

Semantische Interoperabilität im Kontext der IIoT durch Verwendung von Ontologien

Masterarbeit
zur Erlangung des akademischen Grades

Master of Arts in Business (MA)

Fachhochschule Vorarlberg
Business Process Management

Betreut von
Prof. Gunther Rothfuss, MSc

Vorgelegt von
Mag. Raimund Porod, BA

Dornbirn, 03.07.2023

Widmung

Die Voraussetzungen für eine solche Arbeit sind Neugierde und die Muße, dieser Neugierde folgen zu dürfen. Beides hätte ich nicht ohne die Einflussnahme und den Rückhalt meiner Familie und die Hilfe vieler besonderer Menschen, die mich auf dem Weg hierher begleitet haben.

Ein besonderer Dank gilt hierbei Prof. Gunther Rothfuss, MSc, der mir nicht nur als Betreuer dieser Arbeit durch seine fachliche Unterstützung und konstruktives Feedback zur Seite stand, sondern mich auch in seinen Kursen motiviert hat, über den Tellerrand hinauszublicken, Zusammenhänge zu erkennen und neue Wege zu gehen.

Ein weiterer Dank gilt meinen Eltern, die mir den Wert von Bildung beigebracht haben und mich in all meinen Schritten bedingungslos unterstützt haben.

Melissa Lajtos bin ich dankbar für ihre Korrekturarbeit und den stets unterhaltsamen wissenschaftlichen Austausch.

Zuletzt darf ich den größten Dank den beiden wichtigsten Frauen in meinem Leben aussprechen: Meiner Tochter Johanna und meiner Frau Carmen, die mir während dieser zwei intensiven Jahre Geduld, Verständnis und Rücksichtnahme entgegenbrachten.

Auch wenn auf dieser Arbeit mein Name steht, so ist der Beitrag meiner Frau mindestens ebenso hoch einzuschätzen. Insbesondere deshalb, weil sie nicht nur alles rundherum organisiert hat und mir so die notwendige Zeit und Ruhe für das Studium verschafft hat, sondern, weil sie auch für fachliche Fragen immer ein offenes Ohr hatte und mir stets mit Rat und Tat zur Seite stand.

Kurzreferat

Semantische Interoperabilität im Kontext der IIoT durch Verwendung von Ontologien

Die Umsetzung von Konzepten der Industrial Internet of Things (IIoT) stellt Unternehmen vor große Herausforderungen. Ein Kernaspekt dabei ist das Vernetzen maschineller und menschlicher Akteure, sodass sich diese gegenseitig verstehen und dadurch sinnvolle Entscheidungen treffen können. Dieses Verstehen setzt eine semantische Interoperabilität der Akteure voraus, die meist mithilfe von Ontologien und Ontologie Matching ermöglicht werden soll. Ziel dieser Arbeit ist es, die Grundkonzepte zum Ermöglichen semantischer Interoperabilität zusammenzufassen und aktuelle Anwendungsbeispiele und Herangehensweisen zu besprechen, um darauf aufbauend bestimmen zu können, wie weit Forschung und Unternehmen noch von der Umsetzung der semantischen Interoperabilität im Kontext der IIoT entfernt sind. Dazu wurde ein theoretisch-konzeptionelles Vorgehen gewählt. Die Ergebnisse der vorliegenden Arbeit weisen darauf hin, dass in naher Zukunft nicht mit einer erfolgreichen Umsetzung zu rechnen ist, welche generisch genug ist, in unterschiedlichen Bereichen der IIoT eingesetzt zu werden. Dafür konnte gezeigt werden, dass sich mehrere, mitunter gegenseitig beeinflussende Trends in der Herangehensweise zur Lösung der semantischen Interoperabilität ausgebildet haben, nämlich hybride Alignmentarchitekturen, ein modularer Aufbau von Ontologien, das Erstellen von maschinenlesbaren Standards und dazu gehörenden Ontologien sowie ein Trend zur Verwendung von Wissensgraph Embeddings. Besonders der letzte Trend ist dabei kritisch zu betrachten. Darüber hinaus konnten in dieser Arbeit offene Problemfelder bestimmt werden, die aus derzeitiger Sicht einer erfolgreichen Umsetzung semantischer Interoperabilität entgegenstehen. Dies betrifft einerseits das Fehlen einer Harmonisierung von IIoT-Standards und andererseits den Mangel an Alignmentdatenbanken und Benchmarks für Ontologie Alignments im Bereich der IIoT. Darüber hinaus bedarf es eines abgestimmten, strategischen Vorgehens der derzeit laufenden und zukünftigen Initiativen, um vorhandene Erkenntnisse besser zu nutzen und um sich wiederholende Probleme nachhaltig zu lösen.

