch zwischen Bedeutung und Ausdruck zu unterscheiden (ibid.: 109). Göpferich (1995: 179) schlägt vor, dass der Fachausdruck, also ein sprachlicher Ausdruck, der in einer Fachsprache verwendet wird und dort eine spezielle Bedeutung besitzt, gegenüber dem Fachwort als geeigneteres Ersatzwort für den Terminus genutzt werden sollte. Diese Erkenntnis ist auch in der Norm DIN 2342 (2022) hervorgehoben, indem die Definition von Benennung um eine Anmerkung ergänzt wurde. Diese besagt, dass in der Praxis Fachwort und Benennung zwar synonym zueinander verwendet werden können, allerdings nur dann, wenn es sich um eine Einwortbenennung handelt (ibid.). Da aber Terminologie als Gesamtbestand der Begriffe und ihrer Bezeichnungen eines Fachgebietes definiert ist (siehe Abschnitt 2.1.1), sind die Einwortbenennungen bzw. Fachwörter allein nicht ausreichend, um alle Gegenstände und Sachverhalte des Faches einzufangen. Ziel dieser Arbeit soll keine Definitionsbegradigung sein, daher wird an dieser Stelle explizit erwähnt, dass wenn in dieser Arbeit Fachwort erwähnt wurde, dann im Sinne eines Fachausdruckes, so wie es in der Norm DIN 2342 verankert ist, d. h. ohne die Begrenzung auf Einwortbenennung. Allerdings wird in der weiteren Folge aus den in diesem Abschnitt erwähnten Gründen der Begriff Terminus bevorzugt.

Unter Benennung wird laut der Norm DIN 2342 eine „sprachliche Bezeichnung eines Allgemeinbegriffes aus einem Fachgebiet“ (DIN 2342 2022) verstanden, wobei auch Symbole und Formeln als Benennungen gelten (ibid.). Wenn die Rede von einer Bezeichnung ist, ist damit die Repräsentationsform eines Begriffes gemeint, die nicht nur auf rein sprachliche Mittel beschränkt ist, sondern durch verschiedene Mittel, wie z.B. Nummern und Gebärden, ausgedrückt werden kann (ibid.). In der Norm DIN 2342 werden ausgewählte Bezeichnungsarten wie Benennung, Symbol, Formel und Eigenname angeführt, allerdings erhebt diese Aufzählung keinen Anspruch auf Vollständigkeit (ibid.). Von Interesse für diese Arbeit ist vor allem die Benennung, da hier die synonyme Verknüpfung zum Begriff Terminus hergestellt werden kann. Die Benennung ist keineswegs, sowie es der Fall von Fachwort ist, auf ein einziges Wort, die sogenannten Einwortbenennungen, beschränkt, sondern bezieht auch Mehrwortbenennungen, Langform- und Kurzformbenennungen und Produktbezeichnungen ein (ibid.). Diese Breite an Möglichkeiten eignet sich auch für diese Terminologiearbeit, daher wenn im praktischen Teil

die Rede von Termini bzw. Fachtermini ist, damit die Benennung gemeint ist, wobei die Eigennamen aus dem Experiment auch nicht ausgeschlossen werden.

2.2.4 Beziehungen zwischen den Grundelementen der Terminologielehre

Die Grundelemente der Terminologielehre stehen in Beziehungen zueinander. Um diese Beziehungen zwischen den oben definierten Grundelementen zu veranschaulichen, bedienen sich Wissenschaftler*innen, nicht nur in der Semiotik selbst, sondern auch fachübergreifend z.B. in der Terminologiewissenschaft (Arntz et al. 2014: 41f.; Drewer & Schmitz 2017: 7), des semiotischen Dreiecks (siehe Abschnitt 2.2.4.1). Die wichtigsten Arten von Beziehungen kann man außerdem in zwei größeren Gruppen erfassen, erstens Beziehungen zwischen einzelnen Begriffen (siehe Abschnitt 2.2.4.2) und zweitens Beziehungen zwischen Begriff und Benennung (siehe Abschnitt 2.2.4.3). Für die Terminologearbeit und -lehre ist vor allem die direkte Beziehung zwischen Begriff und Benennung von Bedeutung, da sie in verschiedenen Ausprägungen (Eindeutigkeit, Synonymie, Polysemie, Homonymie) vorkommt und so einige Probleme bei der eindeutigen Zuordnung entstehen können.

2.2.4.1 Semiotisches Dreieck

Drewer und Schmitz (2017) zählen das semiotische Dreieck (Abb. 7), das von den Linguisten Ogden und Richards (1956) eingeführt wurde, sogar zu „einem idealen Werkzeug“ (ibid.: 7), um die Beziehungen zwischen Begriff, Benennung und Gegenstand, zu veranschaulichen. Sie bedienen sich einer Variation (Abb. 8) dieser ursprünglichen Darstellung, wobei sie gleichzeitig einräumen, dass diese zu sehr vereinfacht ist und die Komplexität der Beziehungen nicht im vollen Umfang einfängt. Allerdings bietet sie einen soliden Ausgangspunkt, um sich die genannten Grundelemente näher anzusehen und sie voneinander zu unterscheiden.

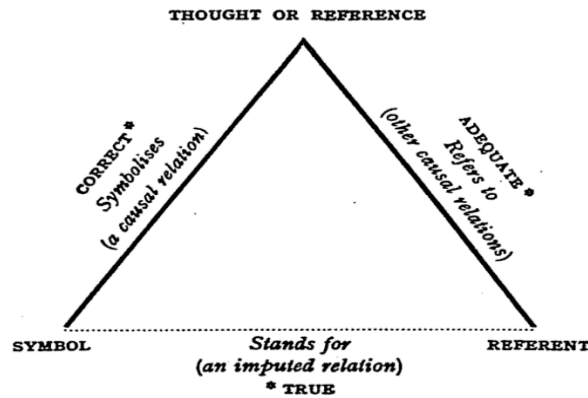


Abbildung 7: Semiotisches Dreieck nach Ogden und Richards (1956: 11)

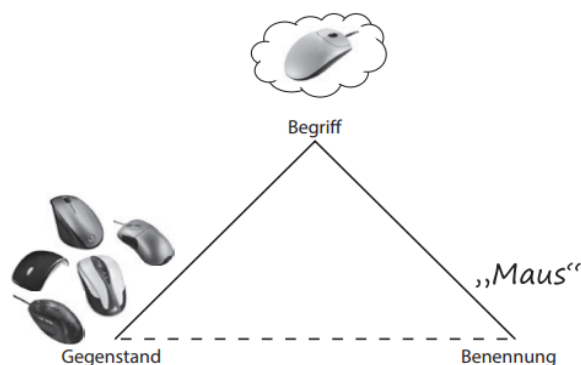


Abbildung 8: Anpassung des semiotischen Dreiecks (Drewer & Schmitz 2017: 7)

Auch wenn die angewendete Variation (Abb. 8) spiegelverkehrt ist, d. h. Gegenstand ist mit „referent“ und Benennung mit „symbol“ gleichzusetzen, ändert es am Verständnis der gegenseitigen Beziehungen zwischen den Elementen nichts Grundlegendes, d. h. der Gegenstand und die Benennung stehen beide in einer direkten Beziehung zum Begriff, wogegen zwischen dem Gegenstand und der Benennung eine indirekte Beziehung besteht, die erst durch den Begriff bzw. die Bedeutung konkret wird. Das bedeutet, dass die Benennung *Maus* mehrere Gegenstände bezeichnen kann und erst durch die Bedeutung bzw. Begrifflichkeit man sie zu einem konkreten Gegenstand zuordnen kann.

2.2.4.2 Beziehungen zwischen Begriffen

Zwischen Begriffen bestehen oder können Beziehungen hergestellt werden, die man entweder als hierarchische oder nichthierarchische Beziehungen klassifizieren kann, wobei Drewer und

Schmitz (2017: 9) hervorheben, dass die hierarchischen Begriffsbeziehungen für die Terminologiearbeit bedeutsamer sind.

Die Norm DIN 2342 (2022) beschreibt die nichthierarchische Begriffsbeziehungen als Beziehungen, die auf keiner Hierarchie, sondern auf verschiedenen thematischen Zusammenhängen zwischen Begriffen beruhen, z.B. auf einer zeitlichen Abfolge, wie im Beispiel *Frühling – Sommer – Herbst – Winter*. In diesem Beispiel handelt es sich um eine sogenannte temporale Begriffsbeziehung. Unter den nichthierarchischen Beziehungen unterscheidet man weiters zwischen sequenziellen Beziehungen, die sich in oben erwähnte temporale und kausale Beziehungen unterteilen, oppositionellen Beziehungen, pragmatischen Beziehungen u.v.m. (Näheres siehe Arntz et al. 2014: 79). Eine vollständige Klassifizierung aller Arten von nicht-hierarchischen Begriffsbeziehungen gestaltet sich schwierig, da die Art der Beziehungen je nach Fachgebiet variieren kann (Arntz et al. 2014; DIN 2342 2022). Ein Beispiel dafür findet man bei Arntz et al. (2014: 79), die als eine spezifische Art der Beziehung in den Sozialwissenschaften und im Recht die schuldrechtliche Beziehung am Beispiel des Begriffspaares Kreditgeber*in-Kreditnehmer*in anführen. Diese Art der Beziehung kann nicht in allen Fachgebieten angewendet werden.

Die hierarchischen Begriffsbeziehungen werden in Abstraktions- und Bestandsbeziehungen unterteilt. Die Abstraktionsbeziehungen, auch generische, logische oder taxonomische Beziehungen genannt, beschreiben, wie ein Begriff auf einer höheren oder abstrakteren Ebene mit anderen Begriffen in Beziehung steht. Durch so entstandene Beziehungen entsteht ein Abstraktionssystem, innerhalb dessen man zwischen Oberbegriff und Unterbegriff unterscheidet, wobei der Unterbegriff alle Merkmale des Oberbegriffs und zusätzlich mindestens ein weiteres Merkmal beinhalten muss (Arntz et al. 2014; DIN 2342 2022; Drewer & Schmitz 2017). Für diese Art der Beziehungen führen Drewer und Schmitz (2017: 9) folgendes Beispiel an: „Eine Rollkugelmaus besteht aus Tasten, Scroll-Rad, Gehäuse, Rollkugel etc.“ Eine weitere Art der hierarchischen Beziehungen sind die Bestandsbeziehungen, auch Teil-Ganzes-Beziehungen oder partitive Beziehungen genannt. Sie zeichnen sich dadurch aus, dass der übergeordnete Begriff, der sogenannte Verbandsbegriff, in seine Einzelteile zerlegt werden kann, wobei sich der untergeordnete Begriff, der sogenannte Teilbegriff, auf einen Teil des übergeordneten Begriffes bezieht (DIN 2342 2022; Drewer & Schmitz 2017: 9). Als konkretes Beispiel führen Arntz et al. (2014: 92) den Verbandsbegriff *Fahrrad* an, der sich in Teilbegriffe wie *Rad*, *Rahmen*, *Sattel*, *Lenker* usw. aufteilen lässt.

In der wissenschaftlichen Literatur, z.B. bei Arntz et al. (2014) oder Drewer und Schmitz (2017) ist oft die Rede von untergeordneten bzw. übergeordneten Begriffen, die nicht

mit Unterbegriff bzw. Oberbegriff zu verwechseln sind. In der Norm DIN 2342 (2022) wird auf diesen Umstand hingewiesen, wobei hervorgehoben wird, dass bei den übergeordneten Begriffen je nach Art der Begriffsbeziehung entweder zwischen Oberbegriffen, d. h. übergeordneten Begriffen in einer Abstraktionsbeziehung, oder Verbandsbegriffen, d. h. übergeordneten Begriffen in einer Bestandsbeziehung, unterschieden wird. Und bei den untergeordneten Begriffen wiederum zwischen Unterbegriff, d. h. untergeordneten Begriffen in einer Abstraktionsbeziehung, und Teilbegriffen, d. h. untergeordneten Begriffen in einer Bestandsbeziehung (ibid.).

Die oben genannten Beziehungen lassen sich je nach Art und Weise in verschiedenen Begriffssystemen graphisch abbilden. Da der Fokus dieser Arbeit rein auf der Extraktion von Termini liegt, wird auf diese Darstellungen nicht näher eingegangen (Näheres siehe bei Drewer und Schmitz 2017: 12ff.).

2.2.4.3 Beziehungen zwischen Begriffen und Benennungen

Die Beziehungen zwischen Begriffen und Benennungen nehmen in der Terminologielehre eine besondere Rolle ein, da vor allem hier Herausforderungen für die praktische Terminologearbeit und daher auch für die automatische Termextraktion entstehen. Wissenschaftler*innen wie Drewer und Schmitz (2017: 15) betonen, dass das geforderte Prinzip der Eineindeutigkeit (siehe Abschnitt 2.2.4.3.1) nur schwierig zu erreichen ist und deshalb einige Problembereiche zu beachten sind. Wissenschaftler*innen (Arntz et al. 2014; Drewer & Schmitz 2017) zählen zu den Problemgruppen Synonymie, Ambiguität bzw. vor allem Polysemie und Homonymie, die im Falle einer mehrsprachigen Terminologearbeit um die Äquivalenzproblematik ergänzt werden müssen. Um die Beziehungsarten zwischen Begriffen und Benennungen zu vervollständigen, führt die Norm DIN 2342 (2022) auch Antonymie auf, d. h. Begriffe, die entgegengesetzt zueinanderstehen. Auf diese Beziehung wird in weiterer Folge nicht näher eingegangen, da es in dieser Arbeit um kein Projekt geht, in dem man von einem Begriff auf andere Begriffe verweist. Im umgekehrten Fall ist die Bestimmung der Antonyme allerdings sinnvoll (Näheres siehe Drewer und Schmitz 2017: 117f.).

2.2.4.3.1 Idealfall Eineindeutigkeit

Wie schon in dieser Arbeit an einigen Stellen erwähnt (siehe Abschnitt 2.1.1 und Abschnitt 2.2.3), ist eine wesentliche Eigenschaft der Terminologie und ihrer Termini die Eindeutigkeit,

d. h. es sollte keine Missverständnisse geben, die entsprechende Bedeutung bzw. Begriff einem entsprechenden Ausdruck bzw. Benennung zuzuordnen und daher eine unmissverständliche Fachkommunikation zu bezwecken. Roelcke (2020) spricht im Zusammenhang mit der Eindeutigkeit von einem der sogenannten „grundsätzlichen Güteeigenschaften fachsprachlicher Lexik“ (ibid.: 92f.), zu denen er auch die Exaktheit und die Eigentlichkeit zählt, und betont den bidirektionalen Charakter der Eindeutigkeit bzw. Eineindeutigkeit, d. h. nur eine Benennung pro Begriff und nur ein Begriff pro Benennung. Von dem Prinzip der Eineindeutigkeit als einem Idealfall sprechen auch Drewer und Schmitz (2017:15) und Hohnhold (1990:51), wobei sie betonen, dass dieser Idealfall der exakten Zuordnung und scharfen Abgrenzung kaum die Wirklichkeit repräsentiert und man die Problemvorkommnisse in der Terminologielehre wie Synonymie und Polysemie, die durch eine uneindeutige Zuordnung zustande kommen, nicht außer Acht lassen darf.

2.2.4.3.2 *Synonymie und Teilsynonymie*

Der Synonymie begegnet man im Falle, wenn ein Begriff zwei oder mehrere Benennungen aufweist bzw. ein Begriff auf verschiedene Art und Weisen bezeichnet werden kann, was die Fachkommunikation im negativen Sinne beeinflussen kann, weil diese Benennungen beliebig austauschbar sind (Arntz et al. 2014: 135; Drewer & Schmitz 2017: 16) und dadurch die Gefahr eines inkonsistenten Gebrauchs erheblich steigt. Außer Synonymie begegnet man in der Terminologielehre und -arbeit auch der Teilsynonymie, auch partielle Synonymie oder Quasisynonymie genannt, die nur in bestimmten Kontexten gegeneinander austauschbare Bezeichnungen für nicht genau denselben Begriff bedeutet, wie z.B. je nach Fahrzeugtyp kann man statt *Kofferraumdeckel* auch *Hecktür* verwenden (DIN 2342 2022). Bei Arntz et al. (2014) wird die Synonymie und die Quasisynonymie, d. h. nicht vollständig synonyme Begriffe, sogar als ein „erhebliches Hindernis für die fachliche Verständigung“ (ibid.: 136) bezeichnet. Laut Drewer und Schmitz (2017) tritt die Synonymie in der Praxis sehr häufig auf und das aus folgenden Gründen: 1. in einem Fachgebiet wurde eine eindeutige Benennung noch nicht etabliert und normiert, 2. durch eine gezielte Firmenpolitik gibt es für ein bestimmtes Produkt mehrere Benennungen, 3. Abkürzungen und Kurzformen sind als Synonyme der Langformen in Verwendung und 4. in der Praxis findet die sogenannte Gebrauchssynonymie statt, d. h. Oberbegriffe werden anstatt komplizierten Unterbegriffen verwendet (ibid.: 16).

2.2.4.3.3 Mehrdeutigkeit

Die Mehrdeutigkeit ist in der Norm DIN 2342 (2022) als eine „Eigenschaft einer Bezeichnung, mehrere Begriffe zu repräsentieren“ (ibid.) definiert und wird in Polysemie, Homonymie, Homographie und Homophonie unterteilt. Da sich die letzteren zwei Kategorien nur auf Schriftbilder und Lautbilder beziehen und die Terminologiearbeit sich vor allem mit geschriebenen Benennungen auseinandersetzt (Drewer & Schmitz 2017: 19), sind für diese Arbeit die ersten zwei Kategorien von Bedeutung.

Mit einer polysemen Benennung werden verschiedene Begriffe bezeichnet, d. h. dass mit nur einer einzigen Benennung mehrere Begriffe mit unterschiedlichen Bedeutungen bezeichnet werden. Die Polysemie entsteht durch eine Übertragung einer Benennung auf einen anderen Begriff und hat verschiedene Auslöser (DIN 2342 2022), wie z.B. einen gemeinsamen etymologische Ursprung, z.B. *Bank* als Sitzgelegenheit und *Bank* als Finanzinstitut, oder eine Ähnlichkeit in der Form, z.B. *Maus* als Nagetier und *Maus* als Computergerät (Abb. 9). Laut Arntz et al. (2014) ist ein sehr häufiger Gebrauch der Polysemie nicht nur in der Gemeinsprache sondern auch in der Fachsprache keineswegs unüblich, weil die Sprache begrenzte Mittel besitzt, um den Bedarf an benötigten Benennungen zu decken (ibid.: 139).

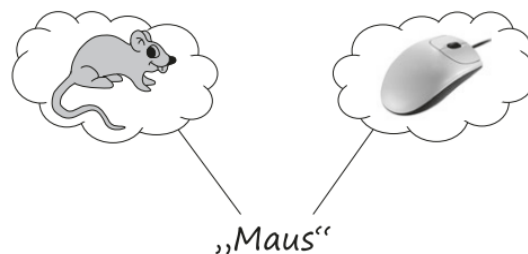


Abbildung 9: Beispiel für Polysemie (Drewer & Schmitz 2017: 18)

Unter Homonymen versteht die Norm DIN 2342 (2022) mehrdeutige Benennungen, die keinen gemeinsamen etymologischen Ursprung als Basis haben, d. h. nicht durch eine Bedeutungsübertragung entstehen, und sich im Gegenteil zu Polysemen oft auch in der Grammatik unterscheiden, z.B. *der/das Schloss*. Reine homonyme Benennungen entstehen durch einen Zufall und sind meistens auf die Übernahme aus anderen Sprachen limitiert (Drewer & Schmitz 2017: 18).

Wie bereits Drewer und Schmitz (ibid.) erwähnen, ist die Unterscheidung zwischen Polysemie und Homonymie für die Terminologiearbeit kaum bedeutend, da diese Kategorien schwer voneinander abzugrenzen sind und es auch vorkommen kann, dass sich aus

Homonymen Polyseme entwickeln. Was aber für die Terminologearbeit wichtig ist, sind die Auswirkungen der Mehrdeutigkeit auf die fachliche Kommunikation, die durch nicht zu vermeidende Mehrdeutigkeit erheblich erschwert wird (vgl. Hohnhold 1990: 52; Arntz et al. 2014: 139f.; Drewer & Schmitz 2017: 18f.). Bei der Erstellung einer Terminologie ist Mehrdeutigkeit nicht zu vermeiden, da bei der Entstehung neuer Begriffe zur Übernahme von Benennungen unter anderem aus Allgemeinsprache oder einer anderen Fachsprache kommt und daher theoretisch zur Verwechslungsgefahr der Begriffe kommen könnte. Diese Verwechslungsgefahr besteht laut Drewer und Schmitz (2017: 20) allerdings nur dann, wenn die Mehrdeutigkeit innerhalb eines Fachgebietes entsteht. Das bedeutet, wenn die Begriffe aus verschiedenen Fachgebieten stammen, kommt es kaum zu solchen Verwechslungen, wie am Beispiel *Maus* als Tier und *Maus* als Computerzubehör ersichtlich ist. Diese zwei Begriffe würden nie in derselben Terminologie erfasst werden.

2.2.4.3.4 Äquivalenz

Die Äquivalenz ist laut Norm DIN 2342 (2022) eine besondere Art der Beziehung „zwischen Bezeichnungen in verschiedenen Sprachräumen für denselben Begriff“ und kommt dann zum Einsatz, wenn man sich von einer einsprachigen zu einer mehrsprachigen Terminologearbeit bewegt. Wie schon Hohnhold (1990) erwähnt, hat die Äquivalenz kaum etwas mit Gleichheit und Identität zu tun, vielmehr mit einer Gleichwertigkeit (ibid.: 56), auch wenn in der Fachliteratur oft die Identität der Begriffe hervorgehoben wird wie z.B. bei Drewer und Schmitz (2017: 20). Fest steht, dass Begriffe je nach Sprachraum unterschiedliche Wirklichkeit abbilden, da sie durch gesellschaftliche und kulturelle Aspekte beeinflusst werden (Arntz et al. 2014: 141; Drewer & Schmitz 2017: 20). Da sich diese Arbeit mit einer einsprachigen Terminologie befasst, wird an die Äquivalenz nicht näher eingegangen. Für nähere Informationen siehe Arntz et. al. (2014) oder Drewer und Schmitz (2017).

2.3 Grundlagen der Terminologiearbeit

Die Terminologielehre und die Terminologiearbeit sind zwei unterschiedliche terminologische Disziplinen. Diese zwei Begriffe sind nicht miteinander zu verwechseln und schon gar nicht synonym zueinander zu verwenden, da wie aus Abb. 10 hervorgeht, die Terminologiearbeit auf der Terminologielehre bzw- wissenschaft basiert. In der Praxis wird der Begriff Terminologiemanagement synonym zu Terminologiearbeit verwendet, obwohl diese Begriffe lediglich Teilsynonyme sind (DIN 2342 2022). Als Beispiel kann hier die Praxis der Unternehmen angeführt werden, die sich häufig des Begriffes Terminologiemanagement im Sinne der Terminologiearbeit bedienen, obwohl sich der Begriff eigentlich nur auf die Verwaltung von Terminologie bezieht. Einfach ausgedrückt, Terminologiemanagement ist ein Teil der Terminologiearbeit, konkret handelt es sich hier um die Terminologieverwaltung, die sich speziell mit Erfassung, Verarbeitung, Pflege und Bereitstellung der Fachtermini befasst (DIN 2342 2022; Drewer & Schmitz 2017: 23f.).

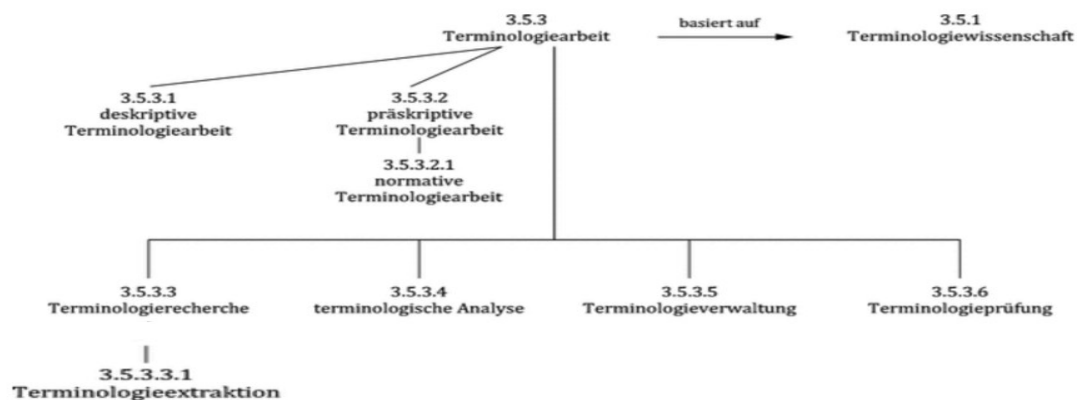


Abbildung 10: Begriffssystem „Terminologiearbeit“ (DIN 2342 2022)

Bei Roelcke (2019) findet man den Hinweis, dass die Terminologiearbeit entsprechend der Terminologielehre erfolgen muss und ihr Ergebnis die Ausarbeitung und der Vergleich von Begriffssystemen ist, wobei ihre Aufgabe „in der beschreibenden und vorschreibenden Aufarbeitung einzelner Fachwortschätze selbst“ (ibid.: 157) besteht. Was das genauer bedeutet kann man in der Norm DIN 2342 (2022) nachlesen, in der die Terminologiearbeit als „auf der Terminologiewissenschaft aufbauende Terminologierecherche, terminologische Analyse, Terminologieverwaltung und Terminologieprüfung“ (ibid.) definiert wird. Noch einfacher

ausgedrückt geht es hier um das Sammeln, Prüfen und Bereitstellen von Fachwortbeständen bzw. Terminologien (Drewer & Schmitz 2017: 24).

2.3.1 Arbeitsformen der Terminologearbeit

Es gibt mehrere Möglichkeiten, wie man Terminologearbeit durchführen kann. Um sich für die richtige Form entscheiden zu können, müssen vor der eigentlichen Terminologearbeit ein paar wesentliche Fragen beantwortet werden. Arntz et al. (2014) fassen diese in 6 Punkten zusammen: 1. Ziel und Umfang, 2. Zielgruppe, 3. verfügbares Personal, 4. Zeitressourcen, 5. Quellenressourcen und 6. Technikressourcen, wobei sie betonen, dass vor allem der Faktor der Wirtschaftlichkeit, d. h. ob man genug Abnehmer*innen für den vorgesehenen Ressourcenaufwand findet, eine ausschlaggebende Rolle spielt (ibid.: 209). Wenn diese Fragen geklärt sind, kann die richtige Form ausgewählt werden. Drewer und Schmitz (2017) unterscheiden zwischen einer deskriptiven und einer präskriptiven, einer punktuellen, textbezogenen und fachgebietsbezogenen und zuletzt einer einsprachigen und mehrsprachigen Terminologearbeit (ibid.: 24). In Empfehlungen der Konferenz der Übersetzungsdienste europäischer Staaten (KÜDES 2018) wird außer den erwähnten Formen noch zwischen der Terminologearbeit a priori (vorgelagert) und posteriori (nachgelagert) unterschieden, allerdings wird auf diese zwei Formen nicht näher eingegangen, da sie beide in die Kategorie sachgebietsbezogene Terminologearbeit fallen (ibid.: 70).

2.3.1.1 Deskriptive und präskriptive Terminologearbeit

Bei der deskriptiven Terminologearbeit geht es um eine reine Beschreibung schon existierenden Wortschatzes bzw. fachsprachlichen Benennungen eines bestimmten Fachgebietes, wobei die Wertung und Einschränkung des Wortschatzes keine Rolle spielen. Laut Drewer und Schmitz (2017) handelt es sich hier „lediglich um ein Zwischenergebnis“ (ibid.: 24) bzw. um den Ist-Zustand und daher um eine Vorstufe der präskriptiven bzw. normativen Terminologierarbeit, die durch das Vorschreiben von Benennungen und Definitionen eine Vereinheitlichung des schon existierenden Wortschatzes verfolgt. In der Norm DIN 2342 wird sie als „Terminologearbeit, bei der die Begriffe, Bezeichnungen und/oder Begriffsbeziehungen verbindlich festgelegt werden“ (DIN 2342 2022) definiert. Beide Formen haben ihre Berechtigung. In der Praxis findet die deskriptive Terminologie Anwendung in verschiedenen Fachwörterbüchern, während die präskriptive Terminologie nicht nur in Sprachorganisationen oder

Normungsinstitutionen, beispielsweise durch die Arbeit der Normungsgremien, sondern auch innerhalb verschiedener Institutionen und Unternehmen Verwendung findet (Arntz et al. 2014: 218; Drewer & Schmitz 2017: 24f.).

2.3.1.2 Punktueller, textbezogene und fachgebietsbezogene Terminologearbeit

Bei diesen Formen spielt der Umfang und die verfügbare Zeit der geplanten Terminologearbeit eine wesentliche Rolle. Bei der systematischen Terminologearbeit spricht man von großangelegten, zeitintensiven Projekten, die die Termini eines klar abgegrenzten Gebietes und ihre Beziehungen zueinander darstellen (KÜDES 2018: 69). In manchen Publikationen wie bei Drewer und Schmitz (2017) ist die Rede von einer fachgebietsbezogenen Terminologearbeit, was sich laut KÜDES (2018: 69) mit der Praxis aus vielerlei Gründen, wie z.B. Zeit und Aufwand das ganze Fachgebiet zu erfassen, nicht ganz deckt und deshalb eher die Rede von einer sachgebietsbezogenen Terminologearbeit sein sollte.

Sowie schon bei der Unterscheidung zwischen deskriptiven und präskriptiven Terminologearbeit sprechen Wissenschaftler*innen wie Arntz et al. (2014: 210) im Falle der punktuellen Terminologearbeit von einer Vorstufe der fachgebietsbezogenen bzw. systematischen Terminologearbeit, wobei man mit dieser Form das Ziel verfolgt, ein gegebenes Problem ad-hoc bzw. sofort zu lösen, d. h. das Fachgebiet kann nicht aus terminologischer Sicht ausreichend erfasst werden. In KÜDES (2018: 66) zählen zu den terminologischen Problemen, die man ad-hoc lösen kann z.B. Benennungen, Neologismen, Fachwendungen, offizielle Bezeichnungen und Äquivalente, die das erste Mal im Text vorkommen und dadurch nicht gesichert sind. Dies hat zu Folge, dass die Fehlerquote bzw. das Fehlerrisiko der unzureichenden Abgrenzung eines Begriffes von seinen möglichen „Schein“-Synonymen auch wesentlich höher als bei einer systematischen Terminologearbeit, die innerhalb eines in sich geschlossenen Fachgebietes durchgeführt wird, ist (Drewer & Schmitz 2017: 26). Zusammenfassend aus KÜDES (2018: 67) sind für die punktuelle Terminologearbeit folgende Punkte charakteristisch: nicht immer zufriedenstellende aber schnelle Ergebnisse, großer Aufwand bzw. umfangreiche Recherche für nur einen Fachausdruck, geeignet für provisorische Einträge bzw. Vorstufe für die systematische Terminologearbeit.

Zwischen diesen zwei Formen befindet sich die textbezogene Terminologearbeit, bei der die Terminologie eines spezifischen Textes erarbeitet wird. Nach Drewer und Schmitz (2017: 26) entscheidet die Qualität und Fachgebietsbezogenheit des Textes, ob die Terminologie zuverlässig und reichlich ausfällt. In KÜDES (2018: 68) wird die Problematik eines

allgemeinen Textes angesprochen und betont, dass die Ergebnisse der textbezogenen Terminologiearbeit wohl der systematischen sehr nahekommen können, wenn der Text entsprechende Länge, Sachgebiets- bzw. Themabezogenheit und Einheitlichkeit aufweist.

2.3.1.3 Einsprachige und mehrsprachige Terminologiearbeit

Bei diesen beiden Formen spielt die Anzahl der Sprachen, wie die Bezeichnungen der Formen schon erahnen lassen, die entscheidende Rolle. Die hohen Anforderungen, die bei der mehrsprachigen Terminologiearbeit zu tragen kommen, hängen nicht nur mit der Tatsache zusammen, dass hier zwei oder mehrere Sprachen, sondern auch Kulturen und Sprachwirklichkeiten, verarbeitet werden (Drewer & Schmitz 2017: 26; KÜDES 2018: 81). Wie Drewer und Schmitz (2017: 26f.) betonen, werden die einzelnen methodischen Schritte der einsprachigen Terminologiearbeit (siehe Abschnitt 2.3.2) auf jede einzelne Sprache angewandt, bevor man die Begriffssysteme miteinander vergleichen kann und die Probleme wie Benennungs- oder Begriffslücken, die infolge völlig oder nur teilweise fehlenden Äquivalenz entstehen, angegangen werden können.

2.3.2 Praktische Terminologiearbeit

In KÜDES (2018: 70) wird betont, dass die Terminologiearbeit „in erster Linie Denk- und nicht Sacharbeit“ ist, die eine genaue Organisation benötigt und sie als eine Projektarbeit mit bestimmten Einzelschritten angesehen werden muss, um erwünschte Ergebnisse zu erzielen. Wenn man sich Abb. 10 (siehe Abschnitt 2.3) noch einmal genau ansieht, stellt man fest, dass die Terminologiearbeit im Allgemeinen aus mehreren Phasen besteht, und zwar aus einer Terminologierecherche, aus einer terminologischen Analyse, einer Terminologieverwaltung und einer Terminologieprüfung.

Die einzelnen Projektschritte einer Terminologiearbeit hängen davon ab, für welche Form man sich entscheidet. Bei einer punktuellen Terminologiearbeit bestehen die Schritte aus einer Phase vor der eigentlichen Recherche resp. Lösungsfindung, die im Wesentlichen darauf abzielt, das terminologische Problem genau zu identifizieren und abzugrenzen und einer eigentlichen Recherchephase (KÜDES 2018: 66f.).

Wie oben erwähnt spricht man im Falle einer systematischen Terminologiearbeit von Projekten und deshalb, wie in KÜDES (2018: 71f.) im Detail nachzulesen ist, finden hier die Arbeitsmethoden des Projektmanagements Anwendung, die man im Wesentlichen in zwei

Abschnitte zusammenfassen kann und zwar in Projektschritte, die für die Vorbereitungsphase, während der das Einarbeiten in das Fachgebiet und das Erstellen von Quellenkorpora stattfinden, und Projektschritte, die für die Terminologieerfassungsphase relevant sind. Vor allem die ersten zwei Schritte der Terminologieerfassungsphase aus KÜDES (2018: 71) sind für diese Arbeit von höchster Relevanz, da sich um folgende Schritte handelt: 1. um die terminologische Auswertung der Quelltexte und 2. um das Erarbeiten von Termkandidatenlisten. In weiterer Folge werden die Begriffssysteme geplant, Arbeitsdateien erstellt, Definitionen, Anmerkungen, Kontexte und Metadaten erfasst, mehrsprachige Terminologiesammlungen durch Erstsprachler*innen überprüft, fachliche Prüfung durchgeführt, Freigabe bzw. Publikation genehmigt (ibid.: 71f.). Drewer und Schmitz (2017: 34) weisen darauf hin, dass ein „pauschaler Super-Prozess“ nicht existiert und nennen neun zentrale Schritte, deren Reihenfolge keiner strikten Einhaltung unterliegt und zwar 1. die Zielsetzung und Planung, 2. Gewinnung der Terminologie, 3. Begriffliche Systematisierung, 4. Sprachliche Bewertung und Bereinigung, 5. Generierung/Schaffung neuer Benennungen, 6. Verwaltung der Terminologie, 7. Darstellung und Verbreitung der Terminologie, 8. Pflege der Terminologie und 9. Kontrolle der Terminologieverwendung.

2.3.2.1 Terminologiegewinnung

Nachdem man sich für eine Form der Terminologearbeit entschieden, ihr Ziel gesetzt und ihre Planung abgeschlossen hat, beginnt die eigentliche Erarbeitung der Terminologie, wobei es bei der Terminologiegewinnung bzw. Terminologierecherche mehrere Formen bzw. Verfahren gibt, die in den folgenden Abschnitten erläutert werden.

Laut Drewer und Schmitz (2017) ist die Recherche von jeglichen terminologischen Informationen zu denen nicht nur die naheliegenden Benennungen, Definitionen und Kontexte sondern auch Abbildungen und anderes multimediale Material gehören, einer der wichtigsten Schritte der Terminologearbeit. Die Autor*innen betonen, dass im Normalfall die benötigten Informationen, wie z.B. Benennungen in einem Fachgebiet, schon vorhanden sind (ibid.: 39) und sprechen daher von einer „terminologischen Bestandsaufnahme“ (ibid.: 35) bzw. Sammlung. Sie unterscheiden dabei drei Terminologien, und zwar 1. die explizite, die z.B. in Form von Wortlisten oder Glossaren vorliegt, 2. die implizite, die z.B. in Texten versteckt und erst extrahiert werden muss und 3. die Mischform aus explizit und impliziert in Form von z.B. Translation-Memory-Systemen (ibid.: 40f.). Man kann daraus folgern, dass die Extraktion von Termini nur eine bestimmte Form der Terminologierecherche darstellt, und zwar die implizite

und daher der Begriff Terminologierecherche nicht synonym zu Termextraktion zu verstehen ist. Dieser Problematik haben sich unter anderen Drewer und Schmitz (2017: 41) gewidmet, die die maschinelle Terminologiegewinnung bzw. Sammlung als Terminologie- bzw. Termextraktion und die menschliche als Terminologierecherche bezeichnen, wobei sie bei den Formen aus Gründen der Übersicht zwischen menschlicher Termextraktion (und nicht Terminologierecherche), menschlicher mit einfacher technischer Unterstützung und maschinellen Termextraktion unterscheiden. Im Unterschied zu Drewer und Schmitz unterscheiden Arntz et al. (2014) zwischen einer humanen und maschinengestützten, in allen ihren Ausprägungen, Termextraktion (ibid.: 245).

2.3.2.2 Menschliche Terminologierecherche

Die menschliche Terminologierecherche, auch humane bzw. manuelle Termextraktion genannt, ist ein rein auf der menschlichen Aktivität basierendes Verfahren der Terminologierecherche, das einige Vor- und Nachteile aufweist. Arntz al. (2014) heben hervor, dass die Qualität der Ergebnisse bei diesem Verfahren von bestimmten Voraussetzungen abhängt, wie z.B. dem Vorhandensein geeigneter Quellen und Bearbeiter*innen mit fundiertem Wissen. Weiter ist es wichtig, parallel zu arbeiten, indem z.B. Synonyme bereits bei der ersten Durchsuchung zugeordnet bzw. eliminiert werden und die gegenseitigen Beziehungen der Begriffe festgelegt werden (ibid.: 245). Drewer und Schmitz (Drewer & Schmitz 2017: 42) gehen sogar so weit, dass sie behaupten bei diesem Verfahren keine Korrekturen und Nachbearbeitungen notwendig sind, da die Ergebnisse im Vergleich zu einer maschinellen Termextraktion „zuverlässiger, vollständiger und hochwertiger“ sind. Weiters heben sie hervor, dass durch die umfangreiche menschliche Terminologie- und Informationsrecherche tiefgehende Kenntnisse erworben werden können und daher die Zuordnung und die Beziehungen zwischen den Begriffen leichter nachvollziehbar sind (ibid.: 42).

Dagegen zählen zu den Nachteilen der zeitliche Aufwand, oft fehlende fachlich fundierte Personalressourcen, hohe Kosten und das Fehlerrisiko Mensch selbst z.B. durch mangelhaftes Wissen und die schwankende Konzentrationsfähigkeit der Bearbeiter*innen (Arntz et al. 2014: 245; Drewer & Schmitz 2017: 42). Drewer und Schmitz (2017) weisen darauf hin, dass unterschiedliche Bearbeiter*innengruppen aufgrund ihrer subjektiven Urteilsvermögen und Wissensstände keine identischen Termkandidatenlisten liefern können. Dies führt dazu, dass es keine einwandfreien Goldstandards (siehe Abschnitt 2.4.2.4) geben kann, die aber eine wesentliche Rolle bei der Auswertung der Qualität der Termkandidatenlisten als Vergleichsmaßstab

spielen (ibid: 42). Bakaric et. al (2021: 18) fassen die Nachteile der manuellen Termextraktion in vier Punkten: Fehleranfälligkeit, Arbeitsintensivität, Zeitintensivität und Subjektivität.

2.3.2.3 Menschliche Terminologierecherche mit einfacher technischer Unterstützung

Auch wenn bei Arntz et al. (2014) dieses Verfahren der technischen Unterstützung anhand der gemeinsamen Merkmale, wie der Notwendigkeit einer Überprüfung und Ergänzung durch den Menschen oder die Maschinenlesbarkeit der Texte, unter die maschinelle Termextraktion fällt, ist diese Form leicht zu identifizieren und von den speziellen maschinellen Termextraktionsprogrammen abzugrenzen, so wie bei Drewer und Schmitz (2017).

Zu den einfachsten Mitteln dieses Verfahrens, das bei Arntz et al. (2014: 246) als einfaches Verfahren mit Bordmitteln bezeichnet wird, zählen die Texteditoren oder Textverarbeitungsprogramme. Das Prinzip der Extraktion beruht auf einer Tabellenkalkulation und hat einige Nachteile. Erstens kann der Zeitaufwand bei der Textvorbereitung je nach Korpus enorm sein, da zuerst alle Leerzeichen und Satzzeichen wie Komma, Gedankenstrich u.v.m. manuell entfernt und durch Zeilenumbrüche ersetzt werden müssen, zweitens werden die flektierten Wortformen als eigenständige Kandidaten vorkommen und drittens werden die Mehrwortbenennungen nicht erkannt bzw. getrennt (Arntz et al. 2014: 246; Drewer & Schmitz 2017: 43)

Um die Mehrwortbenennungen einzubeziehen, werden in der ATE zwei weitere Extraktionswerkzeuge verwendet, die in KÜDES (2018) als Konkordanz- und Indexierungsprogramme bezeichnet sind. Es ist jedoch zu beachten, dass diese Werkzeuge ursprünglich nicht für die Termextraktion, sondern für die Korpusanalyse entwickelt wurden. Nach der Vorbereitung des Korpus durch Lemmatisierung identifizieren Konkordanzprogramme eine bestimmte, wiederkehrende Zeichenkette im Korpus und generieren dabei nicht nur sprachliche, sondern auch definitorische Kontexte (ibid.: 107). Laut Drewer und Schmitz (2017) liefern die Konkordanzprogramme etwas bessere Ergebnisse als die oben erwähnten einfachen Mittel, da sie nicht nur mit alphabetischen Listen und Tabellen arbeiten. Sie arbeiten unter anderem mit Grundformen, statistischen Werten wie Häufigkeiten, den Type-Token-Relationen, oder mit einem sogenannten Keyword in Context Index (KWIC-Index), indem Schlüsselwörter zusammen mit ihrem unmittelbaren Kontext erfasst werden und dadurch können Kollokationen und Kontexte angezeigt werden (Arntz et al. 2014: 246f.; Drewer & Schmitz 2017: 43).

Ähnlich wie die Konkordanzprogramme liefern die Indexierungsprogramme definitorische Kontexte sowie Begriffsbeziehungen, Beispiele und Wortkombinationen mittels wiederkehrender Schlüsselwörter (KÜDES 2018: 107).

2.4 Maschinelle bzw. (halb-)automatische Termextraktion (ATE)

In der Fachliteratur und in der Praxis stößt man auf unterschiedliche Benennungen des Verfahrens der ATE. Bei Arntz et al. (2014: 245) ist die Rede von einer maschinengestützten Terminologieextraktion, bei Drewer und Schmitz (2017: 41ff.) von einer maschinellen Termextraktion. Dieses Verfahren schließt je nach Ansatz mehrere Verfahren ein (Arntz et al. 2014: 245) und auch wenn in einigen Fachpublikationen wie bei Arntz et al. (2014) und Drewer und Schmitz (2017) in weiterer Unterteilung von unterschiedlichen Verfahren die Rede ist, wird zwecks Übersichtlichkeit in dieser Arbeit die Bezeichnung Ansatz bevorzugt.

Bei ATE-Ansätzen handelt es sich laut Arntz et. al (2014: 245) um eine automatische Gewinnung von Termini, die im Allgemeinen durch Rechner bzw. unterschiedlich ausgeprägte Softwareprogramme stattfindet. Die Einbeziehung der Softwareprogramme stellt das markanteste Unterscheidungsmerkmal zu manuellen Terminologierecherche dar, wobei bei Drewer und Schmitz (2017) zusätzlich die Ausprägung der Softwareprogramme eine wichtige Rolle spielt. Sie führen eine Zwischenkategorie ein, die sie als „einfache Unterstützungsverfahren bei der menschlichen Termextraktion“ bezeichnen (siehe Abschnitt 2.3.2.3) und wenn sie von maschinellen Termextraktion sprechen, beziehen sie sich auf spezielle Programme, die nahezu automatische Ergebnisse liefern (ibid.: 42f.). Die Norm DIN 2343 (2022) definiert die automatische Termextraktion als ein „Verfahren zur Unterstützung der Terminologierecherche, bei dem maschinell Zeichenfolgen aus Texten oder Textkorpora identifiziert werden, (...)“, wobei nicht näher spezifiziert wird, wie einfach oder speziell die verwendeten Werkzeuge sein müssen.

Was die erwähnten Definitionen gemeinsam haben, ist die Erwähnung der technischen Komponente, wobei die DIN-Definition die Verwendung des Adjektivs „automatisch“ umgeht, da dieses möglicherweise unrealistische Erwartungen im Sinne der völligen Automatisierung ohne die Notwendigkeit menschlichen Eingreifens hervorrufen könnte. Im Gegensatz dazu ist in KÜDES (2018) von einer automatischen Termextraktion die Rede, wobei erwähnt werden muss, dass man in den meisten Fällen von einer halbautomatischen Termextraktion ausgehen sollte, da der menschliche Faktor bei der Überprüfung und Überarbeitung der Ergebnisse immer noch eine wesentliche Rolle spielt (Arntz et al. 2014: 246; Drewer & Schmitz 2017: 44).

Wie diese kurze Einführung zur ATE und der Abschnitt über Terminologiegewinnung (siehe Abschnitt 2.3.2.1) aufzeigen, tritt terminologische Inkonsistenz in der Fachliteratur bei der Bezeichnung des ATE-Ansatzes häufig auf. Paradoxerweise betrifft dies einen Ansatz,

dessen Ziel es ist, Termini zu gewinnen, für die Eineindeutigkeit eine der wichtigsten Eigenschaften ist (siehe Abschnitt 2.2.1). Ausgerechnet bei der Bezeichnung des ATE-Ansatzes mangelt es an dieser Eineindeutigkeit, was auch durch einige wissenschaftliche Arbeiten belegt wird. Konkret wird bei Astrakchantsev (2018) von einer sogenannten automatischen Termerkennung oder englisch Automated Term Recognition (ATR) und bei Rigouts Terryn et al. (2020) von einer automatischen Termextraktion oder englisch Automated Term Extraction (ATE) gesprochen. Darüber hinaus gibt es auch einige Forscher*innen die beide Begriffe synonym verwenden wie Šajatović et al. (2019), die von einer Automatic Term Extraction (or Recognition) sprechen. Heylen und De Hertog (2015) führen weiter aus, dass die ATE auch unter den Bezeichnungen „Terminology Extraction, Terminology Mining, Term Recognition, Glossary Extraction, Term Identification and Term Acquisition“ (ibid.: 203) bekannt ist. Zwangsläufig stellt sich die Frage, ob in diesen Fällen von ein- und demselben Ansatz die Rede ist.

Obwohl der Begriff ATE oben erläutert wurde, ist es wichtig, die in den Forschungsarbeiten verwendeten Definitionen näher zu betrachten, um weitere wichtige Grundelemente der ATE zu identifizieren. Für diesen Zweck werden folgende Definitionen herangezogen:

1. Astrakchantsev (2018: 854): “Automatic terminology recognition (ATR) aims at extraction of terms – words and collocations designating domain-specific concepts – from a collocation of text documents belonging to that domain”,
2. Rigouts Terryn et al. (2020: 85): „Automatic Term Extraction (ATE) can be defined as the automated process of identifying terminology from a corpus of specialised texts.”,
3. Šajatović et al. (2019: 149): „The aim of Automatic Extraction (or Recognition) (ATE) is to extract terms – single words or multiword expressions (MWEs) representing domain-specific concepts – from a domain specific corpus.”.

In allen, bis jetzt erwähnten, Definitionen wird das Ziel der ATE, Extraktion von Termini oder zumindest nicht näher definierten Zeichenfolgen (siehe DIN-Definition oben), explizit erwähnt. Was unter der Bezeichnung Termini die jeweiligen Forscher*innen verstehen, wird in jeder Forschungsarbeit näher spezifiziert, da man die unterschiedlichen Konzepte des Terminus nicht außer Acht lassen kann (siehe Abschnitt 2.1.4 und 2.2.3). Die ersten Hinweise lassen sich in der Regel am Anfang der jeweiligen Forschungsarbeiten anhand der Definitionen der ATE (siehe oben Definition 1 und 3) feststellen. So werden z.B. in der Forschungsarbeit von

Astrakhantsev (2018) fachspezifische Wörter und Kollokationen, d. h. häufig vorkommende semantisch motivierte Wortverbindungen, wie z.B. *Umsatz steigern*, oder im Fall von Šajatović et al. (2019) Ein- und Mehrwortbenennungen, die fachspezifische Konzepte repräsentieren, extrahiert. Die genaue Spezifizierung, was alles als Terminus in die jeweilige Forschungsarbeit aufgenommen wird, ist nicht nur im theoretischen Teil, sondern auch im praktischen Teil der Forschungsarbeiten und dazugehörigen Anhängen ersichtlich, z.B. in der Form von spezifischen Annotationsschemata und Richtlinien, wie sie z.B. von Rigouts Terryn (2021a) erstellt wurden.

Ein weiteres gemeinsames Merkmal der oben erwähnten Definitionen ist die Hervorhebung zusätzlicher Merkmale bzw. Grundelemente, die bei der ATE auftreten. Dazu gehören unter anderen spezifische (Teil-)Fachgebiete (siehe Abschnitt 2.1.1) und spezifische Korpora. Einige Wissenschaftler*innen (Drouin et al. 2020: 1) betonen die Wichtigkeit des Korpus, da dieser die Qualität der Ergebnisse direkt beeinflusst. Der Korpus besteht aus einzelnen Quellen bzw. Dokumenten, die zwar von Ziel und Inhalt der Termextraktion abhängen, allerdings lassen sich bestimmte Qualitätsmerkmale festlegen, die im Allgemeinen für jeden Ansatz der Gewinnung von Termini gültig sind. Drewer und Schmitz (2017) identifizieren insgesamt vier solche Merkmale. Als erstes Merkmal nennen sie den Zuverlässigkeitsgrad, d. h. die Quelle muss allgemein bekannt und anerkannt sein, z.B. eine Richtlinie, als zweites Merkmal die Aktualität der Quelle, wobei gilt, dass die aktuelleren Quellen zu bevorzugen sind, als drittes Merkmal das Erstsprachenprinzip, das besagt, dass die Dokumente in der Erstsprache der Autor*innen verfasst werden müssen (ibid.: 51). Als das vierte Merkmal wird bei Drewer und Schmitz (ibid.: 52) „die Fachkompetenz des Verfassers“ genannt, die zwar die fachliche Richtigkeit gewährleistet, jedoch auch beachtet werden muss, dass die Fachexpert*innen aus spezifischen Forschungsgruppen stammen können. Daher kann es durch ihre subjektiven Entscheidungen zu terminologischer Inkonsistenz kommen. Drouin et al. (2020: 1) sprechen in Bezug auf Verfasser*innen von sogenannten „ideal requirements“ und betonen, dass die einzelnen Dokumente im Korpus von Expert*innen des jeweiligen Fachbereiches verfasst werden müssen. Drewer und Schmitz (2017: 52) nennen die wichtigsten Quellentypen, wie Normen und Richtlinien, Fachzeitschriften, Fachlexika u.v.m., wobei sie Fachbücher, geschrieben von Fachexpert*innen mit großem Terminologiebewusstsein, als die ideale, auch wenn seltene, terminologische Quelle bezeichnen.

Diese Quellen sind sowohl auf traditionellem Wege zugänglich wie beispielsweise in Bibliotheken, als auch im Internet, was schneller und effizienter ist. Allerdings ist bei den Unmengen an im Internet verfügbaren Dokumenten zu beachten, dass dieselben

Qualitätsmerkmale wie für traditionell zugängliche Quellen gelten (für weitere Informationen siehe Drewer und Schmitz 2017: 55ff.).

Auf dieser Stelle muss angemerkt werden, dass diese idealen Anforderungen an ATE-Forschung und ihre Quellen in der Praxis kaum anzutreffen sind und deshalb der Fokus der Wissenschaftler*innen auch auf atypischen terminologischen Quellen wie auf Einzeldokumenten statt Gesamtkorpora (Šajatović et al. 2019), Korpora aus Zeitungsartikeln (Bakaric et al. 2021) oder allgemeinen Korpora bestehend aus Texten, die das jeweilige Thema nicht direkt ansprechen (Drouin et al. 2020), liegt. In dieser Arbeit wird der Korpus aus Zeitungsartikeln zusammengestellt, die zwar in der ATE-Forschung keine Unbekannte mehr sind, trotzdem aber als atypische Quellen für die ATE gelten, da diese Textsorte an die breite Öffentlichkeit gerichtet ist (Drouin et al. 2020: 1) und dementsprechend für diese Zielgruppe lesbar gemacht wird (siehe Abschnitt 2.1.8.2).

Zusammenfassend steht fest, dass ungeachtet dessen wie man den ATE-Ansatz nennt, das Ziel immer dasselbe ist: die Gewinnung von Termini, indem ein Computerprogramm den aus geeigneten Quellen bestehenden Korpus analysiert.

Wie schon oben erwähnt, den einen ATE-Ansatz gibt es nicht, da es innerhalb dieser Kategorie unterschiedliche Ansätze bzw. Methoden gibt. Die unterschiedlichen Ansätze haben Einfluss auf den gesamten Ablauf bzw. Workflow der Termextraktion, d. h. wie ein System einen Terminus extrahiert, hängt von seiner jeweiligen Architektur bzw. Ansätzen ab, die den Systemen zugrunde liegen (siehe Abschnitt 2.4.1.1 und 2.4.1.2). Deshalb kann man in der ATE keinen allgemein gültigen Arbeitsablauf beschreiben. Wie eine ATE in einzelnen Schritten aussieht, hängt nicht nur von den ATE-Ansätzen allein ab, sondern auch von der unterschiedlichen Perspektiven und Interessensgruppen. Zum Beispiel beschreibt Yee (2018) den gesamten Arbeitsablauf aus der Sicht der Anwender*innen in fünf Schritten: 1. Vorbereitung des Inputs je nach der beabsichtigten Extraktion (ein- bzw. mehrsprachig); 2. Durchlaufen des Algorithmus, was für die meisten Anwender*innen als Blackbox-Prozess, d. h. ohne Kenntnisse der Zusammensetzung des Algorithmus, stattfindet; 3. Erhaltung des Outputs in Form einer Kandidatenliste erhalten; 4. Verifizierung der vorläufigen Termkandidaten durch die Anwender*innen; 5. Speicherung bzw. Export der Ergebnisse (2018: 102f.).

Etwas detaillierter aus der Sicht der Wissenschaftler*innen beschreiben Heylen und De Hertog (2015) den Workflow, indem sie die ATE in mehrere Teilprozesse unterteilen, wobei zwei der wichtigen Grundbegriffe der ATE hier ihre Anwendung finden, und zwar Termhood und Unithood. Diese zwei Begriffe ermöglichen eine präzise Erkennung von Termini bzw. werden anhand denen Termkandidaten extrahiert. Diese Begriffe gehen auf Kageura und Umino

(1996: 260) zurück, die Unithood als Grad der Stärke bzw. Stabilität syntagmatischer Kombinationen und Kollokationen und Termhood als Grad der Zugehörigkeit einer linguistischen Einheit zu einem bestimmten Fachgebiet definieren. Einfacher ausgedrückt, geht es um Zusammenhörigkeit komplexer linguistischer Einheiten (Unithood) und die Termhaftigkeit (Termhood). Die von Heylen und De Hertog (2015) beschriebenen Teilprozesse werden üblicherweise Schritt für Schritt in folgender Reihenfolge ausgeführt: 1. Zusammenstellung des fachgebietsspezifischen Korpus; 2. Identifizierung der Unithood; 3. Identifizierung der Termhood; 4. Identifizierung der unterschiedlichen Realisierungen desselben Konzepts; 5. Evaluation und Validierung im Vergleich zur manuellen Termextraktion (Bakaric et al. 2021: 19; Heylen & De Hertog 2015: 204). Allerdings entsprechen nicht alle ATE-Ansätze dieser Vorstellung des Ablaufs, da nicht alle Ansätze häufigkeitsbasiert sind.

Einen weiteren Versuch um eine Beschreibung des allgemeinen ATE-Ablaufs unternimmt Astrakhantsev (2018), wobei er explizit anführt, dass seine ATE-Ablaufbeschreibung nicht alle Methoden einschließt, da z.B. regelbasierte Methoden nicht in das Toolkit aufgenommen wurden. Er identifiziert und erläutert im Detail folgende Schritte: 1. Textvorverarbeitung 2. Sammlung von Textkandidaten, 3. Scoring-Verfahren und das 4. Ranking-Verfahren. Dabei sieht Astrakhantsev (2018: 855) das Scoring-Verfahren als den wichtigsten und am fortgeschrittensten Teilprozess, infolgedessen jedem einzelnen Termkandidaten eine Nummer zugeschrieben wird, die besagt, wie wahrscheinlich es ist, dass der Termkandidat ein Terminus ist. Dieses Verfahren bildet die Grundlage für die Kategorisierung von ausgewählten Methoden, die in das Open-Source Toolkit ATR4S (Astrakhantsev 2018) implementiert wurden, um einen Vergleich dieser Methoden, die zu diesem Zeitpunkt als state-of-the-art im Bereich der Termerkennung galten, durchzuführen (siehe Abschnitt 2.4.2). Beim Ranking-Verfahren handelt es sich um einen Prozess, bei dem verschiedene potenzielle Termkandidaten bewertet werden, um die geeignetsten Kandidaten auszuwählen bzw. Termini von Nicht-Termini zu unterscheiden. Zur Bewertung werden unterschiedliche Methoden herangezogen, die von einfachen Punktzahlen bis zu komplexen Algorithmen des maschinellen Lernens reichen (Näheres siehe Astrakhantsev 2018: 863f.)

Auch wenn sich die Beschreibung des Prozesses je nach Autor*in und Ansatz unterscheidet, bleiben die wesentlichen Schritte wie Bereitstellung des Inputs, Sammlung der vorläufigen Kandidaten und deren anschließende Verifizierung identisch.

2.4.1 Arten und Ansätze der ATE

Die ATE-Tools können aus unterschiedlichen Perspektiven, wie z.B. Anzahl der Sprache, Art der Ansätze u.v.m., betrachtet und dementsprechend unterteilt werden. In der früheren und heutigen ATE-Forschung begegnet man je nach Perspektive unterschiedlichen Unterteilungen, wie z.B. je nach Umgebung und Art der Zugänglichkeit der Tools wird laut Costa et al. (2016) zwischen folgenden Kategorien unterschieden. Die erste Kategorie beinhaltet die standalone Anwendungen, die auch als Desktop-Anwendungen bekannt sind, für die Costa et al. (2016) das Tool Termsuite als Beispiel anführen. In die zweite Kategorie gehören die webbasierten Anwendungen, die innerhalb eines Browsers arbeiten wie z.B. das Tool Sketch Engine, und in die dritte die sog. Frameworks, d. h. wiederverwendbare Softwares oder Bibliotheken wie Java Automatic Term Extraction (JATE), die innerhalb eines anderen Produktes oder umfangreicherer Software verwendbar sind. Für weitere Informationen zu den angeführten Beispielen siehe z.B. Costa et al. (2016).

Je nach Anzahl der Sprachen wird zwischen zwei Arten der ATE nach einsprachig/monolingual und zwei/bi- oder multilingual unterschieden, wobei Drewer und Schmitz (2017) darauf hinweisen, dass die mehrsprachige Extraktion meist eine zweisprachige Termextraktion ist, die mehrmals mit derselben Ausgangssprache und verschiedenen Zielsprachen durchgeführt wird. Als Ausgangspunkt für die zweisprachige Termextraktion stehen parallele zweisprachige, segmentierte und alignierte Texte zur Verfügung, vor allem in der Form von Translation-Memories (ibid.: 45). Unter alignierten Texten versteht man Texte in der Ausgangs- und in der Zielsprache, die zuerst in Segmente aufgeteilt und danach einander zugeordnet werden, wobei Drewer und Schmitz (ibid.) betonen, dass diese Textpaare eine sehr ähnliche Struktur aufweisen müssen.

In der Fachliteratur (Arntz et al. 2014; Drewer & Schmitz 2017) und in manchen wenigen noch vor Kurzem veröffentlichten Forschungsarbeiten (Bakaric et al. 2021), begegnet man der typischen Zwei- bzw. Dreiteilung der ATE-Ansätze in linguistisch, statistisch und hybrid. Auch wenn diese konventionelle Unterteilung nicht mehr der aktuellen ATE-Forschung gerecht wird, wird in dieser Arbeit kurz auf diese drei Kategorien eingegangen, da sie nicht nur einen Überblick über die früheren Phasen der ATE bieten, sondern auch, weil einige Tools, die heutzutage im Gebrauch sind, linguistische und statistische Methoden in ihren Architekturen implementieren, wie zum Beispiel das in dieser Arbeit verwendete Tool OneClick Terms (siehe Abschnitt 4.3). Heutzutage betont die Mehrheit der Wissenschaftler*innen (Drouin et al. 2020; Rigouts Terryn et al. 2020 u.v.m.), dass die Entwicklungen der letzten Jahre im Bereich der

ATE verknüpft mit den Entwicklungen des maschinellen Lernens, vor allem im Bereich des Deep Learning, solche simplen Unterteilungen sehr schwierig und daher irrelevant machen. Der Grund dafür ist die Unmenge an unterschiedlichen ATE-Methoden, die so divers geworden sind, dass sie keiner klar definierten Kategorie zuzuordnen sind. Unter anderen schlussfolgern Rigouts Terryn et al. (2020: 86f.), dass heutzutage keine simple binäre Unterteilung der Methoden bzw. der Systeme ausreichend ist. Sie präsentieren Beispiele an möglichen Unterteilungen der Methoden mit dazugehörigen Forscher*innen wie z.B. Systeme, die auf einem einzelnen statistischen Wert basieren, merkmalsbasierte Systeme oder sogar merkmallöse Systeme (ibid.). Trotz dieser Diversität werden Versuche unternommen, die Methoden der ATE in Kategorien zusammenzufassen, obwohl diesen Unterteilungen kein Anspruch an Vollständigkeit zugrunde liegt. Zum Beispiel findet sich bei Šajatović et al. (2019: 150) ein kurzer Überblick der ATE-Methoden nach Astrakhantsev (2018), der zur Kategorisierung der ATE-Methoden die Art der Information, die zur Bewertung bzw. dem sogenannten Scoring der Termkandidaten dient, in Betracht zieht (siehe Abschnitt 2.4.2). Einen Überblick über die neuesten neuronalen Methoden in der ATE, vor allem auf Deep Learning basierten Ansätze mit Fokus auf sog. Transformatorbasierte neuronale Modelle, fassen Hanh et al. (2023) zusammen, indem sie die neuesten neuronalen Ansätze mit ausgewählten nicht neuronalen Methoden vergleichen (siehe Abschnitt 2.4.2.9).

2.4.1.1 Konventionelle Ansätze

Der statistische Ansatz gehört zu den zwei ursprünglichen Basisansätzen der ATE (Bakaric et al. 2021: 19; KÜDES 2018: 106) und basiert auf dem Prinzip der Auftretenshäufigkeit eines Terminus im Text bzw. Korpus, die mittels verschiedenen mathematischen Verfahrens wie z.B. Chi-Quadrat-Verteilung-Methode, d. h. Wahrscheinlichkeitsverteilung, berechnet werden kann (Bakaric et al. 2021: 19). Vereinfacht ausgedrückt je höher die Auftretenshäufigkeit, desto wahrscheinlicher ist es, dass es sich um einen Terminus handelt. Allerdings betonen Arntz et al. (2014), dass diese Vermutung nicht richtig ist, da vor allem die Wörter der Gemeinsprache als vermutliche Termini öfters im Korpus identifiziert werden können und deshalb eine Stoppwortliste benötigt wird, um diese im Voraus auszuklammern (ibid.: 247).

In KÜDES (2018) werden die ATE-Hilfsmittel bzw. ATE-Tools, die auf einem statistischen Ansatz basieren, als Werkzeuge beschrieben, die im Stande sind, absolut identische Zeichenketten, die im Korpus immer wieder vorkommen, zu erkennen, wobei die Morphologie und Syntax keine Rolle spielen (ibid.: 106). Folglich ist der statistische Ansatz an keine

bestimmte Sprache gebunden und daher universell einsetzbar, was vor allem Vorteile für seltene Sprachen hat, von denen keine linguistischen Regeln bestehen (Drewer & Schmitz 2017: 49).

Nicht nur die Chi-Quadrat-Verteilung, sondern etliche andere Methoden sind für die statistische Ansätze relevant wie z.B. die Maße Term Frequency (TF), Inverse Document Frequency (IDF) u.v.a. (siehe Abschnitt 2.4.2.1).

Auch wenn der statistische Ansatz einige Vorteile wie z.B. die Sprachunabhängigkeit mit sich bringt, gibt es auch Nachteile. Drewer und Schmitz (2017) betonen vor allem die Fehleranfälligkeit bei der Identifizierung mancher Mehrwortbenennungen, vor allem derer, die aus einem Adjektiv und einem Nomen bestehen, und die benötigte Datenmenge, da für den statistischen Ansatz gilt „[j]e größer die Textmenge, desto besser das Ergebnis“ (ibid.: 49), was die Anwendungsbereiche der Termextraktion einschränken. Zu weiteren Nachteilen dieses Verfahrens zählt der langwierige Vor- bzw. Nachverarbeitungsprozess, da alle Flexionsformen einer Benennung als vermutliche Kandidaten identifiziert werden können, was mit einem zusätzlichen Lemmatisierungsprogramm zwar im Vorfeld reduziert, allerdings nicht ganz eliminiert werden kann, vor allem bei Sprachen, die sehr flexionsreich sind (Arntz et al. 2014: 248; KÜDES 2018: 106). Und Zuletzt ist die Problematik von Silence bzw. Noise der Ergebnisse zu erwähnen, zu der es kommt, wenn man die Schwellenwerteinstellungen der Häufigkeit zu hoch bzw. zu niedrig einstellt (Arntz et al. 2014: 247). Silence tritt auf, wenn relevante Termini nicht oder nur teilweise erkannt werden, was zur Folge hat, dass das Tool potenziell wichtige Begriffe übersieht bzw. laut Drewer und Schmitz (2017: 47) „still“ bleibt. Noise tritt wiederum auf, wenn das Tool falsche oder unnötige Termkandidaten vorschlägt, die keine tatsächlichen Termini sind und dadurch laut Drewer und Schmitz (2017: 47) „Lärm“ in den extrahierten Ergebnissen verursachen. Alle diese Nachteile und auch die Schlussfolgerungen der Wissenschaftler*innen wie Arntz et al. (2014: 248) verdeutlichen, dass die Ergebnisse der rein statischen Ansätze hinter den Ergebnissen anderer Ansätze zurückbleiben. Andererseits liefern die erwähnten Vorteile viel Nützliches für die (Weiter-)Entwicklungen in der ATE-Forschung.

Die zweite Gruppe der Basisansätze ist die linguistische, die sich die Analyse der Morphologie und der Syntax zunutze macht und dadurch zuverlässigere Ergebnisse als die Gruppe der statistischen Ansätze liefert. Voraussetzung für diese Analyse ist ein sprachspezifisches Regelwerk, das laut Arntz et al. (2014) im Idealfall aus Wörterbüchern, morphologischen und syntaktischen Regeln und Musterstrukturen besteht (ibid.: 248).

Sowie bei den statistischen, gibt es auch bei den linguistischen Ansätzen verschiedene Methoden, wie die Analyse durchgeführt werden kann. Witschel (2004) zählt die wichtigsten

Methoden auf, zu denen er 1. die vollständige syntaktische Analyse bzw. die Zerlegung des Satzes in seine grammatikalischen Einheiten, sog. Parsing, 2. NP-Chunking, d. h. die Identifizierung von Nominalphrasen (NPs) bzw. die Zerlegung des Satzes in kleinere Teile, sog. Chunks, 3. Extraktion von Part-Of-Speech-Mustern (POS-Mustern) bzw. POS-Tagging, d. h. Zuordnung der Wörter zu Wortkategorien, zählt (Näheres siehe Aubin & Hamon 2006: 383f.).

Linguistische Ansätze bieten gegenüber den statistischen mehrere Vorteile. Wissenschaftler*innen wie Arntz et al. (2014) und Drewer und Schmitz (2017) betonen insbesondere vier dieser Vorteile. Erstens ermöglicht eine auf einem Regelwerk gestützte Analyse eine präzise Zuordnung der Termkandidaten zu ihren entsprechenden Wortarten. Zweitens werden Termkandidaten auf ihre Grundform zurückgeführt, die zur Konsistenz und Vergleichbarkeit der extrahierten Termini führt. Drittens können durch das Anbinden von Wörterbüchern und lexikalischen Ressourcen viele unerwünschte Noise-Termkandidaten eliminiert werden, was zu einer höheren Präzision und Zuverlässigkeit der extrahierten Termini führt und viertens ermöglicht die syntaktische Analyse die Erkennung komplexer Sprachstrukturen wie Phrasen, Mehrwortbenennungen und Ellipsen. Zusammenfassend haben diese Vorteile einen erheblichen Einfluss auf die Nachbearbeitungsphase der Kandidatenlisten, indem sie den Aufwand der manuellen Überprüfung und Bereinigung der extrahierten Termini reduzieren und die Qualität der Ergebnisse verbessern (Arntz et al. 2014: 248).

Was auf den ersten Blick wie ein Vorteil aussieht, birgt auch einen der größten Nachteile der linguistischen Ansätze, und zwar die Notwendigkeit eines sprachgebundenen Regelwerkes. Wie schon Arntz et al. (2014) betonen, ist ein umfassendes einzelsprachspezifisches Regelwerk unabdingbar, wodurch die Problematik für ressourcenarme Sprachen sichtbar wird. Es wird sogar explizit angeführt, dass es nur wenige ressourcenreiche Sprachen gibt, wie z.B. Deutsch und Englisch, die über umfassende Regelwerke verfügen (ibid.: 248).

Trotz vieler Vorteile bleibt die Sprachabhängigkeit eine enorme Einschränkung der auf linguistischen Ansätzen basierenden Tools und deshalb war die Entwicklung Richtung hybrider Ansätze nur eine Frage der Zeit.

Hybride Ansätze entstanden durch eine Kombination der Vorteile der statistischen und linguistischen Methoden, allerdings, sowie Drewer und Schmitz (2017) betonen, nicht ohne den größten Nachteil der Sprachabhängigkeit (ibid.:49). Im Allgemeinen werden zunächst die Termkandidaten aus dem Korpus mittels Methoden der linguistischen Ansätze extrahiert, z.B. mittels POS-Mustern, und anschließend werden so generierte Termkandidaten durch eine statistische Methode gefiltert. Witschel (2004: 7) stellt fest, dass auch wenn die linguistischen Methoden klare Vorteile gegenüber den statistischen Methoden aufweisen, ist die Frequenz für

das Filtern geeignet, da Fachtermini lexikalische Einheiten sind, die sich durch das „wiederholte und unvariierte Auftreten“ im Gegensatz zu nicht-terminologischen Einheiten auszeichnen.

Wie schon im Abschnitt 2.4.1 erwähnt, ist diese konventionelle Aufteilung in heutiger ATE-Forschung nicht mehr relevant. Schon vor der Dominanz der neuronalen Ansätze und daher der absoluten und dezidierten Systeme in der ATE, betonten Wissenschaftler*innen (Bakaric et al. 2021; Rigouts Terryn et al. 2020), dass die Klassifizierung in linguistische und statische Ansätze veraltet ist. Erstens scheinen die hybriden Ansätze diese zwei Kategorien immer zu übertreffen (Rigouts Terryn et al. 2020: 86) und zweitens kommen die linguistische und statistische Ansätze laut Bakaric et al. (2021: 19) als bloße Filter, konkret z.B. „linguistic filters and term-frequency-based-filters“, in einem komplexen Prozess in diversen hybriden Tools zum Einsatz.

2.4.1.2 Neue Ansätze in der ATE

Das maschinelle Lernen, vor allem mit seinem Teilbereich Deep Learning, hat die Termextraktion in den letzten paar Jahren grundlegend verändert. Laut Mahesh (2020) wird das maschinelle Lernen als ein Prozess definiert, der die Maschinen befähigt, selbst zu lernen, also selbst Muster zu erkennen, ohne dass diese explizit programmiert werden (ibid.: 381). Einfach ausgedrückt lernen Maschinen mithilfe des entsprechenden Algorithmus selbst, wie sie effizient mit den Eingangsdaten umgehen können. Ähnlich wie bei den traditionellen Ansätzen der ATE gibt es auch beim maschinellen Lernen verschiedene Arten, wie in Abb. 11 übersichtlich dargestellt wird, die zur Lösung des jeweiligen Problems bzw. der Aufgabe führen, wobei es den „one-size-fits-all“ (Mahesh 2020: 381) Ansatz nicht gibt. Vereinfacht gesagt gibt es mehrere Arten des maschinellen Lernens.

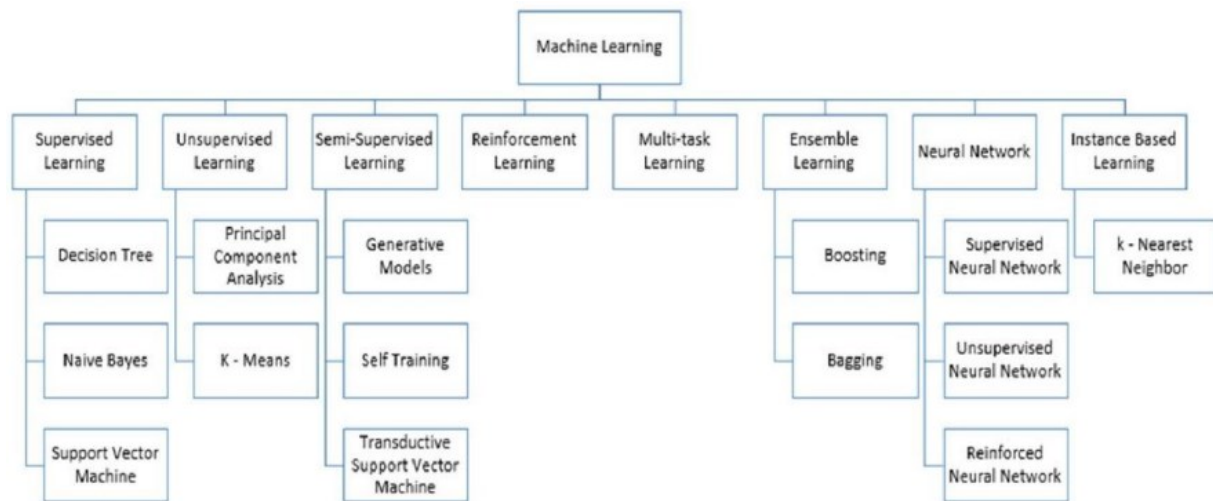


Abbildung 11: Übersicht der Arten des maschinellen Lernens (Mahesh 2020: 381)

Wie aus Abb. 11 hervorgeht, gibt es überwachtes Lernen, unüberwachtes Lernen, semiüberwachtes Lernen (Englisch Semi-Supervised Learning), Reinforcement-Lernen (Englisch Reinforcement Learning), Multitask-Lernen, Ensemble-Lernen, künstliche neuronale Netzwerke (Englisch Artificial Neural Network) und Instance Based-Lernen. Außerdem schließen diese Kategorien einander nicht aus, da heutige Ansätze zum Training von neuronalen Netze oft auf eine Form des überwachten Lernens zurückgreifen. Im Bereich der ATE sind vor allem das überwachte und unüberwachte Lernen von Interesse ist. Zu diesem Schluss kommen Wissenschaftler*innen in einer der neuesten Studie zur ATE (Tran et al. 2023), wobei sie betonen, dass die Ansätze des Deep Learning und des überwachten Lernens beliebter als die des unüberwachten Lernens sind (ibid.: 2). Der überwachte Ansatz basiert auf Datensätzen, die eine explizite Annotation der zu lernenden Merkmale beinhalten. Im Fall der ATE sind diese Merkmale die zu extrahierenden Termini. Im Unterschied zum überwachten Lernen gibt es beim unüberwachten Lernen keine annotierten Datensätze bzw. korrekten Antworten, um der Maschine beizubringen, was ein Terminus ist. Die Modelle sind auf sich selbst gestellt, Muster und interessante Zusammenhänge in den Daten ohne externe Hilfe zu erkennen (Mahesh 2020). Das heißt, dass die Merkmale der zu extrahierenden Termini selbst, ohne Annotation, von der Maschine generiert werden. Auch bei der überwachten Lernmethode wird nur auf Basis von annotierten Daten trainiert und die resultierenden Modelle sind in der Lage neue, zuvor ungesehene Texte ohne Annotation zu verarbeiten.

2.4.2 Zeitgemäße Unterteilung der ATE-Methoden

Ein Versuch um eine gerechte zeitgemäße Kategorisierung der ATE-Methoden kommt von Astrakhansev et al. (2015), die die Methoden erstens nach der Art der Informationen, die dem Scoring-Verfahren (siehe Abschnitt 2.4), und zweitens Methoden, die dem Ranking-Verfahren (siehe Abschnitt 2.4 und 2.4.2.8) zugrunde liegen. Dazu gehören Methoden, die auf Statistiken und Kontexten über das Auftreten der Termini basieren, Methoden, die Themenmodellierung verwenden, und Methoden, die auf externen Ressourcen wie Textsammlungen aus anderen Bereichen, Ontologien oder Wikipedia basieren und die sog. „Methods of Feature-Based Inference“ (Astrakhansev et al. 2015: 345). In den folgenden Abschnitten werden einzelne Kategorien nähergebracht, wobei zu beachten ist, dass jede Kategorie aus etlichen Methoden besteht, deren detaillierte Beschreibung den Rahmen dieser Arbeit überschreiten würde und deshalb nur auf einige Methoden, die stellvertretend für ganze Gruppe stehen, näher eingegangen wird.

2.4.2.1 Methoden basierend auf Statistiken über das Auftreten der Termini

Zu dieser Gruppe gehören Methoden, die laut Witschel (2004) schon im früheren Entwicklungsstadium der ATE-Forschung zu den meistgenutzten Methoden in der ATE gehörten, wie die Maße TF, d. h. die Häufigkeit eines Terminus im Korpus, IDF, d. h. das Vorkommen seltener Begriffe, und die Kombination aus beiden TF-IDF. Die letztere Methode zählt Astrakhansev (2018) zu den klassischen statistischen Methoden der Informationsgewinnung.

Stellvertretend für diese Gruppe wird im Folgenden am Beispiel der TF-IDF-Formel (6) und C-Value-Formel (6) veranschaulicht, was solchen Maßen zugrunde liegen kann.

$$TF \cdot IDF_{(t)} = TF_{(t)} \cdot \log \frac{D}{DTF_{(t)}} \quad (1)$$

Die Formel 1 kombiniert zwei Standardmaße der automatischen Termextraktion TF und IDF, wobei t für den jeweiligen Termkandidaten, D für die Anzahl der Dokumente im Korpus und folglich $DTF(t)$ für die Anzahl der Dokumente, die den Termkandidaten beinhalten steht (ibid.).

$$C - value(a) = \begin{cases} w(a) \times f(a) & \text{if } a \notin nested \\ w(a) \times \left(f(a) - \frac{1}{|S_a|}\right) \times \sum_{b \in S_a} f(b) & \\ otherwise & \end{cases} \quad (2)$$

Der C-Value gehört laut Astrakhantsev (2018) zu einer der populärsten statistischen Methoden, die ursprünglich nur für die Identifizierung von Mehrwortbenennungen gedacht war und später durch die Ausweitung durch Lossio-Ventura et al. (2013) auch die Einwortbenennungen einbezogen. In der Formel 2 sind linguistische und statistische Informationen vereint, wobei a für den Termkandidaten, $|a|$ für die Anzahl der Wörter im Termkandidaten, f_a für die Häufigkeit des Termkandidaten im Dokument, S_a für die Menge der Begriffe, die den Termkandidaten beinhaltet und $|S_a|$ für die Anzahl der Begriffe in der Menge S_a steht. Einfach ausgedrückt, die Anzahl des Vorkommens eines Begriffes im Text und die Anzahl der Begriffe, die den Begriff enthalten werden miteinander in Beziehung gesetzt und dadurch die Relevanz der Begriffe festgestellt (Lossio-Ventura et al. 2013: 46f.).

Zu weiteren Maßen, die auf Häufigkeiten basieren, gehören z.B. Average TF (ATF), Residual Inverse Document Frequency (RIDF), C-Value, Basic, ComboBasic etc. (Näheres in Astrakhantsev 2018: 858f.).

2.4.2.2 Methoden basierend auf Kontextinformationen über das Auftreten der Termini

Die Methoden dieser Gruppe gehen von der Annahme aus, dass die Kontexte, in denen Termini und allgemeine Wörter vorkommen, unterschiedlich sind. Laut Astrakhantsev et al. (2015) ist die grundlegende Methode dieser Gruppe der sog. NC-Wert (Englisch NC-Value), der für die Relevanz und die Kohärenz eines Terminus innerhalb eines Textkorpus steht. Die Berechnung dessen, ob sich um einen Terminus oder Nicht-Terminus handelt, erfolgt in drei Schritten. Zuerst werden die besten 200 Termkandidaten unter anderem mithilfe von C-Wert (siehe Abschnitt 2.4.2.1) erkannt. Im zweiten Schritt werden Gewichte für Kontextwörter berechnet, basierend auf der Anzahl der Termkandidaten, in deren Kontext das Wort vorkommt, im Verhältnis zu ihrer Gesamtzahl im Korpus. Im dritten Schritt wird der endgültige Wert eines Termkandidaten berechnet, indem die Frequenz von Kontextwörtern im Kontext des Termkandidaten und deren Gewichte berücksichtigt werden (Astrakhantsev et al. 2015: 341f.). Einfach ausgedrückt basiert der NC-Wert auf der Häufigkeit, mit der ein Terminus im Verhältnis zur Gesamtzahl

der Wörter im Korpus auftritt. Eine weitere Methode, die durch eine Modifizierung des NC-Wertes entstand, ist die sog. DomainCoherence, die durch spezifische Einschränkungen des Kontextwortes, wie z.B. das Kontextwort muss ein Nomen, Verb oder Adjektiv sein, hervorgegangen ist (Näheres siehe Astrakhantsev et al. 2015: 341). Anschließend erwähnen Astrakhantsev et al. (2015) die sog. PostRankDC-Methode, die beide oben erwähnten Methoden miteinander kombiniert und dadurch die besseren Ergebnisse als jede Methode einzeln erzielt. PostRankDC zielt darauf ab, die Vorteile beider Ansätze zu nutzen, um eine verbesserte Bewertung der Relevanz und Kohärenz von Begriffen innerhalb einer bestimmten Domäne zu erreichen (Näheres siehe Bordea et al. 2013).

2.4.2.3 Methoden basierend auf Themenmodellierung

Diese Methodengruppe basiert auf der Themenmodellierung, d. h. auf einer statistischen Methode zur Erkennung von Themen aus großen Textkorpora, indem eine Extraktion von semantisch verwandten Wörtern und die Gruppierung von Dokumenten nach thematischer Ähnlichkeit durchgeführt wird. Wie von Astrakhantsev (2018: 860) betont wird, führen diese Methoden zu weniger falschen oder unnötigen Vorschlägen an Termkandidaten, was einen enormen Vorteil gegenüber z.B. häufigkeitsbasierten Methoden darstellt. Wenn die Rede von Methoden dieser Gruppe ist, ist meistens damit eine konkrete Methode gemeint, die sog. Novel Topic Model-Methode, die von Li et al. (2013) entwickelt wurde und auch eine Extraktion der längeren Termini ermöglicht. Sie basiert auf der Identifizierung der Themen im Korpus und auf der anschließenden Zuordnung jedes Termkandidaten zum jeweiligen Thema (Astrakhantsev et al. 2015; Li et al. 2013).

2.4.2.4 Methoden basierend auf externen Korpora

Diese Gruppe umfasst mehrere Methoden, die gemeinsam haben, dass zur Feststellung, ob es sich um einen Terminus oder Nicht-Terminus handelt, ein externer Referenzkorpus herangezogen wird, der in den meisten Fällen nicht fachspezifisch ist (Astrakhantsev et al. 2015: 860). Je nachdem, in welchem der beiden (fachspezifischen oder allgemeinen) Korpora der Begriff häufiger auftritt, handelt es sich entweder um einen Terminus oder Nicht-Terminus. In diese Gruppe fassen Astrakhantsev et al. (2015) Methoden wie DomainPertinence, Weirdness und Relevance zusammen. Unter DomainPertinence versteht sich die einfachste Methode, die auf

dem Vergleich der Häufigkeit von Begriffen in spezifischen Fachbereichstexten mit der Häufigkeit in allgemeinen Referenzkorpora basiert, wobei die Begriffe mit höherer relativer Häufigkeit im Fachbereich als relevant betrachtet werden (ibid.: 342). Mit der Weirness-Methode werden auch die Größe der Textsammlung und seltene Begriffe berücksichtigt und mit der Relevance-Methode wird die Relevanz eines Begriffes in Bezug auf den gesamten Text ermittelt, wobei die Begriffe bevorzugt werden, die oft im bestimmten Fachbereich vorkommen und zentral für das Verständnis sind (ibid.: 342f.).

In dieser Methodengruppe spielen Korpora bzw. Referenzkorpora, auch Vergleichskorpora genannt, eine wichtige Rolle. Man unterscheidet zwischen spezifischen, wie z.B. Annotated Corpora for Term Extraction Research (ACTER) von Rigouts Terryn (2021a), oder allgemeinen Referenzkorpora z.B. Wörterbücher. Ein konkretes Beispiel eines allgemeinen Referenzkorpus findet sich bei Astrakhansev (2018), der für sein ATR4S-Projekt den Korpus für historisches Englisch (English Corpus of Historical American English) verwendet. Zwar wurden im Abschnitt 2.4 bereits die Qualitätsanforderungen an Korpora im Allgemeinen beschrieben, jedoch wurden die Arten der Korpora, die für die ATE erforderlich sind, nicht näher erläutert. Bis jetzt wurden in dieser Arbeit Korpora aus der Sicht des für die Extraktion benötigten Ausgangsmaterials beschrieben. Allerdings können Korpora in der ATE auch zu anderen Zwecken eingesetzt werden, z.B. als allgemeine Referenzkorpora für bestimmte Methoden (siehe oben), als Trainingsmaterial für auf neuronalen Ansätzen basierenden Systeme (Rigouts Terryn et al. 2022) oder bei der Evaluation als sog. Goldstandards (Kim et al. 2003; Rigouts Terryn et al. 2020). Unter Goldstandard verstehen Drewer und Schmitz (2017: 42) „Ideal- oder Musterlösungen“ anhand denen die Qualität von der ATE-Systemen gemessen werden kann.

Tran et al. (2023) unterteilen die wichtigsten Korpora je nach Anzahl der Sprachen in monolinguale und multilinguale, wobei in der monolingualen Kategorie nur Korpora berücksichtigt werden, die bestimmte Kriterien, wie z.B. das Vorhandensein der Annotationen oder Status von Goldstandard u.v.a., erfüllen (Näheres siehe Tran et al. 2023: 5). Einer der bekanntesten monolingualen Korpora in der ATE ist laut Tran et al. (2023) GENIA (Kim et al. 2003), der aus englischen medizinischen Artikeln aus dem Fachgebiet Molekularbiologie erstellt wurde. Einer der ersten multilingualen Korpora, der nicht nur multiple Sprachen, konkret Englisch, Französisch, Niederländisch, sondern auch unterschiedliche Fachgebiete, konkret Korruption, Dressur, Herzinsuffizienz, Windenergie, erfasst, gehört der oben kurz erwähnte ACTER, der aus der heutigen ATE-Forschung nicht mehr wegzudenken ist (Rigouts Terryn et al. 2020). Diese unterstreichen einige Forschungsarbeiten wie die von Rigouts Terryn et al. (2022), die die Daten aus ACTER verwendet haben, um das System D-Terminer (Rigouts Terryn et al.

2022) zu trainieren und zu evaluieren oder von Wachowiak et al. (2021), die unter anderem auch ACTER für das Trainieren von Transformer-basierten Systemen (siehe Abschnitt 2.4.2.9) verwendet haben.

2.4.2.5 Methoden basierend auf Suchmaschinen

Astrakhantsev et al. (2015) haben sich verschiedene experimentelle Studien, in denen Methoden basierend auf Suchmaschinen angewendet wurden, angesehen. Dabei heben sie die Studien von Golomazov (2012) und Dobrov und Loukachevitch (2011) hervor, die sich mit der Extraktion von Zweiworttermen befassen. Die Grundlage der Methoden bilden Suchmaschinen, die auf bestimmte Suchanfragen Ergebnisse liefern. Die Methode von Golomazov (2012) basiert auf speziellen Suchanfragen an Suchmaschinen und klaren Bedingungen für die Relevanz eines Begriffs. Die Gültigkeit eines Terminus hängt von der Anzahl der auf Suchanfrage zurückgegebenen Seiten und festgelegten Parametern ab, wobei mindestens eine der festgelegten Bedingungen erfüllt sein muss. Diese Methode bewertet die Wichtigkeit und Spezifität eines Terminus im Kontext seines Fachgebiets oder Konzepts und sichert so die Auswahl relevanter Termini (Astrakhantsev et al. 2015; Golomazov 2012). Im Gegensatz zur Betrachtung der Häufigkeit des Auftretens von Termkandidaten im gesamten Web verwenden Dobrov und Loukachevitch (2011: 712) kurze Textfragmente, sog. „Snippets“, die von Suchmaschinen als Ergebnisse einer Suchanfrage zurückgegeben werden und aus gesamten Termini sowie einzelnen Wörtern des Terminus bestehen. Anschließend werden die Frequenzen der Termkandidaten in Snippets, die Anzahl vorab bestimmter für das Fachgebiet charakteristischen Wörter, die Anzahl von Wörtern, die häufig in Wörterbuchdefinitionen vorkommen, die Verwandtschaft der Snippets, die für den gesamten Kandidaten und für seine einzelnen Wörter erhalten werden und so weiter analysiert (Näheres siehe Dobrov und Loukachevitch 2011: 712f.). Solche umfassenden Analysen ermöglichen eine tiefere Einsicht in die Relevanz und die Bedeutung von Termini innerhalb eines spezifischen Fachgebiets, basierend auf ihrer Präsenz und Darstellung im Internet.

2.4.2.6 Methoden basierend auf Ontologien

Weitere Ressourcen, die in der automatischen Termextraktion, obwohl seltener als externe Ressourcen (siehe Abschnitt 2.4.2.4), eingesetzt werden, sind Ontologien. Laut Drewer und Schmitz (2017: 8f.) versteht man unter Ontologien ein spezifische graphisches Begriffssystem,

das das Wissen über formal geordnete Konzepte und deren Beziehungen in einem bestimmten Fachbereich erfasst. Die Gründe, warum Ontologien selten in der ATE angewendet werden, bestehen laut Astrakhantsev et al. (2015: 343) darin, dass die allgemeinen Ontologien Fachgebiete nur unzureichend abdecken, d. h. sie enthalten nur die grundlegendsten Termini, während die fachspezifischen Ontologien nur für wenige Fachbereiche verfügbar sind. Als erwähnenswerte Forschungsarbeit, die auf solchen Methoden basiert, nennen Astrakhantsev et al. (2015) die Arbeit von Dobrov und Loukachevitch (2011), die eine spezifische Art von Begriffssystem, den sog. Thesaurus, für die Informationssuche verwenden. Ziel dieser Untersuchung war festzustellen, ob man mithilfe bekannter Termini aus einem Thesaurus neue, unbekannte Termini vorhersagen kann, wobei bestimmte Eigenschaften wie Ähnlichkeit zu einem bekannten oder unbekannten Terminus aus dem Thesaurus oder die sog. „Completeness of Description“, d. h. wie viele Synonyme und Beziehungen von einzelnen Wörtern einer Wortkombination es im Thesaurus gibt, in Betracht gezogen werden (für Näheres siehe Dobrov und Loukachevitch 2011: 713f.).

2.4.2.7 Methoden basierend auf Wikipedia

Alle Methoden, die nicht auf externe Ressourcen zurückgreifen, wie C-Value (siehe Abschnitt 2.4.2.1) oder die Novel Topic Model-Methode (siehe Abschnitt 2.4.2.3), erfordern eine große Sammlung an Textdokumenten, in denen der Termkandidat oftmals vorkommt, um diesen verlässlich zu identifizieren. Um dieses Problem zu überwinden, benötigt es externe Ressourcen, die laut Astrakhantsev (2018) zwei Anforderungen erfüllen müssen. Erstens müssen sie spezifisch genug sein, um fachspezifische Informationen zu enthalten, die zur Unterscheidung zwischen Termini und Nicht-Termini erforderlich sind und zweitens müssen sie allgemein genug sein, um in der Praxis auf viele Fachgebiete anwendbar zu sein (ibid.: 861). Wikipedia erfüllt diese Anforderungen, da diese Enzyklopädie viele Fachgebiete umfasst und mehrsprachig ist.

Eine Methode, die Astrakhantsev et al. (2015) hervorheben, ist die LinkProbability-Methode, die Termkandidaten bewertet, die spezifisch genug sind, um mit einem Hyperlink versehen zu werden. Die Link Probability Formel (3) stellt dar, wie die Häufigkeit des Vorkommens von Termkandidat in einem Hyperlink $H(t)$ mit der Gesamtzahl des Vorkommens von Termkandidat auf den Wikipedia-Seiten $W(t)$ miteinander in Beziehung gesetzt werden, wobei der Parameter T benötigt wird, um zu kleine Werte herauszufiltern bzw. Noise-Effekte zu verringern.

$$LinkProb_T(t) = \begin{cases} 0 & \text{if } t \text{ is not contained} \\ & \text{in Wikipedia or } \frac{H(t)}{W(t)} < T \\ \frac{H(t)}{W(t)} & \text{otherwise,} \end{cases} \quad (3)$$

Mithilfe dieser Methode können Termini von allgemeinen Wörtern und Wortverbindungen unterschieden werden, allerdings nicht von Termini aus anderen Fachgebieten.

Wie Astrakhantsev (2018) betont, umfasst diese Gruppe noch andere unterschiedliche Methoden, die jedoch schlechtere Ergebnisse liefern als eine der effektivsten Methoden mitsamt ihren Modifikationen (für Näheres siehe Astrakhantsev 2018: 862f.), die Key Concept Relatedness-Methode. Mithilfe dieser Methode werden Termini aus einer Textsammlung extrahiert, wobei Schlüsselkonzepte identifiziert und ihre Beziehungen zueinander analysiert werden, d. h. dass diese Methode auf der semantischen Verwandtschaft zwischen Begriffen basiert (Astrakhantsev 2018: 861).

2.4.2.8 Methoden basierend auf nicht neuronalem maschinellem Lernen

Abschließend gibt es noch eine Methodengruppe zu nennen, die nach der Aufteilung von Astrakhantsev et al. (2015: 345) dem Ranking-Verfahren (siehe Abschnitt 2.4.) zugrunde liegt und die sie als „Methods of Feature-Based Inference“ bezeichnen. Astrakhantsev et al. (2015) nennen im Zusammenhang mit dieser Methodengruppe mehrere Methoden, die näher betrachtet manche oben erwähnten ATE-Methoden als einzelne Merkmale in einem komplexen System anwenden, wie z.B. die sog. PU-ATR-Methode, die unter anderen z.B. den NC-Wert berücksichtigt (siehe Abschnitt 2.4.2.2). Die Merkmale werden laut Astrakhantsev et al. (2015: 345) über eine der drei folgenden Arten ins System implementiert: erstens über eine lineare Kombination der Merkmale wie bei PostRankDC (siehe Abschnitt 2.4.2.2), zweitens über einen Algorithmus, d. h. der Algorithmus wird erweitert, oder drittens über überwachtes maschinelles Lernen (2.4.1.2). Die ersten Ansätze des maschinellen Lernens in der ATE waren laut Tran et al. (2023) als Ergänzungsmethoden für die erfolgreichen traditionellen Systeme gedacht, d. h. verschiedene statistische und linguistische Methoden bzw. Merkmale wurden in verschiedene Algorithmen, z.B. in den Entscheidungsbaum-, Support-Vector-Machine-Algorithmus etc.,

eingespeist (Für Näheres zu überwachten Lernalgorithmen siehe Mahesh 2018). Die meisten Systeme, die auf überwachtem maschinellen Lernen basieren, befolgen laut Tran et al. (2023) drei Schritten: 1. den Vorverarbeitungsschritt, in dem die Daten für die Weiterverwendung in einem Modell vorbereitet werden, wie z.B. durch die Wortartmarkierung, 2. den Merkmalerstellungsschritt, d. h. aus vorhandenen Daten werden Merkmale erstellt und 3. den Klassifikationsschritt, d. h. die Klassifikatoren werden mit den erstellten Merkmalen aus den Trainingsdaten trainiert, um Vorhersagen für neue Daten treffen zu können (ibid.: 9). In weiterer Folge kann das Gelernte auf unbekannte Daten angewandt werden.

2.4.2.9 Methoden basierend auf neuronalen Netzen bzw. Deep Learning

Die Verwendung neuronaler Netzwerke, vor allem von neuronalen Sprachmodellen, zur ATE ist in den letzten Jahren enorm gestiegen. Sie werden entweder genutzt, um Textinformationen mittels Wortembeddings zu repräsentieren, wie GloVe Wortembeddings (Pennington et al. 2014), oder um durchgängige tiefe Architekturen als Klassifikatoren zu verwenden (wie bei Rigouts Terryn et al. 2022; Wachowiak et al. 2021), wobei vor allem die Modelle basierend auf Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al. 2018) die ATE grundlegend revolutioniert hat. Unter BERT versteht sich ein auf der Transformer-Architektur basiertes Modell, das darauf trainiert ist, den Kontext von Wörtern in Texten zu kennen, wobei im Gegensatz zu früheren Modellen, die Text sequentiell verarbeiten, BERT den gesamten Textblock gleichzeitig, d. h. es ist bidirektional, betrachtet (ibid.).

In Bezug auf ATE-Klassifikatoren werden neuronale Netzwerke als sog. End-to-End-Termextraktoren verwendet, wobei Tran et al. (2023: 11f.) drei Haupttypen hervorheben: Sequenzklassifikatoren, Token-Klassifikatoren und Sequenz-zu-Sequenz-Generierungsmodelle. Die Sequenzklassifikatoren identifizieren relevante Begriffe oder Phrasen auf der Ebene ganzer Sätze oder Abschnitte, indem sie den Kontext zur Klassifizierung nutzen. Die Token-Klassifikatoren klassifizieren jedes einzelne Wort oder Token in einem Text, um festzustellen, ob es Teil eines Fachbegriffs ist und die Sequenz-zu-Sequenz-Generierungsmodelle wandeln Texteingaben direkt in eine Liste relevanter Fachbegriffe um, indem sie den gesamten Text als Grundlage für die Generierung verwenden (Näheres siehe Tran et al. 2023).

Anfänglich wurden laut Tran et al. (2023) die neuronalen Modelle als Sequenz- oder Binärklassifikationsaufgabe eingesetzt, allerdings wird heutzutage der Fokus auf Sprachmodelle und sequenzielle Methoden gelegt, bei denen Termkandidaten in ihrem ursprünglichen Kontext erkannt werden, indem jedem Token in der Textsequenz eine (Teil-eines-Terms oder

Nicht-Term) Markierung zugewiesen wird. Die jüngsten Untersuchungen, z.B. von Wachowiak et al. (2021) zeigen, dass Ansätze, die vortrainierte Sprachmodelle als Token-Klassifikatoren verwenden, besonders effizient sind.

Wie aus dieser kurzen Einleitung der neuronalen Ansätze in der ATE hervorgeht, gibt es heutzutage etliche Modelle bzw. Methoden, die Tran et al. (2023) versuchen zu bewerten, vergleichen und kategorisieren. Dabei liegt der Fokus insbesondere auf den Deep-Learning-Methoden, vor allem auf Sprachmodellen, d. h. Computerprogrammen, die trainiert wurden, um menschliche Sprache zu „verstehen“, indem durch verschiedene Muster und Zusammenhänge erkannt werden. Das zeigt beispielsweise das Sprachmodell Cross-lingual Language Modeling-RoBERTa (X-LMR), das z.B. bei Wachowiak et al. (2021) zum Einsatz kommt.

Abb. 12 bietet einen Überblick der ausgewählten Methoden von Tran et al. (2023), wobei bei dieser Unterteilung das Kriterium die Art und Weise, wie die Systeme mit Sprachen im Trainingskorpus umgehen, die ausschlaggebende Rolle spielt. Daraus folgend gibt es monolinguale, multilinguale und cross-linguale Methoden bzw. Modelle. Die monolingualen Deep-Learning-Modelle zeichnen sich dadurch aus, dass sie ausschließlich auf Daten einer einzelnen Sprache trainiert werden und Vorhersagen für neue, nicht gesehene Daten derselben Sprache generieren. Im Gegensatz dazu werden multilinguale Deep-Learning-Modelle auf mehreren Sprachen gleichzeitig feinabgestimmt, indem sie auf einer Kombination von Daten aus verschiedenen Sprachen trainiert werden. Diese Modelle generieren Vorhersagen für neue, nicht gesehene Daten aus allen trainierten Sprachen. Die cross-lingualen Deep-Learning-Modelle hingegen werden auf einer oder mehreren Sprachen trainiert, generieren jedoch Vorhersagen für eine neue Sprache, die während des Trainings nicht im Korpus vertreten war.

Methods	ACTER					
	ANN			NES		
Non-neural models	EN	FR	NL	EN	FR	NL
RACAI [87]	39.3	-	-	41.3	-	-
e-Terminology [87]	21.4	20.6	15.3	20.1	19.7	14.4
NYU [87]	31.5	-	-	30.6	-	-
HAMLET [88]	54.2	60.2	66.1	55.4	60.8	66.0
Feature-based XGBoost [45]	-	-	-	33.6	50.9	34.1
NMF [78]	33.5	30.9	30.1	33.7	30.7	30.3
Monolingual deep learning models						
TALN-LS2N [44]	45.0	45.9	-	46.7	48.1	-
NLPLab UQAM [59]	17.8	12.9	18.6	18.1	13.2	18.7
XLM-R sequence Classifier [58]	-	-	-	45.2	46.0	48.5
XLM-R Token Classifier [58]	-	-	-	58.3	52.9	69.6
mBART NMT [58]	-	-	-	53.2	55.9	65.2
Vanilla-biLSTM-CRF [45]	-	-	-	8.17	6.53	7.50
BERT [45]	-	-	-	48.2	-	-
CamemBERT [45]	-	-	-	-	51.1	-
BERT (NER) [45]	-	-	-	37.4	-	51.0
CamemBERT (NER) [45]	-	-	-	-	51.1	-
BERT-biLSTM-CRF [45]	-	-	-	29.5	25.6	27.4
mBERT [45]	-	-	-	45.8	47.2	-
ALBERT [96]	49.9	-	-	54.5	-	-
DistilBERT [96]	46.5	-	-	53.9	-	-
ELECTRA [96]	51.2	-	-	52.2	-	-
RoBERTa [96]	54.3	-	-	59.1	-	-
XLNet [96]	55.2	-	-	57.8	-	-
FlauBERTa [96]	-	38.9	-	-	51.4	-
Dutch BERT [96]	-	-	65.6	-	-	66.8
robBERTa [96]	-	-	48.2	-	-	62.2
robBERTa-v2 [96]	-	-	48.3	-	-	63.4
mInfoXLM [96]	56.1	50.4	68.5	57.6	58.0	68.6
mBERT [96]	43.0	48.4	66.3	55.0	57.4	67.8
mDistilBERT [96]	49.8	51.9	65.3	55.6	55.8	67.8
Multilingual deep learning models						
XLM-R Sequence Classifier [58]	-	-	-	46.0	46.7	56.0
XLM-R Token Classifier [58]	-	-	-	56.2	55.3	67.8
mBART NMT [58]	-	-	-	55.3	57.6	64.9
mBERT [45]	-	-	-	45.4	44.9	51.0
Cross-lingual deep learning models						
XLM-R Sequence Classifier [58]	-	-	-	44.7	48.1	58.0
XLM-R Token Classifier [58]	-	-	-	58.3	57.6	69.8
mBART NMT [58]	-	-	-	55.2	57.4	59.6
XLM-R Token Classifier [95]	56.0	59.9	68.5	58.0	61.1	68.3

Abbildung 12: Unterteilung der neuesten Methoden nach Art und Weise, wie die Systeme mit Sprachen im Trainingskorpus umgehen (Tran et al. 2023: 18)

Die in der Abb. 12 angeführten Systeme sind komplex und daher wurde eine detaillierte Beschreibung jedes einzelnen Systems den Rahmen dieser Arbeit übersteigen. Deshalb wird auf sie verzichtet, allerdings wird im Abschnitt 4.3 auf die einzelnen Tools, die für diese Arbeit relevant sind, kurz eingegangen.

2.4.3 Vor- und Nachteile der ATE

Die ATE hat einige Vorteile, die Bakaric et al. (2021) in ihren Beitrag kurz ansprechen. Als Vergleichsgröße ziehen sie dabei die manuelle Terminologierecherche, die im Idealfall so aussieht, dass zuerst eine Kandidatenliste von Terminolog*innen erstellt und diese dann mit Spezialist*innen von jeweiligem Fachgebiet überprüft wird (ibid.: 18), d. h. dass hier zwangsläufig Fragen geklärt werden müssen, z.B. wie viel Zeit die Terminolog*innen für das gründliche Lesen und Identifizieren von Termini des jeweiligen Korpus benötigen u.v.m. Die Klärung solcher Fragen nimmt viel Zeit und zusätzliche Personalressourcen in Anspruch, was bei einer ATE nicht in solchem Ausmaß der Fall ist und dadurch dieses Verfahren in der Phase der Extraktion zeitsparender ist (Bakaric et al. 2021: 18). Heylen und De Hertog (2015: 203) bezeichnen die ATE als ersten Filter, der den Anwender*innen eine extensive Untersuchung des ganzen Textes bzw. Textsammlung erspart.

Ein weiterer Vorteil ist die Objektivität, die durch eine Computeranalyse gewährleistet ist, wogegen man bei einer manuellen Terminologierecherche mit einem gewissen Risiko des subjektiven menschlichen Urteilens und des menschlichen Versagens rechnen muss (Heylen & De Hertog 2015: 203).

Einen der weiteren Vorteile der automatischen Termextraktion sprechen Drewer und Schmitz (2017: 35) an, die betonen, dass am Ende der Termextraktion Kandidatenlisten vorliegen sollten, die idealerweise mit Metadaten versehen sein sollen. Zu solchen zusätzlichen Informationen gehören z. B. Häufigkeit des Vorkommens eines Terms im Korpus oder Kontext- bzw. Verwendungsinformationen.

Zu erwähnen ist auch die Möglichkeit mancher Tools die Ergebnisse zu verfeinern, indem sog. Stoppwortlisten definiert werden, z.B. die Artikel werden ausgeklammert, oder indem durch das Vergleichen mit den vorhandenen Einträgen die Wiederholungen nicht angezeigt werden (Drewer & Schmitz 2017: 46).

Wo die ATE ihre Schwächen hat, ist die Identifikation der Termkandidaten auf der semantischen und extralinguistischen Ebene, z.B. Formeln und Symbole, wogegen ein Mensch diese Ebenen wohl in Betracht ziehen kann (Bakaric et al. 2021: 18). Der menschliche Faktor ist auch nicht bei der finalen Überprüfung der Ergebnisse auszuschließen und daher betrachtet man die extrahierten Termini als vorläufige Kandidaten, die noch überprüft werden müssen (ibid.). Drewer und Schmitz (2017) weisen auf die „unerwünschten Nebeneffekte“ (ibid.: 44) der ATE hin, und zwar unter der Bezeichnung „Silence“, d. h. dass relevante Kandidaten teilweise oder gar nicht erkannt werden, und „Noise“, d. h. dass Kandidaten generiert werden, die

keine Termini sind. Nicht alle verfügbaren Tools haben die Einstellungsmöglichkeit über einen Qualitätsfilter diese Nebeneffekte zu minimieren (Drewer & Schmitz 2017: 48f.).

Man muss anmerken, dass die automatische Termextraktion von der Entwicklung von tiefen neuronalen Netzwerken und der Transformer-basierter Modelle erheblich profitiert, allerdings gibt es noch gewisse Nachteile, die Tran et al. (Tran et al. 2023) eher als Herausforderungen ansehen, die es auf dem Weg zur völlig automatischen Extraktion zu bewältigen gilt. Diese umfassen unter anderen die Erkennung der verschachtelten Termini, d. h. dass der Klassifikator aus dem Kontext erschließen muss, ob ein bestimmter Terminus Teil eines längeren Terminus ist oder nicht, oder die sog. Generalisierbarkeit, d. h. dass Modelle die Fähigkeit haben müssen, Termini aus neuen Sprachen und Domänen zu extrahieren, für die keine spezifischen Trainingsdaten vorhanden sind (Für weitere Herausforderungen siehe Tran et al. 2023: 19f.).

2.4.4 Anwendungsbereiche

Wie Wissenschaftler*innen (Hernani Costa et al. 2016; Rokas et al. 2020) in ihren Studien stets betonen, tragen die ATE-Tools zur Entwicklung von Terminologiedatenbanken und Ontologien bei, die wiederum zu einer unverzichtbaren Ressource im Bildungs-, Forschungs- und Wirtschaftssektor geworden sind. Insbesondere haben sich die ATE-Tools bei den Aufgaben, bei denen die Verarbeitung der natürlichen Sprache im Vordergrund steht, bewährt, und sind für viele Berufsgruppen wie Translator*innen, Terminolog*innen und technische Redakteur*innen ein festes Werkzeug in ihren Arbeitsabläufen geworden.

Ein konkretes Beispiel kommt aus der Translationsindustrie, in der die ATE-Tools bei der Vorbereitung auf Übersetzungs- und Dolmetschprojekte zum Einsatz kommen, bei denen sehr spezifische Themen eine intensive Terminologievorbereitung benötigen, vor allem wenn es sich um Projekte handelt, an denen mehrere Übersetzer*innen mitwirken und bei denen die Konsistenz gewährleistet werden muss (Bakaric et al. 2021: 19). Heylen und De Hertog (2015) und Bakaric et al. (2021) fügen hinzu, dass die ATE-Tools eine wesentliche Rolle bei der (Weiter-)Entwicklung und der Verbesserung der bestehenden Übersetzungsprogramme spielen (ibid.: 19).

Ein weiterer Bereich, in dem die ATE-Tools ihre Anwendung finden, ist der Wirtschaftsbereich. Für immer mehr Unternehmen ist durch ihr globales Auftreten die interne sowie externe Kommunikation, die im Idealfall einheitlich sein muss, der Schlüssel zum Erfolg. Um diesen Erfolg zu erzielen, setzen viele Unternehmen auf ein professionelles

Terminologiemanagement und daher auch auf automatische Termextraktion, die dem Unternehmen viel Zeit und Geld ersparen kann. Drewer und Schmitz (2017: 1f.) nennen etliche positiven Effekte, die von Konsistenz der Unternehmensdokumentation bis zum Wissenstransfer im Unternehmen reichen.

Zu erwähnen ist auch der Bereich der Terminographie, ein Teil der Lexikographie, in dem die Erstellung der Wörterbücher und Terminologiedatenbanken im Vordergrund steht, wobei die Termkandidatenlisten als deren Basis dienen (Heylen & De Hertog 2015: 204).

Die erwähnten Anwendungsbeispiele bilden keine vollständige Liste der Anwendungsmöglichkeiten der ATE, allerdings sollen diese Beispiele verdeutlichen, dass die ATE in vielen Bereichen, nicht nur in der Terminologiewissenschaft selbst, einzusetzen ist.

2.5 Evaluation in der ATE

Die Evaluation von Modellen für die ATE ist aus unterschiedlichen Gründen äußerst problematisch, wie in einigen Studien explizit hervorgehoben wird. Zum Beispiel wird in Jerdhaf et al. (2022) die Evaluation als „notoriously difficult“ (ebd.: 30) und in Hazem et al. (2020) als „very difficult“ (ibid.: 95) bezeichnet.

Eine dieser Herausforderungen besteht darin, dass es keinen Konsensus darüber gibt, was genau unter dem Begriff Terminus zu verstehen ist und welche Charakteristika ein Terminus haben sollte (siehe Abschnitt 2.1.4 und 2.2.3). Dieses Problem wird laut Rigouts Terryn et al. (2020) auf zwei Ebenen deutlich: zum einen auf der praktischen Ebene, die Aspekte wie Länge oder Wortartmuster der Termini umfasst, und zum anderen auf der theoretischen Ebene, auf der die Unterscheidung zwischen einem Terminus und einem einfachen Wort thematisiert werden muss (ibid.: 85). Konkret bedeutet es, dass Modelle Termkandidaten extrahieren können, die irrelevant sind und wiederum relevante Kandidaten nicht bemerkt werden.

Zu der weiteren Herausforderung zählt die Unvorhersehbarkeit vieler maschineller Ansätze, die sich vor allem in der Inkonsistenz der Ergebnisse äußert, was laut Rigouts Terryn et al. (2020) ein wesentliches Hindernis für die praktische Nutzung darstellt, da die Kriterien Konsistenz, Vorhersehbarkeit und Transparenz die Nutzbarkeit der Tools in der Praxis bestimmen sollten (ibid.: 90).

Ein weiteres Problem bei der Bewertung der Qualität von ATE-Modellen besteht in der Abwesenheit klar abgegrenzter Fachgebiete oder Teilfachgebiete und in fehlenden Goldstandards (siehe Abschnitt 2.4.2.4). Ebenso spielt die Art und Weise, wie die jeweiligen Goldstandards erstellt werden, eine entscheidende Rolle bei der Bewertung. Zum Beispiel ist es ist

wichtig zu unterscheiden, ob es sich um allgemeine oder fachgebietsspezifische Standards handelt, da dies die Qualität der Ergebnisse unter bestimmten Bedingungen beeinflussen kann. Jerdhaf et al. (2022) zeigen auf demselben Testdatensatz, welche Einwirkungen auf die Ergebnisse ein Sprachmodell hat, der einmal mit einem fachspezifischen und einmal mit einem allgemeinen Datensatz trainiert wurde (ibid.: 30f.).

Warum man viele Studien auch nicht direkt miteinander vergleichen kann, hängt auch von dem Umfang der Kandidatenlisten, der mit dem Goldstandard verglichen wird, ab. Laut Tran et al. (2023) wird in einigen Studien dieselbe Bewertung wie bei der Keyword-Extraktion angewandt, indem die besten k Kandidatenbegriffe für die Auswertung gewählt werden, z.B. die Top-100-Kandidatenbegriffe (siehe Abb. 13).

Diese Herausforderungen unterstreichen die Notwendigkeit kontinuierlicher Forschung, um effektive und zuverlässige Bewertungsmethoden für die ATE zu entwickeln, wie es Rigouts Terryn et al. (2020) im Rahmen des gemeinsamen Wettbewerbs TermEval 2020 versuchen, indem verschiedene Teams mit verschiedenen Modellen unter Verwendung derselben Daten und desselben Evaluationsrahmen direkt verglichen werden. Dieser Wettbewerb führte zu dem Ergebnis, dass es aufgrund der Unvorhersehbarkeit von Modellen, die auf komplexen neuronalen Netzwerken basieren, und aufgrund der stark variierenden Kandidatenlisten, neue Evaluationsmethoden und Metriken (siehe Abschnitte 2.5.1 und 2.5.2) für zukünftigen Studien bedarf (ibid.: 93).

2.5.1 Evaluationsmethoden

Es gibt verschiedene Methoden (siehe Abb. 13), wie die Leistung der ATE-Modelle geprüft und bewertet werden kann. In Bakaric et al. (2021) ist die Rede von zwei Arten, den direkten (intrinsischen) und den indirekten (extrinsischen) Evaluationsmethoden.

Bei den ersteren handelt es sich um Methoden, die bestimmte Eigenschaften des Modells bewerten, indem die Bewertung unter Anwendung von zwei unterschiedlichen Techniken, entweder manuell durch Expert*innen oder durch einen Vergleich der Kandidatenlisten mit einem Goldstandard, stattfindet (ibid.:20).

Der Goldstandard ist auch unter der Bezeichnung „Reference Term List“ (ibid.) bekannt und stellt einen von einer Fachexpert*in bzw. Fachexpert*innengruppe manuell erstellten Referenzkorpus dar, der als eine Vergleichsgrundlage für die Feststellung der Qualität der Ergebnisse dient und der in verschiedenen Formen vorliegen kann. Tran et al. (2023) unterscheiden zwischen a) einem bereits existierenden und weit verbreiteten Referenzkorpus wie GENIA

(Kim et al. 2003) und ACTER (Rigouts Terry et al. 2020), auf die man frei zugreifen kann, b) einem Referenzkorpus, der aus sog. „seed terms“, d. h. fachspezifischen Schlüsselwörter, besteht und c) einer vollständigen Termkandidatenliste, die direkt aus den Korpusdaten annotiert und zum Maßstab für die Bewertung wird (ibid.: 15).

In Tran et al. (2023) ist zusätzlich zu den zwei oben erwähnten Techniken noch die wörterbuchbasierte Bewertung erwähnt, deren Vorteil in der Reduktion des menschlichen Aufwandes besteht. Laut Tran et. al (2023: 15) handelt es sich um einen Abgleich von Wörterbuch mit den extrahierten Termini, wobei das Wörterbuch in Form einer klassischen Liste von Wörtern oder vordefinierten Kriterien sein kann. Der menschliche Aufwand bezieht sich nicht nur auf den zeitlichen (Dauer der Bewertung) und personellen (Fachexpert*in vs. Terminolog*in) Faktor, sondern auch auf den konsensuellen Faktor unter den Expert*innen, d. h. auf dem laut Bakaric et al. (2021: 20) sog. „inter-annotator agreement“, das in zwei Modi auftreten kann, dem sog. strengen Modus, d. h. alle Annotator*innen müssen sich über den Terminus einig sein, und dem nachsichtigen Modus, d. h. eine Annotator*in reicht aus, um einen Kandidaten als Terminus zu bestimmen.

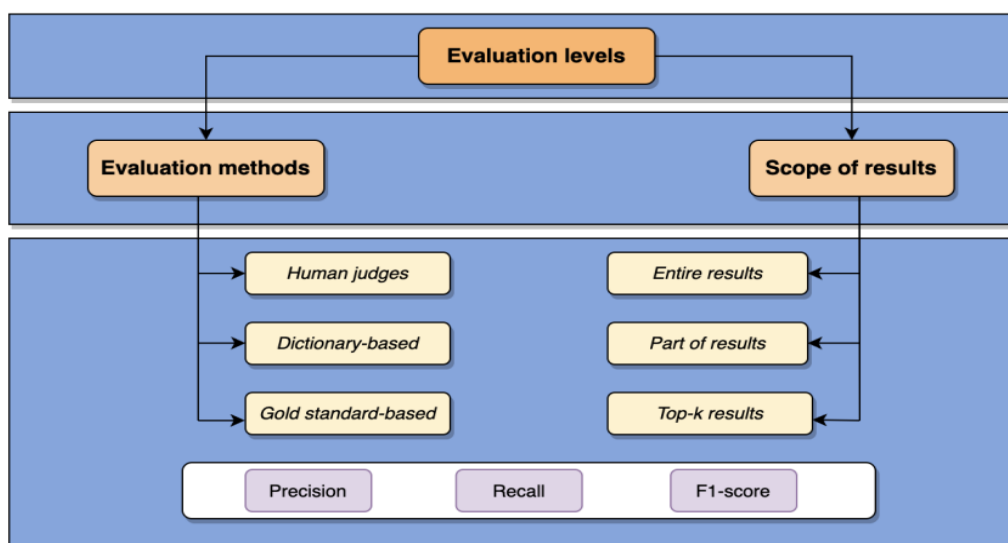


Abbildung 13: Überblick der Evaluationsmethoden in der ATE (Tran et al. 2023: 16)

Bei der indirekten Evaluationsmethoden handelt es sich um Methoden, die die Nützlichkeit der Ergebnisse der Termextraktion in einem anderen System bewerten, d. h. man schaut sich die Verbesserung in diesem System an, z.B. wie sich die Informationssuche in einer Datenbank verbessert hat (Bakaric et al. 2021: 21).

2.5.2 Überblick der gängigen Metriken in der ATE

Die gängigste Methode zur Bewertung der ATE ist der Vergleich der vom jeweiligen System extrahierten Termkandidatenlisten mit einem Goldstandard, indem vor allem drei Metriken bzw. Bewertungskriterien zur Anwendung kommen, und zwar die Genauigkeit (Engl. Precision), d. h. wie viele der extrahierten Kandidaten tatsächlich Termini sind (4), die Trefferquote (Engl. Recall), d. h. wie viele der im Text enthaltenen Termini tatsächlich extrahiert wurden (5), und F1-Score, d. h. das harmonische Mittel zwischen Genauigkeit und Trefferquote (6) (Tran et al. 2023: 16).

$$Precision = \frac{\text{Number of extracted terms}}{\text{All the candidate terms}} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (4)$$

$$Recall = \frac{\text{Number of extracted terms}}{\text{All terms in the corpus}} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negative}} \quad (5)$$

$$F1 - score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (6)$$

Um die explizite Behauptung von Tran et al. (2023: 17) zu überprüfen, dass die meisten ATE-Modelle heutzutage nach allen drei Metriken ausgewertet werden, wurde eine stichprobenartige Überprüfung durchgeführt, bei der Tabelle 1 generiert wurde.

Studie von...	Genauigkeit	Durchschnittliche Genauigkeit	Trefferquote	F1-Score
Wachowiak et al. 2021	Ja	Nein	Ja	Ja
Rigouts Terryn et al. 2020	Ja	Nein	Ja	Ja
Rokas et al. 2020	Ja	Nein	Ja	Ja
Bakaric et al. 2021	Nein	Ja	Ja	Ja
Šajatović et al. 2019	Ja	Ja	Ja	Ja

Tabelle 1: Übersicht der verwendeten Metriken in ausgewählten Studien

Tabelle 1 zeigt nicht nur, dass Genauigkeit, Trefferquote und F1-Score in ausgewählten Studien zum Einsatz gekommen sind, sondern sie zeigt auch, dass in der Studie von Bakaric et al. (2021) und Šajatović et al. (2019) auch die sog. Average Precision, d. h. durchschnittliche Genauigkeit,

angewendet wurde, die dann zum Einsatz kommt, wenn eine Rangliste bzw. die Reihenfolge der Begriffe relevant ist. Sie wird berechnet, indem man die Genauigkeit an verschiedenen Schwellenwerten für die Anzahl der generierten Begriffe berechnet und dann den Durchschnitt dieser Werte nimmt. Zum Vergleich dazu, handelt es sich bei der Genauigkeit nur um die korrekt extrahierten Begriffe, ohne die Reihenfolge zu berücksichtigen (Bakaric et al. 2021: 20; Šajatović et al. 2019: 151).

3 Aktueller Stand der Forschung

Die ATE ist in der Forschung keine Unbekannte mehr, was auch die zahlreichen Forschungsarbeiten wie z.B. von Astrakhantsev (2018), Šajatović et al. (2019), Bakaric et al. (2021) u.v.a., diverse Projekte und Workshops deutlich aufzeigen. Unter den Workshops sind vor allem die folgenden zwei hervorzuheben: der Internationale Workshop COMPUTERM 2020 im Rahmen der internationalen Konferenz LREC 2020 (Daille et al. 2020), wo das gemeinsame Projekt TermEval2020 eingeleitet wurde und dessen Ergebnisse in diversen Beiträgen, unter anderem in Hazem et al. (2020) und Rigouts Terryn et al. (2020), veröffentlicht wurden und der Workshop Terminology in the 21st century: many faces, many places (TERM21), deren Ergebnisse z.B. in Jerdhaf et al. (2022) oder Rigouts Terryn et al. (2022) präsentiert wurden. Trotz intensiver Forschung und zahlreichen Beiträgen zu diesem Thema, stellt man fest, dass die ATE einen enormen Raum und Möglichkeiten für weitere Forschung bietet. Vor allem seit dem Einzug des maschinellen Lernens, spezifischer des Deep Learnings, hat sich die ATE sehr stark verändert. Obwohl es eine Menge an wissenschaftlichen Arbeiten zu diesem Thema gibt, werden hier nur die präsentiert, die ausschlaggebend für die vorliegende Masterarbeit sind, d. h. dieser Abschnitt erhebt keinen Anspruch auf einen vollständigen Überblick aller bestehenden Studien in der ATE.

Wie schon oben erwähnt spielt das Projekt TermEval2020 eine wesentliche Rolle in der Evolution der ATE, dessen Ergebnisse richtungsweisend für die weitere Entwicklungen sind. Um nur ein paar zu erwähnen, geht es bei Hazem et al. (2020) um eine richtungsweisende Studie für die ATE, da hier die Grundlagen für die Anwendung von Transformers-Modellen in der Termextraktion gelegt wurden. In der TermEval-Studie wurden unter anderen ein merkmalsbasierter und ein Deep Learning-basierter Ansatz untersucht und für die Termiextraktion optimiert. Die Ergebnisse zeigen, dass BERT-Modelle (RoBERTa für Englisch und CamemBERT für Französisch) andere Systeme mit der französischen und englischen Sprache übertreffen (ibid.). Eine kurze Vorstellung der teilnehmenden Teams und jeweiligen Ansätzen am Projekt TermEval2020 und Vergleich deren Ergebnisse findet sich in Rigouts Terryn et al. (2020), wobei die Ergebnisse nicht nur das enorme Potenzial von Deep-Learning-Techniken im Bereich der ATE zeigen, sondern auch die Wichtigkeit der großen Datensätze (wie z.B. ACTER) und traditionellen Ansätze, die nach einer Aktualisierung z.B. der Implementierung der Wortembeddings, brauchbare Ergebnisse liefern können.

Als eine der Motivationsquellen für die vorliegende Masterarbeit fungiert die Arbeit von Bakaric et al. (2021), in der die Leistung ausgewählter statistischen und hybriden ATE-

Methoden verglichen wurde. In der Studie wurde ermittelt, welcher Ansatz der ATE unter einem statistischen und zwei hybriden Methoden die höchste Genauigkeit und Trefferquote unter denselben experimentellen Bedingungen erzielt. Die Ergebnisse haben die Annahmen der Forscher*innen bestätigt, dass eine vollautomatische Termextraktion noch nicht realisierbar ist, und die Wahl der Methoden davon abhängen sollte, ob man bei der Verwendung mehr das Augenmerk auf die Genauigkeit oder die Trefferquote legt. Um möglichst den Aufwand für die Translator*innen gering zu halten, zeigt diese Studie, dass die Genauigkeit entscheidender für die Wahl einer Anwendung ist (ibid.).

Drei weitere Studien, die den aktuellen Stand der Forschung widerspiegeln, da sie die neuen Zugänge zur ATE bzw. die neuronalen Ansätze/Methoden, die sich neuronaler Netze und in den meisten Fällen vortrainierter Sprachmodelle bedienen, näherbringen, stammen von Drouin et al. (2020), Jerdhaf et al. (2022) und Rigouts Terryn et al. (2022). Drouin et al. (2020), prüfen eine vielversprechende Methode, die auf Worteinbettungen basiert. von Jerdhaf et al. (2022) stellen Experimente mit Methoden vor, die auf vortrainierten Sprachmodellen wie BERT basieren und Rigouts Terryn et al. (2022) stellen ein neues Tool zur monolingualen und mehrsprachigen ATE aus parallelen Korpora vor, wobei die einsprachige Termextraktion auf einem rekurrenten neuronalen Netzwerk trainiert mit einer überwachten Methodik, die auf vortrainierten Einbettungen beruht, basiert.

Im Rahmen des TERM21-Workshops wurde das Thema monolinguale automatische Termextraktion um weitere Forschungsperspektiven vertieft, wie die Studie von Jerdhaf et al. (2022) zeigt, in der das Thema domänenspezifische vortrainierte Sprachmodelle angesprochen wird. Es wird untersucht, ob generalistisch vortrainierte Sprachmodelle für domänenspezifische Aufgaben wettbewerbsfähige Ergebnisse liefern können, und dadurch viel Zeit und Kosten gespart werden können, da man für domänenspezifische Aufgaben kein spezifisches Sprachmodell braucht (ibid.). Die Ergebnisse zeigen, dass domänenspezifische Modelle unter bestimmten Bedingungen die Genauigkeit bei der Terminologieextraktion verbessern können. Das bedeutet, dass der Einsatz solcher Modelle die Anzahl der irrelevanten Begriffe im Vergleich zu generalistischen Modellen erheblich verringern kann (ibid.: 31).

Nicht nur die monolinguale automatische Termextraktion stand auf der Agenda des Workshops, sondern auch der bis jetzt weniger erforschten bi- resp. multilingualen Termextraktion wurde Platz geboten. Vor allem der Beitrag von Rigouts Terryn et al. (2022) widmet sich der Methodik der multilingualen Termextraktion, ganz konkret dem System D-Terminer, welches auch eine Online-Demo zur Verfügung stellt, dass die bilinguale automatische Termextraktion aus parallelen Korpora mit Paaren dieser Sprachen unterstützt.

Zu erwähnen ist auch eine der aktuellsten Studien von Tran et al. (2023), in der sowohl die klassischen merkmalsbasierten maschinellen Sprachmodelle als auch die neueren auf neuronalen Netzwerken basierenden Modelle untersucht werden und in der ein Versuch unternommen wird, einen umfassenden Überblick über die aktuellen Entwicklungen in der ATE darzubieten wie z.B. ein Überblick über gut annotierte monolinguale und multilinguale Korpora oder eine Kategorisierung der Metriken nach Methodik und Umfang. Die Ergebnisse dieser Studie unterstreichen die überlegene Leistung der neuronalen Modelle im Vergleich zu den klassischen maschinellen Lernansätzen. Vor allem die Transformer-basierten Modelle haben sich sowohl in verschiedenen Domänen als auch über verschiedene Sprachen hinweg als äußerst leistungsfähig erwiesen (ibid.).

Eine weitere, für die ATE zentrale, Erkenntnis der Überblicksstudie von Tran et al. (2023) war die Effektivität des Token-Klassifikators auf Satzebene, insbesondere in domänen- und sprachübergreifenden Anwendungen, ein Ansatz der von Wachowiak et al. (2021) vorgestellt wurde. Dieser Ansatz bringt bessere Ergebnisse im Vergleich zu dem herkömmlichen monolingualen Extraktionsansatz, d. h. ermöglicht eine präzisere Extraktion von Begriffen (Tran et al. 2023: 18).

Was diese Überblicksstudie zusätzlich hervorhebt, ist der Trend hin zu multilingualen Modellen in der ATE in den letzten Jahren, die gegenüber den monolingualen Modellen deutliche Vorteile, vor allem wenn es um Termextraktion inklusive Eigennamen geht, aufweisen (ibid.). Dies unterstreicht die Bedeutung von multilingualen Ansätzen in der ATE-Forschung, die unter anderen auch in der Studie von Wachowiak et al. (2021) vorgestellt und untersucht wurden, indem ein Vergleich zwischen zwei Varianten (dem XLM-R-basierte Sequenz-Klassifikator und dem XLM-R-basierter Token-Klassifikator) des mehrsprachigen vorab trainierten Sprachmodells XLM-RoBERTa (XLM-R) und dem mehrsprachigen vorab trainierten neuronalen mBART für ATE-Aufgaben durchgeführt wurde. Diese mehrsprachigen Modelle haben eine bessere Leistung als ein monolinguales Modell erzielt, wobei die beste Leistung das XLM-R-basierte Token Klassifikator erzielte (ibid.). Als Ergebnis dieser Studie bzw. dieses Projektes namens Text2TCS ist eine einfach zu bedienende Anwendung, die auf der Plattform für Sprachtechnologien European Language Grid (ELG) zur freien Nutzung bereitgestellt wurde, hervorgegangen, die nicht nur gute Ergebnisse in Sprachen in denen sie trainiert wurde, sondern auch gute Ergebnisse in Sprachen in denen sie nicht spezifisch trainiert wurde erzielt, d. h. die Anwendung zeigt eine gute „Zero-Shot“-Fähigkeit (ibid.: 2021: 3615). Zudem wurde das System Text2TCS nicht nur zur Extraktion von Termini, sondern zur Extraktion eines ganzen terminologischen Begriffssystems oder auf Englisch Terminological Concept System (TCS),

entwickelt. Gromann et al. (2023: 289) definieren ein TCS als „a terminological resource that conceptually structures domain-specific terms and provides hierarchical and nonhierarchical relations between them“. Da es in dieser Arbeit vorrangig um die monolinguale Termextraktion geht, wird nicht näher auf dieses Konzept bzw. auf die Beziehungsextraktion eingegangen. In der Benutzeroberfläche der Online-Demo sowie der entsprechenden Python-Softwarebibliothek ist es möglich ausschließlich Termini zu extrahieren und den Schritt der Extraktion eines vollständigen Begriffssystems wegzulassen. Wie schon oben erwähnt stützt sich der Ansatz von Text2TCS bei der Termextraktion auf die Klassifikation jedes Tokens, also jedes Wortes oder Wortbestandteils, in einem Satz als entweder Terminus oder Bestandteil eines Terminus, wodurch Mehrwortbenennungen extrahiert werden können (Gromann et al. 2023: 291f.). Zu diesem Zweck wird das vortrainierte mehrsprachige neuronale Sprachmodelle XLM-R (Conneau et al. 2019), ein Modell beruhend auf der Transformer-Architektur (Vaswani et al. 2017), auf eigens für diesen Zweck erstellte annotierte fachspezifische Texte auf Englisch und Deutsch für die Aufgabenstellung der Termextraktion gefinetuned. Die Ergebnisse der Extraktion werden in einem Textformat für eine vereinfachte Lesbarkeit als graphisch visualisiertes sowie speicherbares TBX-Dateiformat ausgegeben. Das System unterstützt eine große Menge an natürlichen Sprachen auf Basis der mehr als 100 Sprachen des darunterliegenden XLM-R Modells.

Wie aus den neuesten Studien hervorgeht, wurden in den letzten Jahren bedeutende Fortschritte in Richtung vollautomatische Begriffsextraktion erzielt, allerdings gibt es noch etliche Herausforderungen, die die Forschung angehen muss, um dem Ziel der Vollautomatik näher zu kommen. In Tran et al. (2023) werden einige dieser Herausforderungen identifiziert: 1. Extraktion der verschachtelten Begriffe, was Forschung an kontextbezogenen Modellen erfordert, 2. Verallgemeinerbarkeit der Modelle, d. h. Wissensübertragung zwischen verschiedenen Kontexten ist ausschlaggebend, um brauchbare Modelle, die auch auf Sprachen und Domänen anwendbar sind, die als ressourcenarm gelten, 3. Integration der Esemblemethode, d. h. Kombinierbarkeit bzw. Integrierbarkeit der linguistischen und statistischen Ansätze in die neuronale Modelle, um dadurch bessere Leistung zu erzielen, 4. Skalierbarkeit der Modelle, d. h. Modelle sollten skalierbar und ressourceneffizient sein, 5. Prompting Forschung auf Token-Ebene als Verbesserungsschritt der Termextraktion (ibid.: 19f.).

4 Methode

Im folgenden Blackbox-Experiment liegt der Fokus auf der ATE, wobei soweit möglich neue Entwicklungen auf diesem Gebiet berücksichtigt werden, indem aktuell am Markt verfügbare ATE-Tools eingesetzt werden. Die Ergebnisse eines neuronalen Systems werden mit den Ergebnissen eines nicht-neuronalen Systems, dem ein kontrastiver Ansatz zugrunde liegt, verglichen. Das heißt, dass die Terminologieextraktion durch einen Vergleich der von Benutzer*innen hochgeladenen Daten mit einer Datenbank der Allgemeinsprache erfolgt. Das Experiment erfolgt aus dem Blickwinkel der Anwender*innen, wobei die Tools OneClick Terms und D-Terminer (Näheres zu diesen Tools siehe Abschnitt 4.3) für das Experiment ausgewählt wurden. Das Experiment besteht aus mehreren Schritten und beinhaltet neben der eigentlichen ATE (siehe Abschnitt 4.3) auch die Erstellung eines spezifischen Korpus, der aus Artikeln zum Thema Energiekrise von ausgewählten österreichischen Zeitungen (siehe Abschnitt 4.1) zusammengestellt ist. Zudem ist die manuelle Erstellung eines Goldstandards, der nach vorher festgelegten Regeln (siehe Abschnitt 4.2) annotiert ist, ein wichtiger Teil dieses Experiments, da dieser den Richtwert und die Vergleichsmaße für die Qualität der extrahierten Ergebnisse der jeweiligen Tools darstellt.

Die Auswahl der Tools erfolgt anhand expliziter und nachstehend genannter Kriterien, wie Zugänglichkeit und aktueller Forschungsstand in der ATE. Weiters spielt auch die Unterstützung der relevanten Sprache eine Rolle.

Nach der Gewinnung der Ergebnisse von einzelnen Tools werden diese miteinander verglichen, um nicht nur festzustellen, welcher Ansatz die besseren Ergebnisse liefert. Um diese Ergebnisse miteinander vergleichen zu können werden sie unter anderem anhand der Metriken Genauigkeit, Trefferquote und F1-Score (siehe Abschnitt 2.5.2) ausgewertet, da diese Metriken in der ATE-Forschung immer wieder verwendet werden, wie z.B. bei Bakaric et al. (2021) oder bei Rokas et al. (2020). Allerdings sind nicht nur diese Metriken für die anschließende Ergebnispräsentation und Diskussion ausschlaggebend, sondern auch die Auswertung der Leistung betreffend die einzelnen Kategorien der Termini, die durch die Annotation im Goldstandard festgelegt wurden.

4.1 Korpuserstellung

Dieses Experiment zielt auf die monolinguale automatische Termextraktion ab, da sich diese gut für den beabsichtigten Vergleich im gegebenen Rahmen eignet, was z.B. auch ein ähnliches Experiment von Bakaric et al. (2021) aufzeigt.

Für das Experiment wurde ein Korpus deutschsprachiger Zeitungsartikel im Fachgebiet Energiekrise aus österreichischen Zeitungen zusammengestellt. Dabei wurden folgende Kriterien berücksichtigt: Herkunftsland, Reichweite, Qualitätsanspruch, Format und Zugänglichkeit. Unter Berücksichtigung des Herkunftslandes beschränkt sich die Auswahl auf Zeitungen aus Österreich, um mögliche Synonym- oder Äquivalenzprobleme bei den Standardvarietäten des Deutschen in den verschiedenen deutschsprachigen Ländern zu vermeiden, wie zum Beispiel bei Wortpaaren wie *Bankomat* (österreichisches Deutsch) vs. *Geldautomat* (deutsches Deutsch). Weiters basiert die Auswahl der Zeitungen auf dem Kriterium der Reichweite, wobei Daten des Vereins Arbeitsgemeinschaft Media (VMA), der jährlich Analysen zur Erfassung der Reichweiten von Printmedien in Österreich durchführt (VMA 2022), verwendet wurden. Dabei wurden Zeitungen bevorzugt, die ihre Leserschaft in ganz Österreich erreichen. Das bedeutet, dass regionale Ausgaben, die zwar eine höhere Reichweite als die für dieses Experiment gewählten Zeitungen erzielt haben, aber nur in bestimmten Regionen erhältlich sind, nicht in Betracht gezogen wurden. Zum Beispiel wird deshalb die Kleine Zeitung, die zwei Ausgaben abhängig von Region anbietet, nicht berücksichtigt (siehe Abb. 14). Das Kriterium Qualitätsanspruch schränkt die Auswahl weiter ein. Es werden Qualitätszeitungen vor Boulevardzeitungen bevorzugt, um mögliche Unterschiede zwischen Dialekt und Hochsprache zu vermeiden. Dabei soll die Wichtigkeit der Boulevardzeitung Kronen Zeitung, die zu den auflagenstärksten Tageszeitungen in Österreich zählt oder der gratis verfügbaren Zeitungen wie Heute etc. (siehe Abb. 14), für die Gesellschaft nicht in Frage gestellt werden. Ein weiteres ausschlaggebendes Kriterium für die Zeitungsauswahl ist das Format, wobei die Verfügbarkeit der Zeitungen im Online- oder ePaper-Format bevorzugt werden, um den Aufwand für die maschinelle Lesbarkeit der Artikel zu minimieren. Das letzte Kriterium ist die Zugänglichkeit der Archive der jeweiligen Zeitungen, wobei darauf geachtet wird, ob die Zeitungen ihre Archive frei zur Verfügung stellen oder ob bereits ein Online-Abo vorliegt, um so die Kosten des Experiments niedrig zu halten. Basierend auf diesen Kriterien wurden zwei Zeitungen ausgewählt.

Media-Analyse 2022	LpA Tageszeitungen																				
	Fälle	NRW Tageszeitungen	Der Standard	Die Presse	Heute (GRATIS)	Kronen Zeitung	Kurier	oe24 (GRATIS)	Österreich	Österreich/oe24-Kombi (KAUF/GRATIS)	Kleine Zeitung gesamt	Kleine Zeitung (Graz)	Kleine Zeitung (Klgt.)	OÖN-OÖ Nachrichten	SN-Saizburger Nachrichten	TT-Tiroler Tageszeitung	Kombi TT/TT Kompakt (KAUF/GRATIS)	Neue Vbg. Tageszeitung	VN-Vbg. Nachrichten	TOP Vorarlberg	
	Total	15.030	52,7	6,8	3,3	8,2	22,2	5,6	3,3	3,7	6,1	8,6	6,0	2,6	4,3	3,1	3,1	3,3	0,4	2,0	2,2
	Geschlecht																				
	Männer	7.360	53,7	7,9	3,8	8,3	21,7	6,2	4,1	4,0	6,8	8,8	6,1	2,7	4,8	3,2	3,1	3,3	0,6	2,1	2,4
	Frauen	7.670	51,7	5,8	2,8	8,1	22,7	4,9	2,6	3,3	5,4	8,4	6,0	2,4	3,9	3,0	3,1	3,3	0,3	1,9	2,1

Abbildung 14: Medien-Analyse 2022 (VMA 2022)

Wie aus Abb. 14 hervorgeht, erreicht die Qualitätszeitung Der Standard 6,8 % der Gesamtreichweite in ganz Österreich, gefolgt vom Kurier mit 5,6 %. Basierend auf diesen Daten erfüllen diese beiden Zeitungen die oben genannten Kriterien am besten und sind daher für diese Arbeit relevant.

Nachdem die Auswahl der geeigneten Printmedien abgeschlossen ist, werden die Zeitungsartikel ausgewählt. Ausschlaggebend für die Auswahl der Artikel sind folgende Kriterien: Fachbereich, Thema bzw. Fachgebiet der Artikel und der Zeitraum, in dem die Artikel erschienen sind. Zusätzlich werden bei der Artikelauswahl aus dem Kurier auch Artikel berücksichtigt, die innerhalb der regionalen Beilage Oberösterreich erschienen sind, die Teil der österreichweiten Ausgabe ist. Das Kriterium für diese Artikel ist, dass der Titel den Begriff Energiekrise beinhalten musste. Das zentrale Thema der Artikel ist die Energiekrise, die unter den Fachbereich Wirtschaft fällt. Der gewählte Zeitraum erstreckt sich von 24.02.2022 bis 24.02.2023. Das Thema und der Zeitraum wurden aus bestimmten Gründen gewählt. Erstens spielen das geopolitische Ereignis und die Aktualität des Themas eine wichtige Rolle. Die aktuelle Energiekrise ist eng an den Ausbruch des Krieges gegen die Ukraine am 24.02.2022 gekoppelt und betrifft nicht nur die Industrie, sondern auch den privaten Haushalt (Statista Research Departement 2023). Zweitens wurde die Energiekrise im Jahr 2022 omnipräsent, was unter anderem auch dadurch belegt ist, dass der Begriff Energiekrise es sogar auf Platz 1 bei der Abstimmung der Gesellschaft für österreichisches Deutsch (GSÖD) zum Unwort des Jahres 2022 in Österreich geschafft hat (GSÖD 2022). Anhand dieser Suchkriterien, die in Tabelle 2 detailliert zusammengefasst sind, wurde in den jeweiligen Online-Archiven der ausgewählten Zeitungen die Suche gestartet.

Der Standard	Kurier
Suchkriterien: • Thema Energiekrise • aus dem Bereich (in der Suchmaschine als Ressorts gekennzeichnet) Wirtschaft • Zeitraum: 24.02.2022-24.02.2023 • Chronologisch sortiert	Suchkriterien: • Energiekrise, Wirtschaft • Zeitraum 24.02.2022-24.02.2023 • Region: Österreich
Gesamtanzahl der Ergebnisse: 319 Treffer	Gesamtanzahl der Ergebnisse: 402 Treffer
Format: TXT-Datei	Format: TXT-Datei

Tabelle 2: Suchkriterien der Artikelsuche und Gesamtergebnisse

Da es im Rahmen dieser Arbeit aus begrenzten zeitlichen und personellen Ressourcen nicht möglich ist, alle Artikel für die manuelle Terminologierecherche in Betracht zu ziehen, sind aus der vorgeschlagenen Menge jeweils 12 Artikel (siehe Tabellen 3 und 4) ausgesucht worden, die möglichst über den ganzen gewählten Zeitraum regelmäßig verteilt sind. In der Endfassung besteht der Korpus insgesamt aus 16.286 Wörtern aus insgesamt 24 Zeitungsartikeln, was bedeutet, dass der Korpus keinen Anspruch auf vollständigen Überblick aller Artikeln der ausgewählten Zeitungen erhebt. Er soll lediglich eine repräsentative Menge darstellen, die das Thema Energiekrise möglichst gut abdeckt. Für ein detailliertes Verzeichnis der ausgewählten Artikel siehe Anhang A.

Artikel von...	Überschrift des Artikels...	Datum des Erscheinens
Hahn und Ungerboeck (2022)	Wie teuer wird die Energiekrise?	07.03.2022
Strobl und Laufer (2022)	Was, wenn kein russisches Gas mehr nach Österreich kommt?	02.04.2022
APA (2022a)	Finanzministerium prüft Gewinnabschöpfung wegen steigender Energiepreise.	06.05.2022
Gepp (2022)	Energiesparen: lässt türkis-grüne Regierung Potenzial ungenutzt?	22.06.2022
APA (2022b)	Mikl-Leitner für Strompreisdeckel – Finanzminister Brunner warnt vor Gefahren.	10.07.2022
APA (2022c)	Nach Preiserhöhung bei Wien Energie und EVN: Andere Landesversorger warten ab.	04.08.2022
Danzer und Ungerboeck (2022)	Die Strompreisbremse kommt: Was sie für die Bevölkerung bedeutet.	05.09.2022
Ungerboeck (2022)	Klimaministerium will mehr Befugnisse für den Energienotfall.	05.10.2022
Strobl (2022a)	Normalisierung bei Energiepreisen frühestens 2024 erwartet.	02.11.2022
Strobl (2022b)	Sanktionen zeigen Wirkung, russisches Rohöl viel weniger wert.	08.12.2022
APA (2023a)	Haushaltsenergiepreise stiegen im November im Schnitt um 42 Prozent.	10.01.2023
APA (2023b)	Österreich und Deutschland einigen sich auf Nutzung des Gasspeichers Haidach.	17.02.2023

Tabelle 3: Auswahl der Artikel aus Der Standard

Artikel	Überschrift des Artikels...	Datum des Erscheinens
Scheidinger (2022a)	Industrie sucht Wege aus der Energiekrise.	03.04.2022
Wobrazek (2022a)	Die Suche nach der Energielösung.	01.05.2022
Arends (2022)	Kein Gas aus Russland: In Bulgarien droht ein kalter und damit heißer Herbst.	10.05.2022
Scheidinger (2022b)	Die Energiekrise richtig einschätzen und bewältigen.	29.05.2022
Wobrazek (2022b)	Die Wege aus der Energiekrise.	04.09.2022
Bachner (2022)	Inflation kratzt an 10-Prozent-Marke.	17.09.2022
Scheidinger (2022c)	Energiekrise trifft die Industrie.	02.10.2022
Unterhuber (2022)	Stromschlag für die Nahversorger.	22.10.2022
Steiner-Gashi und Meyrath (2022)	Warum der Gaspreis nicht mehr sinken wird.	24.11.2022
Kotrba und Meyrath (2022)	Stromschlag für die Nahversorger	20.12.2022
Pressberger (2023)	„Verstehe diese Doppelmoral nicht“.	04.01.2023
Bachner (2023)	Sonne, Wind, Wasser – und Atomkraft.	09.02.2023

Tabelle 4: Auswahl der Artikel aus Kurier

Die ausgewählten Artikel wurden direkt aus dem Browser kopiert und als TXT-Dateien gespeichert, da alle für dieses Experiment ausgewählten Tools dieses Format unterstützen. Im letzten Schritt wurden die TXT-Dateien alle einzeln durchgesehen, auf Vollständigkeit geprüft und wenn notwendig kleine Korrekturen vorgenommen, wie z.B. Entfernung der mitkopierten Graphiken oder Formatierungssymbole, Entfernung der Erscheinungsdaten und der Namen der Autor*innen aus dem jeweiligen Artikel, um dadurch mögliche ungültige Termkandidaten in der Kategorie Named Entities (siehe Kapitel 2.1.4) im Voraus zu reduzieren. Hier muss angemerkt werden, dass diese Korrekturen nur deshalb durchgeführt wurden, weil der Korpus aus einer beschränkten Artikelanzahl besteht, d.h. sie kaum Zeit in Anspruch genommen haben. Bei einem umfangreicheren Korpus wären solche Korrekturen zeitaufwendig und daher nicht empfehlenswert.

4.2 Erstellung eines Goldstandards

Nachdem der Korpus erstellt ist, folgt die Erstellung des Goldstandards (siehe Anhang C), der auch entsprechend der festgelegten Kategorien manuell annotiert ist. Schon die kurze Beschreibung dieses Schrittes wirft einige wesentliche Fragen auf, die sich kurz als drei W-Fragen zusammenfassen lassen, konkret: was wird annotiert, wie wird annotiert und wer annotiert. Für das beabsichtigte Experiment ist es unabdingbar festzulegen, was alles als Terminus überhaupt in den Goldstandard aufgenommen wird. Im zweiten Schritt werden die ausgewählten Termini anhand vordefinierter Bestimmungen ihrer jeweiligen Kategorie zugeordnet. Bevor man diese Schritte ausführen kann, muss noch die Frage nach geeigneten Annotator*innen geklärt werden. Da es sich um eine manuelle Terminologierecherche handelt, gelten für die Auswahl der Annotator*innen die schon im Abschnitt 2.3.2.2 erwähnten Kriterien, die man kurz unter der Bezeichnung fachlich fundierte Personalressourcen bzw. Bearbeiter*innen mit fundiertem Wissen zusammenfassen kann (siehe Abschnitt 2.3.2.2). Für diese Arbeit sind zwei Annotator*innen vorgesehen, die zwar nicht direkt aus dem Fachbereich Wirtschaft kommen, allerdings durch ihre Ausbildung und erfolgreichen Abschluss des Moduls Fachübersetzen in Recht und Wirtschaft am Zentrum der Translationswissenschaft (ZTW) mit der Terminologie im Fachbereich Wirtschaft in der Sprachkombination Tschechisch/Deutsch vertraut sind.

Nachdem die Frage wer geklärt ist, kommt die Frage was an die Reihe. Die unterschiedlichen Definitionen von Terminus, die in Abschnitten 2.1.1 und 2.2.3 erwähnt wurden, und Merkmalbeschreibungen von Terminus wie sprachliche Einheit in einem Fachgebiet, Eineindeutigkeit, Terminus etc., sind für die Praxis und für das beabsichtigte Experiment wenig zielführend, da sie Raum für unterschiedliche Interpretationen lassen. Zu diesem Schluss kommt unter anderen auch Rigouts Terryn (2021a) und schlägt konkrete Annotationsregeln vor, um die Annotator*innen bei ihrer Entscheidung, ob eine sprachliche Einheit in einem Text ein Terminus oder Nicht-Terminus ist, zu unterstützen. Zusätzlich hilft dieses Annotationsregelwerk auch die Frage wie zu klären, indem Kategorien der Termini bestimmt und genau beschrieben werden. Anhand dieser Regeln, der terminologischen Besonderheiten in der Wirtschaftssprache (siehe Abschnitt 2.1.8.3) und der gängigen Klassifizierung der Termini in Hinsicht auf die ATE (siehe Abschnitt 2.1.4) wurde das folgende Annotationsschema bzw. die folgenden Richtlinien zusammengestellt, um dadurch möglichst objektive Annotation in diesem Experiment zu gewährleisten.

4.2.1 Generelle Richtlinien für Annotation von Termini

Anhand des Richtlinienbeispiels von Rigouts Terry (2021a) gilt es beim Bestimmen der Termini folgende Kriterien zu beachten: Länge der Termini, rekursive Annotation, Sonderfälle wie gesplittete und fremdsprachige Termini, Wortarten, Abkürzungen, Named Entities und ihre Kurzformen.

Erstens gibt es keine Mindest- oder Höchstlänge für Termini, allerdings sollten die meisten nicht länger als ein paar Wörter sein. Das heißt Termini können ein einzelnes Wort sein, zwei Wörter oder mehrere Wörter umfassen, wobei in dieser Arbeit zwischen zwei Längen unterschieden wird, nämlich Einwort- und Mehrwortbenennung. Beispiele für Einwortbenennungen z.B. aus dem Fachgebiet Finanzpolitik sind Termini wie *Inflation*, *Steuern* oder *Wachstum*, für Mehrwortbenennungen Termini wie *fiskale Stimulierung*, *öffentliche Schulden* oder *Kauf von Staatsanleihen am Sekundärmarkt*.

Zweitens sollten Termini in einer verschachtelten oder rekursiven Art und Weise annotiert werden. Das bedeutet, dass, wenn ein Terminus Teil eines anderen Terminus ist, beide Termini annotiert werden, und dies auf allen Ebenen der Verschachtelung erfolgt. Deutlich wird es am Beispiel vom Terminus *Kauf von Staatsanleihen am Sekundärmarkt*. Dieser Terminus besteht aus weiteren Termini wie *Sekundärmarkt* und *Kauf von Staatsanleihen*, der wiederum den Terminus *Staatsanleihen* beinhaltet. Alle diese Termini gilt es als individuelle Termini zu erfassen (für weitere Beispiele siehe Rigouts Terry 2021: 7). Hier ist zu beachten, dass Komposita, die aus zwei oder mehreren Wörtern bestehen, als ein Terminus gelten und nicht in ihre Einzelteile zerlegt werden. Das bedeutet, wenn zum Beispiel ein Terminus wie *Inflationsrate* vorliegt, wird er nicht in *Inflation* und *Rate* zerlegt.

Drittens sind Sonderfälle wie z.B. gesplittete Termini zu beachten. Hierbei handelt es sich um Termini, die durch andere Wörter oder Satzzeichen, wie Klammern etc., unterbrochen werden können. Im Falle von Anglizismen, anderen fremdsprachigen Termini und Lehnwörtern gilt es diese in den Goldstandard separat aufzunehmen, auch wenn deutsche Äquivalente vorliegen. Ähnlich wird auch mit Synonymen umgegangen, welche separat annotiert werden.

Weiters werden nicht nur Substantive, sondern auch andere Wortarten wie Adjektive, Verben und Adverbien, wie z.B. *fiskalpolitisch*, *subventionieren*, *monetär-politisch* etc., und ihre Kombinationen wie Funktionsverbgefüge, wie z.B. *den Vertrag auflösen*, als Termini betrachtet.

Abkürzungen werden separat und in die gleiche Kategorie als ihre ausgeschriebenen Formen aufgenommen, was auch bei englischsprachigen Abkürzungen gilt, wie bei

Bruttoinlandsprodukt und seiner deutschen Abkürzung *BIP* und der englischen *GDP*. Dies ist nicht immer offensichtlich, da Abkürzungen weniger bekannt sein können als die Vollformen.

Im Falle von Named Entities, d.h. Eigennamen wie etwa von Organisationen, Orten, Produkten usw. kann es vorkommen, dass diese nur durch einen kurzen Teil des gesamten Terminus im Korpus vorkommen. Ein Beispiel dafür wäre *Deklaration zur Steuertransparenz*, die in weiterer Folge im Artikel nur als *Deklaration* weiterverwendet werden kann. In solchen Fällen wird die abgekürzte Form in der Regel nicht annotiert, da solche Abkürzungen zu ungenau sind.

Weiters werden auch keine Website-Adressen annotiert, ebenso wenig wie allgemeine Substantive, die in Kombination mit adjektivischen Termini auftreten. Die Einheiten von Maßeinheiten sollten denselben Prozess wie andere Termini durchlaufen. Rechtschreibfehler und Tippfehler werden annotiert, solange sie noch leicht erkennbar sind.

4.2.2 Kategorien der Termini

Anhand der im Abschnitt 2.1.4.1 erwähnten Typologien von Termini bieten sich unterschiedliche Klassifizierungen der Termini an, und zwar entweder nach Roelcke (2020: 71), der die Fachspezifizität der Termini in den Fokus stellt, oder nach Adamzik (2018: 299), die die Erkennbarkeit in Nachschlagwerken bzw. Lexika in den Fokus stellt. Wie schon im Abschnitt 2.1.4 erwähnt, wird in der ATE-Forschung häufig eine Mischung (siehe Abb. 15) aus diesen beiden Klassifikationen angewendet, die beide Aspekte Fachspezifizität und die Erkennbarkeit in den Lexika in Betracht zieht, wie z.B. bei Rigouts Terryn et al. (2022).

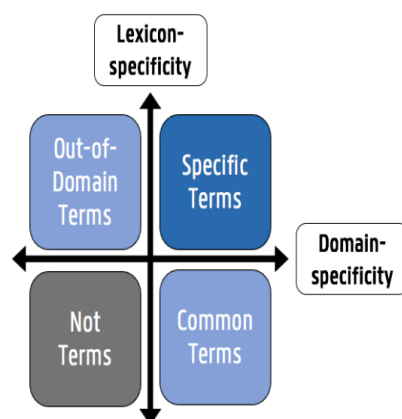


Abbildung 15: Annotationsschema (Rigouts Terryn 2021a: 4)

Die einzelnen Kategorien bzw. Typen von Termini wurden schon im Abschnitt 2.1.4 kurz erwähnt. Allerdings werden die einzelnen Kategorien an dieser Stelle detaillierter erläutert und mit konkreten, von Rigouts Terryn (2021a) übernommenen Beispielen, versehen. Dieser Schritt soll die Annotation erleichtern.

Unter Specific Terms bzw. (fach-)spezifischen Termini werden die Termini verstanden, die eng mit einem bestimmten Fachgebiet verbunden sind und daher spezielles Fachwissen erfordern, um verstanden zu werden. Zum Beispiel ist *Myokardinfarkt* ein spezifischer Terminus aus dem Fachbereich Medizin.

Wiederum werden unter den Common Terms bzw. allgemeinen Termini solche Termini verstanden, die zwar stark mit einem Fachgebiet verbunden sind, aber auch von Personen außerhalb des Fachgebiets verstanden werden können, wie zum Beispiel der Terminus *Blutdruck*.

Weiters gibt es die Out-of-Domain (OOD) Terms bzw. nicht-fachspezifische Termini, die nichts mit einem bestimmten Fachgebiet zu tun haben, aber in einem speziellen Kontext auftreten können. Zum Beispiel könnte das Wort *Algorithmus* aus der Informatik stammen, aber in der Medizin verwendet werden, um bestimmte Prozesse zu beschreiben.

Außer dieser drei Kategorien, wird in der ATE auch die Kategorie Named Entities bzw. Eigennamen verwendet. Es handelt sich hier um Eigennamen von Personen, Orten, Produkten oder Organisationen, die nicht unbedingt fachspezifisch sind. Zum Beispiel ist *Albert Einstein* ein Eigenname, der in verschiedenen Kontexten auftreten kann.

Wie Rigouts Terryn (2021a) in ihrem Annotationsschema betont, ist in manchen Fällen die Unterscheidung zwischen den einzelnen Kategorien nicht immer eindeutig. Aus diesem Grund schlägt die Wissenschaftlerin vor, dass wenn Unsicherheit besteht, ob ein Terminus überhaupt als Terminus betrachtet werden sollte, ihn besser nicht in den Goldstandard aufzunehmen (ibid.: 12). Außerdem kann bei einer Unsicherheit Google-Suche eine Abhilfe schaffen, da man hier nachsehen kann, wie der fragliche Terminus im bestimmten Fachgebiet verwendet wird. Außerdem ist es auch ratsam, eine Liste von schwierigen Fällen zu führen (ibid.: 12). Dies wird auch bei diesem Experiment den Annotator*innen empfohlen, um mögliche Unstimmigkeiten nachträglich besprechen zu können.

Anhand der oben erwähnten Kategorisierung von Termini wurde eine tabellarische Übersicht der Kategorien mit einer kurzen Eigenschaftsbeschreibung der jeweiligen Kategorie für die Annotator*innen bereitgestellt, um die Annotation möglichst objektiv zu gestalten (siehe Anhang B).

4.3 Systemwahl und Vorgehensweise der Automatischen Termextraktion

Den Kern des Experiments bildet die ATE bzw. die verwendeten Systeme, die nach bestimmten Kriterien wie Zugänglichkeit, Benutzerfreundlichkeit, unterstützte Sprachen, Stand der Forschung bzw. implementierte Ansatz und unterstützte Datei-Formate ausgewählt worden sind. Unter Beachtung aller dieser Kriterien wurden folgende Systeme ausgewählt: OneClick Terms (Baisa et al. 2017) und D-Terminer (Rigouts Terryn 2021b).

Das System OneClick Terms ist als kostenlose Testversion, allerdings mit beschränkter Funktionalität, als Open-Source-System im Internet verfügbar. Um auf die volle Funktionalität des Systems zugreifen zu können, benötigt man entweder einen institutionellen oder kommerziellen Zugang. Für dieses Experiment ist ein institutioneller Zugriff über die Universität Wien verfügbar. OneClick Terms wurde innerhalb des Sketch Engine-Frameworks (Kilgarriff et al. 2004) entwickelt, um die Termextraktion zu vereinfachen und dadurch auch Benutzer*innen anzusprechen, die Sketch Engine für diesen Zweck zu komplex finden. Wie Baisa et al. (2017) betonen, können Benutzer*innen ohne jegliche Kenntnisse in NLP, Korpora, Tagging oder Korpusplattform und -analysissoftware eine Datei in die Benutzeroberfläche ziehen, einmal klicken und eine Termkandidatenliste erstellen. Das System funktioniert, indem es statistische und linguistische ATE-Methoden kombiniert, die durch den Vergleich des hochgeladenen Korpus mit einem externen, häufig allgemeinen, Korpus unterstützt werden (siehe Abschnitt 2.4.2.4). Dieser Vergleich von fachspezifischem Text und allgemeinen Texten ermöglicht es festzustellen, welche lexikalischen Einheiten Termini und welche nur häufig auftretende Phrasen ohne fachspezifische Bedeutung sind. Mit anderen Worten kommt das Konzept Termhood zur Anwendung (siehe Abschnitt 2.4). Außerdem stützt sich das System auf einen zweiten in der ATE bekannten Begriff und zwar Unithood (siehe Abschnitt 2.4), welcher in Form der sogenannten term grammar (Baisa et al. 2017) im System realisiert wird, d.h. linguistische Regeln bzw. Definitionen eines Terminus in der jeweiligen Sprache werden festgelegt. Was das System für die Praxis unter anderem nützlich macht, ist die Unterstützung mehrerer Sprachen, verbesserte Extraktionsalgorithmen und unterstützte Formate. Konkret unterstützt das System die Termextraktion aus Dateiformaten wie TMX, XLIFF, PDF, DOC, DOCX, HTML und TXT und den Export der Ergebnisse in CSV-, XLS- oder TBX-Formaten (Baisa et al. 2017). Für den Zweck dieses Experiments ist vor allem die Unterstützung der deutschen Sprache wichtig und die Unterstützung der Termextraktion aus dem Dateiformat TXT. Es ist anzumerken, dass auch wenn die Methoden des Systems OneClick Terms kontinuierlich weiterentwickelt und aktualisiert werden, dieses System aufgrund seiner spezifischen Eigenschaft höchstwahrscheinlich keine

Wettbewerbsfähigkeit gegenüber Systemen mit neuronalen Ansätzen hat. Konkret handelt es sich um die Unterscheidung zwischen Keyword, d.h. Schlüsselwort und Term, d.h. Terminus. Auf der Website, über die das Tool zugänglich ist (Sketch Engine o. J.) wird betont, dass Schlüsselwörter aus einzelnen Tokens bestehen, bzw. aus einzelnen Wörtern, während sich Termini aus mehreren Tokens zusammensetzen. Daraus resultieren zwei separate Kandidatenlisten. Diese Unterscheidung zwischen Schlüsselwort und Terminus legt nahe, dass nur die Kandidatenliste mit den Mehrwortbenennungen als Termkandidatenliste berücksichtigt werden sollte. Dies steht jedoch im Widerspruch zur allgemeingültigen Definition von Terminus (siehe Abschnitt 2.2.3), da Termini sowohl Einwort- als auch Mehrwortbenennungen umfassen können. Daher muss im Verlauf des Experiments auf diese Unterscheidung verzichtet werden, d.h. die Kandidatenlisten werden bloß als Einwort- und Mehrwortbenennungslisten und nicht als Terminus- und Schlüsselwortlisten betrachtet. Das führt dazu, dass beide Kandidatenlisten miteinander verknüpft werden. Dabei ist es zu beachten, dass nicht alle Einwortbenennungen als Termini gelten, was zu einer hohen Anzahl von ungültigen Kandidaten führen kann. Umgekehrt muss man bei der Ausschließung von Einwortbenennungen mit einer hohen Anzahl von gültigen Kandidaten rechnen, die nicht erkannt werden. Um diese Fehlerquelle zu reduzieren, ist es in der ATE üblich, bei klassischen, statistischen und hybriden Methoden die Anzahl an Termini festzulegen (Bakaric et al. 2021; Šajatović et al. 2019), welche die Kandidatenlisten einschränkt. Solche Abschätzung einer richtigen Kandidatenanzahl war immer schon schwierig und ist auch nicht Gegenstand dieser Arbeit. Deshalb wird in dieser Arbeit nicht mit unterschiedlichen Schwellenwerten experimentiert. Ausschlaggebend für die Kandidatenlisten ist, dass das System OneClick Terms Termini aus dem Text gereiht nach berechneter Wichtigkeit für den jeweiligen Fachbereich anzeigt.

Das zweite in dieser Arbeit verwendete System ist das schon im Abschnitt 3 kurz erwähnte D-Terminer (Rigouts Terryn 2021b), das als Online-Demo im Internet frei verfügbar ist und das außer der monolingualen auch die bilinguale ATE aus parallelen Korpora unterstützt. Das Tool nutzt ein sog. Recurrent Neural Network (RNN) und eine überwachte Trainingsmethode, die auf vorab trainierten Word Embeddings (siehe Abschnitt 2.4.2.9) beruht, um monolingual Termextraktion durchzuführen. Obwohl die derzeit dominierende neuronale Architektur jene des Transformers (Vaswani et al. 2017) ist, beruht D-Terminer auf der davor entwickelten Architektur der RNNs. Laut der Autorin eignet sich ein RNN ideal für die Textanalyse bzw. Termextraktion, da sie auf einem Rückkoppelungsprinzip basiert, was konkret bedeutet, dass vorherige Informationen in die Verarbeitung und den Lernprozess späterer Sequenzen einbezogen werden können (Asudani et al. 2023: 10348f.). Die unterstützten

Dateiformate umfassen reinen Text, d.h. TXT-Dateien, für monolinguale Texte und TMX-Dateien für zweisprachige Texte. D-Terminer unterstützt mehrere Sprachen, wenn auch immer nur einzeln, darunter Niederländisch, Englisch, Französisch und Deutsch.

Nachdem die Auswahl der Systeme abgeschlossen ist, werden die einzelnen ausgewählten Texte mithilfe der gewählten Systeme verarbeitet. Um die Extraktionszeit der Systeme zu verkürzen und um bessere Übersichtlichkeit der Termkandidaten zu gewährleisten, werden die Artikel als separate Dateien gleichzeitig auf Termini untersucht. Auch wenn diese gleichzeitig mögliche Untersuchung Zeit erspart, sollte man bei umfangreicheren Korpora beachten, dass je nach System die Kapazität an gleichzeitig hochgeladenen Artikeln variieren kann. An dieser Stelle ist noch anzumerken, dass beim System mit einstellbaren Systemeinstellungen kleine Veränderungen durchgeführt werden. Konkret werden beim OneClick Terms folgende Einstellungen vorgenommen: die Anzahl der angezeigten Kandidaten beider Kandidatenlisten wird auf den Wert 500 (siehe Abb. 16) gesetzt, es werden auch seltene Termini erlaubt und zuletzt werden die Kandidaten in ihrer ursprünglichen Wortform dargestellt. Diese Einstellungen werden mit dem Wissen vorgenommen, dass andere Werte möglicherweise zu besseren Ergebnissen führen können. Im Gegenteil zu diesen klassischen Methoden versuchen neuronale Methoden alle Termini zu extrahieren, d.h. es müssen keine bestimmten Einstellungen vorgenommen werden.

SETTINGS

Extract single-words as

☐ lemma (basic form) ? ☒ word form ?

Give preference to ?

rare words ? common words

Single words

Results to view on screen (1 - Minimum frequency ?
100) All results can be downloaded

Multi-words

Results to view on screen (1 - Minimum frequency ?
100) All results can be downloaded

Abbildung 16: Einstellung des Schwellenwertes von System OneClick Terms

Die Kandidatenliste der ATE für das System OneClick Terms befindet sich im Anhang D und für das System D-Terminer im Anhang E.

4.4 Evaluierung

In dieser wissenschaftlichen Arbeit wird die Leistung der ausgewählten Systeme anhand der generierten Termkandidatenlisten mittels verschiedener Metriken evaluiert. Die Auswertung erfolgt sowohl quantitativ durch Berechnungen als auch qualitativ durch eine manuelle Überprüfung der Ergebnislisten.

Die quantitative Auswertung der Termkandidatenlisten erfolgt anhand der in der ATE gängigen Metriken Genauigkeit, Trefferquote und F1-Metrik (Näheres zu den Metriken siehe Abschnitt 2.5.2). Die jeweilige Metrik wird mithilfe der Programmiersprache Python berechnet. Zusätzlich zu dieser Auswertung erfolgt auch eine Auswertung nach der im Abschnitt 4.2.2 vorgestellten Kategorien der Termini, d.h. dass die Ergebnisse in mehreren Schritten präsentiert werden.

Zunächst werden die quantitativen Ergebnisse als Zahl dargestellt, d.h. wie viele Termini von jeweiligem System überhaupt erkannt wurden, wie viele davon im Goldstandard vorkommen und wie viele davon nicht im Goldstandard vorkommen. Dann werden die quantitativen Ergebnisse für die Genauigkeit, Trefferquote und F1-Metrik des jeweiligen Systems im Prozentsatz angegeben. Diese Metriken geben einen ersten Überblick über die Leistungsfähigkeit des jeweiligen Systems. Alle diese Ergebnisse werden in einer Tabelle dargestellt. Im weiteren Schritt folgen die Ergebnisse als Zahl je nach Kategorie, d.h. wie viele Termini aus dem Goldstandard je Kategorie nicht gefunden wurden. Diese Ergebnisse werden in einer Tabelle dargestellt.

Anschließend erfolgt eine qualitative, manuelle Auswertung der Ergebnisse. Hierbei werden die vom jeweiligen System nicht erkannten Termini und die fehlenden Termini im Goldstandard analysiert. Wobei eine besondere Aufmerksamkeit den im Goldstandard erhaltenen Termini, die nicht vom jeweiligen System erkannt wurden, gilt. Es wird überprüft, ob bestimmte Muster erkennbar sind, wie z. B. Probleme bei der Identifizierung von Termini mit Bindestrichen oder bei Mehrwortbenennungen. Weiters werden auch die im Goldstandard nicht erhaltene Termini analysiert, d.h. die Termini, die vom jeweiligen System zwar erkannt wurden, jedoch nicht im Goldstandard enthalten sind. Diese Liste wird manuell überprüft, um zu ermitteln, ob diese Termini tatsächlich relevante Fachbegriffe darstellen und eventuell in den Goldstandard hätten aufgenommen werden sollen. Gleichzeitig werden auch hier Muster identifiziert, etwa ob bestimmte Arten von Wörtern oder Phrasen fälschlicherweise erkannt wurden. Diese Ergebnisse zeigen, ob ein bestimmtes System häufig Kandidaten identifiziert, die eigentlich zum Goldstandard gehören sollten, oder ob bestimmte Typen von Fehlern auftreten (z. B. falsche Identifizierung von Eigennamen oder allgemeinen Wörtern als Fachbegriffe). Die

manuelle Überprüfung der Ergebnislisten ist für die ATE wesentlich, da ihr Ziel ist, die Fehler des Systems zu identifizieren und deren Ursachen zu analysieren. Diese Analyse hilft dabei, die Stärken und Schwächen des Systems zu identifizieren und mögliche zukünftige Optimierungspotenziale aufzuzeigen.

Zusammenfassend werden folgende Auswertungsschritte durchgeführt: 1. Automatische Zählung der gefundenen und nicht gefundenen Termini je System gesamt und je Kategorie, 2. Automatische Berechnung der Metriken, 3. Qualitative Analyse, um Muster und Fehlerquellen zu identifizieren. Durch diese Kombination aus automatisierter und manueller Auswertung kann eine umfassende Evaluation gewährleistet werden.

5 Ergebnisse

Um eindeutige Aussagen zur Leistung eines Systems tätigen zu können, müssen die Ergebnisse aus beiden Perspektiven, der quantitativen und qualitativen, analysiert werden. Deshalb werden zuerst die quantitativen und darauffolgend die qualitativen Ergebnisse getrennt präsentiert.

5.1 Quantitative Analyse der ATE-Ergebnisse

Im ersten Schritt werden die quantitativen Ergebnisse dargestellt, wobei diese aus einem automatisch ausgewerteten Vergleich der einzelnen Kandidatenlisten mit dem Goldstandard, der insgesamt aus 1250 Termini besteht, hervorgehen. Dieser Vergleich hat folgende Tabelle 5 hervorgebracht.

	D-Terminer	OneClick Terms (1000)	OneClick Terms (2000)
1. Anzahl Termini	1441	1000	2000
2. Anzahl Termini in Goldstandard	734	377	616
3. Anzahl der nicht gefundenen Termini	516	873	634
4. Anzahl der gefundenen Termini, die nicht im Goldstandard auftreten	407	623	1384

Tabelle 5: Quantitative Ergebnisse der Evaluation von Termini in Zahlen

In der ersten Zeile der Tabelle 5 befindet sich die Anzahl aller Termini, die das jeweilige System gefunden hat, wobei man hier anmerken muss, dass beim System OneClick Terms der Schwellenwert 1000 im Voraus eingestellt werden musste, da ohne diese Begrenzung eine große Anzahl an ungültigen Kandidaten geliefert worden wäre (siehe Abb. 17) und dadurch möglicherweise noch schlechtere Ergebnisse zustande kommen würden. Da es aber nicht das Ziel dieser Arbeit ist, den besten Schwellenwert für das ausgewählte System zu eruieren, wird der Vergleich zwischen den konkurrierenden Systemen vor allem mit dem Schwellenwert 1000 vorgenommen, wobei 500 Termini der Liste Einwortbenennungen und 500 Termini der Liste Mehrwortbenennungen entnommen werden. Jedoch enthalten die Ergebnistabellen in dieser Arbeit auch die Ergebniswerte mit dem Schwellenwert 2000, um aufzuzeigen, wie sich die Ergebnisse je nach Schwellenwert in jeweiligen Messkategorien verschlechtern oder verbessern können.

4447	Von	1
4448	war	7
4449	selbst	2
4450	habe	4
4451	Seite	1
4452	Stadt	1
4453	Leben	1
4454	z.	1
4455	B.	1
4456	Sie	8
4457	konnte	1
4458	Ich	4

Abbildung 17: Auszug der ATE-Ergebnisse von OneClick Terms ohne Schwellenwert

Der zweite Messwert zeigt die Anzahl der gefundenen Termini, die sich im Goldstandard befinden. Hier zeichnet sich ab, dass D-Terminer mit 734 Termini in dieser Messkategorie das System OneClick Terms mit 377 Termini deutlich übertrifft. Der dritte Messwert richtet sich auf die Termini, die im Goldstandard vorkommen, aber vom System nicht als solche identifiziert worden sind. Hier liegt das System D-Terminer auch deutlich vorne. Der letzte Wert dieser Tabelle zeigt die Anzahl der von den jeweiligen Systemen gefundenen Termini, die nicht im Goldstandard auftreten. Die Aussagekraft dieses Wertes zeigt erst die qualitative Analyse, wobei der Wert 623 beim OneClick Terms vermuten lässt, dass viele ungültigen Termini beinhaltet sind.

Die automatische Berechnung von Metriken hat die Ergebnisse in Tabelle 6 erbracht:

	D-Terminer	OneClick Terms (1000)	OneClick Terms (2000)
1. Genauigkeit	64,33%	37,7%	30,8%
2. Trefferquote	58,72%	30,16%	49,28%
3. F1-Metrik	61,4%	33,51%	37,91

Tabelle 6: Quantitative Ergebnisse berechnet mittels der Metriken Genauigkeit, Trefferquote und F1-Metrik

Wie aus der Tabelle 6 hervorgeht, liegt in jeder Messkategorie das System D-Terminer vor dem System OneClick Terms, wobei hier anzumerken ist, dass die Trefferquote beim OneClick Terms mit Schwellenwert 2000 deutlich bessere Ergebnisse zeigt als mit Schwellenwert 1000.

Die bisherig eindeutigen Ergebnisse deuten darauf hin, dass das System D-Terminer auch in den einzelnen Kategorien der Termini bessere Ergebnisse als das System OneClick Terms liefern wird. Diese Vermutung wird anhand aller Termini aus dem Goldstandard untersucht, die von den jeweiligen Systemen in den einzelnen Kategorien nicht gefunden wurden.

Art der Termini	D-Terminer	OneClick Terms (1000)	OneClick Terms (2000)
Fachspezifische	228	308	211
Allgemeine	157	266	184
nicht-fachspezifische	44	55	35
Eigennamen	87	244	204

Tabelle 7: Anzahl der Termini je Kategorie, die nicht gefunden worden sind

Wie die Tabelle 7 zeigt, lässt sich diese Vermutung zumindest gegenüber dem System OneClick Terms mit Schwellenwert 1000 bestätigen. Allerdings zeigt sich hier das System OneClick Terms mit Schwellenwert 2000 ein erstes Mal in einigen Kategorien effektiver als das System D-Terminer ist. Konkret zeigt das System bessere Ergebnisse in den Kategorien fachspezifische und nicht-fachspezifische Termini.

5.2 Qualitative Analyse der ATE-Ergebnisse

Während der qualitativen Analyse werden zuerst die Ergebnisse des Systems D-Terminer und dann des Systems OneClick Terms mit Schwellenwert 1000 näher betrachtet, wobei jede einzelne Kategorie gesondert angesehen wird.

5.2.1 Qualitative Analyse der Ergebnislisten von D-Terminer

In der Kategorie fachspezifische Termini fällt auf, dass die Einwortbenennungen bis auf einige wenigen Ausnahmen, wie z.B. *Gasknappheit*, *Rohstoffpreise*, *Rohstoffexperte* usw. erkannt werden. Das Einzige, was bei den Einwortbenennungen auffällt ist, dass die aus dem Englischem entlehnten Termini mit einer deutschen Mehrzahlendung nicht erkannt werden, wie z.B. *E-Fuels*, *Pipelines*. Allerdings gibt es im untersuchten Korpus nur wenige solche Beispiele, um hier ein eindeutiges Muster festzustellen. Bei weiterer Analyse der Termini tauchen einige deutlichere Muster auf, die auf potenzielle Schwächen des Systems D-Terminer hindeuten. Erstens werden viele Termini, die aus mehreren Wörtern bestehen und durch Bindestrich verbunden

nicht erkannt. Beispiele hierfür sind *Strom- und Gaspreise*, *Atom-, Gas-, Kohle- und Laufkraftwerke* oder *CO₂- und Energiepreise*. Dies deutet darauf hin, dass das System Schwierigkeiten mit der Erkennung von zusammengesetzten fachspezifischen Termini hat, insbesondere wenn diese durch Bindestrich getrennt sind. Ähnlich wie bei Bindestrich-Wörtern scheint das System auch Schwierigkeiten mit Mehrwortbenennungen zu haben, die aus mehr als zwei Wörtern bestehen. Beispiele wie *Abkehr von fossilen Ressourcen*, *Transformation des Energiesystems* oder *Steuerbelastung bei den Energiekosten* wurden nicht erkannt. Dies könnte darauf hindeuten, dass das System entweder einzelne Begriffe innerhalb der Benennung separat identifiziert hat oder nicht in der Lage ist, komplexe Phrasen als zusammenhängende Entität zu erkennen. Weiters gibt es einige Beispiele für Termini, die in variierenden Formen vorkommen und nicht erkannt wurden, wie *energieintensiven Betrieben* und *energieintensiven Unternehmen*. Diese Unterschiede in der Flexion oder der grammatikalischen Form könnten darauf hindeuten, dass das System Schwierigkeiten hat, ähnliche Begriffe mit variierenden Endungen oder Präfixen zu erkennen, allerdings kann es sich hier auch darum handeln, dass diese gar nicht als Termini unabhängig von der Endung erkannt werden. Weiters fällt noch auf, dass einige fachspezifische Termini, die sich auf sehr spezifische Konzepte beziehen, wie *E-Fuels* oder *Merit Order*, ebenfalls nicht erkannt werden. Dies könnte darauf hindeuten, dass das System bei der Erkennung von spezialisierten Termini in bestimmten Bereichen Schwierigkeiten hat, insbesondere wenn diese nicht häufig vorkommen oder besonders fachspezifisch sind. Zuletzt fällt auf, dass das System Schwierigkeiten hat, Eigennamen in Verbindung mit spezifischen Titeln oder Rollen korrekt zu identifizieren, wie *Chef der Internationalen Energieagentur* oder *Geschäftsführer der Österreichischen Energieagentur*.

Ähnlich wie bei den fachspezifischen Termini hat das System Schwierigkeiten auch bei den allgemeinen Termini den Bindestrich und Mehrwortbenennungen zu erkennen, wie *(preis-)treibende*, *Stahl-, Zement- oder Glasöfen*, *Haupt- und Nebenwohnsitz* oder *Drei-Personen-Haushalts*. Was zusätzlich bei den allgemeinen Termini auffällt ist, dass es Schwierigkeiten bei der Erfassung weit verbreiteten, aber eher unspezifischen Termini gibt, die eng mit dem Thema Energiekrise in Verbindung stehen. Beispiele hierfür beziehen sich auf Haushalte und private Konsumkontexte, z.B. *Privathaushalte* oder *Haushalte* oder auf allgemeine und weit verbreitete wirtschaftliche Termini, wie *Marktpreis*, *Inflation*, *Entlastungspakete*, *Einmalzahlungen* u.v.a. Auch allgemeine Energie- und Umweltbegriffen wie *autofreier Tag* und *Temporeduktion* wurden nicht erkannt. Weiters fällt auf, dass allgemeine Termini, die keine Nomen sind, nur schwer erkannt werden, wie *inflationstreibend*, *wachstumsabschwächend*, *eingelagert* oder *importabhängig*. Man sollte hier ergänzen, dass bei diesen Termini die Grenze zwischen Kein Terminus

und allgemeiner Terminus in einigen Fällen nicht eindeutig ist und deshalb diese Ergebnisse auch zu erwarten waren.

Aus der Kategorie nicht-fachspezifische Termini werden viele Termini, die sich auf politische und rechtliche Konzepte, insbesondere im Zusammenhang mit Genehmigungsverfahren, politischen Maßnahmen und rechtlichen Rahmenbedingungen, Managementthemen, Finanzen oder BWL beziehen, nicht erkannt. Beispiele hierfür sind *Körperschaften öffentlichen Rechts*, *Genehmigungspolitik Österreichs*, *EU Sanktionen*, *Regierungsvorstoß*, *gefördertes Kontingent*, *Sicherheitspolster*.

Bei den Eigennamen gibt es ähnliche Erkennungsschwierigkeiten wie in den oben genannten Kategorien. Viele der nicht erkannten Termini bestehen aus mehreren Wörtern, wie *Faserhersteller Lenzing* oder aus mehreren möglich als einzelne Wörter erkennbaren Wörtern, wie *Pipeline Nord Stream 1* oder *Müllverbrennung Wien-Spittelau*. Weiters werden folgende Termini nicht erkannt: Eigennamen und spezifische Bezeichnungen, wie *Voestalpine*; politische und wirtschaftliche Begriffe, wie *Österreichs Außenminister Alexander Schallenberg*, *Papier- und Zellstoffkonzern Sappi* oder *türkis-grüne Regierung*; Abkürzungen und Akronyme, wie *ZiB 2*, *SPÖ* oder *UBS*; geografische und temporale Bezeichnungen, wie *Nordseesorte Brent*, *9. Mai*, *Kraftwerk Thieß*; sowie Bindestrich-Termini, wie *Werkzeugmaschinen-Hersteller Trumpf Austria*. Auffällig bei oben erwähnten Beispielen sind auch Schwierigkeiten bei der Erkennung eines Terminus, wenn dieser eine Zahl beinhaltet. Das wird auch durch Beispiele wie *Tempo 100*, *Tempo 130* oder *Global 2000* bestätigt.

5.2.2 Qualitative Analyse der Ergebnislisten von OneClick Terms

Beim System OneClick Terms mit Schwellenwert 1000 sind ähnlich wie beim D-Terminer einige Muster festzuhalten.

In der Kategorie fachspezifische Termini hat das System Schwierigkeiten mit der Erkennung von komplexeren Mehrwortbenennungen, die aus mehreren Wörtern bestehen, wie *Lieferstopp von russischem Erdgas*, *Abhängigkeit von russischen Gaslieferungen*, *Reduzierung der Energiekosten* oder *Verfügbarkeit von Energie*. Ein weiteres wiederkehrendes Muster scheinen die Bindestrich-Termini zu sein, wie *Atom-, Gas-, Kohle- und Laufkraftwerke*, *Atom-, Gas-, Kohle- oder Kohle- oder Atomkraftwerke*. Allerdings kann man hier beobachten, dass wenn der Bindestrich ein Teil eines Kompositums ist, dieser Terminus wohl erkannt wird, wie *Erneuerbaren-Energie-Ausbau* oder *Erneuerbaren-Ausbau*. Weiters fällt auf, dass das System

einige sehr spezifische Einwortbenennungen nicht erkennt, wie *Terawattstunden*, *Megawattstunden*, *Energiewende*, *Flüssiggas* oder *Energieeffizienz*. Was hier zusätzlich festzustellen ist, dass im Falle *Megawattstunde* zwar dieser Terminus erkannt wird, allerdings die Mehrzahl *Megawattstunden* nicht, was darauf deutet, dass das System Probleme mit Flexion und verschiedenen Wortformen haben könnte, wie auch das Beispiel *energieintensive Betriebe* vs. *energieintensiven Betrieben/energieintensiven Betriebe* zeigt. In unflektierter Form wird der Terminus erkannt, in flektierter nicht mehr. Ein weiteres deutlich zu erkennbares Muster bezieht sich auf Abkürzungen und Einheiten, wie *LNG*, *MWh*, *kWp* oder *LED*. Diese werden vom System One-Click Terms nicht erkannt.

In der Kategorie allgemeine Termini werden ähnliche Schwierigkeiten in Bezug auf Bindestrich-Termini, Mehrworttermini, Flexionen und verschiedene Wortformen der Termini festgestellt. Was hier noch deutlicher als in der Kategorie fachspezifische Termini wird, ist die Problematik der nicht erkannten Einwortbenennungen. Viele Termini, wie *Erdgas*, *Strom*, *Fotovoltaik*, *Windräder*, *CO₂-Emissionen*, *Kernkraft*, *Energieprojekte*, oder *Speicheranlage* sind für präzise Diskussionen über technische Details und Energiepolitik für den Energiesektor wichtig, werden aber vom System nicht als Termini erkannt. Was hier zusätzlich auffällt sind die Komposita, die in einer größeren Anzahl nicht erkannt werden, wie *Investitionspaket*, *Industrieunternehmen*, *Umweltschutzorganisation*, *Klimastrategie*, *Energieprojekte* u.v.m.

In der Kategorie nicht-fachspezifische Termini kommen Schwierigkeiten im Erkennen von Mehrwortbenennungen auf, wie *gefördertes Kontingent* oder *Preisanpassung an Indizes*, wobei hier insbesondere die Kombinationen von Adjektiven und Substantiven auffallen. Beispiele hierfür sind *konjunkturellen Normallage* oder *gedeckelter Großhandelspreis*. Ein weiteres Muster scheint bei Lehnwörtern aus dem Lateinischen aufzukommen, wie *Votum*, *Inflation*, *Kontingent*, *Dividenden* und *Stagflation*, welche nicht erkannt werden.

In der Kategorie Eigennamen fällt vor allem die Problematik der Nicht-Erkennung der Abkürzungen auf, wie *EU*, *GPA*, *ÖGB*, *ACER* oder *GIS*. Aus diesem Grund werden auch Termini, die aus mehreren Wörtern bestehen und eine Abkürzung enthalten, nicht erkannt, wie *Industriellenvereinigung OÖ*, *CEO Thilo Preß*, *EU-Kommission* oder *EU-Agentur ACER*. Die Kombination von Eigennamen und einer Funktion stellt auch Schwierigkeiten dar, wie *CEO Martin Bergsmann* oder *Vorstandsvorsitzender Peter Untersperger*, sowie auch die Kombination von Unternehmen und Produkten, wie *Salinen Austria*, *Pichler Ziegel* oder *Wienerberger Building Solutions*. Named Entities, die aus einem Wort bestehen, werden auch nicht erkannt, wie *Voestalpine*, *Sutterlüty*, *Spectra*, *Adeg* oder *Spar*. Weitere Erkennungsschwierigkeiten beziehen sich auf geographische Bezeichnungen, wie *Österreich* oder *Russland*, und historische

Referenzen, wie *Nazi-Deutschland* oder *Sowjetunion*. Zuletzt gibt es auch bei den Kombinationen mit einer Zahl Schwierigkeiten bei der Erkennung der Termini, wie *Energieeffizienzgesetzes 2020* oder *Pipeline Nord Stream 1*.

5.2.3 Qualitative Analyse der nicht im Goldstandard vorkommenden Termini

Die letzte Analyse betrifft die Ergebnislisten der Termini bzw. der möglichen Termini, die von den verwendeten Systemen gefunden wurden, aber sich nicht im Goldstandard befinden. Die quantitative Analyse hat hervorgebracht, dass das System D-Terminer 407 und das System OneClick Terms 623 potenzielle Termini erkannt haben. Bei diesen Zahlen stellt sich die Frage, ob sie eher auf die Fehlerquote des jeweiligen Systems oder des Goldstandards hinweisen, wobei die Termkandidaten, die beide Systeme als potenzielle Kandidaten erkannt haben, am interessantesten sind. Das aus dem Grund, weil hier am wahrscheinlichsten mögliche Termini vorkommen können, die bei der Erstellung des Goldstandards nicht annotiert resp. übersehen worden sind. In dieser Kategorie können am wahrscheinlichsten Termini vorkommen, die bei der Erstellung des Goldstandards nicht annotiert resp. übersehen wurden. Im ersten Schritt werden daher die beiden Ergebnislisten der Systeme zusammengeführt und mithilfe des Microsoft Excel Programms auf Dubletten untersucht. Diese Untersuchung hat hervorgebracht, dass es insgesamt 74 Kandidaten (siehe Tabelle 8) gibt, die von beiden Systemen als mögliche Kandidaten in Betracht gezogen werden.

türkis-grüne	Spar-Kaufmann	Harvard-Absolvent
Energieallianz	Kirchmayr-Schliesselberger	Warenkontrollen
IfW-Volkswirt	Spritpreis	Geschäftsbasis
Dominoeffekte	Vor-Krisenniveau	Temporeduktion
Nahversorgern	Preischaos	Ofenanlage
Blitzumfrage	Neongas	Abzuschöpfen
Lohnrunde	Dachziegelwerken	Gashandel
Gashahn	Megaprogramm	Stromhunger
Metallern	Stahlerzeugung	Autobahntankstellen
Reststoffverwertungskessel	STANDARD-Anfrage	Wocheneinkauf
Mauerziegelwerken	Energiebedingte	Oepping
SPÜO-Landesparteivorsitzenden	Konsumentenebene	Ölembargo
SPO-Bundesgeschäftsführer	STANDARD-Interview	Baustoffkonzern
Drei-Personen-Haushalts	Flachgläser	Erdgaslieferungen
Uralsöl	Energieintensives	Nahrungsmittelbereich
Doppel-Wumms	Papierkonzern	Großteilen
Wünsch-Dir-was-Politik	Kohlekessel	Greißler
Russland-Freundlichkeit	EU-Territorium	EU-Agentur
Salz-Gewinnung	Faserhersteller	Illwerke
Salz-Produktion	Industriekapitäne	Karl Nehammer
Metaller-Lohnverhandlungen	Marmeladegläser	Chinas Comeback
Konsumenten-Ebene	Energieversorgungen	Kampf gegen Gasknappheit
Frisch-Kaufmann	Technologieoffenen	Erdgastarif der Linz
Gasembargo	Primorsk	Europa Energieversorgung
ADEG-Märkte	Förderwesen	

Tabelle 8: Termini, die von beiden Systemen erkannt worden sind, sich aber nicht im Goldstandard befinden.

Nach näherer Betrachtung der Tabelle 8 wird ersichtlich, dass sich im Falle dieser Vorschläge größtenteils um keine Termini handelt. Lediglich bei 17 Kandidaten (grau markiert in der Tabelle 8) handelt es sich um Termini, die teilweise bei der Erstellung des Goldstandards zwar als Termini diskutiert worden sind, allerdings aufgrund keiner Übereinstimmung der Annotator*innen in den Goldstandard aufgenommen werden könnten, wie *Uralsöl*, oder ein Ergebnis gesplitteter Annotation sind, die durch einen Rechtschreibfehler zustande kommt, wie bei der Mehrwortbenennung *EU Gasembargo*, die bloß als *Gasembargo* von den Systemen erkannt ist. Nur in seltenen Fällen handelt es sich tatsächlich um bei der Erstellung des Goldstandards

übersehene Termini, wie *Kampf gegen Gasknappheit*, *Drei-Personen-Haushalts*, *Ölembargo*, *Energieintensives* oder *Energiebedingte*.

Was bei den einzelnen Ergebnislisten der potenziellen Termini einzeln betrachtet auffällt ist, dass das System D-Terminer überwiegend Einwortbenennungen als potenzielle Termini erkennt. Bei der näheren Betrachtung dieser Eiwortbenennungen stellt man fest, dass es sich meistens um allgemeine, in der Alltagssprache oft vertretene, Begriffe handelt, wie *Hemden*, *Jacken*, *Rasenmäher*, *Polizei*, *Freitag* u.v.m. Bei den Mehrwortbenennungen befinden sich genauso allgemeine Begriffe, wie *Kilo Mehl* oder *öffentliche Wahrnehmung*, oder nur unvollständig erkannte Termini, wie *Rohöl der Sorte*.

Was beim System OneClick Terms auffällt, ist die hohe Anzahl an Termini, die aus einer Kombination von Adjektiv und Nomen bestehen, wie *hohe Energiekosten*, *hoher Gaspreis*, *niedriger Deckel*, *zuständiges Klimaschutzministerium*, *gekommener Welthandel* oder *vorgezogene Preiserhöhung*. Solche Kombinationen mit allgemeinen Adjektiven wurden in den Richtlinien (siehe Abschnitt 4.2.1) ausgeschlossen und sind daher in dieser Arbeit als kein Terminus zu betrachten. Ein weiterer Problempunkt sind falsch segmentierte bzw. falsch zusammengefügte Mehrwortbenennungen. Es gibt einige Vorschläge an Termini, die keinen Sinn ergeben, wie *Erdgasabhängigkeit der nächste Punkt*, *Erdgastarif der Linz*, *Industrie von den gesunkenen Großhandelspreisen*, *Inflationswert bis Ende*, *Energieversorgung auf Platz*, *Energieversorgung besonderen Wert* oder *Energiewelt zu Verknappung*. An dieser Stelle ist anzumerken, dass obwohl das System OneClick Terms hier viele unbrauchbare Termkandidaten liefert, sich hier auch komplexere Termini bzw. Konzepte befinden, die durchaus interessant sein könnten, wenn man den Goldstandard um längere Termini erweitern würde. Hierzu gehören Termkandidaten wie *Kampf gegen Gasknappheit und Energiepreise*, *Engpass aufgrund des russischen Aggressionskriegs*, *Gewährleistung der sicheren Gasversorgung*, *Energieverbrauch je Fahrzeug*, *Espo-pipeline an die Pazifikküste bei Wladiwostok*, *Genehmigungsverfahren für Infrastruktur und Energieprojekte*, *Export als wirtschaftliches Erfolgsmodell* u.v.a.

5.2.4 Zusammenfassung der Ergebnisse

Aus den Ergebnissen der quantitativen Analyse geht hervor, dass das System D-Terminer bessere Leistung in den Kategorien Genauigkeit, Trefferquote und F1-Metrik erzielt (siehe Tabelle 6). Dies deutet darauf hin, dass D-Terminer eine insgesamt bessere Erkennung von Termini bietet, insbesondere in den Kategorien allgemeine Termini und Eigennamen (siehe Tabelle 7). Obwohl das System OneClick Terms dem System D-Terminer im direkten Vergleich in jeder Kategorie unterlegen ist, zeigt das System auch gewisse Stärken. Vor allem, wenn beim

OneClick Terms der Schwellenwert von 1000 auf 2000 verändert wird, erzielt das System in Kategorien fachspezifische Termini und nicht-fachspezifische Termini sogar bessere Ergebnisse als das System D-Terminer. Dies lässt vermuten, dass eine größere Anzahl von extrahierten Termini zu einer erhöhten Erkennungsrate spezifischer und seltener Termini führen kann, allerdings auf Kosten der Genauigkeit.

Im Laufe der Analyse der Ergebnisse könnte man einige Bereiche feststellen, in denen beide Systeme ihre Schwierigkeiten haben, allerdings unterschiedliches Grades. Die erste Problemstellung bei der Erkennung der Termini betrifft die Mehrwortbenennungen, insbesondere Genitiv-Phrasen oder Phrasen mit Präposition *von*. Solche Termini werden oft als nicht zusammengehörige Einheiten erkannt, was darauf schließen lässt, dass die Systeme dazu neigen, Wörter isoliert zu betrachten. In diesem Vergleich hat vor allem das System D-Terminer Schwierigkeiten, komplexere Mehrwortbenennungen als zusammenhängende terminologische Einheiten zu erkennen. Insbesondere, wenn diese aus mehr als zwei Wörter bestehen. Hier zeigt sich Verbesserungspotenzial bei der Erfassung von solchen Termini. Das System OneClick Terms hat hier ebenfalls Schwierigkeiten, allerdings werden hier, auch wenn nicht immer richtig, zu gewissem Grad auch komplexere Konzepte erkannt, die aus mehr als zwei Wörtern bestehen. Dagegen hat dieses System Schwierigkeiten beim Erkennen von Termini, die aus einer Kombination von Adjektiv und Nomen bestehen, da oft allgemeine Adjektive fälschlicherweise als Teil eines Terminus erkannt werden. Deshalb benötigen beide Systeme hier eine Optimierung, entweder in Anpassung der term grammar-Regeln (siehe Abschnitt 4.3) beim OneClick Terms oder Erweiterung der Trainingsdaten beim D-Terminer.

Ein weiterer Bereich, in dem beide Systeme Schwierigkeiten haben, ist die Erkennung von Termini, deren Bestandteil ein Bindestrich ist. Diese werden nicht immer erkannt, wobei beim System OneClick Terms man beobachten könnte, dass Komposita mit Bindestrich in der Regel erkannt werden, was beim System D-Terminer nicht immer der Fall ist, wie *Erneuerbaren-Energie-Ausbau* oder *Erneuerbaren-Ausbau*. Dies könnte durch Regelanpassungen oder Machine-Learning-Ansätze, die die Bindestrich-Konstruktionen explizit trainieren, verbessert werden.

Ein weiteres wiederkehrendes Problem stellen häufig Abkürzungen und Einheiten dar, da sie oft als nicht-terminologische Elemente betrachtet werden, wobei das System D-Terminer weniger Probleme mit der Erkennung von Abkürzungen und Einheiten hat, was darauf hindeutet, dass es eine größere Datenbank oder bessere Algorithmen für diese Art von Terminologie verwendet werden. Beim OneClick Terms treten hier häufiger Probleme auf, da das System Abkürzungen und Einheiten oft nicht erkennt, werden auch Mehrwortbenennungen oder

Komposita, die Abkürzungen beinhalten ignoriert. Eine gezielte Erweiterung der Datenbank mit Abkürzungen könnte hier helfen.

Eine weitere Fehlerquelle stellen Einwortbenennungen dar. Das System D-Terminer zeigt bessere Leistung in der Erkennung von Einwortbenennungen der Kategorien fachspezifische und allgemeine Termini als das System OneClick Terms. Vor allem bei der Kategorie allgemeine Termini kommt es zu häufigen Fehlern, was auf mangelnde Abdeckung dieser Art von Terminologie hinweist. Eine Verbesserung bzw. Erweiterung der Datenbank wäre hier deshalb sinnvoll.

Zuletzt haben beide Systeme Probleme mit flexierten Termini und Lehnwörtern, wobei vor allem das System OneClick Terms mehr Schwierigkeiten bei der Verarbeitung von Flexionen aufweist. Im Gegensatz zu D-Terminer hat das System OneClick Terms Probleme nicht nur mit Lehnwörtern aus Englischem aber auch aus Lateinischem, wie *Votum*, *Inflation*, *Kontingent* und *Stagflation*.

6 Diskussion

Die Ergebnisse dieser Arbeit decken sich mit bisherigen Erkenntnissen der ATE-Forschung überein, wie z.B. mit denen der Studie von Rigouts Terryn et al. (2022), die ebenfalls die Überlegenheit neuronaler Systeme gegenüber nicht-neuronalen Systemen im Großen und Ganzen hervorhebt. Das Experiment mit dem Tool D-Terminer zeigt, dass dieses neuronale System in allen gemessenen Metriken bessere Leistungen erzielt als das nicht-neuronale System OneClick Terms. Dieses Ergebnis bestätigt die Annahme, dass Deep-Learning-Ansätze bei der Terminologieextraktion insgesamt leistungsfähiger sind. Gleichzeitig weisen die Ergebnisse von OneClick Terms bei höheren Schwellenwerten darauf hin, dass auch nicht-neuronale Systeme unter bestimmten Bedingungen konkurrenzfähig sein können, z.B. bei der Erkennung spezifischer Termini durch Anpassung des Schwellenwertes. Die bedingte Konkurrenzfähigkeit nicht-neuronaler Systeme wird nicht nur durch die TermEval 2020 Studie von Rigouts Terryn et al. (2022) bestätigt, in der ein nicht-neuronales System den zweiten Platz belegte, sondern insbesondere auch durch die Studie von Hazem et al. (2022), die darauf hinweist, dass nicht-neuronale Systeme in der französischen Sprache ähnliche Ergebnisse wie neuronale Systeme liefern können. Schließlich ist auch die Überblicksstudie von Tran et al. (2023) zu beachten, die zeigt, dass es bei nicht-neuronalen Systemen durchaus Ansätze gibt, die manche neuronalen Systeme sogar übertreffen. Während Rigouts Terryn et al. (2022) und Drouin et al. (2020) in ihren Studien auf die Leistungsfähigkeit von BERT-Modellen hinweisen, zeigt unser Experiment, dass auch das auf der zuvor entwickelten RNN-Architektur basierende System bessere Ergebnisse im Vergleich zu einem nicht-neuronalen System erzielen kann. Im Gegensatz dazu bleibt das Tool OneClick Terms zwar hinter dem neuronalen System zurück, zeigt jedoch eine erhöhte Flexibilität bei der Erkennung seltener bzw. fachspezifischer Termini.

Eine der zentralen methodischen Herausforderungen dieser Arbeit bestand in der Erstellung eines Goldstandards. Die manuelle Annotation der Termini war durch die uneinheitliche Definition des Begriffs *Terminus* in der ATE geprägt. Es gibt nicht nur keine einheitliche Definition, sondern auch die Besonderheiten der Wirtschaftssprache, wie z.B. komplexe Nominalphrasen mit Genitiv, erschweren die ATE. Diese terminologische Komplexität spiegelt sich in den Ergebnissen beider Systeme wider. Während D-Terminer insbesondere bei der Erkennung komplexerer Termini Schwächen zeigte, war es bei OneClick Terms die Genauigkeit, die dem System Schwierigkeiten bereitete. Dies verdeutlicht die Notwendigkeit weiterer Forschung zur Modellierung komplexer Terminologieeinheiten, wie etwa bei Genitiv-Phrasen oder Komposita mit Bindestrich, sowie spezifischeren Trainingsmethoden für neuronale Systeme.

Eine weitere Herausforderung bei der Erstellung des Goldstandards zeigte sich bei der Kategorisierung der Termini, insbesondere bei der Unterscheidung zwischen fachspezifischen und allgemeinen Termini in journalistischen Texten. Diese Kategorisierung machte die manuelle Annotation zeitaufwendig, war jedoch notwendig, um die Ergebnisse der Systeme aus unterschiedlichen Perspektiven miteinander vergleichen zu können. Die Unterscheidung von Termini in spezifische Kategorien, wie von Rigouts Terryn (2021a) vorgeschlagen, erwies sich als anspruchsvoll. Für die Weiterentwicklung von ATE-Systemen wäre es hilfreich, die Annotation noch stärker zu standardisieren und klare Grenzen zwischen den verschiedenen Termkategorien zu ziehen. Darüber hinaus könnte eine stärkere Berücksichtigung kontextbezogener Informationen im Goldstandard zu noch präziseren Ergebnissen führen.

Zusammenfassend ist festzuhalten, dass das auf neuronalen Ansätzen basierte System D-Terminer bessere Ergebnisse als das auf nicht-neuronalen Ansätzen basierte System One-Click Terms liefert, was nicht nur die Annahme dieser Arbeit bestätigt, sondern auch die Entwicklungen in der ATE der letzten Jahre spiegelt. Da sich dieses Experiment auf zwei Systeme begrenzt, wäre es in weiterer Folge spannend dieses Experiment, um weitere Systeme zu erweitern. Insbesondere wäre ein Vergleich anderer neuronalen Systeme untereinander von Interesse, vor allem der Systeme, die auf der Transformer-Architektur (Vaswani et al. 2017) beruhen, wie das Text2TCS (Wachowiak et al. 2021).

Eine weitere Erweiterung dieses Experiments könnte den Korpus betreffen. Da der Korpus in dieser Arbeit aus Zeitungsartikeln besteht, also aus Texten, die an ein breites Publikum gerichtet sind, lässt sich vermuten, dass mit einem Korpus aus hochwissenschaftlichen Texten bessere Ergebnisse erzielt werden können. Es wäre interessant, wie ein bestimmtes System mit verschiedenen Korpora umgeht.

7 Fazit

Diese Arbeit befasste sich mit der ATE von deutschen Termini zum Thema Energiekrise, wobei zwei verschiedene Systeme, D-Terminer und OneClick Terms, miteinander verglichen wurden. Das Ziel dieser Arbeit bestand darin, die Leistungsfähigkeit beider Systeme anhand gängigen Metriken, die in der ATE häufig angewendet werden: Genauigkeit, Trefferquote und F1-Metrik zu evaluieren. Allerdings hat sich herausgestellt, dass die Evaluierung der Systeme bloß mithilfe dieser Metriken nicht ausreichend ist und deshalb wurde auch eine qualitative Auswertung der Kandidatenlisten vorgenommen. Erst die Zusammenführung beider Auswertungsmethoden, quantitativ und qualitativ, hat gezeigt, welches System unter den gegebenen Bedingungen die besten Ergebnisse liefert. Die Evaluierung erfolgte anhand eines spezifisch erstellten Korpus von Zeitungsartikeln und eines manuell annotierten Goldstandards, der die Grundlage für den Vergleich bildete.

Die Haupte Erkenntnisse dieser Arbeit zeigen, dass das neuronale System D-Terminer die besseren Ergebnisse in allen Metriken erzielt. Dies bestätigt die zentrale Annahme dieser Arbeit, dass Systeme mit neuronalen Ansätzen leistungsfähiger als nicht-neuronale Systeme sind. Diese Ergebnisse decken sich mit der aktuellen Forschung, die die Überlegenheit von Deep-Learning-basierten Ansätzen bei der Terminologieextraktion betont.

Allerdings zeigte trotz der insgesamt besseren Leistung von D-Terminer das nicht-neuronale System OneClick Terms seine Stärken bei der Erkennung bestimmter Termini, insbesondere bei höheren Schwellenwerten. Bei einem Schwellenwert von 2000 konnte OneClick Terms in spezifischen Kategorien, wie den fachspezifischen Termini, sogar bessere Ergebnisse erzielen. Dies deutet darauf hin, dass nicht-neuronale Systeme durch gezielte Anpassungen, wie die Wahl des optimalen Schwellenwerts, in bestimmten Anwendungsfällen konkurrenzfähig sein können. Somit zeigt sich, dass beide Ansätze je nach Einsatzgebiet und Zielsetzung ihre Berechtigung haben.

Die zentrale Erkenntnis, dass neuronale Ansätze wie D-Terminer in vielen Bereichen überlegen sind, bietet wertvolle Einblicke für die Entwicklung zukünftiger ATE-Systeme. Für die Zukunft lassen sich aus den Ergebnissen dieser Arbeit mehrere interessante Ansätze ableiten. Eine Erweiterung des Experiments auf weitere neuronale Systeme, insbesondere solche, die auf fortschrittlichen Architekturen wie Transformer-Modellen basieren, könnte neue Erkenntnisse liefern. Ebenso wäre es wertvoll, die Leistung der neuartigen Systeme auf anderen Arten der Korpora, wie wissenschaftlichen Texten, zu evaluieren. Dadurch ließe sich

besser verstehen, welche Leistungen die unterschiedlichen Systeme in verschiedenen Kontexten liefern.

Ein weiterer möglicher Ausblick wäre die Verbesserung der aktuellen Systeme in Bezug auf die Erkennung von Mehrwortbenennungen und komplexen Termini. Die Evaluierung könnte ebenfalls durch die Einbeziehung einer größeren Vielfalt an Kategorien der Termini und einer Verfeinerung des Goldstandards, etwa durch die stärkere Berücksichtigung von Kontextinformationen, erweitert werden.

Zusammenfassend lässt sich sagen, dass die Arbeit gezeigt hat, dass neuronale Systeme wie D-Terminer in der ATE überlegen sind, jedoch nicht-neuronale Systeme, wie OneClick Terms, in bestimmten Situationen weiterhin relevant sein können. Festzuhalten ist, dass die ATE viele Forschungsthemen liefert und weiterhin eine herausfordernde Aufgabe bleibt.

Literaturverzeichnis

- Adamzik, Kirsten (2018). Fachsprachen: die Konstruktion von Welten. Tübingen: A. Francke Verlag
- Arntz, Reiner; Picht, Heribert & Mayer, Felix (2009). Einführung in die Terminologearbeit. Hildesheim Zürich: Olms. 6., verb. Aufl
- Arntz, Reiner; Picht, Heribert & Schmitz, Klaus-Dirk (2014). Einführung in die Terminologearbeit. Hildesheim: Olms. 7., vollst. überarb. u. aktual. Auflage
- Astrakhantsev, N. A.; Fedorenko, D. G. & Turdakov, D. Yu (2015). Methods for automatic term recognition in domain-specific text collections: A survey. *Programming and Computer Software*, 41, 336–349
- Astrakhantsev, Nikita (2018). ATR4S: toolkit with state-of-the-art automatic terms recognition methods in Scala. *Language Resources and Evaluation*, 52 (3), 853–872
- Asudani, Deepak Suresh; Nagwani, Naresh Kumar & Singh, Pradeep (2023). Impact of word embedding models on text analytics in deep learning environment: a review. *Artificial Intelligence Review*, 56 (9), 10345–10425 DOI: 10.1007/s10462-023-10419-1
- Aubin, Sophie & Hamon, Thierry (2006). Improving term extraction with terminological resources. Paper präsentiert bei der „International Conference on Natural Language Processing (in Finland)“, Springer. 380–387
- Baisa, Vít; Michelfeit, Jan & Matuška, Ondřej (2017). Simplifying terminology extraction: Oneclick terms. Paper präsentiert bei der „Corpus Linguistics 2017 Conference, University of Birmingham“
- Bakaric, Marija Brkic; Babic, Nikola & Matetic, Maja (2021). Application-based Evaluation of Automatic Terminology Extraction. *International Journal of Advanced Computer Science and Applications*, 12 (1)
- Bordea, Georgeta; Buitelaar, Paul & Polajnar, Tamara (2013). Domain-independent term extraction through domain modelling. Paper präsentiert bei der „TIA 2013-10th International Conference on Terminology and Artificial Intelligence“
- Buhlmann, Rosemarie & Fearn, Anneliese (2000). Handbuch des Fachsprachenunterrichts: unter besonderer Berücksichtigung naturwissenschaftlich-technischer Fachsprachen. Gunter Narr Verlag
- Costa, Hernani; Zaretskaya, Anna; Corpas Pastor, Gloria & Seghiri Domínguez, Míriam (2016). Nine terminology extraction Tools: Are they useful for translators?
- Costa, Rute; Carvalho, Sara; Anić, Ana Ostroški & Khan, Fahad (2022). Proceedings of the Workshop on Terminology in the 21st century: many faces, many places. Paper präsentiert bei der „Proceedings of the Workshop on Terminology in the 21st century: many faces, many places“
- Crestani, Valentina (2010). Wortbildung und Wirtschaftssprachen: Vergleich deutscher und italienischer Texte. Bern: Peter Lang
- Daille, Béatrice; Kageura, Kyo & Terryn, Ayla Rigouts (2020). Proceedings of the 6th International Workshop on Computational Terminology. Paper präsentiert bei der „Proceedings of the 6th International Workshop on Computational Terminology“
- Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton & Toutanova, Kristina (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*,
- DIN 2342 (2022). Terminologie und Terminologielehre - Begriffe. Berlin: Beuth
- Dobrov, Boris V. & Loukachevitch, Natalia (2011). Multiple evidence for term extraction in broad domains. Paper präsentiert bei der „Proceedings of the international conference recent advances in natural language processing 2011“. 710–715

- Drewer, Petra & Schmitz, Klaus-Dirk (2017). Terminologiemanagement: Grundlagen - Methoden - Werkzeuge. Berlin [Heidelberg]: Springer Vieweg
- Drouin, Patrick; Morel, Jean-Benoît & L’Homme, Marie-Claude (2020). Automatic Term Extraction from Newspaper Corpora: Making the Most of Specificity and Common Features. Paper präsentiert bei der „Proceedings of the 6th International Workshop on Computational Terminology“. 1–7
- Filipec, Josef (1969). Zur Spezifik des spezialsprachlichen Wortschatzes gegenüber dem allgemeinen Wortschatz. na
- Gläser, Rosemarie (1990). Fachtextsorten im Englischen. Gunter Narr Verlag. Bd. 13
- Golomazov, DD (2012). Methods and tools for managing scientific information with ontologies. Cand. Sci.(Phys.-Math) Dissertation,
- Göpferich, Susanne (1995). Textsorten in Naturwissenschaften und Technik: pragmatische Typologie-Kontrastierung-Translation. Narr
- Gromann, Dagmar; Wachowiak, Lennart; Lang, Christian & Heinisch, Barbara (2021). Multilingual Extraction of Terminological Concept Systems. Deep Learning and Neural Approaches for Linguistic Data, 5
- GSÖD (2022). Das österreichische Wort des Jahres 2022. [<https://oewort.at/wort-des-jahres/2022/>] (Stand: 8.12.2023)
- Gyáfrás, Edit (1999). Fachlexik der Wirtschaft. Linguistische Studien aus Szeged,
- Hazem, Amir; Bouhandi, Mérieme; Boudin, Florian & Daille, Béatrice (2020). Termeval 2020: Taln-ls2n system for automatic term extraction. Paper präsentiert bei der „6th International Workshop on Computational Terminology (COMPUTERM 2020)“
- Heine, Carmen & Schubert, Klaus (2013). Modellierung in der Fachkommunikation. Fachsprache, 35 (3–4), 100–117
- Heylen, Kris & De Hertog, Dirk (2015). Automatic term extraction. Handbook of terminology, 1 (01)
- Hohnhold, Ingo (1990). Übersetzungsorientierte Terminologearbeit: eine Grundlegung für Praktiker. Stuttgart: InTra, 1. Fachübersetzer-genossenschaft eG
- Jerdhaf, Oskar; Santini, Marina; Lundberg, Peter; Bjerner, Tomas; Al-Abasse, Yosef; Jönsson, Arne & Vakili, Thomas (2022). Evaluating Pre-Trained Language Models for Focused Terminology Extraction from Swedish Medical Records. Paper präsentiert bei der „Proceedings of the Workshop on Terminology in the 21st century: many faces, many places“. 30–32
- Kageura, Kyo (2015). Terminology and lexicography. Handbook of terminology, 1 (45–59)
- Kageura, Kyo & Umino, Bin (1996). Methods of automatic term recognition: A review. Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication, 3 (2), 259–289
- Kilgarrieff, Adam; Rychlý, Pavel; Smrz, Pavel & Tugwell, David (2004). The Sketch Engine. In: Proceedings of the 11th EURALEX International Congress. Lorient, France. 105–115 <https://api.semanticscholar.org/CorpusID:13188196>
- Kim, J. D.; Ohta, Tomoko; Tateisi, Yuka & Tsujii, Jun’ichi (2003). GENIA corpus—a semantically annotated corpus for bio-textmining. Bioinformatics, 19 (suppl_1), i180–i182
- KÜDES – Konferenz der Übersetzungsdienste europäischer Staaten (Hrsg.) (2018). Empfehlungen für die Terminologearbeit. Schweizerische Bundeskanzlei
- Li, Sujian; Li, Jiwei; Song, Tao; Li, Wenjie & Chang, Baobao (2013). A novel topic model for automatic term extraction. Paper präsentiert bei der „Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval“. 885–888
- Lossio-Ventura, Juan Antonio; Jonquet, Clement; Roche, Mathieu & Teisseire, Maguelonne (2013). Combining c-value and keyword extraction methods for biomedical terms extraction. Paper präsentiert bei der „LBM: Languages in Biology and Medicine“

- Maghețiu, Anca (2015). Zur Didaktik der Fachsprachen: mit besonderer Berücksichtigung des Deutschen als Wirtschaftssprache. Hamburg: Kovač
- Mahesh, Batta (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], 9 (1), 381–386
- Ogden, Charles K. & Richards, Ivor A. (1956). *The meaning of meaning : a study of the influence of language upon thought and of the science of symbolism*. London: Routledge & Kegan. 10. ed. (4. impr.) <https://ubdata.univie.ac.at/AC03188457>
- Pennington, Jeffrey; Socher, Richard & Manning, Christopher D. (2014). Glove: Global vectors for word representation. Paper präsentiert bei der „Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)“. 1532–1543
- Rigouts Terryn, Ayla (2021a). ACTER Terminology Annotation Guidelines. <http://hdl.handle.net/1854/LU-8503113>
- Rigouts Terryn, Ayla (2021b). D-TERMINE: data-driven term extraction methodologies investigated.
- Rigouts Terryn, Ayla; Hoste, Véronique; Drouin, Patrick & Lefever, Els (2020). Termeval 2020: Shared task on automatic term extraction using the annotated corpora for term extraction research (acter) dataset. Paper präsentiert bei der „6th International Workshop on Computational Terminology (COMPUTERM 2020)“, European Language Resources Association (ELRA). 85–94
- Rigouts Terryn, Ayla; Hoste, Veronique & Lefever, Els (2022). D-Terminer: Online Demo for Monolingual and Bilingual Automatic Term Extraction. Paper präsentiert bei der „Proceedings of the TERM21 Workshop“, Language Resources and Evaluation Conference (LREC 2022). 33–40
- Roelcke, Thorsten (2019). Fachsprache. Vortrag, Ludwig-Maximilians-Universität, München. Verfügbar unter: https://epub.ub.uni-muenchen.de/61967/1/Roelcke_Fachsprache.pdf (12.04.2023),
- Roelcke, Thorsten (2020). Fachsprachen. Berlin: Erich Schmidt Verlag. 4., neu bearbeitete und wesentlich erweiterte Auflage
- Rokas, Aivaras; Utkla, Andrius & Rackevičienė, Sigita (2020). Automatic extraction of lithuanian cybersecurity terms using deep learning approaches. Paper präsentiert bei der „Human language technologies–the baltic perspective: proceedings of the ninth international conference baltic HLT 2020“. Bd. 328. 39
- Šajatović, Antonio; Buljan, Maja; Šnajder, Jan & Bašić, Bojana Dalbelo (2019). Evaluating automatic term extraction methods on individual documents. Paper präsentiert bei der „Proceedings of the Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019)“. 149–154
- Shuttleworth, Mark & Cowie, Moira (1997). *Dictionary of translation studies*. Manchester, UK: St. Jerome Pub
- Sketch Engine (o. J.) About OneClick Terms. <https://terms.sketchengine.eu/about> (Stand 21.01.2023).
- Statista Research Departement (2023). Energiekrise in Österreich: Herausforderungen und Perspektiven. <https://de.statista.com/themen/11741/energiekrise-in-oesterreich> (Stand: 8.12.2023)
- Tran, Hanh Thi Hong; Martinc, Matej; Caporusso, Jaya; Doucet, Antoine & Pollak, Senja (2023). The Recent Advances in Automatic Term Extraction: A survey. *arXiv preprint arXiv:2301.06767*,
- Vaswani, Ashish; Shazeer, Noam; Parmar, Niki; Uszkoreit, Jakob; Jones, Llion; Gomez, Aidan N.; Kaiser, Łukasz & Polosukhin, Illia (2017). Attention is all you need. *Advances in neural information processing systems*, 30
- VMA (2022). MA 2022 Presse. <https://www.media-analyse.at/table/3887> (Stand: 8.12.2023)

- Wachowiak, Lennart; Lang, Christian; Heinisch, Barbara & Gromann, Dagmar (2021). Towards learning terminological concept systems from multilingual natural language text. Paper präsentiert bei der „3rd Conference on Language, Data and Knowledge (LDK 2021)“, Schloss Dagstuhl-Leibniz-Zentrum für Informatik
- Wendt, Susanne (1997). *Terminus, Thesaurus, Text: Theorie und Praxis von Fachbegriffssystemen und ihrer Repräsentation in Fachtexten*. Tübingen: G. Narr
- Witschel, Hans Friedrich (2004). Text, Wörter, Morpheme–Möglichkeiten einer automatischen Terminologie-Extraktion. Vortrag, Universität Leipzig, Institut für Informatik, Leipzig. Verfügbar unter: <http://wortschatz.uni-leipzig.de/~fwitschel/papers/GLDVPreis.pdf> (25.9. 2009),
- Wüster, Eugen (1959). Das Wort in der Welt, schaubildlich und terminologisch dargestellt. *Sprachforum*, 3 (4), 183–203
- Yee, Olivia Kwong Oi (2018). Evaluating term extraction tools: System performance vs user perception. In: *The Human Factor in Machine Translation*. Routledge. 100–117

Anhang A

Verzeichnis der Zeitungsartikel:

- APA (2022a). Finanzministerium prüft Gewinnabschöpfung wegen steigender Energiepreise. Der Standard, 06.05.2022. <https://www.derstandard.at/story/2000135494990/finanzministerium-stellt-kopplung-von-strom-und-gaspreis-infrage> (Stand: 18.04.2024)
- APA (2022b). Mikl-Leitner für Strompreisdeckel – Finanzminister Brunner warnt vor Gefahren. Der Standard, 10.07.2022. <https://www.derstandard.at/story/2000137313967/mikl-leitner-fuer-strompreisdeckel-finanzminister-brunner-warnt-vor-gefahren> (Stand: 18.04.2024)
- APA (2022c). Nach Preiserhöhung bei Wien Energie und EVN: Andere Landesversorger warten ab. Der Standard, 04.08.2022. <https://www.derstandard.at/story/2000138034677/nach-preiserhoehung-bei-wien-energie-und-evn-andere-landesversorger-warten> (Stand: 18.04.2024)
- APA (2023a). Haushaltsenergiepreise stiegen im November im Schnitt um 42 Prozent. Der Standard, 10.01.2023. <https://www.derstandard.at/story/2000142418920/haushaltsenergiepreise-stiegen-im-november-im-schnitt-um-42-prozent> (Stand: 18.04.2024)
- APA (2023b). Österreich und Deutschland einigen sich auf Nutzung des Gasspeichers Haidach. Der Standard, 17.02.2023. <https://www.derstandard.at/story/2000143662932/oesterreich-und-deutschland-einigen-sich-auf-nutzung-von-gas-speicher> (Stand: 18.04.2024)
- Arends, Johannes (2022). Kein Gas aus Russland: In Bulgarien droht ein kalter und damit heißer Herbst. Kurier, 10.05.2022. <https://epaper.kurier.at/issue.act?issueId=115918&newsItemId=9905413&issueDate=20220510&issueMutation=12> (Stand: 18.04.2024)
- Bachner, Michael (2022). Inflation kratzt an 10-Prozent-Marke. Kurier, 17.09.2022. <https://epaper.kurier.at/issue.act?issueId=117817&newsItemId=10025675&issueDate=20220917&issueMutation=12> (Stand: 18.04.2024)
- Bachner, Michael (2023). Sonne, Wind, Wasser – und Atomkraft. Kurier, 09.02.2023. <https://epaper.kurier.at/issue.act?issueId=119969&issueDate=20230209&issueMutation=12&newsItemId=10162684> (Stand: 18.04.2022)
- Danzer, Andreas; Ungerboeck, Luise (2022). Die Strompreisbremse kommt: Was sie für die Bevölkerung bedeutet. Der Standard, 05.09.2022. <https://www.derstandard.at/story/2000138804166/die-strompreisbremse-kommt-was-sie-fuer-die-bevoelkerung-bedeutet> (Stand: 18.04.2024)
- Gepp, Joseph (2022). Energiesparen: lässt türkis-grüne Regierung Potenzial ungenutzt? Der Standard, 22.06.2022. <https://www.derstandard.at/story/2000136760235/energiesparenlaesst-es-die-tuerkis-gruene-regierung-potential-ungenutzt-liegen> (Stand: 18.04.2024)
- Hahn, Alexander; Ungerboeck, Luise (2022). Wie teuer wird die Energiekrise? Der Standard, 07.03.2022. <https://www.derstandard.at/story/2000133911991/wie-teuer-wird-die-energiekrise> (Stand: 18.04.2024)
- Kotrba, David; Meyrath, Martin (2022). Stromschlag für die Nahversorger. Kurier, 20.12.2022. <https://epaper.kurier.at/issue.act?issueId=119263&newsItemId=10117916&issueDate=20221220&issueMutation=12> (Stand: 18.04.2024)
- Pressberger, Thomas (2023). „Verstehe diese Doppelmoral nicht“. Kurier, 04.01.2023. <https://epaper.kurier.at/issue.act?issueId=119462&newsItemId=10130072&issueDate=20230104&issueMutation=12> (Stand: 18.04.2024)

- Scheidinger, Herta (2022a). Industrie sucht Wege aus der Energiekrise. Kurier, 03.04.2022. <https://epaper.kurier.at/issue.act?issueId=115371&newsitemId=9873735&issueDate=20220403&issueMutation=ooe-beilage> (Stand: 18.04.2024)
- Scheidinger, Herta (2022b). Die Energiekrise richtig einschätzen und bewältigen. Kurier, 29.05.2022. <https://epaper.kurier.at/issue.act?issueId=116184&newsitemId=9921647&issueDate=20220529&issueMutation=ooe-beilage> (Stand: 18.04.2024)
- Scheidinger, Herta (2022c). Energiekrise trifft die Industrie. Kurier, 02.10.2022. <https://epaper.kurier.at/issue.act?issueId=118008&newsitemId=10036977&issueDate=20221002&issueMutation=ooe-beilage> (Stand: 18.04.2024)
- Steiner-Gashi, Ingrid; Meyrath, Martin (2022). Warum der Gaspreis nicht mehr sinken wird. Kurier, 24.11.2024.
- Strobl, Günther; Laufer Nora (2022). Was, wenn kein russisches Gas mehr nach Österreich kommt? Der Standard, 02.04.2022. <https://www.derstandard.at/story/2000134618526/was-wenn-kein-russisches-gas-mehr-nach-oesterreich-kommt> (Stand: 18.04.2024)
- Strobl, Günther (2022a). Normalisierung bei Energiepreisen frühestens 2024 erwartet. Der Standard, 02.11.2022. <https://www.derstandard.at/story/2000140482976/normalisierung-bei-energiepreisen-fruehestens-2024-erwartet> (Stand: 18.04.2024)
- Strobl, Günther (2022b). Sanktionen zeigen Wirkung, russisches Rohöl viel weniger wert. Der Standard, 08.12.2022. <https://www.derstandard.at/story/2000141635247/sanktionen-zeigen-wirkung-russisches-rohoel-viel-weniger-wert> (Stand: 18.04.2024)
- Ungerboeck, Luise (2022). Klimaministerium will mehr Befugnisse für den Energienotfall. Der Standard, 05.10.2022. <https://www.derstandard.at/story/2000139723123/klimaministerium-will-mehr-befugnisse-fuer-den-energienotfall> (Stand: 18.04.2024)
- Unterhuber, Wolfgang (2022). Stromschlag für die Nahversorger. Kurier, 22.10. 2022. <https://epaper.kurier.at/issue.act?issueId=118332&newsitemId=10057424&issueDate=20221022&issueMutation=l2> (Stand: 18.04.2024)
- Wobrazek, Sandra (2022a). Die Suche nach der Energielösung. Kurier, 01.05.2022. <https://epaper.kurier.at/issue.act?issueId=115783&newsitemId=9897739&issueDate=20220501&issueMutation=ooe-beilage> (Stand: 18.04.2024)
- Wobrazek, Sandra (2022b). Die Wege aus der Energiekrise. Kurier, 04.09.2022. <https://epaper.kurier.at/issue.act?issueId=117596&newsitemId=10011002&issueDate=20220904&issueMutation=ooe-beilage> (Stand: 18.04.2024)

Anhang B

Tabellarische Übersicht der Kategorien von Termini

Ist die Benennung eng mit dem Fachgebiet Energiewirtschaft bzw. dem Thema Energiekrise verbunden?

Ja			Nein		
Eigenschaften	Anmerkungen	Kategorie	Eigenschaften	Anmerkungen	Kategorie
Die Benennung wird ausschließlich von Spezialist*innen verwendet. • lexikonspezifisch	In Medien erscheint sie meistens mit einer Erklärung.	Specific Term	Die Benennung wird ausschließlich von Spezialist*innen verwendet. • lexikonspezifisch	Die Benennung kommt aus einem anderen Fachgebiet und hat dort eine bestimmte Bedeutung, die jedoch für das aktuelle Fachgebiet nichtzutreffend ist.	Out-Of-Domain Term
Die Benennung wird auch von Lai*innen, die eine vage Vorstellung von ihrer Bedeutung haben, verwendet. • lexikonspezifisch • dem Fachgebiet/Thema meistens auch ohne Kontext zuzuordnen	Zwar erscheint sie in den Medien ohne Erklärung, aber es benötigt bestimmtes Fachwissen, um sie richtig anzuwenden.	Specific Term	Die Benennung wird auch von Lai*innen, die eine vage Vorstellung von ihrer Bedeutung haben, verwendet. • lexikonspezifisch • oft einem anderen konkreten Fachgebiet/Thema zuzuordnen	Erscheint in den Medien ohne Erklärung, ist aber einem anderen Fachgebiet oder Thema zuzuordnen, z.B. der Politik, den Finanzen, dem Umweltschutz etc.	Out-Of-Domain Term
Die Benennung wird von Lai*innen verstanden und verwendet. Sie ist so eng mit dem Thema Energiekrise verbunden, dass sie zweifelsfrei dem Thema zuzuordnen ist.	Erscheint in den Medien ohne Erklärung. Sie scheint als Common Term, ist aber sehr eng mit dem Thema verbunden.	Specific Term			
Die Benennung wird von Lai*innen häufig verwendet. • nicht-lexikonspezifisch	Die Benennung wird im allgemeinen Wortschatz oft angewendet bzw. das beschreibende Konzept ist allgemein bekannt.	Common Term	Die Benennung wird von Lai*innen häufig verwendet. • nicht-lexikonspezifisch	Sie wird im allgemeinen Sprachgebrauch verwendet.	No Term
Die Benennung bezieht sich auf ein real existierendes Objekt wie eine Person, einen Ort, eine Organisation oder ein Produkt.	Wenn es sich um einen konkreten und präzisen benannten Sachverhalt handelt.	Named Entity	Die Benennung bezieht sich auf ein real existierendes Objekt wie eine Person, einen Ort, eine Organisation oder ein Produkt	Wenn es sich um einen konkreten und präzisen benannten Sachverhalt handelt.	Named Entity
	Wenn es sich um einen Oberbegriff handelt, der eine ganze Gruppe von Personen, Produkten etc. zusammenfasst.	Common Term		Wenn es sich um einen Oberbegriff handelt, der eine ganze Gruppe von Personen, Produkten etc. zusammenfasst.	No Term

Anhang C

Goldstandard

Benennung	Kategorie	Häufigkeit
Energiekrise	ST	19
Energiekrisen	ST	1
Lieferstopp von russischem Erdgas	ST	1
Lieferstopp	ST	2
russischem Erdgas	ST	2
Erdgas	CT	19
Energieversorgung	ST	11
Ukraine-Krieges	NE	2
Strom	CT	28
Wasser	CT	3
Joachim Haindl-Grutsch	NE	3
IV OÖ	NE	6
Genehmigungspolitik	OOD	2
Genehmigungspolitik Österreichs	OOD	1
Österreichs	NE	9
Österreich	NE	62
Energiemenge	ST	1
Energiepolitik Österreichs	ST	1
Energiepolitik	ST	4
Energie- und Mobilitätssystems	ST	1
Genehmigungsverfahren	OOD	5
Haindl-Grutsch	NE	16
Infrastruktur für die Energieversorgung	ST	1
Infrastruktur	CT	5
fossilen Energieträgern	ST	1
Industriellenvereinigung OÖ	NE	1
Markt- und Meinungsforschungsinstitut Spectra	NE	1
Spectra	NE	1
Bewältigung der Krise	CT	1
Russland	NE	28
Haushalte	CT	12
Energiequellen	ST	2
Wasserkraft	ST	2
Sonnenenergie	ST	1
Windkraft	ST	4
Biomasse	ST	1
fossilen Energieträger	ST	1
Energie	CT	28
Preisgünstigkeit	OOD	1
Umweltschonung	CT	1
Umweltschutz	CT	1
Erdgasabhängigkeit	ST	1
Abhängigkeit von russischen Gaslieferungen	ST	1
Gaslieferungen	ST	8
Energiesystem	ST	2
Windräder	CT	2
Energiekosten	ST	18
energieintensiven Betrieben	ST	1

energieintensiven	ST	7
Votum	OOD	1
energieintensiven Betriebe	ST	2
Grundversorgung der Bevölkerung	ST	1
Grundversorgung	ST	1
Wertschöpfungsketten	OOD	2
Klimaziele	CT	2
Stromerzeugung	ST	3
erneuerbar zu machen	ST	1
Ausstieg aus der fossilen Energie	ST	1
fossilen Energie	ST	1
Energie- und Klimapolitik	ST	2
Energielösung	CT	1
Industrieunternehmen	CT	6
Salinen Austria	NE	2
Energiepreissteigerungen	ST	1
energieintensives Industrieunternehmen	ST	1
Energieintensives	ST	1
Energiepreisen	ST	5
Vorstandsvorsitzender Peter Untersperger	NE	1
Peter Untersperger	NE	2
Gasembargos gegenüber Russland	ST	1
Gasembargos	ST	1
Wirtschaftsablauf	ST	1
Fahrerpersonal	CT	1
Frächtern	CT	1
Materialbeschaffung	CT	1
Produktion	CT	13
Investitionspaket	CT	1
Beschichtungsspezialist Hueck Folien	NE	1
Hueck Folien	NE	1
europäischen Markt	CT	1
Internationalen Märkten	CT	1
CEO Martin Bergsmann	NE	1
Martin Bergsmann	NE	2
Erhöhungen der Rohstoffpreise	ST	1
Rohstoffpreise	ST	1
Transportkapazitäten	ST	1
Aufrechterhaltung der Lieferketten	ST	1
Lieferketten	CT	7
Bergsmann	NE	3
Lieferausfälle	ST	2
Werkzeugmaschinen-Hersteller Trumpf Austria	NE	1
Trumpf Austria	NE	1
Instabile	CT	1
Lieferzeiten	CT	2
CEO Thilo Preß	NE	1
Thilo Preß	NE	2
Kunststoffverarbeiter Röchling	NE	1
Röchling	NE	5
Rohstoffen	CT	3
Krieg in der Ukraine	NE	5
Röchling Industrial Oeppling	NE	1
Peter Eckerstorfer	NE	2

Röchling Industrial	NE	2
nachhaltige Energie	ST	1
Hueck	NE	1
Diversifizierung der Erdgasversorgung	ST	1
Diversifizierung	OOD	2
Erdgasversorgung	ST	1
Stärkung der Wasserkraft	ST	1
Industrialisierung von Technologien	OOD	1
Ersatz von fossilen Brennstoffen	ST	1
fossilen Brennstoffen	ST	1
nachhaltige Energieträger	ST	1
Energieträger	ST	8
Energieträgern	ST	3
energieintensive Industrie	ST	3
Steuerbelastung bei den Energiekosten	ST	1
Energieforecasts	ST	2
Klimastrategie	CT	1
Energielösungen	ST	1
Fotovoltaikanlage	ST	1
kWp	ST	1
MWh	ST	4
Laserhybridanlage	OOD	1
Vorwärme	ST	1
Gasverbrauch	ST	3
Einkaufsvereinigungen	CT	1
Reduzierung der Energiekosten	ST	1
Verfügbarkeit von Energie	ST	1
energieintensiven Unternehmen	ST	1
Umstellung auf alternative und nachhaltige Energien	ST	1
Gasverknappung	ST	1
Gasengpässe	ST	1
energieintensiven Industrieunternehmen	ST	1
Energieformen	ST	1
Verteuerung	ST	2
Pichler Ziegel	NE	1
CEO Ekkehart Pichler	NE	1
Teuerung	ST	6
Pichler	NE	4
Energieabhängigkeit	ST	1
Energieeinsatz	ST	1
Stromverbrauch	ST	5
Tonzerkleinerungsmaschinen	OOD	1
Brennertechnologie	OOD	1
Hilfsaggregat	ST	1
Hallenbeleuchtung	CT	1
LED	ST	1
Fotovoltaik-Anlage	ST	1
Gasbezugsquellen	ST	3
Voestalpine	NE	2

voestalpine AG	NE	1
Wärmebehandlungsaktivitäten	ST	1
Walzwerke	OOD	1
Gaslieferstopp	ST	1
EU Gasembargo	ST	1
Speicher	CT	5
Gasbezugsquellen	ST	3
Herbert Eibensteiner	NE	1
LNG-Terminals	NE	1
Einspeicherung	CT	1
Alternative Energien	ST	2
Christian Klinger	NE	2
Internorm	NE	3
Nachfrage nach energieeffizienten Fenstern und Türen	ST	1
Energieeffizienten	ST	1
Klinger	NE	3
thermische Sanierung	ST	1
Gaskrise	ST	2
Energiebedarf	ST	2
Abhängigkeit von Rohstoffen	ST	1
Gespart	CT	1
Eingespart	CT	1
Umwelt	CT	1
energieintensiver Lebensmittelproduzent	ST	1
energieintensiver	ST	1
energieintensive	ST	5
CEO Peter Weidinger	NE	1
Peter Weidinger	NE	2
Weidinger	NE	3
Biogasanlagen	ST	1
Fotovoltaik	CT	1
Pipelines	ST	2
EU Sanktionen	OOD	2
Krise	CT	10
Europäische Pipelines	ST	1
intraeuropäische Pipelines	ST	1
Ekkehart Pichler	NE	2
Merit-Order	ST	3
Merit Order	ST	1
Abrechnungssysteme	OOD	1
Marktversagen	ST	1
Entlastungspakete	CT	2
Hochreiter Fleischwaren	NE	2
Entkoppelung des Strompreises vom Gaspreis	ST	1
Strompreises	ST	2

Gaspreis	ST	7
Strom- und Gaspreis	ST	3
stillgelegten Kraftwerken	ST	1
Kraftwerken	CT	1
Atomkraft	CT	5
energieintensiven Branchen	ST	1
Metall-, Papier-, Kunststoff-, Nahrungsmittel-, Baustoff- oder Chemieindustrie	CT	1
Metall-	CT	1
Papier-	CT	2
Kunststoff-	CT	1
Nahrungsmittel-	CT	1
Baustoff-	CT	1
Chemieindustrie	CT	2
Stornierungen von Investitionsprojekten	ST	1
Industriellenvereinigung Oberösterreich	NE	1
Anstieg der Energiekosten	ST	1
Produktionsstopps	ST	1
Fixpreise	ST	1
energieintensive Betriebe	ST	2
Energiepreise	ST	14
Energiekostenanteil	ST	1
Gewinnmargen	OOD	1
Energie-Kostenniveaus	ST	1
Adaptierung der Merit Order	ST	1
Entkopplung von Strom- und Gaspreis	ST	2
Bundesregierung	NE	6
Energiekostenzuschüsse	ST	1
Stillstände der Produktion	ST	1
Kurzarbeit	CT	2
Stromproduktion	ST	2
Erneuerbarer Energieträger	ST	2
Ökostromanlagen	ST	1
Grundlastkapazitäten	ST	2
Atom-, Gas-, Kohle- und Laufkraftwerke	ST	1
Atom-	ST	1
Gas-	ST	2
Kohle-	ST	3
Laufkraftwerke	ST	1
Photovoltaik	CT	1
Windstrom	CT	1
Elektrifiziert	CT	1
CO2-neutralem Strom	ST	1
Wasserstoff	CT	2
E-Fuels	ST	1

Gas	CT	67
Russlands	NE	5
9. Mai	NE	1
Tag des Sieges	NE	1
Roten Armee	NE	1
Nazi-Deutschland	NE	1
Deutschland	NE	17
Ukraine	NE	21
Witoscha	NE	1
Sowjetunion	NE	1
Wessela Wladkowa	NE	1
Ostalgie	OOD	1
ostalgotischer Bulgaren	OOD	1
Kiril Petkow	NE	1
Vier-Parteien-Koalition	OOD	1
Petkows pro-westliche Partei	OOD	1
DB	NE	1
Staatspräsident Rumen Radew	NE	1
Rumen Radew	NE	1
Österreichs Außenminister Alexander Schallenberg	NE	1
Alexander Schallenberg	NE	1
russischen „Erpressungsversuche“	CT	1
Privathaushalte	CT	3
Privathaushalten	CT	1
Gas-Stopp	CT	1
Reserven	CT	1
Obergrenze beim Strompreis	ST	1
Pipeline	ST	2
Wladkowa	NE	2
Ostalgotischen	OOD	1
Lieferkettenprobleme	ST	2
Industriebundesland	CT	1
Anstieg der Energiepreise	ST	1
Ukraine-Konflikts	NE	1
Energie	CT	32
Importiert	CT	3
Vorprodukte	CT	1
Rohstoffe	CT	6
Autoindustrie	CT	2
Automobil- und Elektronikindustrie	CT	1
Automobil-	CT	1
Lieferkettenproblematik	ST	1
Vorprodukt	CT	1
Vorprodukte	CT	1
Gesamtprodukt	CT	1
Fossile Energie	ST	1
Lieferungen von fossiler Energie	ST	1
fossiler Energie	ST	1
Frühwarnstufe	ST	1
Notfallplan für die Gasversorgung	ST	1
Gasversorgung	ST	9
Rubel	NE	1
Energielenkungsmaßnahmen	ST	2

Rationierungen	ST	1
Embargo auf russisches Gas	ST	1
russisches Gas	ST	5
Embargo	CT	1
Gasmengen	ST	1
Diversifiziert	OOD	1
Abhängigkeit von russischem Erdgas	ST	1
russischem Erdgas	ST	1
Energiepreiswelle	ST	1
Papierindustrie	CT	1
Automotive Industrie	CT	1
Energiepreisentwicklung	ST	1
Vor-Krisenniveau	CT	1
konjunkturellen Normallage	OOD	1
Stromkosten	ST	6
Energiesektor	ST	1
Modell der „Strompreiskompensation“	NE	1
Strompreiskompensation	ST	2
CO2- und Energiepreise	ST	1
stromintensiven Unternehmen	ST	1
stromintensiven	ST	1
Energiepreiskrise	ST	1
Transformation des Energiesystems	ST	1
Energiesystems	ST	1
Energie-	ST	4
Energieinfrastruktur	ST	1
Ausbau der erneuerbaren Energien	ST	2
erneuerbaren Energien	ST	4
Energieprojekte	CT	1
Energieautark	ST	1
Energiemengen	CT	1
Inflation	CT	2
Benzin	CT	6
Diesel	CT	9
Teuerung	CT	6
ÖGB	NE	3
Inflationsdaten	CT	1
Statistik Austria	NE	2
Treibstoffe	CT	2
Treibstoffen	CT	2
Haushaltsenergie	CT	1
Inflationsrate	CT	3
August-Inflation	NE	1
Spritpreise	ST	4
Strompreiserhöhungen	ST	2
Wien Energie	NE	7
EVN	NE	14
Inflationswerte	OOD	1
Mikrowarenkorb	OOD	1
Miniwarenkorb	OOD	1
Jahresvergleich	OOD	5
GPA	NE	1
Inflationswert	OOD	1
Einmalzahlungen	CT	1

Teuerungsprämien	CT	1
EU	NE	18
Energiewende	ST	4
IEA-Chef Fatih Birol	NE	1
IEA-Chef	ST	2
Fatih Birol	NE	3
Chef der Internationalen Energieagentur	ST	2
Internationalen Energieagentur	NE	2
IEA	NE	4
Klimawandel	CT	2
Energiesicherheit	ST	2
Birol	NE	8
TU Wien	NE	1
erneuerbare Energieträger	ST	1
Energie-Souveränität	ST	1
importabhängig	CT	1
Ausbau von Wind- und Solarkraftwerken	ST	1
Wind-	ST	1
Solarkraftwerken	ST	1
Fahrzeugen mit Elektroantrieb	ST	1
Elektroantrieb	ST	1
E-Autos	CT	1
Preissprünge für Konsumentinnen und Konsumenten	ST	1
Konsumentinnen	CT	1
Konsumenten	CT	1
Speichern	CT	4
Flüssiggas-Abnehmer	ST	1
Weltmarkt	CT	3
Kubikmeter	CT	2
Einsparungen	CT	5
KURIER	NE	4
europäische Solidarität	ST	2
russischen Energielieferungen	ST	1
Energielieferungen	ST	1
Erneuerbaren-Energie-Ausbau	ST	1
Sonne	CT	1
Wind	CT	1
Erneuerbare	ST	2
Versorgungssicherheit	ST	3
Preisanstiegen	CT	1
Preisanstiege	CT	2
Gas-Großhandel	ST	1
Blackouts	ST	1
Auffüllung der Speicher	ST	1
Preissteigerungen der Stromanbieter	ST	1
Stromanbieter	ST	1
Energiekostenhilfen	ST	1
Kernkraft	CT	2
Engpässen	CT	1
russische Öl- und Gas-Exporte	ST	1
Öl- und Gas-Exporte	ST	1

Öl-	ST	3
Gas-Exporte	ST	1
EU-Importverbote für russisches Erdöl	ST	1
EU-Importverbote	ST	1
russisches Erdöl	ST	1
Erdöl	CT	2
Heizöl	CT	3
Schmierstoffe	CT	1
Handelsblatt	NE	2
Angebotszuwachs	CT	2
Klimakrise	CT	1
Ausbau der Erneuerbaren	ST	1
Strommarktbericht	ST	1
Industrie	CT	42
Sparaufrufe	CT	1
Ausbau von Windkraft und Solaranlagen	ST	1
Solaranlagen	ST	1
Atomausstieg	ST	1
Energiebedarfs	ST	1
IEA-Prognosen	ST	1
Strombedarf	ST	2
Stromnachfrage	ST	2
Energieagentur	ST	4
Stromverbrauchs	ST	1
Kohle- oder Atomkraftwerke	ST	1
Atomkraftwerke	ST	1
Emissionen	CT	1
CO2-Emissionen	CT	2
Energiemix	CT	1
Gesamt-Energie	CT	1
fossilen Quellen	ST	1
Kohle	CT	4
Öl	CT	9
Stromerzeugungsmix	ST	1
kohle- und gasbefeuelten Stromerzeugung	ST	1
gasbefeuelten Stromerzeugung	ST	1
CO2-Ausstoß	CT	1
Staatshilfe	CT	1
Schließungswelle	CT	1
Wirtschaftskammer	NE	2

Energiekostenzuschuss	ST	2
Mehrkosten für Energie	ST	1
Adeg	NE	1
Nah&Frisch	NE	1
Spar	NE	1
Unimarkt	NE	2
Sutterlüty	NE	1
Wolfgang Benischko	NE	1
Benischko	NE	4
Strom- und Energiepreise	ST	1
Strom-	ST	9
Christian Prauchner	NE	1
Prauchner	NE	2
Gewinns	CT	1
Verlust	CT	3
Peter Buchmüller	NE	1
Salzburg AG	NE	3
Kostensteigerungen bei Strom	ST	1
Kostensteigerungen	ST	2
Buchmüller	NE	2
Senkung des Strompreises	ST	1
Christof Kastner	NE	1
Kastner	NE	2
Kühlbedarfs	CT	1
Unimarkt-Chef Andreas Haider	NE	1
Andreas Haider	NE	1
Doppelmoral	CT	2
Transformation	ST	6
Georg Knill	NE	2
Deindustrialisierung Europas	ST	2
Deindustrialisierung	ST	3
IV-Chef Knill	NE	1
Georg Knill	NE	2
Ausbau der erneuerbaren Energie	ST	1
erneuerbaren Energie	ST	1
Ukraine-Krieg	NE	3
energieintensiven Bereichen	ST	1
Rezession	CT	1
Liquidität	CT	1

Kapital	CT	1
Insolvenzen	CT	1
Österreichischen Kontrollbank	NE	1
Putin	NE	4
Putins	NE	1
Transformation im Energiebereich	ST	1
Energiebereich	ST	1
Elektroautos	CT	1
Gaspreisdeckel	ST	2
Börsenbremse	OOD	1
EU-Energieminister	ST	4
Europäischen Union	NE	1
Einkaufskartell	ST	1
Handel mit Gas	CT	2
Subventionsmodell	ST	1
Marktpreis	CT	2
Marktpreise	CT	1
fossilen Energiewirtschaft	ST	1
Energiewirtschaft	ST	2
Preisdeckel für Gas	ST	2
Preisdeckel	ST	10
Title Transfer Facility	NE	1
TTF	NE	3
bilateralen Handel	CT	1
Energiebörse	ST	1
Beruhigung des europäischen Gasmarktes	ST	1
europäischen Gasmarktes	ST	1
Gasmarktes	ST	1
EU-Kommission	NE	2
Megawattstunde	ST	7
Megawattstunde Gas	ST	2
Megawattstunde Strom	ST	2
Megawattstunden	ST	1
Weltmarktpreis für Flüssiggas	ST	1
Weltmarktpreis	CT	4
Flüssiggas	ST	2
Preisschocks	CT	1
Margin-Zahlungen	OOD	1
Energieunternehmen	ST	4
Liquiditätsprobleme	ST	1

Endverbrauchermarkt	CT	1
Haushalts- und Gewerbekunden	CT	1
Haushalts-	CT	1
Gewerbekunden	CT	2
Preissteigerung	CT	1
Preissteigerungen	CT	2
Preisanpassung an Indizes	OOD	1
Großhandelspreis	CT	2
Großhandelspreisen	CT	1
Großhandelspreise	CT	4
Rekordpreise	CT	1
gedeckelter Großhandelspreis	OOD	1
Endverbraucherpreisen	CT	1
Deckel	CT	2
fixer Preisdeckel	ST	1
Versorgungsengpässen	ST	1
LNG	ST	1
Endverbraucherpreise	CT	1
Vorkrisenniveau	CT	3
Stromsparen	ST	1
Spitzenzeiten	CT	1
Zufallsgewinne von Stromunternehmen	ST	1
Zufallsgewinne	ST	2
Stromunternehmen	ST	3
Zufallsgewinnen	ST	1
Solidaritätsabgabe	CT	1
Gaseinkauf	ST	1
Speicherbestände	ST	1
Rohstoff	CT	3
Genehmigungsverfahren für den Ausbau erneuerbarer Energieträger	ST	1
Ausbau erneuerbarer Energieträger	ST	1
erneuerbarer Energieträger	ST	1
Aufstellen eines Windrades	CT	1
Windrades	CT	1
Aufstellen	CT	3
Solidaritätsmaßnahmen	ST	1
bilateralen Vertrag	CT	1
EU-Agentur ACER	NE	1
ACER	NE	1

Preissprung	CT	1
Preisausschlägen	CT	1
Entkoppeln des Gaspreises vom Strompreis	ST	1
Gaspreises	ST	1
Strompreis	ST	4
Strompreisbremse	ST	10
Ersparen	CT	2
Regierung	CT	23
Entlastung der Bevölkerung	ST	1
türkis-grüne Regierung	NE	4
Ministerrat	NE	1
gefördertes Kontingent	OOD	1
Kontingent	OOD	1
Verbrauchs	CT	1
Drei-Personen-Haushalts	CT	1
Drei-Personen-Haushalt	CT	1
Kilowattstunden	ST	3
kWh	ST	3
Regulierungsbehörde E-Control	NE	2
E-Control	NE	2
Jahresverbrauch	ST	3
persönlichen Verbrauch	CT	1
Verbrauch	CT	2
Jahresabrechnung des Energieversorgers	ST	1
Jahresabrechnung	CT	2
Energieversorgers	ST	2
Durchschnittsverbrauch	CT	1
Erspart	CT	1
Arbeitspreis	ST	1
verbrauchten Strom	CT	1
Energiesparen	ST	8
Wifo	NE	2
Energiekontingent	ST	1
Jahresverbrauchs	ST	1
Energieversorger	ST	8
eingefrorene Preis	CT	1
Nationalratsbeschluss	CT	1
Bremse	CT	3
Haushalt	CT	10
Zweitwohnsitze	CT	1

Nebenwohnsitze	CT	1
Stromversorger	ST	3
Anschlüssen	CT	1
Haupt- und Nebenwohnsitz	CT	1
Haupt-	CT	1
Nebenwohnsitz	CT	1
soziale Staffelung	OOD	2
Staffelung	OOD	2
Haushaltsgröße	CT	1
individuellen Stromverbrauch	ST	1
Zentralen Melderegister	NE	1
Zählpunkten des Stromanschlusses	ST	1
Stromanschlusses	ST	1
Hilfen für einkommensschwächere Haushalte	ST	1
sozial differenzieren	OOD	1
Mehrpersonenhaushalte	CT	1
Rundfunkgebühr	NE	1
GIS	NE	1
Datenlage	OOD	3
Mindestsicherungs- und Pflegegeldbezieher	CT	1
Mindestsicherungs-	CT	1
Pflegegeldbezieher	CT	1
Pensionisten	CT	1
Rente	CT	1
Ausgleichszulage	OOD	1
Eco-Austria-Chefin Monika Köppl-Turyna	NE	1
Monika Köppl-Turyna	NE	1
ZiB 2	NE	1
IHS-Chef Klaus Neusser	NE	1
Klaus Neusser	NE	1
NÖ-Strompreisrabatt	ST	1
niederösterreichische Strompreisrabatt	ST	1
Haushaltsverbrauchs	CT	1
blau-gelbe Förderung	ST	1
Stromrechnung	CT	2
Opposition	OOD	2
Neos	NE	6
ÖVP	NE	12
Grüne	NE	9
FPÖ	NE	4

Grünen	NE	1
Energiesparanreize	ST	1
Neos-Energiesprecherin Karin Doppelbauer	NE	1
Neos-Energiesprecherin	ST	1
Karin Doppelbauer	NE	1
SPÖ	NE	7
Klubobmann Jörg Leichtfried	NE	1
Jörg Leichtfried	NE	1
Landespartei- und Klubobmann Udo Landbauer	NE	1
Udo Landbauer	NE	2
niederösterreichischen Landtag	NE	1
Arbeiterkammer	NE	1
ÖGB	NE	3
Unterstützungsmaßnahmen für die energieintensive Industrie	ST	1
Industriebetriebe	CT	2
Industriebetrieben	CT	2
CO2-Kosten	CT	1
Gewerbe	CT	1
Handel	CT	5
EU-Beihilfevorgaben	OOD	1
Klein- und Mittelbetriebe	CT	1
Klein-	CT	1
Mittelbetriebe	CT	1
Betriebe	CT	15
Großunternehmen	CT	1
Unternehmen	CT	39
Faserhersteller Lenzing	NE	1
Lenzing	NE	1
AMS	NE	1
Gaspreise	ST	8
Stopp russischer Gaslieferungen	ST	1
russischer Gaslieferungen	ST	1
Pipeline Nord Stream 1	NE	1
Nord Stream 1	NE	1
Rally	OOD	1
europäische Future	OOD	1
Rekordpreis	CT	1
Potenziale	CT	4
Tempo 100	NE	4
Energiebonus	ST	1

Tempolimits	ST	4
Energieeinsparungen	ST	1
Energiekrisen	ST	1
Rohölpreise	ST	2
kollektive Sparanstrengungen	CT	1
Temperaturbegrenzung für Heizungen	ST	1
Temperaturbegrenzung	ST	1
Heizungen	CT	2
Grad Celsius	CT	1
Autobesitzer	CT	1
autofreien Tag	CT	1
Energieferien	ST	1
Sparen	CT	5
Sparens	CT	1
Bundeskanzler Bruno Kreisky	NE	1
Bruno Kreisky	NE	1
Energiespargründen	ST	1
Nassrasur	CT	2
elektrischen Rasierer	CT	1
elektrischer Rasierer	CT	1
Gasknappheit	ST	1
Flüssiggasimporte aus Katar	ST	1
Flüssiggasimporte	ST	1
Katar	NE	2
Ausbau erneuerbarer Energien	ST	1
erneuerbarer Energien	ST	1
Kohlekraftwerk	CT	1
Mellach	NE	2
Verbund-Großanlage Mellach	NE	1
Klimaschutzministerin Leonore Gewessler	NE	2
Leonore Gewessler	NE	4
Klimaschutzministerin	NE	3
Fernsehsender Puls4	NE	1
Franz Angerer	NE	1
Geschäftsführer der Österreichischen Energieagentur	ST	1
Österreichischen Energieagentur	NE	2
AEA	NE	1
Einsparung	CT	3
Investitionen	CT	4

Energiesparkampagne	ST	1
Stromkunden	ST	1
Stromkundinnen	ST	1
Tempo 130	NE	1
Einsparerfolgen	CT	1
Johannes Wahlmüller	NE	1
Energieexperte der Umweltschutzorganisation Global 2000	ST	1
Energieexperte	ST	1
Umweltschutzorganisation	CT	1
Global 2000	NE	1
Temporeduktion	CT	1
Kraftfahrzeugen	CT	2
Gassektor	ST	1
Verbrennungsmotoren	CT	2
Verbrauch von Ölprodukten	ST	1
Ölprodukten	ST	1
Umweltbundesamt	NE	1
Wahlmüller	NE	2
Kühlvittrinen	CT	2
Energie einzusparen	ST	1
Kühlschränke	CT	1
Energiefresser	CT	1
Kühlschrank	CT	1
Jacuzzi	CT	1
Energieverbrauch	ST	2
Energieagentur-Chef Angerer	NE	1
Energieagentur-Chef	ST	1
Angerer	NE	3
Initiativen	CT	2
Energiesparpotenzial	ST	1
Betriebsoptimierung	CT	1
Terawattstunde	ST	1
Terawattstunden	ST	1
neu dämmen	CT	1
Energieeffizienz	ST	3
Energieeffizienzgesetzes	ST	2
Energieeffizienzgesetz	ST	1
Energie einsparen	CT	1
Gutschriften	CT	1
Energiesparsames Verhalten	ST	1

Energiesparsames	ST	1
Energieverschwendung	ST	1
energiefressenden Geräte	CT	1
Energiefressenden	CT	1
Heizung	CT	2
Aufzudrehen	CT	1
Energieeffizienzgesetzes 2020	NE	1
Klimaschutzministerin Gewessler	NE	1
Gewessler	NE	9
Klimaschutzministerium	NE	3
Energiesparmaßnahme	ST	1
Wiener Wirtschaftsforschungsinstitut	NE	1
Energiesparbonus	ST	1
Konsument	CT	2
Strom- und Gasversorger	ST	1
Gasversorger	ST	1
Bonus	CT	2
Ministerin Gewessler	NE	1
Gewessler-Sprecherin Martina Stemmer	NE	1
Martina Stemmer	NE	1
Sparkampagne	CT	1
Einsparungen bei Gas	ST	1
Stemmer	NE	3
Finanzministerium	NE	2
Finanzministeriums	NE	1
Gewinnabschöpfung	OOD	3
Entkoppelung von Strom- und Gaspreis	ST	1
Gaspreisen	ST	2
Kraftwerk	CT	3
Kanzler Karl Nehammer	NE	1
Bundeskanzler Karl Nehammer	NE	1
teilstaatlichen Stromunternehmen	ST	1
Verbund	NE	4
ÖVP-geführte Ministerium	NE	1
ÖVP-geführte	NE	1
Preisentwicklung	CT	3
erneuerbarer Energie	ST	2
Stroms	CT	1
Strom-Gaspreis-Kopplung	ST	1
Bund	NE	1

APA	NE	1
Strompreisbildung	ST	1
Kanzleramts	NE	1
Strompreisen	ST	3
Energieministerin Leonore Gewessler	NE	2
Energieministerin	ST	2
geschnürte Hilfspakete	OOD	1
Vizekanzler Werner Kogler	NE	2
Werner Kogler	NE	2
Übergewinne	OOD	2
Übergewinnen	OOD	1
Regierungsvorstoß	OOD	1
Überschussgewinne	OOD	1
soziale Ausgleichsmaßnahmen	OOD	1
Kanzler Nehammer	NE	1
Nehammer	NE	6
Nehammers	NE	1
Energiefragen	ST	1
Neos-Wirtschaftssprecher Gerald Loacker	NE	1
Gerald Loacker	NE	1
FPÖ-Chef Herbert Kickl	NE	1
Herbert Kickl	NE	1
Dividenden	OOD	1
teilstaatlichen börsenorientierten Stromversorger	ST	1
Marktwert	OOD	1
öffentlichen Hand	OOD	1
öffentlichem Eigentum	OOD	1
Instituts für Finanzrecht	NE	1
Universität Wien	NE	1
Sabine Kirchmayr-Schliesselberger	NE	1
ORF	NE	1
teilstaatliche Unternehmen	OOD	1
teilstaatlichen Energieversorger	ST	1
Sondersteuer	OOD	1
STANDARD	NE	2
Anleger	OOD	2
Florian Beckermann	NE	1
Interessenverbands für Anleger	NE	1
IVA	NE	1
öffentliche Hand	OOD	1

Aktionärsstruktur	OOD	1
Kapitalmarkt	OOD	1
Martin Jaksch-Fliegenschnee	NE	1
IG Windkraft	NE	1
Investitionssicherheit	CT	1
Jaksch-Fliegenschnee	NE	2
Energieunabhängigkeit	ST	1
Öl- und Gasfirmen	ST	1
Gasfirmen	ST	1
Gewinne	CT	1
Strombörse	ST	1
Auktionsverfahren	OOD	1
Deckung der Stromnachfrage	ST	1
Gaskraftwerk	ST	1
Gaskraftwerken	ST	1
Strompreise	ST	5
Haushaltsenergiepreise	ST	1
Haushaltsenergiepreisen	ST	1
Sprit	CT	2
Holzpellets	CT	1
Haushaltskundenpreise	ST	1
Großhandelssituation	CT	1
Großhandelsmärkten	CT	1
Preiseinbruch	CT	1
Nachfrage nach Gas	ST	1
Kund:innen	CT	1
Spotpreis-	OOD	1
Float-Tarifen	OOD	1
Preisgarantie	CT	3
Preisanpassungsklausel	OOD	1
Preisanpassung	CT	3
Preisanpassungen	CT	1
Haushaltspreise für Strom	ST	1
Haushaltspreise	ST	1
Fernwärme	ST	4
Brennholz	CT	1
Superbenzin	CT	2
Klimaministerium	NE	3
Energienotfall	ST	3
Notlage bei Strom und Gas	ST	1

Energieministerium	ST	1
Energienotfall	ST	3
Energielenkungsgesetz	ST	2
Störungen der Energieversorgung	ST	2
Störungen	CT	2
Energie- und Umweltpolitik	ST	1
Umweltpolitik	CT	1
Maßnahmen zur Prävention von Störungen der Energieversorgung	ST	1
Maßnahmen zur Einsparung oder Reduktion von festen oder flüssigen Energieträgern	ST	1
flüssigen Energieträgern	ST	2
Elektrizität	CT	3
Wärme	CT	4
Vorbegutachtungsentwurf	OOD	1
Maßnahmen zur Steigerung der Energieeffizienz	ST	1
Betriebsweise von Anlagen	ST	1
Anlagen	CT	2
Abweichungen von Emissionsgrenzwerten	ST	1
Emissionsgrenzwerten	OOD	1
Erwerbsfreiheit	OOD	1
Marktwirtschaft	CT	1
Energielenkungsfall	ST	1
Eigentums- und Persönlichkeitsrechte	OOD	1
Marktstörungen	ST	1
Marktstörung	ST	2
Novellierung	OOD	1
Multi-Energiekrise	ST	1
Umrüstung von Gas auf Öl	ST	1
Gasnotfall	ST	1
Entschädigung	CT	3
Entschädigungen	CT	2
Ausnahmefällen	CT	1
Entschädigung bei Enteignung	ST	1
Gesetzentwurf	CT	1
Körperschaften öffentlichen Rechts	OOD	1
Vermögensvorteile von Antragsstellern	OOD	1
Marktlage	CT	1
Subventionen	OOD	1
Krisengewinner	CT	1
Mikl-Leitner	NE	8

Strompreisdeckel	ST	2
Finanzminister Brunner	NE	1
Johanna Mikl-Leitner	NE	3
Begrenzung der Stromkosten	ST	1
Preisdeckel für Strompreise	ST	1
Koalition	OOD	1
Bundeskanzler Karl Nehammer	NE	1
Nationalrat	NE	1
Presse	NE	1
Stromzone	ST	1
Finanzminister Magnus Brunner	NE	1
Magnus Brunner	NE	1
Marktmechanismen	CT	2
Preisdeckelung	ST	1
LH Mikl-Leitner	NE	1
Merit-Order System	ST	1
Gas- und Energiemärkten	ST	1
Energiemärkten	ST	2
Hahn	NE	2
EU-Kommissar Johannes Hahn	NE	1
Ö1-Mittagsjournal	NE	1
einkommensschwache	CT	1
SPÖ-Vorsitzender Pamela Rendi-Wagner	NE	1
Pamela Rendi-Wagner	NE	1
SPÜO-Landesparteivorsitzenden Franz Schnabl	NE	1
Franz Schnabl	NE	1
SPO-Bundesgeschäftsführer Christian Deutsch	NE	1
Christian Deutsch	NE	1
Niederösterreichs	NE	4
Niederösterreichs Neos	NE	1
Oberösterreichs	NE	6
Oberösterreich	NE	12
Oberösterreicher	NE	1
Strompreisdeckel-Forderung	ST	1
Landtag	NE	2
niederösterreichischen Landtag	NE	1
Landespartei- und Klubobmann Udo Landbauer	NE	1
Udo Landbauer	NE	2
Landessprecherin Indra Collini	NE	1
Indra Collini	NE	2

Collini	NE	3
Deckeln	ST	1
Übergewinne der Energiekonzerne	ST	1
Energiekonzerne	ST	1
soziale Härtefälle	OOD	1
Transparenz der Notfallpläne	ST	1
Transparenz	CT	2
Notfallpläne	CT	2
Gasverteilung	ST	1
Preiserhöhung	CT	4
Preiserhöhungen	CT	3
Energieallianz Austria	NE	3
Indexanpassung	OOD	4
Indexanpassungen	OOD	2
Standardtarifen	CT	2
Preisdeckel bei Energie, Treibstoffen, Grundnahrungsmitteln	ST	1
Landesenergieversorger	ST	1
EAA	NE	2
Österreichische Gewerkschaftsbund	NE	1
Burgenland Energie	NE	3
Energie AG	NE	1
Preisgarantie für Strom und Gas	ST	1
Linz AG	NE	2
Tarifierhöhung bei Bestandskunden	ST	1
Tarifierhöhungen	CT	1
Bestandskunden	CT	4
Fernwärme-Tarife	ST	1
städtischen Energieversorgers	ST	1
Erdgastarif	ST	1
Energiemarkt	ST	1
Gaspreisindex-Entwicklung	ST	1
Energieversorger Salzburg AG	NE	1
Salzburg AG	NE	3
Strom- und Gaspreise	ST	1
Energie Steiermark	NE	1
Strom- und Gaspreisanpassungen	ST	1
Gaspreisanpassungen	ST	1
Mehrkosten	CT	5
Energiekonzerns	ST	1
Kelag	NE	3

Erdgasverbrauchspreis	ST	1
Tiroler Landesversorger Tiwag	NE	1
Vorstandsvorsitzender Erich Entstrasser	NE	1
Erich Entstrasser	NE	1
Neukunden	CT	1
Österreichische Strompreisindex	ST	1
Öspi	ST	2
Strompreiserhöhung	ST	1
Entstrasser	NE	2
Landesstatthalterin Barbara Schöbi-Fink	NE	1
Barbara Schöbi-Fink	NE	1
Strompreisdeckelung	ST	1
Illwerke/VKW	NE	1
Andreas Neuhauser	NE	1
Marktsituation	CT	1
EVN-Management	ST	1
EVN-Mehrheitseigentümer	ST	1
Dagmar Belakowitsch	NE	1
Normalisierung bei Energiepreisen	ST	1
E-Control-Chef Wolfgang Urbantschitsch	NE	1
E-Control-Chef	ST	1
Wolfgang Urbantschitsch	NE	2
EU-27	NE	1
Strom- und Gasmärkte	ST	1
Gasmärkte	ST	1
Pinken	NE	1
Urbantschitsch	NE	4
Klub der Wirtschaftspublizisten	NE	1
Strom- oder Gashandel	ST	1
Energiewelt	CT	1
Verknappung	CT	3
Gasspeicher	ST	2
Haushaltskunden	CT	2
Beschaffungsstrategie der Energieversorger	ST	1
Preisnachlässe	CT	1
Gasspeichers	ST	1
Haidach	NE	6
bilateraler Gasdurchleitungsvertrag	ST	1
Gasdurchleitungsvertrag	ST	2
7Fields	NE	5

Mangellage	CT	1
bilateralen Gasdurchleitungsvertrag	ST	1
Befüllung der Erdgasspeicheranlagen	ST	1
Erdgasspeicheranlagen	ST	1
Wirtschaftsminister Robert Habeck	NE	1
Robert Habeck	NE	1
Schloss Schönbrunn	NE	1
Abkommen	CT	5
Absichtserklärung	OOD	1
unionsrechtlicher Vorgaben	NE	1
Speicheranlagen	CT	1
Befüllungsziel	CT	2
Erdgasspeicher	ST	1
deutsche Netz	CT	1
Transit für Gas	ST	1
Transit	CT	3
Gasmangellage	ST	4
Eingelagert	CT	1
Speichermengen	ST	1
Gewährleistung der Versorgungssicherheit	ST	1
Eingespeichert	ST	2
Energielenkung	ST	1
Sicherheitsnetz	CT	1
Reserve	CT	1
Sicherheitspolster	OOD	1
Krisenfall	CT	3
Putins Erpressungsversuche	NE	1
Wirtschaftsminister Habeck	NE	1
Solidarität	CT	5
Solidaritätsabkommen zur Gewährleistung der sicheren Gasversorgung	ST	1
Energiesicherheit und -versorgung	ST	1
europäischen Energiekooperation	ST	1
Energiekooperation	ST	1
Sanktionen	OOD	9
russisches Rohöl	ST	1
G7	NE	2
Boykott von russischem Öl	ST	1
russischem Öl	ST	1
Sorte Urals	NE	3

Treibstofflager	CT	1
Lukoil	NE	1
Rohöls	CT	1
Sanktionen gegenüber Russland	ST	1
Rohöl	CT	5
Importstopp für russisches Öl	ST	1
Importstopp	CT	2
russisches Öl	ST	1
Exportpreise	CT	1
Argus Media	NE	1
Uralspreis	ST	1
Nordseesorte Brent	NE	1
Nordseesorte	CT	1
Brent	NE	4
Nordsee-Sorte	CT	1
Espo-Pipeline	NE	1
Argus	NE	2
Dollar pro Fass	ST	3
Preisobergrenze	CT	3
Giovanni Staunovo	NE	1
Rohstoffexperte	ST	1
Schweizer Großbank UBS	NE	1
UBS	NE	1
Importembargo	CT	1
Staunovo	NE	4
Handelsblatt	NE	2
Stabilisieren	CT	2
Ölhandel	CT	1
Ölpreise	CT	1
Tanker	CT	1
UBS-Experte Staunovo	NE	1
Produktionseinschränkungen	CT	1
Lagerkapazitäten	CT	1
Ölexporte	ST	2
Ölindustrie	ST	1
Präsident Putin	NE	1
Gegenmaßnahmen Russlands	ST	1
Rabattgrenze	ST	1
Wirtschaftszeitung Wedomosti	NE	1
Urals	NE	4

Ölgesellschaften	ST	1
Gasboykott	ST	1
Gasboykotts	ST	1
Energiedichte	ST	1
Energieform	ST	1
Chemie	CT	2
Dämmmaterial	CT	1
Beeinträchtigung der Versorgung	ST	1
Stopp der Gasimporte	ST	1
Gasimporte	ST	1
Kubikmeter	CT	2
TWh	ST	6
Kraftwerk Thieß	NE	1
Strom erzeugen	CT	1
Wärme auskoppeln	ST	1
Fernwärmenetze	ST	1
Einspeisen	CT	1
Müllverbrennung Wien-Spittelau	NE	1
Müllverbrennung	CT	1
Wien-Spittelau	NE	1
Lebensmittelindustrie	CT	1
Red Bull	NE	1
Agrana	NE	1
Mars	NE	1
NÖM	NE	1
Gaszufuhr	ST	1
Norske Skog	NE	1
Kosteneinsparung	CT	1
Papier- und Zellstoffkonzern Sappi	NE	1
Sappi	NE	1
Biomassekraftwerk	ST	1
Gasstopp	CT	1
RHI Magnesita	NE	1
Stahl-, Zement- oder Glasöfen	CT	1
Stahl-	CT	2
Zement-	CT	1
Glasöfen	CT	1
Stefan Borgas	NE	1
Umstellung auf alternative Brennstoffe	ST	1
alternative Brennstoffe	ST	1

Brennstoffe	ST	2
Brennstoffen	ST	1
Borgas	NE	3
Rohstoffgewinnung	ST	1
Baustoffkonzern Wienerberger	NE	1
Wienerberger	NE	2
Brennöfen	CT	1
Trockner	CT	1
Gasvolumens	ST	1
Solveig Menard-Galli	NE	1
Wienerberger Building Solutions	NE	1
Menard-Galli	NE	2
Stahl- und Metallverarbeitungsindustrie	CT	1
Metallverarbeitungsindustrie	CT	1
Stopp der Gaslieferungen	ST	1
Eisenerz	OOD	2
Eisenschwamm	OOD	1
Elektroofen	ST	1
Energiezufuhr	ST	1
Gasversorgung aufrechterhalten	ST	1
Stopp russischer Gaseinfuhren	ST	1
Gaseinfuhren	ST	1
Einbruch der Wirtschaftsleistung	ST	1
Wirtschaftsleistung	CT	1
Umstieg auf Gas	ST	1
Totalausfall russischer Lieferungen	ST	1
Zwangsrationierungen	ST	1
gespeichertes Erdgas	ST	1
Versorgung von Endkunden	ST	1
Endkunden	CT	1
Verbrauchsreduktion	ST	1
Energielenkungsbeirat	ST	1
Erdöl- und Gaswirtschaft	ST	1
Erdöl-	ST	1
Gaswirtschaft	ST	1
Transitstreits	CT	1
Versorgung in Notfällen	ST	1
Corona	CT	1
Öl- und Gaspreise	ST	1
Finanzkrise	CT	1

Schweröl	CT	1
Kerosin	CT	1
Chemieprodukte	CT	1
Ölsorten	CT	1
WTI	NE	1
Energie getriebene Inflation	ST	1
Ölmarkt	CT	1
Russische Föderation	NE	1
Ölgeschäft	ST	1
Abkehr von fossilen Ressourcen	ST	1
fossilen Ressourcen	ST	1
Erdölressourcen	ST	1
Bundesanstalt für Geowissenschaften und Rohstoffe	NE	1
BGR	NE	1
Corona-Krise	CT	1
Institut für Weltwirtschaft	NE	1
IfW	NE	1
Warenkontrollen zur Einhaltung der Sanktionen	ST	1
IfW-Volkswirt Vincent Stamer	NE	1
Vincent Stamer	NE	1
inflationstreibend	CT	1
wachstumsabschwächend	CT	1
Stagflation	OOD	1
Wifo-Chef Gabriel Felbermayr	NE	1
Gabriel Felbermayr	NE	1
Belastung der Verbraucher	ST	1
Ausfällen in der Produktion	ST	1
IfW-Volkswirt Stamer	NE	1
Stillständen	CT	1
fossile Brennstoffe	ST	1
Ausfall des Handels	ST	1
Ölembargo gegen Russland	ST	1
Wifo-Chef Felbermayr	NE	1
Ölverkauf	CT	1
Erdgasverkauf	CT	1
Verknappung des Angebots	ST	1
ÖAMTC	NE	2
Super	CT	1
Nordsee-Sorte Brent	NE	1
(preis-)treibende	CT	1

Raffineriemargen	OOD	1
Aussetzen des CO2-Preises	ST	1
CO2-Preises	CT	1
Mineralölsteuer	ST	1
Pendlerpauschale	CT	1
Kilometergeld	CT	1
Aufschieben der CO2-Aufschläge	ST	1
CO2-Aufschläge	OOD	1
CO2-Preis	CT	1
Klimabonus	CT	1
CO2-Bepreisung	CT	1
Wifo-Chef Felbemayr	NE	1
Senkung der Elektrizitätsabgabe	ST	1
Elektrizitätsabgabe	ST	1
Härtefallregel	OOD	1
Wien	NE	4
Berlin	NE	2
Wien / Berlin	NE	1
USA	NE	9
China	NE	6
Chinas	NE	2
Italien	NE	2
Griechenland	NE	2
Indien	NE	1
Paris	NE	1
Südostasien	NE	1
Saudi-Arabien	NE	1
Europe Machined Components	NE	1
Kongo	NE	1
Aschach	NE	1
Kapfenberg	NE	1
Berlins	NE	1
Afrika	NE	1

Anhang D

Termkandidatenliste OneClick Terms (2000)

Termkandidaten	Häufigkeit
Haindl-Grutsch	16
Preisdeckel	10
Energiekrise	19
Strompreisbremse	10
Haidach	6
Gewessler	9
Birol	8
Megawattstunde	7
Staunovo	4
Gasmangellage	4
Benischko	4
EU-Energieminister	4
Urbantschitsch	4
Gaslieferungen	8
Indexanpassung	4
Energienotfall	3
Gasbezugsquellen	3
Klimaministerium	3
Wifo-Chef	3
Klimaschutzministerium	3
türkis-grüne	4
Klimaschutzministerin	3
Energieallianz	3
Landesversorger	3
Bergsmann	3
EVN	14
Merit-Order	3
Ukraine-Krieg	3
teilstaatlichen	3
energieintensiven	7
Urals	4
Borgas	3
Nehammer	6
Gaspreis	8
Stromunternehmen	3
Gaspreisdeckel	2
Befüllungsziel	2
Wladkowa	2
Gasdurchleitungsvertrag	2
Energieforecasts	2
Öspi	2

Grundlastkapazitäten	2
IfW-Volkswirt	2
Energielenkungsmaßnahmen	2
Strompreisdeckel	2
Entlastungspakete	2
Menard-Galli	2
Jaksch-Fliegenschnee	2
Großhandelspreise	4
Lieferkettenprobleme	2
Prauchner	2
Energielenkungsgesetz	2
Energiekostenzuschuss	2
Ukraine-Krieges	2
Gewinnabschöpfung	3
Zufallsgewinne	2
Röchling	5
Angebotszuwachs	2
Übergewinne	2
IEA-Chef	2
Entstrasser	2
Internorm	3
Genehmigungspolitik	2
zudreht	3
Energieeffizienzgesetzes	2
Preisobergrenze	3
Standardtarifen	2
Marktstörung	2
Indexanpassungen	2
Strompreiskompensation	2
Weltmarktpreis	4
Energiepreise	16
Gaskrise	2
Kühlvitrienen	2
Collini	3
Untersperger	2
Dominoeffekte	2
Mikl-Leitner	8
Produktionseinschränkungen	2
Lieferausfälle	2
Kelag	3
Gasversorgung	9
Eckerstorfer	2

Nahversorgern	2
Blitzumfrage	3
energieintensive	5
Industriellenvereinigung	5
Vorkrisenniveau	3
Knill	3
Unimarkt	2
Mellach	2
Lohnrunde	2
Felbermayr	3
Wahlmüller	2
Energieministerin	2
Energiepreisen	4
Importstopp	2
Gashahn	3
Erpressungsversuche	2
Großhandelspreis	2
Gaspreise	8
Preß	2
Ölexporte	2
Lieferstopp	2
Energiemärkten	2
Strompreiserhöhungen	2
Landessprecherin	2
Energieagentur	8
Deindustrialisierung	3
Gasspeicher	2
Metallern	2
Oberösterreichs	6
Landbauer	2
Stemmer	3
Wifo	2
Stamer	2
TWh	6
verteuerte	3
Hochreiter	2
Buchmüller	2
Preisanstiege	2
Espo-Pipeline	1
Energiepreiswelle	1
Energiekostenhilfen	1
Wärmebehandlungsaktivitäten	1
hinuntergedrehte	1
Umsgerechnet	1
NÖ-Strompreisrabatt	1
Erdgasverbrauchspreis	1
wachstumsabschwächend	1

SPO-Bundesgeschäftsführer	1
Mauerziegelwerken	1
Strompreisdeckel-Forderung	1
Tonzerkleinerungsmaschinen	1
Gaspreisindex-Entwicklung	1
Eco-Austria-Chefin	1
Dauerbeflegelungen	1
Energie-Kostenniveaus	1
Gewessler-Sprecherin	1
Uralsöl	1
Multi-Energiekrise	1
Reststoffverwertungskessel	1
Drei-Personen-Haushalts	1
Flüssiggas-Abnehmer	1
Erneuerbaren-Energie-Ausbau	1
Verbund-Großanlage	1
Neos-Energiesprecherin	1
Vorbegutachtungsentwurf	1
Uralspreis	1
Strom-Gaspreis-Kopplung	1
Energielenkungsfall	1
Teuerungsprämien	1
Haushaltsenergiepreisen	1
SPÜO-Landesparteivorsitzenden	1
EVN-Mehrheitseigentümer	1
IfW	2
August-Inflation	1
Dezember-Lieferungen	1
EU-Beihilfevorgaben	1
Strommarktbericht	1
Börsenbremse	1
Wohnraumerrichtung	1
Laserhybridanlage	1
EU-Importverbote	1
Weiterverarbeitungsstandorte	1
Großhandelssituation	1
Copy-Paste-Plagiat	1
Gas-Großhandel	1
Transitstreits	1
Strompreisdeckelung	1
Doppel-Wumms	1
Gasverknappung	1
Gasembargos	1
Energieagentur-Chef	1
Gasboykotts	1
Russland-Freundlichkeit	1
Erholungssignal	1

Wünsch-Dir-was-Politik	1
Gasstopp	1
Unimarkt-Chef	1
Zwangsrationierungen	1
Gasengpässe	1
Gasnotfall	1
Energiesparsames	1
Fernwärme-Tarife	1
Solidaritätsabkommen	1
Float-Tarifen	1
Sparaufrufe	1
Energiekostenzuschüsse	1
Haushaltsenergiepreise	1
Salz-Gewinnung	1
Energie-Souveränität	1
EVN-Management	1
Strompreisrabatt	1
Energielenkungsbeirat	1
Salz-Produktion	1
Energiesparanreize	1
Lieferkettenproblematik	1
Metaller-Lohnverhandlungen	1
Haushaltsverbrauchs	1
Gas-Stopp	1
Sparkampagne	1
Gaspreisen	2
Konsumenten-Ebene	1
CO2-Aufschläge	1
Zufallsgewinnen	1
Erdgasspeicheranlagen	1
preisführende	1
intraeuropäische	1
Komplettstopp	1
Speicherbestände	1
Petkows	1
Gaspreisanpassungen	1
Haushaltskundenpreise	1
Frisch-Kaufmann	1
Stromzone	1
IEA-Prognosen	1
Übergewinnen	1
Totalstopp	1
Gasembargo	1
Energiepreiskrise	1
Erdgasabhängigkeit	1
Regierungsvorstoß	1
Vorwärme	1

Vier-Parteien	1
Gasboykott	1
nichtverfügbare	1
Vormaterialkosten	1
Energiesparbonus	1
Überschussgewinne	1
Einkaufskartell	1
Frühwarnstufe	1
ADEG-Märkte	1
Energiekontingent	1
Einsparerfolgen	1
verdienbar	1
Gas-Exporte	1
Spar-Kaufmann	1
Kirchmayr-Schliesselberger	1
Spritpreis	2
Vor-Krisenniveau	1
Neos-Wirtschaftssprecher	1
Haushaltspreise	1
Zellstoffkonzern	1
E-Control-Chef	1
Rabattgrenze	1
russisches	9
Wedemosti	1
Mindestsicherungs	1
Einkaufsvereinigungen	1
Aggressionskriegs	1
Flüssiggasimporte	1
Gaslieferstopp	1
Preischaos	1
Neongas	1
Dachziegelwerken	1
Preisverwerfungen	1
Köppl-Turyna	1
Gaseinführen	1
Megaprogramm	1
Erdgasverkauf	1
Gasknappheit	1
ostalgerischer	1
Importembargo	1
Industriebundesland	1
Stahlerzeugung	2
inflationstreibend	1
STANDARD-Anfrage	1
Machined	1
Rohölpreise	2
Beschichtungsspezialist	1

Stromerzeugungsmix	1
Formgebungsprozess	1
Gesamt-Energie	1
IV-Chef	1
Metallverarbeitungsindustrie	1
Inflationswerte	1
Inflationswert	1
Landesstatthalterin	1
Fahrerpersonal	1
Raffineriemargen	1
Energiebedingte	1
Nehammers	1
Härtefallregel	1
Markteingriffen	1
Subventionsmodell	1
Hilfsaggregat	1
Konsumentenebene	1
Brancheninformationsdienstes	1
ÖVP-geführte	1
importabhängig	1
Preisanpassung	3
Preisdeckelung	1
Brennertechnologie	1
Wessela	1
Energiesparkampagne	1
STANDARD-Interview	1
Herbstlohnrunde	1
Kühlbedarfs	1
Solidaritätsmaßnahmen	1
Stromkundinnen	1
Stromnachfrage	2
Preisausschlägen	1
Witoscha	1
Ekkehart	2
Werkzeugmaschinen-Hersteller	1
Energiebonus	1
Energiekostenanteil	1
Dominoeffekten	1
Erdgastarif	1
Datenproblem	1
Glaswanne	1
Mikrowarenkorb	1
Gasspeichers	1
Terawattstunde	1
Gaseinkauf	1
TTF	3
Nationalratsbeschluss	1

durchzutauchen	1
Energiekooperation	1
Flachgläser	1
Miniwarenkorb	1
Strompreisbildung	1
Hueck	2
Frächtern	1
Preisschocks	1
IHS-Chef	1
ostalgischen	1
Strompreisindex	1
Elektrizitätsabgabe	1
Landesenergieversorger	1
Zählpunkten	1
energieintensives	1
Teuerung	6
Wirtschaftspublizisten	1
Siebenfachen	1
Irreparable	1
Neos	6
Nordsee-Sorte	1
Hilfsmöglichkeit	1
Papierkonzern	1
Sechzigstel	1
Kohlekessel	1
gedeckelter	1
Mobilitätssystems	1
aufgelassenes	1
Energieferien	1
Großhandelsmärkten	1
EU-Territorium	1
Solidaritätsabgabe	1
Gasvolumens	1
Markteingriff	1
IEA	4
Petkow	1
hitzeresistenten	1
Faserhersteller	1
Gasfirmen	1
mitzumarschieren	1
Klimabonus	1
Pflegegeldbezieher	1
Wirtschaftsablauf	1
Lohnrunden	1
Industriekapitäne	1
Erdölressourcen	1
Ölverkauf	1

Marmeladegläser	1
Energieversorgungen	1
Laufkraftwerke	1
Mangellage	1
Endverbrauchermarkt	1
umsetzungsstarke	1
Energiesparmaßnahme	1
unleistbar	1
Energielenkung	1
Nebenwohnsitze	1
Schließungswelle	1
energiefressenden	1
Gasverteilung	1
zweckgewidmet	1
Endverbraucherpreisen	1
CO2-neutralem	1
Radew	1
Sozialsprecherin	1
Wiederverstaatlichung	1
technologieoffenen	1
Energiepreissteigerungen	1
CO2-Kosten	1
Lebensmittelproduzent	1
Folgeeffekte	1
Landeshauptfrau	3
EAA	2
zigtausender	1
Energiespargründen	1
Primorsk	1
teilstaatliche	1
Temperaturbegrenzung	1
gasbefeuerten	1
Energiekrisen	1
Jahresverbrauchs	1
Interessenverbands	1
Quartalsbeginn	1
technologiegetriebene	1
Speichermengen	1
Drei-Personen-Haushalt	1
Strompreises	2
Energiepreisentwicklung	1
Symbolhandlungen	1
Förderwesen	1
finanzkräftig	1
Harvard-Absolvent	1
Eisenschwamm	1
Ö1-Mittagsjournal	1

Luftwaffengeneral	1
Warenkontrollen	1
Sonntagspredigten	1
Markteingriffe	1
Realloohnerhöhung	1
Geschäftsbasis	1
Preisgünstigkeit	1
herunterzudrücken	1
Gasimporte	1
Glasöfen	1
Spritpreise	4
Emissionsgrenzwerten	1
Temporeduktion	1
Tempolimits	4
Schöbi-Fink	1
LNG-Terminals	1
eingespeichert	2
Marktstörungen	1
Ofenanlage	1
E-Control	2
Haushaltskunden	2
Erscheine	1
Verfahrensdauern	1
Gasmärkte	1
abzuschöpfen	2
Strompreisen	3
Gashandel	1
Inflationsdaten	1
Verbrauchsreduktion	1
Energielösung	1
Rohstoffexperte	1
Preiseinbruch	1
Stromhunger	1
Skog	1
Autobahntankstellen	1
Mehrpersonenhaushalte	1
Energiebörse	1
Ölprodukten	1
Energieeffizienzgesetz	1
Exportpreise	1
russischem	4
Energiesparpotenzial	1
Erwerbsfreiheit	1
Stromanschlusses	1
niederösterreichische	3
Wocheneinkauf	2
KURIER	4

Gasverbrauch	3
Rationierungen	1
Treibstofflager	1
Finanzwerten	1
Weidinger	3
Oepping	1
auskoppeln	1
verkraftbar	3
Ölembargo	1
pro-westliche	1
Baustoffkonzern	1
Erdgaslieferungen	1
Nassrasur	2
Nahrungsmittelbereich	1
rationiert	2
Chemieprodukte	1
Angriffskriegs	1
Energieunabhängigkeit	1
Stillständen	1
NÖM	1
CO2-Preises	1
durchleiten	1
stromintensiven	1
Eingriffsmöglichkeiten	2
Endverbraucherpreise	1
einkommensschwächere	1
Kriegsbedingt	1
Gasmarktes	1
Preisanstiegen	1
Ölhandel	1
Fixpreise	1
Großteilen	1
Energieabhängigkeit	1
Greißler	1
kostendeckender	1
EU-Agentur	1
Kraftakte	1
Tiwag	1
Krisengewinner	1
Gaspreises	1
versiebenfacht	1
Illwerke	1
Sappi	1
Produktionsstopps	1
Sicherheitspolster	1
Preisankpassungsklausel	1
Hilfspakete	2

Rekordwerten	1
zusperren	2
Sutterlüty	1
Rekordpreise	1
Energiekosten	18
Leonore	4
Eibensteiner	1
Erdgasspeicher	1
Beschaffungsstrategie	1
MWh	4
vervielfachen	1
Entkoppeln	1
Fotovoltaik-Anlage	1
unionsrechtlicher	1
Technologieoffenheit	1
Adeg	1
Gebietsgewinne	1
Energiesicherheit	2
Industriestandortes	1
Normallage	1
Superbenzin	2
Speichieranlagen	1
Vormaterialien	1
VKW	1
Energiesparen	8
Puls4	1
Landespartei	2
Haushaltsenergie	1
Pinken	1
Marktmechanismen	2
Vorprodukt	1
Auktionsverfahren	1
AEA	1
Vermögensvorteile	1
Biomassekraftwerk	1
Hallenbeleuchtung	1
Betriebsoptimierung	1
Nahversorger	3
Kontrollbank	1
Großhandelspreisen	1
Gassektor	1
Elektroofen	1
Metallerzeugung	1
Rohöls	1
Einkaufstempel	1
Rekordmarken	1
Maximalpreis	1

Deviseneinnahmen	1
Ländervertreter	1
Umweltschonung	1
Preisgarantie	3
Fuels	1
Spittelau	1
Energieunternehmen	4
Vermieden	1
Ökostromanlagen	1
Niederösterreichs	4
börsenotierten	1
Kunststoffverarbeiter	1
Finanzrecht	1
Magnesita	1
Ölsorten	1
Glaserhersteller	1
Fernwärmenetze	1
blau-gelbe	1
energieintensiver	1
Spotpreis	1
Notfallpläne	2
Ukraine-Konflikts	1
zukunftsfit	1
Lukoil	1
Preissprünge	1
Explodierende	1
Menschenrechtsverstöße	1
Eigeninitiativen	1
oberösterreichische	3
Preissprung	1
Stagflation	1
Energielieferungen	1
Zweitwohnsitze	1
Herauskommen	1
Respektabstand	1
Klubobmann	3
Erdgasversorgung	1
Hausverstand	2
zuzusichern	1
Preisrückgänge	1
Gasmengen	1
profitierenden	1
Vorgeben	1
Kiril	1
Gasunternehmen	1
Donawitz	1
Expertenschätzungen	1

durchgesickerten	1
Belakowitsch	1
Investitionspaket	1
Solarkraftwerken	1
Wirtschaftsforschungsinstitut	1
Krisenfall	3
Rekordpreis	1
Auszahlen	1
Doppelbauer	1
frühsommerlichen	1
Ölgesellschaften	1
petrochemische	1
Abrechnungssysteme	1
Jahresvergleich	5
Strompreiserhöhung	1
Rekordhöhen	1
Ärmere	1
Investitionsprojekten	1
Standortnachteil	1
Walzwerke	1
Baumgartenberg	1
hiesigem	1
Meinungsspektrum	1
energieautark	1
Norske	1
Treibstoffen	2
Entkoppelung	2
Häuslbauer	1
Rumen	1
Stillstände	1
Industriestandort	3
Kanzleramts	1
Gaszufuhr	1
Pasching	2
Gaswirtschaft	1
Hautcremes	1
umzuverteilen	1
Beckermann	1
Frühindikator	1
Sparanstrengungen	1
Aktionärsstruktur	1
Energieversorgers	2
beherrschendes	1
geschnürte	1
Vorständin	1
Nordseesorte	1
Vorschreibungen	1

Handelsorganisation	1
Einspeicherung	1
Gesamtprodukt	1
Produktionswert	1
Elektronikindustrie	2
prorussische	1
Heizsaison	1
abzufedern	3
Gaskraftwerk	1
Agrana	1
ZiB	1
Windrades	1
Salinen	2
Klimastrategie	1
Preisgrenze	1
Rohöl	5
Energiehunger	1
Rohölpreis	1
Lieferketten	6
Ölgeschäft	1
Branchenkreisen	1
einberechnen	1
Brennöfen	1
klimaschädliches	1
Energieinfrastruktur	1
Standesvertretung	1
Unterstützungsmaßnahmen	2
ausverhandelt	1
Energiewelt	1
Gesetzesentwurfs	1
Energieprojekte	1
Ausgangslagen	1
Auswegen	1
Energieexperte	1
abgegraben	1
Kommunikationsabteilung	1
Wienerberger	2
gespeichertes	1
Gaskraftwerken	1
gegengerechnet	1
umweltbewusster	1
Energieministerium	1
konzerneigenen	1
Strompreise	6
Unternehmenssprecher	2
Nebenwohnsitz	1
Containerhafen	1

Megawattstunden	1
Liquiditätsprobleme	1
EU-27	1
abschätzbar	2
weitertransportiert	1
Merit	2
Stahlwerken	1
Transportkapazitäten	1
Versorgungsengpässen	1
Stromversorger	3
Jahresabrechnung	2
Jahresverbrauch	2
westeuropäischer	1
realitätsnahe	2
Ostalgie	1
verschifft	3
ACER	1
Reduktionsmittel	1
Energiefresser	1
Terawattstunden	1
Einheitswährung	1
Spitzenvertreter	1
Energielösungen	1
Argus	2
Kabelbäume	1
rascheren	1
Kilometergeld	1
Protestpartei	1
zerbombt	1
gegenzusteuern	2
explosionsartigen	1
Worst-Case-Szenario	1
Fleischwaren	2
Energiefragen	1
bilanziell	1
Wirtschaftsforscher	1
Sondersteuer	1
Flachglas	1
weiterzuverkaufen	1
Gratkorn	1
Spectra	1
Vorprodukte	1
inflationbereinigt	1
explodierenden	2
Angerer	3
Standardprodukte	1
Preisanpassungen	1

Chemieindustrie	2
aufzudrehen	1
RHI	1
Fatih	3
Rückgängen	1
Kriegskasse	1
ersetzbar	2
Osterreich	1
IVA	1
Energieversorger	8
Sparens	1
einpendelt	1
verfeuert	1
leistbare	1
Erdgas	19
Strompreis	4
Schurkenstaaten	1
Haushaltsgröße	1
voestalpine	2
Unterzeichnet	1
Überregulierung	1
subventionieren	2
Abschöpfung	1
Fotovoltaikanlage	1
ÖGB	3
schrumpften	1
nostalgisches	1
Energiekonzerns	1
Industrieunternehmen	6
Bulgaren	3
Industriebetrieben	2
Treibstoffe	2
auszuhelfen	1
Branchenvertreter	1
SPO	1
Parteilinie	1
Arbeitspreis	1
deckeln	1
Ol	1
Schallenberg	1
Brent	4
eingefrorene	1
Gasversorger	1
leistbaren	1
Glasindustrie	1
STANDARD	2
Umfrageergebnisse	2

Auskleidung	1
Blackouts	1
Lagerkapazitäten	1
Massiv	3
Welthandel	3
prorussischen	1
Müllverbrennung	1
Ölmarkt	1
Stromkosten	6
Vorsorgen	1
Staatseigentum	1
Verteuerung	2
Höchstgeschwindigkeiten	1
Wirtschaftsvertreter	1
CO2-Bepreisung	1
Tarifierhöhung	1
Aufschieben	1
GPA	1
Wettbewerbsnachteil	1
Genehmigungsverfahren	5
Loacker	1
Wirtschaftsräume	1
angestoßene	1
Schnabl	1
Franchisenehmer	2
Datenlage	3
erneuerbar	1
Kostensteigerungen	2
Ausgleichszulage	1
Gemüsehändler	1
Gewerbekunden	2
Staffelung	2
traditionellerweise	1
Großhandels	1
Rohstoffgewinnung	1
Indra	2
prescht	1
Energiemengen	1
Materialbeschaffung	1
niederösterreichischen	3
Regulierungsbehörde	3
Energieverschwendung	1
Aschach	1
Marktversagen	1
aufetrete	1
Staatshilfe	1
novelliert	1

Schweröl	1
Theiß	1
intelligentere	1
Pipelines	2
Energieträgern	3
draufzahlen	1
Zündstoff	2
Stromsparen	1
Klinger	3
Gegenmaßnahme	1
Adaptierung	1
Kaffeetassen	1
Gas	69
Haftungen	1
herausforderndes	1
Pichler	4
Bauleistung	1
einkommensschwache	1
prioritär	1
CO2-Preis	1
Wirtschaftszeitung	1
Preiserhöhung	4
Fass.	1
oberösterreichischen	3
Leichtfried	1
Sozialeinrichtungen	1
Südburgenland	1
Investitionssicherheit	1
Strombörse	1
Privathaushalte	3
EU-Durchschnitt	1
Weitergeben	1
Fossile	1
bilateraler	1
Grundnahrungsmitteln	1
FPÖ-Chef	1
Einmalzahlungen	1
halbjährlichen	1
Auffüllung	1
abfließt	1
Hochfilzen	1
Energieträger	8
sondiert	1
auszutrocknen	1
Gewinnmargen	1
Steigerungen	3
Stromrechnung	2

versorgung	1
Wirksam	1
Kernkraft	2
Fernwärme	4
verbrauchten	3
hissen	1
geglückte	1
Betriebsweise	1
Direktiven	1
nachgefragte	1
treffsicher	2
feststellbaren	1
verzehnfacht	1
umweltschonend	2
Energieeinsatz	1
ÖAMTC	2
Supermarktketten	1
Fehlendes	1
Wertschöpfungsketten	2
Verknappung	2
Belastungsprobe	1
Sicherheitsnetz	1
Verkehrsbereich	1
Energieform	1
abgeschöpft	1
zurückfahren	2
Diversifizierung	2
Mittelbetriebe	1
liebäugeln	1
WTI	1
Energiesystems	1
Solveig	1
Traunsee	1
Zeitenwende	2
Geldentwertung	1
Melderegister	1
diversifiziert	1
Ebensee	1
Überdenken	1
Halbierung	2
Preiserhöhungen	3
flächendeckender	1
Produktionskette	1
Mineralölsteuer	1
voriger	1
Strombedarf	2
verteuert	2

Rasierklinge	1
Gewerbebetriebe	2
Lebensmittelpreise	1
Energiemix	1
unwichtiger	1
klarerweise	1
fossilen	7
Meinungsbild	2
westlichsten	1
Stromproduktion	2
Energiepolitik	4
Papierindustrie	1
Weihnachtspause	1
Kilowattstunde	3
Schlüsselfigur	1
Ölindustrie	1
abschöpfen	1
vertragsgemäß	1
Bestandskunden	4
Härtefälle	1
Rundfunkgebühr	1
Industriebetriebe	2
Fertigungsprozesse	1
Lebensmittelhändler	1
Meinungsforschungsinstitut	1
Schmierstoffe	1
Doppelmoral	2
Dämmmaterial	1
Umweltschutzorganisation	1
gefördertes	1
Fotovoltaik	1
abgefedert	1
Eisenerz	2
Sowjetarmee	1
OÖ	7
bearbeitung	1
Energieversorgung	11
Kastner	2
Rendi-Wagner	1
geeint	1
Marktmacht	2
Marktpreise	1
Kogler	2
Produktionsprozessen	1
Freiheitliche	1

Baubereich	1
Ausgleichsmaßnahmen	1
Kilowattstunden	3
Neuhauser	1
Marktpreis	2
Atomkraft	4
Düngemitteln	1
harschen	1
Feldversuch	1
Parteilpolitik	1
wettbewerbsfähig	3
Kosteneinsparung	1
Versorgungssicherheit	3
Wladiwostok	1
Pendlerpauschale	1
Oberösterreicher	1
zurückgehende	1
Flüssiggas	2
getriebene	1
Durchschnitts	1
Investitionsentscheidungen	1
bulgarische	2
Solang	1
abmildern	1
Konsumentinnen	1
Notfallplan	1
energieintensiver Betrieb ¹	5
türkis-grüne Regierung	4
hoher Energiepreis	6
Burgenland Energie	3
Megawattstunde Gas	3
Sorte Urals	3
Joachim Haindl-grutsch	3
Energieallianz Austria	3
Strom- und Gaspreis	4
Leonore Gewessler	4
hoh Energiekosten	5
Fatih Birol	3
Dollar pro Fass	3
russisches Gas	6
hoher Gaspreis	3
russisches Öl	3
Wien Energie	7
fester oder flüssiger Energieträger	2

¹2. Teil der Termkandidatenliste (Mehrwortbenennungen)

Peter Eckerstorfer	2
Zündstoff für die Lohnrunde	2
Preisdeckel für Gas	2
bilateraler Gasdurchleitungsvertrag	2
Martin Bergsmann	2
Standardtarif der Wien	2
niedriger Deckel	2
vorgezogene Preiserhöhung	2
Indexanpassung in den Standardtarifen	2
Oberösterreichs Industrieunternehmen	2
Landessprecherin Indra	2
Wifo-chef Felbermayr	2
zuständiges Klimaschutzministerium	2
Speicher Haidach und 7fields	2
Geschäftsführer zu den Ergebnissen	2
Energieministerin Leonore	2
gekommener Welthandel	2
Forderung nach einem Preisdeckel	2
Landessprecherin Indra Collini	2
in Fahrt gekommener Welthandel	2
Industriestandort Oberösterreich	2
Kühlvitrine ohne Türen	2
Hochreiter Fleischwaren	2
Energiekrise des Jahres	2
Energieministerin Leonore Gewessler	2
Ekkehart Pichler	2
Speicher Haidach	2
deutliches Meinungsbild	2
Weg aus der Energiekrise	2
russischer Angriffskrieg	2
Landespartei- und Klubobmann	2
russischer Ölexport	2
Deindustrialisierung Europas	2
Treffen der Eu-energieminister	2
oberösterreichische Industrie	2
Handel mit Gas	2
Chef der internationalen Energieagentur	2
Wolfgang Urbantschitsch	2
Störung der Energieversorgung	2
niederösterreichische Landeshauptfrau Johanna	2

Mensch in einem Haushalt	2
leichte Entspannung	3
Prozent der Mehrkosten	2
energieintensives Industrieunternehmen	2
niederösterreichische Landeshauptfrau	2
Hilferuf an die Politik	2
Klimaschutzministerin Leonore Gewessler	2
Klimaschutzministerin Leonore	2
Indra Collini	2
Christian Klinger	2
Peter Untersperger	2
Regulierungsbehörde E-control	2
Peter Weidinger	2
Georg Knill	2
russisches Erdgas	3
teures Kraftwerk	2
schnelles Genehmigungsverfahren	2
Eu Sanktion	2
rasche Maßnahme	2
Megawattstunde Strom	2
gestiegene Kosten	4
Euro pro Megawattstunde	2
flüssiger Energieträger	2
soziale Staffelung	2
Sicherheit der Energieversorgung	2
Sanktion des Westens	2
russische Gaslieferung	2
Landeshauptfrau Johanna Mikl-Leitner	3
Udo Landbauer	2
wöchentlicher Einkauf	2
Landeshauptfrau Johanna	3
elektrischer Rasierer	2
energieintensive Industrie	3
österreichische Energieagentur	2
weite Preissteigerung	2
Vizekanzler Werner Kogler	2
Vizekanzler Werner	2
beteiligter Betrieb	2
andere Börse	2
Salzburg Ag	3
rascher Ausbau	2
bulgarische Hauptstadt	2

durchschnittlicher Jahresverbrauch	2
Dezember 2022	2
Karl Nehammer	2
Kubikmeter Gas	2
steigender Energiepreis	4
Ausbau der Erneuerbaren	2
Wien –	7
Energie- und Klimapolitik	2
Gesamt-energie Chinas	1
Geschäftsführer zu den Ergebnissen der Studie	1
Eingriffsmöglichkeit in Erwerbsfreiheit	1
Gewährleistung der sicheren Gasversorgung	1
Strommarktbericht der Iea	1
Rohstoffgewinnung in der Produktionskette	1
Großhandel am Montag	1
Embargo auf russisches Erdgas	1
Schock auf den Energiemärkten	1
Reaktion auf die Sanktionen des Westens	1
Haushaltskundenpreise auf Vorkrisenniveau	1
Verzicht auf hohen Gasverbrauch	1
Trumpf Austria	1
Rückgang beim Verbrauch	1
Prozent im November gegenüber Oktober	1
Aufschieben der Co2-aufschläge	1
Schulterschluss in der Koalition	1
Abweichung von Emissionsgrenzwerten	1
Welt gegen den Aggressor	1
Berechnung der österreichischen Energieagentur	1
Schritt Energiekostenzuschuss	1
Berechnung des Instituts für Weltwirtschaft	1
Wettbewerbsnachteil für den Industriestandort	1
Grundsatz der Entschädigung bei Enteignungen	1
Hafenstadt Primorsk	1
Hand aufgrund der Aktionärsstruktur	1
Vorständin des Instituts für Finanzrecht	1

Beruhigung des europäischen Gasmarktes	1
Prävention von Störungen der Energieversorgung	1
Ankündigung der Energieallianz	1
aktuelle Gaskrise	1
Energie getriebene Inflation	1
Schätzung 1.000	1
Herstellung von Vormaterialien	1
Energie-kostenniveau in den nächsten Monaten	1
Umrüstung von Gas	1
Hintergrund der hohen Gaspreise	1
Energieagentur in einer Studie	1
Risiko unerwünschter Effekte	1
Vielzahl von Eu-ländern	1
Ifw-volkswirt Stamer	1
Ihs-chef Klaus	1
Vergeltung für die harschen Sanktionen	1
Staat mit der Rasierklinge	1
Importstopp zu einem Einbruch	1
Energieexperte der Umweltschutzorganisation	1
Bio- oder Gemüsehändler	1
Industrie von den gesunkenen Großhandelspreisen	1
Energiekonzern in eine Art	1
Energiekosten an den Gesamtkosten von Unternehmen	1
Westen Geschäft	1
Inflationsdatum der Statistik	1
Ausbau von Windkraft und Solaranlagen	1
Ansicht des Chefs der Regulierungsbehörde	1
Spüo-landespartei vorsitzende Franz	1
Regierung gegen die Preisanstiege	1
Ausbruch des Ukraine-krieges	1
aufgelassenes –	1
abgeschlossenes Abkommen zu den Speichern	1
Versorger als Krisengewinner	1
Salz-produktion seit Jänner	1
Rabatt gegenüber dem Weltmarktpreis	1
Branchenvertreter Alarm	1
Putins Erpressungsversuch	1

Beitrag zur Reduzierung der Energiekosten	1
Jahr gesicherte Energiepreise	1
Energielösung Herausforderung	1
Verfügbarkeit von Stahl	1
Strompreisdeckelung in den vergangenen Wochen	1
Ansicht von E-control-chef	1
Energienotfall klare Direktiven	1
Sanktion des Westens nach dem Einmarsch	1
Energiepolitik Österreichs	1
Energiepreisen Salz	1
Zufallsgewinne von Stromunternehmen	1
Teil der Eaa	1
Kampf gegen Gasknappheit und Energiepreise	1
Energiesparen in Planung	1
Auskleidung von Stahl	1
Solidaritätsabkommen zur Gewährleistung	1
Energiesparsames Verhalten im Betrieb	1
Auslastung durch Industrie	1
Woche zu einem Preiseinbruch	1
Klimaschutzministerium Energielenkungsmaßnahmen	1
Energieversorger am Donnerstag	1
Auslastung durch Industrie und Private	1
Kohle statt Gas	1
Engpass aufgrund des russischen Aggressionskriegs	1
Entkoppeln des Gaspreises vom Strompreis	1
altes energiefressendes Gerät	1
alternative und nachhaltige Energie	1
Zählpunkt des Stromanschlusses	1
Strompreiserhöhung für Bestandskunden	1
Konsequenz auf die Lieferketten	1
Ausnahmefall mittels Verordnung	1
Ukraine-krieg –	1
Kosten für den Arbeitspreis	1
Entspannung bei den Haushaltsenergiepreisen	1
Umwandlung von Eisenerz zu Eisenschwamm	1

Teuerung zu bundesweiten Protesten	1
Krise durch Anreize	1
Ceo Martin Bergsmann	1
Standard-anfrage im Klimaschutzministerium	1
Anteil an den globalen Co2-emissionen	1
Aussetzen des Co2-preises	1
Unternehmen Hochreiter	1
Rahmen der konzerneigenen Klimastrategie	1
Land geförderte Rohöl	1
Erfolgsmodell Österreichs	1
Ergebnissen Auskunft	1
Vertrag mit einer Preisgarantie	1
Landesversorger Tiwag	1
Stärkung der Wasserkraft	1
Zündstoff für die Lohnrunde ab Montag	1
Technologie zum Ersatz	1
Espo-pipeline an die Pazifikküste	1
Ausweg aus dem Preischao seit Monaten	1
Lieferung im Winter	1
Eu-27 trotz unterschiedlicher Ausgangslagen	1
Eu-energieminister der Be-schluss	1
ausgelegtes Energielenkungs-gesetz auf Stand	1
andere russische Großstadt	1
aktuelle Marktstörung	1
Vorschlag aus den Reihen der Wissenschaft	1
Regierung seit Kriegsbeginn	1
Putin drei	1
Prozent in öffentlichem Eigen-tum	1
Christian Prauchner	1
Vergleich zu den internationalen Wettbewerbern	1
Seite der Haushalte	1
Euro an Marktwert	1
Strompreisbremse der Bundesre-gierung	1
Rohöl im Einkauf	1
Co2-ausstoß der weltweiten Stromerzeugung	1
Marktstörung in Milliardenhöhe	1

Martina Stemmer	1
Anteil der Energiekosten an den Gesamtkosten	1
Auswirkung des Ukraine-konflikts	1
Unternehmen auf alternative Energieformen	1
Suche nach der Energielösung	1
Maßnahme mit den Eu-beihilfe-vorgaben	1
Sache Spritpreis	1
Bemessung von Entschädigungen	1
Eigentums- und Persönlichkeitsrecht	1
Bemühung in die Richtung	1
Zweck der Gewährleistung der Versorgungssicherheit	1
Beobachtung des Öamtc	1
Geschäft an anderen Börsen	1
Aufstellen eines Windrades	1
alter Kohlekessel	1
Abkehr vom Grundsatz der Entschädigung	1
Gewessler-sprecherin Martina	1
Auftrag aufgrund fehlender Wettbewerbsfähigkeit	1
Stopp der Gaslieferungen	1
Einsparung in Industrie	1
Einsparung in Industrie und Haushalten	1
Auftrag des Kanzleramts	1
ausgelegtes Energielenkungs-gesetz	1
Bereich Chancen	1
Unterbrechung der Produktion von Kaffeetassen	1
Adaptierung der Merit	1
Rekordpreis von voriger Woche	1
Emission bei einem chinesischen Anteil	1
Strompreisbremse wegen der komplizierten Datenlage	1
Rente mittels Ausgleichszulage	1
aktuelle Energiepreiskrise	1
Haushalt Mehrkosten	1
Haushalt mit der Strompreis-bremse	1
Endverbraucherpreis in absehbarer Zeit	1

Verbrauch eines durchschnittlichen österreichischen Drei-personen-haushalts	1
Beschichtungsspezialist Hueck	1
Heiligenkreuz im Südburgenland	1
Betrieb Oberösterreichs	1
auf Gas anderer Lieferant	1
Anordnung der Einsparung	1
Sicherheitsnetz für die stabile Gasversorgung	1
Betrieb von abgesicherten Preisen	1
Thema Teuerung	1
Abkehr von fossilen Ressourcen	1
Hilfe für einkommensschwächere Haushalte	1
Betriebsweise von Anlagen	1
Woche Stefan Borgas	1
Anpassung für Bestandskunden	1
Rohstoff zur Herstellung von Düngemitteln	1
Bevölkerung größtmöglichen Schaden	1
Viertel der teilnehmenden Betriebe	1
Energieeffizienz per Gesetz	1
Rückkehr der Haushaltskundenpreise auf Vorkrisenniveau	1
Energieeffizienzgesetz 2020	1
Zweifel an Versorgungssicherheit	1
Energieforecasts Trumpf	1
Energieforecasts aus der Industrie	1
Industrie Kostenfaktor	1
Werkzeugmaschinen-hersteller Trumpf Austria	1
Blitzumfrage teilnehmenden Betriebe	1
Verbrauch von Ölprodukten	1
Energiekosten der österreichischen Industrie	1
andere Hälfte der Unternehmen	1
Boykott von russischem Öl	1
Ausbleiben flächendeckender Blackouts	1
Brancheninformationsdienst Argus	1
Unterstützungsmaßnahme für die energieintensive Industrie	1
Branchenverband im Namen	1

Energiekostenzuschuss der Regierung	1
Energiekostenzuschuss für Betriebe	1
beschlossener Maximalpreis	1
Bremse ab Dezember	1
Rückgang beim Verbrauch von Ölprodukten	1
Ausfall des Handels	1
Rückkehr auf Vorkrisenniveau	1
Brennstoff durch nachhaltige Energieträger	1
Jahresabrechnung zur Anwendung	1
Jahresbeginn 2023	1
Warenkontrolle zur Einhaltung der Sanktionen	1
Ausgangslage auf ein gemeinsames Drehbuch	1
Wirtschaftszeitung Wedemosti	1
Ag für eine Kilowattstunde	1
Stillstand der Produktionen	1
Ausgleich für soziale Härtefälle	1
Sabine Kirchmayr-schliesselberger	1
Kampagne im Herbst für Energiesparen	1
am Mittwoch publizierter Strommarktbericht	1
Energiespargrund zur Nassrasur	1
Kanzler Karl Nehammer	1
Abkommen zu den Speichern	1
Steigerung in einer Aussendung	1
Energieunternehmen beim Handel	1
Energieverbrauch je Fahrzeug	1
Bundeskanzler Karl Nehammer	1
aktuelle Strom-gaspreis-koppelung	1
Bundesland von fundamentaler Bedeutung	1
Energieversorgung besonderen Wert	1
Bundesregierung als ersten Schritt	1
Umformung von Metallen	1
Auslaufen des alten Energieeffizienzgesetzes	1
Kommunikationsabteilung des Energiekonzerns	1
Entkoppelung des Strompreises	1

aufgrund des Krieges außergewöhnliche Entwicklung	1
Entkoppelung von Strom	1
Entkopplung von Strom	1
Kosten der Bauleistung	1
Wirtschaft seit Beginn des Ukraine-krieges	1
Ceo Ekkehart	1
Kraftakte bei der Suche	1
Krieg außergewöhnliche Entwicklung	1
Versorgung des süddeutschen Raums	1
Erdgas mit massiv steigenden Kosten	1
Kunststoffverarbeiter Röchling	1
Erdgasabhängigkeit der nächste Punkt	1
Treffen der Eu-energieminister der Beschluss	1
Erdgastarif der Linz	1
Ceo Peter Weidinger	1
Ceo Thilo	1
allfälliger Vermögensvorteil	1
Anteil aus erneuerbarer Energie	1
Reduktionsmittel für die Umwandlung von Eisenerz	1
Verantwortung zur Nutzung	1
Wolfgang Benischko	1
Stromkosten bei Unimarkt	1
Wirtschaftsminister Robert	1
Ausweg aus dem Preischaos	1
altes Energieeffizienzgesetz	1
Espo-pipeline an die Pazifikküste bei Wladiwostok	1
Eu Gasembargo	1
Chinas Comeback	1
September für einen durchschnittlichen Haushalt	1
Anteil beim Verbund	1
Lieferung zu Verzögerungen und Einschränkungen	1
Eu-energieminister seit Beginn	1
Eu-energieminister seit Beginn der Energiekrise	1
Abhängigkeit von Internorm von Gas	1
Ländern Vorprodukt	1
Strompreis vom Gaspreis	1
Verkauf von Fahrzeugen mit Elektroantrieb	1

Aluminium für den Baubereich	1
Machined Components	1
Sanktion im Raum	1
Unimarkt-chef Andreas	1
Co2-emission aus dem Verkehrssektor	1
Veranstaltung in der Tu	1
Euro für Superbenzin	1
Stromproduktion zu einem überwiegenden Anteil	1
Euro gedeckelter Großhandelspreis	1
Comeback als Importeur	1
Containerhafen St.	1
Zeit der Einsparungen in Industrie	1
Anbetracht der aktuellen Großhandelssituation	1
Gasspeicher Haidach	1
Maßnahme zur Prävention von Störungen	1
Auswirkung mit Jahresbeginn	1
Stromsparen in Spitzenzeiten	1
Strom 365	1
Euro über dem Weltmarktpreis für Flüssiggas	1
Datenlage bezüglich der Möglichkeiten	1
Mehrkosten für Energie	1
Auto- oder Fensterscheibe	1
Datenlage bezüglich der Möglichkeiten des Sparens	1
Melderegister mit den Zählpunkten	1
Verfügbarkeit von Stahl und Beton	1
Melderegister mit den Zählpunkten des Stromanschlusses	1
Evn angekündigten Strompreiserhöhungen	1
Datum aus dem zentralen Melderegister	1
Transformation im Energiebereich	1
Regelwerk beim Strom	1
Export als wirtschaftliches Erfolgsmodell	1
Senkung der Elektrizitätsabgabe	1
Militär von Hunderten	1
Datum der deutschen Bundesanstalt	1

Datum des Brancheninformationsdienstes	1
Prozentsatz an Rabatt	1
Gaspreis vom Strompreis	1
Minderheit der oberösterreichischen Industriebetriebe	1
außergewöhnlicher Preisausschlag	1
ausgebaute Energieinfrastruktur	1
Vorstand des Interessenverbands	1
Thema Preisgünstigkeit	1
Dauerbeflegelung durch die Opposition	1
Miniwarenkorb für den wöchentlichen Einkauf	1
Anweisung für die Betriebsweise	1
September 2023. Frage	1
Mittwoch im Klub	1
Abschöpfung sogenannter Zufallsgewinne	1
Spo Niederösterreichs	1
Modell im Ministerrat	1
Fall einer Energielenkung	1
Fall einer Gasmangellage	1
Monat nach Ausbruch des Ukraine-krieges	1
Dekade Geldentwertung	1
Montag Analyse	1
Weltmarktpreis für Flüssiggas	1
Verbot für Verträge	1
Sinn der europäischen Werte	1
Fall einer Notlage bei Strom	1
Anweisung für die Betriebsweise von Anlagen	1
Schweißen von Großteilen	1
März Plan	1
Möglichkeit für eine kurzfristige Veränderung	1
Verbund-großanlage Mellach	1
Rally beim Gaspreis	1
Detail zur Strompreisbremse	1
Deutlich erhöhter Preis	1
Nahversorger Krise	1
Fass für Lieferungen an Drittstaaten	1
Nassrasur statt elektrischer Rasierer	1
abrupter Stopp der Gasimporte	1
Vorgabe für Tempolimits	1

Vielfache über dem Vor-krisen-niveau	1
Nehammers Ankündigung	1
Nationalrat gegen Preisdeckel	1
Deutliche Verteuerung	1
Trockner Gas	1
Neos-wirtschaftssprecher Gerald	1
Feldversuch mit Stromkunden	1
Transit für Gas	1
Feldversuch mit Stromkunden und Stromkundinnen	1
Deutsch am Sonntag	1
Fernsehsender Puls4	1
Fernwärme-tarif des städtischen Energieversorgers	1
Baustoff- oder Chemieindustrie	1
Sicht in einem Jahr	1
Wirtschaftsminister Robert Habeck	1
Superbenzin neue Höchstwerte	1
Statistik von Ende	1
Deviseneinnahme aus dem Öl-verkauf	1
Finanzminister Brunner	1
Notlage bei Strom und Gas	1
Finanzminister Magnus	1
Dezember geltende Importem-bargo	1
Standort Pasching	1
Sozialsprecherin der Freiheitli-chen	1
Finanzministerium im Gespräch	1
November geformten Regierung	1
Wärme in Unternehmen	1
Finanzwert der Fall	1
Fixpreis mit Jahresende	1
besonders kleines Einkommen	1
ausgelöste Multi-energiekrise	1
alternative Energie Christian	1
abgesicherter Preis	1
Obergrenze beim Strompreis für Privathaushalte	1
Flachglas für Auto	1
Director Europe Machined	1
Bedarf an Co2-neutralem	1
Vergleich mit August	1
Rohstoff an Länder	1
Flüssiggas-abnehmer der Welt	1
Stopp russischer Gaseinfuhren	1

Regierung im Kampf gegen Gas-knappheit	1
Direktive für den Fall einer Not-lage	1
Diskurs von Gegensätzen	1
Papierkonzern mit Hauptsitz	1
Folge eines Gasboykotts	1
Diskussion bezüglich eines Gasembargos	1
Zweifel an der Versorgungssi-cherheit	1
Welt laut Daten	1
Rückgang bei den Strompreisen	1
Folge neben der Mobilität	1
Genehmigungspolitik Öster-reichs	1
Genehmigungsverfahren für Inf-rastruktur	1
Genehmigungsverfahren für Inf-rastruktur und Energieprojekte	1
Genehmigungsverfahren für al-ternative Energien	1
Wifo-chef Gabriel	1
Einbruch seit Frühjahr	1
aktuelles Energie-kostenniveau	1
Eindämmung der Gaspreise	1
Gesamtkosten von Unternehmen	1
Einfamilienhaus mit einem Jah-resverbrauch	1
Vielfache des langjährigen Durchschnits	1
Eingriff in Eigentums	1
Steuerbelastung bei den Energie-kosten	1
beeinträchtigte Lieferkette	1
Eingriffsmöglichkeit in Erwerbs-freiheit und Marktwirtschaft	1
Gespräch mit den Kurier	1
Gewerbebetriebe Kühlvitrine	1
Solidaritätsmaßnahme der Eu-staaten	1
Gewicht auf den Weltmarkt	1
Gewinnabschöpfung wegen stei-gender Energiepreise	1
Einsparung bei Gas	1
Strompreisbremse der Bundesre-gierung in Kraft	1
Gigant im Ölgeschäft	1
Giovanni Staunovo	1
Gipfel des Witoscha	1

Stahl- und Metallverarbeitungs-industrie	1
Einstellung zur Atomkraft	1
Griff nach Zufallsgewinnen	1
Eisenerz für die Stahlerzeugung	1
begrenzte Solidaritätsabgabe	1
Großhändler Kastner	1
Eisenerz zu Eisenschwamm	1
Vorstand des Interessenverbands für Anleger	1
Elektroofen zu Stahl	1
Gurken- oder Marmeladeglas	1
Gutschriften –	1
Welt mit vollem Gewicht	1
Haftung der österreichischen Kontrollbank	1
Halbierung der Gaszufuhr	1
Hallenbeleuchtung auf Led	1
Werkzeugmaschinen-hersteller Trumpf	1
Hand gehaltene Anteile	1
Treffen der Eu-energieminister seit Beginn	1
Diskussion bezüglich eines Gasembargos gegenüber Russland	1
Art alternative Börse	1
Forderung nach einem Preisdeckel für Energie	1
Tempolimit auf Straßen	1
Haushalt gemeldeten Personen	1
Transparenz der Notfallpläne	1
Haushalt unter den hohen Energiepreisen	1
Umstieg auf andere Energiequellen	1
Beschaffungsstrategie der Energieversorger	1
Verhandlungstisch mit Diplomaten	1
Haushaltspreis für Strom	1
Energie Christian	1
Energie Christian Klinger	1
Gasverbrauch seitens der Industrie	1
Energie Martin	1
Heizung in öffentlichen Gebäuden	1
Energie Martin Bergsmann	1
besagte Kampagne im Herbst	1
Betrieb der Brennöfen	1

Staatspräsident Rumen Radew	1
Herbst für Energiesparen	1
Tag nach Unterzeichnung in Kraft	1
Energie gewonnenen Stroms	1
Rede vom Sparen	1
Spitzenvertreter der Erdöl	1
Pkw E-autos	1
Diskussion um eine mögliche Strompreisdeckelung	1
Zittern um die Gasversorgung in Europa	1
Herz der österreichischen Salzgewinnung	1
Hilfe Lieferkette	1
Hilfe für die Wohnraumerrichtung	1
Simulation für die Versorgung	1
Route in Kauf	1
allgemeine Sondersteuer	1
Energieagentur mit Sitz	1
Bevölkerung Oberösterreichs	1
Energieagentur-chef Angerer	1
Vorstandsvorsitzende Erich Entstrasser	1
Diskussion über Deals	1
Höchstgeschwindigkeit von Kraftfahrzeugen	1
Energiebörse zur Beruhigung	1
Prüfung einer Gewinnabschöpfung	1
Befindlichkeit der einzelnen Branchen	1
Ifw-volkswirt Vincent	1
Ifw-volkswirt Vincent Stamer	1
Energieeffizienz über Energiesparen	1
Thema Energieunternehmen	1
Ihs-chef Klaus Neusser	1
Importembargo der Eu	1
Importstopp für russisches Öl	1
beeinträchtigte Lieferkette durch den Krieg	1
Importstopp zu einem Einbruch der Wirtschaftsleistung	1
Diskussion über die Funktionsweise	1
Industrialisierung von Technologien	1
Abhängigkeit von Internorm	1

Suche nach einer raschen Lösung	1
Industrie Oberösterreichs	1
Zellstoffkonzern Sappi	1
Industrie vor extremen Herausforderungen	1
Energiekosten an den Gesamtkosten	1
Zusammenarbeit im Parlament statt Dauerbeflegelungen	1
Wochenbeginn verschärften Sanktionen	1
Umwandlung von Eisenerz	1
Prozent petrochemische Produkte	1
Politik gegenüber dem Standort	1
Fpö-chef Herbert	1
Industrieunternehmen Oberösterreichs	1
Rohstoff für die Transformation	1
auch europäische Einkaufsvereinigung	1
Inflationswert bis Ende	1
Energiekosten von einer drohenden Deindustrialisierung	1
Installation der Ökostromanlagen	1
Energiekostenanteil an den Gesamtkosten	1
Energiekostenanteil an den Gesamtkosten des Unternehmens	1
Instrument der Überbrückung	1
Energiekostenhilfe im Ausmaß	1
Vorsitzender Pamela	1
Internorm von Gas	1
Energiekostenzuschuss I	1
Investition zu Einsparerfolgen	1
Investitionsprojekt auf Firmen	1
Investitionssicherheit im Land	1
Iv-chef Knill	1
Energiekrise Maßnahme	1
beteiligter Betrieb von abgesicherten Preisen	1
Energiekrise Wettbewerbsfähigkeit	1
Verantwortung der Politik gegenüber dem Standort	1
Verantwortung für das Befüllungsziel	1
Energielenkungsgesetz auf Stand	1

Rabatt gegenüber dem Weltmarktpreis für Brent	1
Energielenkungsgesetz für den Fall	1
Vincent Stamer	1
Energielenkungsgesetz für den Fall von Störungen	1
Jahr zeitlicher Abstand	1
Jahresabrechnung des Energieversorgers	1
Temperaturbegrenzung für Heizungen	1
Befugnis für den Energienotfall	1
Diskussion über die Funktionsweise der Marktmechanismen	1
Art der Strompreisbildung	1
Energieministerium –	1
Journalistin Wessela	1
Journalistin Wessela Wladkowa	1
Zweifel für die Industrie	1
Juli Diesel	1
Energiepolitik neue Wege	1
Junge der Umweltschutz	1
alternative Gasbezugsquelle	1
Energiepreis in Rekordhöhen	1
Spüo-landespartei-vorsitzende Franz Schnabl	1
Jänner 2023.	1
Walzwerk in den Stahlwerken	1
Sicherheit der Energieversorgung auf Platz	1
Simulation für die Versorgung in Notfällen	1
Kampf gegen Gasknappheit	1
beliebter Politiker Bulgariens	1
Kampf gegen die Preisentwicklung	1
Kampf mit dem rasanten Anstieg	1
Energiesparkampagne für den Herbst	1
Stornierung von Investitionsprojekten	1
Kanzler Nehammer	1
Energiesparsames Verhalten	1
Totalausfall russischer Lieferungen	1
Kilowattstunde bei der nächsten Jahresabrechnung	1
Kiril Petkow	1
Klaus Neusser	1
Spritpreis im Vergleich	1

Energieversorger Salzburg	1
Diskussion über die Marktmechanismen	1
Franchisenehmer einer Handelsorganisation	1
Solveig Menard-galli	1
Sicherheitspolster für den Krisenfall	1
Klimastrategie auf neue Energielösungen	1
Energieversorgung auf Platz	1
Klubobmann Jörg	1
Klubobmann Jörg Leichtfried	1
Energiewelt zu Verknappung	1
Thema günstiger Preis	1
Kohle im Gasnotfall	1
Umrüstung von Gas auf Öl	1
Entkoppeln des Gaspreises	1
Kohle- und Laufkraftwerk	1
Kombination mit den europaweiten Anstrengungen	1
Verbrauchsreduktion in der Industrie	1
Komplettstopp der Produktion	1
Entkoppelung des Strompreises vom Gaspreis	1
besonders hitzeresistentes Material	1
Konsum gegenüber dem Durchschnitt	1
Konto der Industrie	1
Konto in Sachen	1
Strom im November	1
Stopp russischer Gaslieferungen	1
Anfang 2023. Haushaltskunden	1
Kosten für Rohöl	1
aktuelle Zeitenwende	1
Entspannung bei Gas	1
Kostensteigerung bei Strom	1
Thema Energieversorgung und Energie	1
Preis im Gas-großhandel	1
Befüllung der Erdgasspeicheranlagen	1
Abkommen zwischen den Ressorts	1
Stemmer –	1
Kriegskasse des Kreml	1
alter zweiter Kühlschrank	1
Krise profitierenden Unternehmen	1

Kunde der Landesversorger	1
Standort in Heiligenkreuz	1
Kurzarbeit als Instrument	1
Kurzarbeit als Instrument der Überbrückung	1
Erdgasspeicheranlage Haidach	1
Unimarkt-chef Andreas Haider	1
Richtung Stagflation	1
Land Energieministerin	1
Land Energieministerin Leonore	1
Reduktionsmittel für die Umwandlung	1
Land Vorprodukte und Rohstoffe	1
Erdöl- und Gaswirtschaft	1
Land der Welt laut Daten	1
Wienerberger Building	1
Preis pro Kilowattstunden	1
Winter bei zweistelligen Minusgraden	1
Teil des europäischen Großhandels	1
Dividende zur Unterstützung	1
Frühjahr 2020. Unsicherheit	1
Westen nach dem Einmarsch	1
Termin für das Gesetz	1
Wifo-chef Gabriel Felbermayr	1
Dividende zur Unterstützung der Menschen	1
Beginn der Energiekrise	1
Stromunternehmen von den steigenden Gaspreisen	1
Umsetzung unionsrechtlicher Vorgaben	1
Lebensmittelhändler auf eigene Rechnung	1
Lebensmittelproduzent mit massiven Kostensteigerungen	1
Frühjahr Strom	1
Frühwarnstufe im Notfallplan	1
Dollar gehandelten Rohstoffe	1
Aspirin 30	1
Lieferant an der Diversifizierung	1
Lieferant für die Medizintechnik	1
Lieferausfall für den Werkzeugmaschinenhersteller	1
Lieferkette durch den Krieg	1
Lieferstopp von russischem Erdgas	1
Lieferung an Drittstaaten	1

Vermögensvorteil von Antragstellern	1
Eu-agentur Acer	1
Lieferung von fossiler Energie	1
Lieferung zu Verzögerungen	1
Reststoffverwertungskessel in Betrieb	1
Lieferung zu Zwangsrationierungen	1
beträchtliche Arbeitslosigkeit	1
Preisdeckel unter dem Weltmarktpreis	1
Eu-importverbot für russisches Erdöl	1
Lohn im Kampf	1
Lohnrunde ab Montag	1
Fußball-wm für Menschenrechtsverstöße	1
Funktionsweise der Marktmechanismen	1
Ausweitung der Kompetenzen des Ministeriums	1
Tarifierhöhung für Bestandskunden	1
Stromversorger bei den Anschlüssen	1
Wärme für Wohnungen	1
Lösung für die Probleme der Betriebe	1
Co2- und Energiepreis	1
Lösung zur Beendigung des Krieges	1
Eco-austria-chefin Monika	1
Euro auf einen niedrigen Deckel	1
Markt tätigen Industrie	1
Wärme für Wohnungen und Büros	1
Markteingriff an den europäischen Gas	1
Marktmacht der Eu	1
Marktpreis zum Teil	1
Volumen der russischen Ölexporte	1
Marktwirtschaft mit großer Skepsis	1
Dollar pro Fass für Lieferungen	1
Martin Jaksch-fliegenschnee	1
Wessela Wladkowa	1
Material zur Auskleidung	1
Material zur Auskleidung von Stahl	1
Materie vertraute Person	1

Theiß der Evn	1
Maßnahme der Politik gegen die Teuerung	1
Stornierung von Investitionsprojekten auf Firmen	1
aufkommende Schwierigkeit bei der Materialbeschaffung	1
anderer Landesversorger	1
Zement- oder Glasofen	1
Solidarität mit dem ukrainischen Volk	1
Schlüsselfigur des prorussischen Lagers	1
Preissprung für Konsumentinnen	1
Beginn des Ukraine-krieges	1
Containerhafen St. Petersburg	1
Solidaritätsabgabe von Unternehmen	1
Strom- und Gaspreisanpassung	1
Euro über dem Weltmarktpreis	1
Ruf nach Wiederverstaatlichung	1
Mehl 9,90	1
Europa Energieversorgung	1
Europe Machined	1
Rekordniveau nach einer Inflationsrate	1
Europe Machined Components	1
Rekordpreis des Jahres	1
alternative Börse	1
Evn ins Spiel	1
Situation von Angebot und Nachfrage	1
Experte Staunovo	1
Unternehmen angestoßene Prozess	1
Mensch unter stark steigenden Strompreisen	1
Fachkraft zum Bau	1
Verlust zigtausender Arbeitsplätze	1
Fahrerpersonal bei den Frächtern	1
Rahmen des Energieeffizienzgesetzes	1
Fahrt gekommene Welthandel	1
Verhandlung mit Sicherheit	1
Fahrt gekommenen Welthandel	1
Begrenzung der Stromkosten	1
bei Energieversorgung besonderer Wert	1
Preissteigerung der Stromanbieter	1

Dominoeffekte für ein energieintensives Industrieunternehmen	1
Fall Erdgas	1
Mittwoch 131	1
Fall der hohen Co2	1
Stromrechnung ab Oktober	1
Mittwoch im Klub der Wirtschaftspublizisten	1
Mittwoch publizierte Strommarktbericht	1
aber wettbewerbsfähiger Energiepreis	1
Modell zur Eindämmung	1
Modell zur Eindämmung der Gaspreise	1
Fall einer Mangellage	1
Tausende Arbeitnehmer	1
Bandbreite in der Blitzumfrage	1
Tonne flüssiges Glas	1
Montag Höhe	1
autofreier Tag pro Woche	1
Montag bei einem Besuch in Wien	1
Fall künftiger Schocks	1
Denkmal zu Ehren der Sowjetarmee	1
Teuerungsprämie für Bezieher	1
Fall von Störungen der Energieversorgung	1
Totalstopp der Gaslieferungen	1
Müllverbrennung Wien	1
Verteuerung Eines	1
Fass für Lieferungen	1
Nachfrage nach energieeffizienten Fenstern	1
Nachteil für die Betriebe	1
Spotpreis- oder Float-tarif	1
Deutlich hoher Inflationswert	1
Unternehmen der fossilen Energiewirtschaft	1
Gas anderer Lieferanten	1
benötigtes Gasvolumen	1
Neos Kritik	1
Umfrage beteiligten Betriebe	1
Neos-energiesprecherin Karin	1
Ziegel Erdgas	1
Neos-wirtschaftssprecher Gerald Loacker	1
Nickel für die Stahlerzeugung	1
Niederösterreichs Freiheitliche	1

aktuell benötigtes Gasvolumen	1
Niederösterreichs Neos	1
Niederösterreichs Spüo-landespartei-vorsitzende	1
Niveau auf ein hohes Niveau	1
Stopp der Gasimporte	1
Siebenfachen einer konjunkturellen Normallage	1
Dramatik der aktuellen Energiepreiskrise	1
Filiale ab dem nächsten Jahr	1
Stromversorger aufgrund der aktuellen Marktstörungen	1
Normalisierung bei Energiepreisen	1
angestoßener Prozess der Umstellung	1
Notfallplan für die Gasversorgung	1
Notlage bei Strom	1
Umweltpolitik zuständige Klimaschutzministerium	1
Novellierung mit dem Bestreben	1
Finanzminister Magnus Brunner	1
November 2021. Heizöl	1
Welt Kraftwerke	1
Bauten auf öffentlichem Grund	1
Unternehmen Planbarkeit	1
November gegenüber Oktober	1
Zweifel an der Versorgungssicherheit in Europa	1
Nutzung des Gasspeichers	1
Obergrenze beim Strompreis	1
besonders kräftige Steigerung	1
Fleischware massive Entlastungspakete	1
Of Industry Management	1
Direktive für den Fall	1
Warenkontrolle zur Einhaltung	1
Option im Auftrag	1
Option im Auftrag des Kanzleramts	1
Folge am Standort	1
Wendepunkt in der Geschichte der Energie	1
Oö Geschäftsführer	1
Folge bei anderen Produktgruppen	1
Pandemie 2,4	1
Strombörse durch ein spezielles Auktionsverfahren	1

Parlament statt Dauerbeflegelungen	1
Person umfassendes Gremium	1
aktuelle Großhandelssituation	1
Drittel des weltweiten Stromverbrauchs	1
Gas auf Öl oder Kohle	1
Atomkraft Energiekrise	1
Forderung nach einem Preisdeckel bei Energie	1
Phase mit extremen Preisen	1
Pichler Ziegel	1
Pipeline Internorm	1
Forderung nach einer Reallohnerhöhung	1
Sicherheit vom Durchschnittsverbrauch	1
Fortsetzung des aktuellen Energie-kostenniveaus	1
Plan zu einem Atomausstieg	1
Fpö 17	1
Realismus in die Energiepolitik	1
Fpö Landespartei	1
Politik Benischko	1
Politik gegen die Teuerung	1
auch deutsches Gasunternehmen	1
Politiker Bulgariens	1
Fpö-chef Herbert Kickl	1
Preis für Dezember-lieferungen	1
Winter zur Herausforderung	1
Preis für Energie und Gasengpässe	1
Frage der Verantwortung der Politik	1
Zeit für Ideologie	1
Beispiel Gaslieferung	1
Preis für Uralsöl	1
Diversifizierung der Erdgasversorgung	1
Preis für die Megawattstunde	1
Sprecher des Finanzministeriums im Gespräch	1
Freitag verbreiteten Statement	1
Teil der russischen Ölindustrie	1
Zeit für Ideologie und Deals	1
Wirtschaftsminister Habeck	1
Unternehmenssprecher von Internorm	1
Sorge unter Politikern und Experten	1

Gas im Fall	1
Preisanpassung an Indizes	1
Preisanstieg 78	1
Preischaos seit Monaten	1
Preisdeckel bei Energie	1
Preisdeckel bei Treibstoffen	1
Preisdeckel für Energie	1
Preisdeckel für Strompreise	1
Preisentwicklung auf den Energiesektor	1
Preisgarantie für Strom	1
Preisgarantie für Strom und Gas	1
Preisgarantie für ein Jahr	1
Preisnachlass im Lauf	1
Preisnachlass im Lauf des Jahres	1
begrenzte Solidaritätsabgabe von Unternehmen	1
Preissprung für Konsumentinnen und Konsumenten	1
Förderung von einem fiktiven Durchschnittswert	1
Preissteigerung bei Öl und Gas	1
Vorbegutachtungsentwurf zur Novelle	1
Preisverwerfung bei Strom	1
Preisverwerfung bei Strom und Gas	1
Gas auf Standardprodukte	1
Produktion benötigte Energie	1
Gas auf Öl	1
Produktion von Kaffeetassen	1
Produktion von Kaffeetassen und Waschbecken	1
Gas aus Speichern	1
Prozent 2022	1
Prozent 2025	1
allfälliger Vermögensvorteil von Antragstellern	1
Durchschnittswert für einen Drei-personen-haushalt	1
Gas im Fall einer Mangellage	1
Beispiel staatliche Haftungen	1
Vorschreibung für das nächste Jahr	1
Gas im Gefolge	1
Spo-bundesgeschäftsführer Christian	1
Prozent an die Industrie	1
Gas in der Produktion	1
Dämmmaterial –	1

Gas zur Lieferung	1
Zittern um die Gasversorgung	1
Wärme in Unternehmen und Haushalten	1
Sorge unter Politikern	1
Rückkehr der Haushaltskundenpreise	1
Gas- und Energiemarkt	1
Regierung Potenzial	1
Aufbrechen der Merit	1
Umstellung auf alternative Brennstoffe	1
Prozent der Gesamtenergie	1
Gasbezugsquelle bei der Voestalpine	1
Prozent der Mehrkosten für Energie	1
Gasboykott in den Raum	1
E-control-chef Wolfgang	1
ausgelegte Entschädigung	1
am Freitag verbreitetes Statement	1
Wachstum der weltweiten Stromnachfrage	1
Schaden in Summe	1
Gaskrise ausgelegte Energielenkungsgesetz	1
Prozent der jeweiligen Speicherbestände	1
Prozent der russischen Ölexporte	1
Prozent der österreichischen Co2-emissionen	1
Gaslieferung aus anderen Ländern	1
Gaslieferung aus Übersee	1
Tonne auf Rang	1

Anhang E

Termkandidaten D-Terminer

Termkandidaten	Häufigkeit
Gas	68
Europa	30
Russland	28
Energie	41
Strom	28
Ukraine	21
EU	20
Energiekrise	19
Energiekosten	18
Erdgas	19
Unternehmen	41
Deutschland	17
Energiepreise	16
EVN	14
ÖVP	12
Oberösterreich	12
Haindl-Grutsch	16
Energieversorgung	11
Strompreisbremse	10
USA	9
Regierung	25
Strom-	9
Energiesparen	8
Wien	21
Öl	9
Gaslieferungen	8
Tirol	8
Grüne	9
Vorarlberg	8
Gasversorgung	9
energieintensiven	7
Österreich	62
Gaspreise	8
Stromkosten	6
China	6
Gaspreis	8
SPÖ	7
Strompreise	6
Politik	16
Diesel	9
Benzin	6
Oberösterreichs	6

Bundesregierung	6
Stahl	5
Energieversorger	8
FPÖ	5
energieintensive	5
Stromverbrauch	5
Energie-	4
Kohle	4
IV OÖ	5
Österreichs	9
erneuerbaren Energien	4
Infrastruktur	5
Steiermark	5
Energieträger	8
Energieunternehmen	4
Energiewende	4
Strompreis	4
Energieagentur	8
Teuerung	6
Montag	10
Urals	4
Sanktionen	11
Brent	4
Russlands	5
Atomkraft	4
EU-Energieminister	4
Energiepreisen	4
Lieferketten	6
MWh	4
Energien	8
7Fields	5
Gasmanigellage	4
Birol	8
Klimaministerium	2
Energienotfall	3
Energieeffizienz	3
Tempolimits	4
Wärme	4
Stromversorger	3
Joachim Haindl-Grutsch	3
Aluminium	3
Westens	3
TWh	6

Kraftwerk	3
Gasverbrauch	3
Speicher	5
Strompreisen	3
Neos	6
Öl-	3
Ukraine-Krieg	3
Megawattstunde	7
Energiepolitik	4
OÖ	7
Windkraft	4
Entspannung	3
Vorkrisenniveau	3
ÖGB	3
Spritpreise	4
Fatih Birol	3
IEA	4
kohle-	3
Wien Energie	7
Burgenland Energie	3
Fernwärme	4
Kelag	3
kWh	3
EU-Staaten	3
KURIER	4
TTF	3
Börsen	3
Sparen	5
Haidach	6
Putin	4
Gasbezugsquellen	3
Johanna Mikl-Leitner	3
Mikl-Leitner	8
Sofia	3
Nahversorger	3
Benischko	4
Energielenkungsgesetz	2
Klimaschutzministerium	3
festen	2
Elektrizität	2
STANDARD	2
Kraftfahrzeugen	2
Gaskrise	2
Gesetz	5
Elektronikindustrie	2
Eisenerz	2
Summe	3

Betriebe	19
Strompreiskompensation	2
Italien	2
Klimapolitik	2
Gashahn	3
Statistik Austria	2
Papier	2
Glas	5
hart	6
fahren	2
Stahl-	2
Unternehmens	3
Borgas	3
Gasspeicher	2
West	2
Gewinnabschöpfung	3
Erneuerbare	3
Gaspreisen	2
Verbund	4
erneuerbarer Energie	2
Stromunternehmen	3
Stromproduktion	2
Werner Kogler	2
Nehammer	6
Merit-Order	3
Stromnachfrage	2
Strombedarf	2
Ukraine-Krieges	2
Energiequellen	2
Windräder	2
Stromerzeugung	3
Sprit	3
Oktober	7
Österreichischen Energieagentur	2
Energiebedarf	2
Heizöl	3
Superbenzin	2
Lohnrunde	2
Strompreiserhöhungen	2
Rohöl	5
G7	2
Brüssel	2
Petersburg	2
Staunovo	4
Erdöl	2
Ölexporte	2
Rohstoff	3

Kernkraft	2
Erneuerbaren	8
CO2-Emissionen	2
Energieallianz Austria	3
Kärnten	2
EAA	2
Energieversorgers	2
AG	7
Salzburg AG	3
Stromrechnung	2
Öspi	2
Indra Collini	2
Preisdeckel	10
Klimawandel	2
Moskau	2
Welthandel	3
Wifo-Chef Felbermayr	2
ÖAMTC	2
Spritpreis	2
Gaspreisdeckel	2
EU-Kommission	2
Flüssiggas	2
Handel	6
erneuerbarer Energieträger	2
Salinen Austria	2
Salz	2
Peter Untersperger	2
Pasching	2
Thilo Preß	2
Röchling	5
handeln	4
Energieforecasts	2
politisch	2
Nassrasur	2
Katar	2
Leonore Gewessler	4
Energieverbrauch	2
Türen	3
Energieeffizienzgesetzes	2
Wifo	2
Herbst	8
Gasdurchleitungsvertrag	2
Speichern	4
Berlin	3
Transit	3
Gewessler	9
Krisenfall	3

Energiesicherheit	2
E-Control	2
Bremse	3
Kilowattstunde	3
Udo Landbauer	2
Kurzarbeit	2
Wolfgang Urbantschitsch	2
Energiemärkten	2
Urbantschitsch	4
Doppelmoral	2
Georg Knill	2
Wettbewerbsfähigkeit	5
Energiekostenzuschuss	2
Ekkehart Pichler	2
Christian Klinger	2
Internorm	3
Peter Weidinger	2
EU-Ebene	2
Strompreises	2
Strompreisdeckel	2
Bulgarien	2
Bulgaren	3
Griechenland	2
Blitzumfrage	3
Unimarkt	2
Energieministerium	1
Übergewinnen	1
Umweltpolitik	1
Gesetzesentwurfs	1
Störungen	2
Energieträgern	3
öffentlichem Grund	1
Emissionsgrenzwerten	1
Falles	1
Erwerbsfreiheit	1
Energielenkungsfall	1
Eigentums-	1
Multi-Energiekrise	1
Gasnotfall	1
Einsparung	3
flüssigen Energieträgern	2
Enteignungen	1
Marktstörungen	1
Gesetzesentwurf	1
Förderungen	2
öffentlichen Rechts	1
allfällige Vermögensvorteile	1

Versorger	1
Krisengewinner	1
Lieferkettenprobleme	2
Industriebundesland	1
wettbewerbsfähig	3
Ukraine-Konflikts	1
Stahlerzeugung	2
Maschinenbau	1
Palladium	1
Automobil-	1
Kobalt	1
Batterien	1
Neongas	1
Nahrungsmittelbereich	1
Lieferkettenproblematik	1
Dominoeffekte	2
Fossile Energie	1
fossiler Energie	1
Erdgaslieferungen	1
Frühwarnstufe	1
Rubel	1
Energielenkungsmaßnahmen	2
Embargo	1
verbrauchten Gasmengen	1
Wertschöpfungsketten	2
Bezug von Energie	1
aufstellen	3
Energiepreiswelle	1
bearbeitung	1
Papierindustrie	1
Energiepreisentwicklung	1
Vor-Krisenniveau	1
konjunkturellen Normallage	1
Nachbarland	1
Energiesektor	1
Industriestandort	3
CO2-	1
EU-Ländern	1
Industrie	42
Export	1
Transformation	6
Energiepreiskrise	1
Energiesystems	1
Energieinfrastruktur	1
Gesetzesvorhaben	1
russischem Gas	1
Energieprojekte	1

energieautark	1
Energiemengen	1
Ideologie	3
Osterreich	1
EU-Länder	1
Gasboykott	1
Nerven	1
Europa Energieversorgung	1
Laufen	6
Energiedichte	1
Energieform	1
Düngemitteln	1
Dämmmaterial	1
Chemie	2
Hemden	1
Hosen	1
Jacken	1
Aspirin	1
Hautcremes	1
Branchenvertreter	1
Alarm	1
vor dem Verlust	1
Gasimporte	1
Terawattstunden	1
Gaskraftwerken	1
Verbund-Großanlage Mellach	1
Theiß	1
Fernwärmenetze einspeisen	1
Müllverbrennung	1
Wien-Spittelau	1
Chemieindustrie	2
Stein	1
Lebensmittelindustrie	1
Red Bull	1
Agrana	1
Mars	1
NÖM	1
klein-	2
Schublade	1
Gaszufuhr	1
Gasboykotts	1
Norske Skog	1
Bruck an der Mur	1
Papierkonzern	1
Oslo	1
Reststoffverwertungskessel	1
Sappi	1

Gratkorn	1
Graz	1
Biomassekraftwerk	1
Kohlekessel	1
Gasstopp	1
RHI Magnesita	1
Auskleidung	1
Zement-	1
Glasöfen	1
Stefan Borgas	1
STANDARD-Interview	1
Hochfilzen	1
Rohstoffgewinnung	1
Rasenmäher	1
Baustoffkonzern	1
Wienerberger	2
Mauerziegelwerken	1
Dachziegelwerken	1
Brennöfen	1
Trockner Gas	1
Gasvolumens	1
Output	1
Solveig Menard-Galli	1
COO	1
Chief Operating Officer	1
Wienerberger Building Solutions	1
Menard-Galli	2
Geschäftsbasis	1
Öfen	1
Eisenschwamm	1
Elektroofen	1
Glasindustrie	1
flüssiges Glas	1
Flaschen	1
Gurken-	1
Marmeladegläser	1
Flachgläser	1
Auto-	1
Energiezufuhr	1
Standesvertretung	1
Kaffeetassen	1
Waschbecken	1
Medizintechnik	1
Gaseinfuhren	1
Lücke	1
Ökonomen	1

Importstopp	2
Wirtschaftsleistung	1
Zwangsrationierungen	1
Private	2
Krankenhäuser	1
Feuerwehr	1
Polizei	1
Klimaschutzministerium Energielenkungsmaßnahmen	1
Endkunden	1
Energielenkungsbeirat	1
Gremium	1
Bundeskanzleramt , Sozialpartner	1
Erdöl-	1
Gaswirtschaft	1
Ost	2
Karl Nehammer	2
teilstaatlichen Stromunternehmen	1
Freitag	4
ÖVP-geführte	1
Stroms	1
Strom-Gaspreis-Kopplung	1
Bund	1
Finanzministeriums	1
APA	1
Strompreisbildung	1
Ministerium	6
schriftlichen	1
Energieministerin Leonore Gewessler	1
Allgemeinheit	1
abzuschöpfen	2
SPÖ für Regierungsvorstoß	1
Überschussgewinne	1
Kanzler	3
Energiefragen	1
Neos-Wirtschaftssprecher Gerald Loacker	1
Herbert Kickl	1
Dividenden	1
börsennotierten	1
öffentlichem Eigentum	1
öffentlichen Hand	1
Instituts für Finanzrecht an der Universität Wien	1
Sabine	1
Kirchmayr-Schliesselberger	1

ORF	1
Sondersteuer	1
Anleger	2
Florian Beckermann	1
Interessenverbands für Anleger	1
IVA	1
öffentliche Hand	1
Aktionärsstruktur	1
Gift	1
Martin Jaksch-Fliegenschnee	1
IG Windkraft	1
Jaksch-Fliegenschnee	2
Energieunabhängigkeit	1
Gasfirmen	1
Strombörse	1
teuerste	2
Gaskraftwerk	1
Megawattstunde Strom	2
Megawattstunden	1
Massiv	12
Wirtschaft	6
Steckdose	1
Wünsch-Dir-was-Politik	1
Energiemenge	1
fossilen Energieträgern	1
Klimaziele	2
Hausverstand	2
Markt- und Meinungsfor- schungsinstitut Spectra	1
Wasserkraft	2
Sonnenenergie	1
Biomasse	1
fossilen Energieträger	1
Preisgünstigkeit	1
Älteren	3
öffentliche Wahrnehmung	1
Umweltschutz	1
Erdgasabhängigkeit	1
Energiesystem	1
Oberösterreicher	1
Lieferstopp	2
Soll	22
Arbeitslosigkeit	1
Durchaus	6
Votum	1
medizinischen	1
Klimaziele erreichbar	1

erreichbar	3
fossilen Energie	1
Ausstieg	2
Meinungsspektrum	1
Schlüsse	1
Haushaltsenergiepreise	1
Holzpellets	1
Großhandelsmärkten	1
Spotpreis-	1
Float-Tarifen	1
Preisgarantie	3
Brennholz	1
Einkauf	4
Arbeitnehmer	2
die Preisentwicklung	3
Metaller-Lohnverhandlungen	1
Haushaltsenergie	1
August-Inflation	1
EU-Durchschnitt	1
Inflationswerte	1
Miniwarenkorb	1
Alltag	2
Wocheneinkauf	2
Tanken	2
Pflege-	1
Gesundheitsberufe	1
GPA	1
Metallern	2
Gewerkschaft	1
Teuerungsprämien	1
russischem Öl	1
Sorte	3
Lukoil	1
Rohöls	1
St	2
Primorsk	1
Rohöl der Sorte	1
Argus Media	1
Uralspreis	1
Fass	5
Nordseesorte	1
Sibirien	1
Espo-Pipeline	1
Pazifikküste	1
Wladiwostok	1
Argus	2
Giovanni Staunovo	1

UBS	1
Importembargo	1
Drittstaaten	1
Kanada	1
Weltmarkt	3
Australien	1
Kreml	1
Ölhandel	1
Ostsee	1
Tanker	1
UBS-Experte Staunovo	1
Lagerkapazitäten	1
Ölindustrie	1
Uralsöl	1
Administration	1
Präsident Putin	1
Wirtschaftszeitung	1
Weltmarktpreis	4
Ölgesellschaften	1
Preisobergrenze	3
Sonne	1
Wind	1
Stromhunger	1
Gas-Großhandel	1
Blackouts	1
Konsumenten-Ebene	1
Stromanbieter	1
Energiekostenhilfen	1
Gas-Exporte	1
EU-Importverbote	1
Schmierstoffe	1
Internationalen Energieagentur IEA	1
Chinas Comeback	1
Klimakrise	1
Strommarktbericht	1
Energiehunger	1
Solaranlagen	1
Atomausstieg	1
Energiebedarfs	1
IEA-Prognosen	1
Pandemie	1
Indien	1
Südostasien	1
Paris	1
Stromverbrauchs	1
Kraftwerke	1

Atomkraftwerke	1
Emissionen	1
Energiemix	1
Gesamt-Energie	1
Chinas	2
fossilen Quellen	1
Stromerzeugungsmix	1
gasbefeuerten Stromerzeugung	1
CO2-Ausstoß	1
Landesenergieversorger	1
Salzburg	5
Österreichische Gewerkschaftsbund	1
Energie AG	1
Linz AG	1
Bestandskunden	4
Erdgastarif der Linz	1
Energiemarkt	1
Gaspreisindex-Entwicklung	1
Energie Steiermark	1
Gaspreisanpassungen	1
Energiekonzerns	1
Indexanpassung	4
Kärntner	1
Erdgasverbrauchspreis	1
Gewerbekunden	2
Tiroler Landesversorger TiwaG	1
Erich Entstrasser	1
Neukunden	1
Österreichische Strompreisindex	1
Strompreiserhöhung	1
Landesstatthalterin Barbara Schöbi-Fink	1
Strompreisdeckelung	1
Illwerke	1
VKW	1
Geschäftsbedingungen	1
Andreas Neuhauser	1
Energieallianz	3
Grünen	2
Donnerstag	3
Collini	3
EVN-Management	1
Land Niederösterreich	1
EVN-Mehrheitseigentümer	1
Dagmar Belakowitsch	1
Corona	1
Autofahrer	1

Kerosin	1
Chemieprodukte	1
Ölsorten	1
WTI	1
Ölmarkt	1
Russische Föderation	1
Ölgeschäft	1
fossilen Ressourcen	1
Erdölressourcen	1
Bundesanstalt für Geowissenschaften und Rohstoffe	1
BGR	1
Saudi-Arabien	1
Industriekapitäne	1
Corona-Krise	1
Fahrt	2
Weltwirtschaft	1
IfW	1
Kiel	1
Unsicherheit	2
Warenkontrollen	1
IfW-Volkswirt	2
Vincent Stamer	1
Güter	1
Stagflation	1
Gabriel Felbermayr	1
Tankstellen	1
IfW-Volkswirt Stamer	1
Kabelbäume	1
fossile Brennstoffe	1
Handels	1
Ölembargo	1
Ölverkauf	1
Erdgasverkauf	1
Ölpreis	1
Stopp	5
Rohölpreis	1
Montagnachmittag	1
Rechnet	3
Super	1
Autobahntankstellen	1
Auto	1
Öffis	1
Konto	2
Nordsee-Sorte	1
Raffineriemargen	1
Rohstoffe	6

Abgaben	1
CO2-Preises	1
Mineralölsteuer	1
Pendlerpauschale	1
CO2-Aufschläge	1
CO2-Preis	1
Klimabonus	1
CO2-Bepreisung	1
Elektrizitätsabgabe	1
EU-Minimum	1
Betrieben	3
Börsenbremse	1
Preischaos	1
Europäischen Union	1
hohen Gaspreise	2
Subventionsmodell	1
Energiewirtschaft	2
niederländischen Handelsplatz Title Transfer Facility	1
bilateralen Handel	1
Energiebörse	1
Gasmarktes	1
Margin-Zahlungen	1
Marktteilnehmer	1
Liquiditätsprobleme	1
Endverbrauchermarkt	1
Großhandelspreis	2
Endverbraucherpreisen	1
Deckel	2
LNG	1
Großhandelspreise	4
Megawattstunde Gas	2
Stromsparen	1
fossilen Energiewirtschaft	1
Gaseinkauf	1
Speicherbestände	1
Zentralstelle	1
geballte	1
Windrades	1
Slowenien	1
EU-Agentur	1
ACER	1
Niederlanden	1
Gaspreises	1
Energielösung	1
Salz-Gewinnung	1
Ebensee	1

Traunsee	1
Energiepreissteigerungen	1
Salz-Produktion	1
energieintensives	1
Gasembargos	1
Wirtschaftsablauf	1
Fahrerpersonal	1
Frächtern	1
Balanceakt	1
Hueck Folien	1
Baumgartenberg	1
europäischen Markt	1
Martin Bergsmann	2
Energiebedingte	1
Rohstoffen	3
Bergsmann	3
Trumpf Austria	1
Lieferausfälle	2
Lieferzeiten	2
Kunststoffverarbeiter Röchling	1
Oepping	1
Peter Eckerstorfer	2
Europe Machined Components	1
Head	1
Röchling Industrial	2
Martin Bergsmann von Hueck	1
Diversifizierung	2
Erdgasversorgung	1
Technologien	1
fossilen	7
Brennstoffen	1
Klimastrategie	1
Energielösungen	1
Fotovoltaikanlage	1
kWp	1
Fertigungsprozesse	1
Laserhybridanlage	1
Vorwärme	1
Großteilen	1
Peter Eckerstorfer von Röchling	1
nichtverfügbare Energie	1
türkis-grüne	4
Energiebonus	1
Energieeinsparungen	1
Energiekrisen	1
Rohölpreise	2
Temperaturbegrenzung	1

Heizungen	2
Autobesitzer	1
Energieferien	1
Bundeskanzler Bruno Kreisky	1
Energiespargründen	1
elektrischen Rasierer	1
elektrischer Rasierer	1
Kampf gegen Gasknappheit	1
technische	2
Flüssiggasimporte	1
klimaschädliches	1
Kohlekraftwerk	1
Mellach	2
Fehlanzeige	1
Klimaschutzministerin	3
Fernsehsender Puls4	1
Riesige Potenziale	1
Potenziale	4
Franz Angerer	1
AEA	1
Klimaschutzministerin Leonore Gewessler	2
Energiesparkampagne	1
Stromkunden	1
Stromkundinnen	1
führen	8
Hebel	1
Johannes Wahlmüller	1
Energieexperte	1
Umweltschutzorganisation	1
Temporeduktion	1
Fahrzeug	1
Gassektor	1
Verbrennungsmotoren	1
Ölprodukten	1
Klima	1
Umweltbundesamt	1
Wahlmüller	2
Kühlvittrinen	2
Kühlschränke	1
Energiefresser	1
Motoren	1
Kühlschrank	1
Keller	1
Jacuzzi	1
Terrasse	1
riesigen	1

Energieagentur-Chef	1
Heizung	2
Gerät	1
Angerer	3
Energiesparpotenzial	1
Terawattstunde	1
Zwar	12
Förderwesen	1
Energieeffizienzgesetz	1
Energiesparsames	1
Energieverschwendung	1
Gewerbebetriebe	2
energiefressenden	1
aufzudrehen	1
Klimaschutzministerin Gewessler	1
Wirtschaftskammer	2
STANDARD-Anfrage	1
Belohnung	1
Energiesparmaßnahme	1
Wiener Wirtschaftsforschungsinstitut	1
Energiesparbonus	1
Gasversorger	1
Konsument	2
Gewessler-Sprecherin Martina Stemmer	1
Sparkampagne	1
Sommer	4
Stemmer	3
Autobahn	2
Kampagne	2
Verpflichtung	1
Gasspeichers	1
Mangellage	1
Wien / Berlin	1
Abkommen	4
Erdgasspeicheranlagen	1
Energieministerin	2
Robert Habeck	1
Schloss Schönbrunn	1
Speicheranlagen	1
Gasunternehmen	1
Erdgasspeicher	1
Speichermengen	1
Energielenkung	1
Putins	1
Wirtschaftsminister Habeck	1

Solidarität	5
Speichern Haidach	2
süddeutschen Raums	1
Energiekooperation	1
Entlastung der Bevölkerung	1
Drei-Personen-Haushalts	1
Drei-Personen-Haushalt	1
verbrauchten Strom	1
Energiekontingent	1
Kilowattstunden	3
Zweitwohnsitze	1
Nebenwohnsitze	1
Haupt-	1
Nebenwohnsitz	1
Zentralen Melderegister	1
Stromanschlusses	1
Mehrpersonenhaushalte	1
Rundfunkgebühr	1
GIS	1
Rente	1
Eco-Austria-Chefin Monika Köppl-Turyna	1
Staffelung	2
ZiB	1
Klaus Neusser	1
NÖ-Strompreisrabatt	1
niederösterreichische Strompreisrabatt	1
Haushaltsverbrauchs	1
blau-gelbe Förderung	1
Energiesparanreize	1
Gießkanne	1
Neos-Energiesprecherin	1
Karin Doppelbauer	1
Jörg Leichtfried	1
Symptome	1
niederösterreichischen Landtag	1
Arbeiterkammer	1
CO2-Kosten	1
Gewerbe	1
EU-Beihilfevorgaben	1
Faserhersteller	1
Lenzing	1
Heiligenkreuz	1
Südburgenland	1
AMS	1
Nord Stream 1	1
E-Control-Chef	1

EU-27	1
Regeln	1
Finanzwelt	1
Gasmärkte	1
Klub der Wirtschaftspublizisten	1
Gashandel	1
Energiewelt	1
Mittwoch	7
Haushaltskunden	2
Deals	2
Schurkenstaaten	1
Deindustrialisierung Europas	2
IV-Chef Knill	1
Pharma	1
Doppel-Wumms	1
Nutznießler	1
Aggressor	1
Deindustrialisierung	3
Liquidität	1
Megaprogramm	1
staatliche Haftungen	1
Österreichischen Kontrollbank	1
Kiew	1
Hilfspakete	2
Eigeninitiativen	1
Zivilbevölkerung	1
Kriegen	1
beenden	1
Westen	2
Energiebereich	1
Elektroautos	1
Kongo	1
Gasverknappung	1
Gasengpässe	1
Energieformen	1
Verteuerung	2
Pichler Ziegel	1
Aschach	1
Tonzerkleinerungsmaschinen	1
Energieabhängigkeit	1
Energieeinsatz	1
Pichler	4
Ofenanlage	1
Brennertechnologie	1
Hilfsaggregat	1
Hallenbeleuchtung	1
LED	1

Fotovoltaik-Anlage	1
voestalpine	2
Wärmebehandlungsaktivitäten	1
Walzwerke	1
Stahlwerken	1
Linz	3
Donawitz	1
Kapfenberg	1
Gaslieferstopp	1
Gasembargo	1
Worst-Case-Szenario	1
Herbert Eibensteiner	1
voestalpine AG	1
Konzern	1
Übersee	1
LNG-Terminals	1
weitertransportiert	1
laufenden Betrieb	1
energieeffizienten	1
Vormaterialien	1
Flachglas	1
Kunststoff	2
Energieversorgungen	1
Klinger	3
thermische	1
Umwelt	1
Hochreiter Fleischwaren	2
Biogasanlagen	1
Fotovoltaik	1
Europäische Pipelines	1
intraeuropäische Pipelines	1
Sparmaßnahmen	2
aktiv	2
rasch	3
Weidinger	3
Kraftwerken	1
Finanzminister Brunner	1
SPO	1
Schulterschluss	1
Koalition	1
Parlament	1
Bundeskanzler Karl Nehammer	1
Rasierklinge	1
Stromzone	1
Finanzminister Magnus Brunner	1
Gesellschaft	2
LH Mikl-Leitner	1

Merit-Order System	1
Hahn	2
EU-Kommissar	1
Johannes Hahn	1
Ö1-Mittagsjournal	1
SPÖ-Vorsitzender	1
Pamela Rendi-Wagner	1
Niederösterreichs	4
SPÜO-Landesparteivorsitzenden	1
Franz Schnabl	1
SPO-Bundesgeschäftsführer	1
Christian Deutsch	1
Strompreisdeckel-Forderung	1
Landtag	2
Niederösterreichs Neos	1
Pinken	1
Energiekonzerne	1
Transparenz	2
Notfallpläne	2
Gasverteilung	1
Roten Armee	1
Nazi-Deutschland	1
Berlins	1
Hundertern	1
EU-Territorium	1
prorussische	1
Witoscha	1
politische	2
Russland-Freundlichkeit	1
Sowjetunion	1
Wessela Wladkowa	1
Ostalgie	1
Harvard-Absolvent	1
Kiril Petkow	1
Petkows pro-westliche Partei	1
DB	1
Sozialisten	2
Protestpartei	1
Bulgariens	1
Rumen	1
Alexander Schallenberg	1
Gas-Stopp	1
Pipeline	2
Heizen	2
bulgarischen	3
Wladkowa	2
Krise	10

Internationalen Energieagentur	2
TU Wien	1
erneuerbare Energieträger	1
Energie-Souveränität	1
Wind-	1
Solarkraftwerken	1
Fahrzeugen	1
Elektroantrieb	1
E-Autos	1
Nahrung	1
Schulden	1
Afrika	1
Europäische Solidarität	2
Preissprünge	1
Konsumentinnen	1
Konsumenten	1
Flüssiggas-Abnehmer	1
IEA-Chef	2
Energielieferungen	1
politischen	1
Metall-	1
Papier-	2
Nahrungsmittel-	1
Baustoff-	1
Maschinen-	1
Räder	2
wirtschaftliche	1
Konsumentenebene	1
Industriellenvereinigung	1
Oberösterreich	
Mittelbetriebe	1
Umfrage	3
Energiekostenanteil	1
Gewinnmargen	1
Energie-Kostenniveaus	1
Weihnachtspause	1
Viertel der an	1
Merit Order	1
Energiekostenzuschüsse	1
Überbrückung	1
Beton	1
Atom-	1
Laufkraftwerke	1
Photovoltaik	1
Windstrom	1
Technologieoffenheit	1
Mobilität	2

elektrifiziert	1
technologieoffenen	1
CO2-neutralem Strom	1
E-Fuels für Industrie und Mobilität	1
Energiepolitik Österreichs	1
Industriestandortes	1
Stromschlag	1
Staatshilfe	1
Kahlschlag	1
Greißler	1
Nahversorgern	2
Lebensmittelhändler	1
Handelsorganisation	1
Adeg	1
Nah & Frisch	1
Spar	1
Sutterlüty	1
Bio-	1

Gemüsehändler	1
Märkte	1
Einkaufstempel	1
Wolfgang Benischko	1
Nah	2
Frisch-Kaufmann	1
Kilo Mehl	1
Christian Prauchner	1
Spar-Kaufmann	1
Niederösterreich	2
Prauchner	2
Peter Buchmüller	1
ADEG-Märkte	1
Lebensmittelpreise	1
Buchmüller	2
Christof Kastner	1
Kastner	2
Lebensmittelgeschäft	1
Unimarkt-Chef Andreas Haider	1