Schlagwörter: IIoT, Industrial Internet of Things, semantische Interoperabilität, Ontologie, Ontologie Matching, Ontologie Alignment

Abstract

Semantic Interoperability in the Context of IIoT by using Ontologies

The implementation of concepts of the Industrial Internet of Things (IIoT) poses a great challenge to companies across all industries. A key aspect of this is the connection of machine and human actors so that they can understand each other and thus make meaningful decisions. This understanding requires a semantic interoperability of the actors, which is usually enabled by ontologies and ontology matching. This work sets out to summarize the basic concepts for enabling semantic interoperability and to discuss current application examples and approaches in order to be able to determine, how far research and companies are from implementing semantic interoperability in the context of the IIoT. For this purpose, a theoretical-conceptual approach was chosen. The results of the present work indicate that a successful implementation, which is generic enough to be used in different areas of the IIoT, is not to be expected soon. However, it could be shown that several mutually influencing trends in the approach to semantic interoperability have emerged, namely hybrid alignment architectures, a modular structure of ontologies, the creation of machine-readable standards and associated ontologies, as well as a trend towards the use of knowledge graph embeddings. Particularly the last trend must be considered critically. In addition, this work has identified several obstacles that currently prevent a successful implementation of semantic interoperability. On the one hand, this concerns the lack of harmonization of IIoT standards and on the other hand the lack of alignment databases and benchmarks for ontology alignments. Furthermore, there is a need for a coordinated, strategic approach of current and future initiatives to better leverage existing knowledge and sustainably solve recurring problems.

Keywords: IIoT, Industrial Internet of Things, Semantic Interoperability, Ontology, Ontology Matching, Ontology Alignment

Vorwort

Auch wenn es der sperrige Titel dieser Arbeit vielleicht nicht suggeriert, so sei darauf hingewiesen, dass der Kern dieser Arbeit nicht nur fundamentale Aspekte der maschinellen und menschlichen Kommunikation anspricht, sondern auch auf Methoden eingeht, die im Moment des Entstehens dieser Arbeit zu einem Hype im Bereich der künstlichen Intelligenz geführt haben. Anders formuliert: Das Lesen dieser Arbeit könnte der Mühe wert sein.

Damit der Lesefluss nicht durch zu viele Details und Querverweise behindert wird, wurden ergänzende Beschreibungen und Verbindungen zu weiteren relevanten Aspekten in Fußnoten abgelegt. Darüber hinaus war es dem Autor ein Anliegen, möglichst verständliche und nachvollziehbare Beispiele und Darstellungen zu finden, die das Thema, trotz der mitunter sehr mathematisch-technischen Aspekte, auf möglichst allgemein verständliche Weise erklären. Zuletzt wurde versucht, den Kernaspekt einiger Kapitel kurz und knapp mit einem passenden Zitat zusammenzufassen, welches sich jeweils am Anfang dieser Kapitel befindet.

Inhaltsverzeichnis

Darstellungsverzeichnis	9
Abkürzungsverzeichnis	11
1 Einleitung	12
1.1 Problemstellung	12
1.2 Aufbau der Arbeit	14
1.3 Methodik	15
1.4 Zielgruppe und Bezug zum Studiengang	16
2 Grundlegende Konzepte	17
2.1 IIoT, CPS und die Bedeutung der semantischen Interoperabilität	17
2.1.1 IIoT	17
2.1.2 Cyber-physical Systems	17
2.1.3 Interoperabilität als grundlegende Herausforderung der IIoT	18
2.1.4 Einordnung semantischer Interoperabilität	21
2.2 Ontologien	24
2.2.1 Versuch einer Definition	24
2.2.2 Formale Repräsentationen von Ontologien	29
2.2.3 Typen von Ontologien	30
2.2.4 Wissensgraph vs. Ontologie	32
2.2.5 Allgemeiner Nutzen von Ontologien innerhalb einer Domäne	34
2.2.6 Beispiele für Ontologien bzw. Wissensgraphen und deren Verwendung	34
2.2.6.1 Ontologien außerhalb des IIoT-Kontexts	34
2.2.6.2 IIoT-Ontologien	38
3 Ontologie Matching	41
3.1 Semantische Heterogenität	41
3.1.1 Semantische Heterogenität anhand dreier Beispiele	41
3.1.2 Arten von Heterogenität	43
3.1.3 Heterogenität im Kontext der IIoT	46
3.2 Begriffsbestimmung und formale Definition des Ontologie Matchings	47
3.3 Vom Alignment zur erfolgreichen Interoperabilität	50
3.4 Alignment Architekturen	51
3.5 Methoden des Ontologie Matching	53

3.5.1	Analytische Methoden	54
3.5.1.1	Auf formalen Ressourcen basierte Methoden	55
3.5.1.2	Auf informellen Ressourcen basierte Methoden	56
3.5.1.3	Auf Zeichenketten (Strings) basierte Methoden	56
3.5.1.4	Sprachen-basierte Methoden	56
3.5.1.5	Auf Einschränkungen basierende Methoden	56
3.5.1.6	Graph-basierte Methoden	56
3.5.1.7	Taxonomie-basierte Methoden	57
3.5.1.8	Instanzen-basierte Methoden	57
3.5.1.9	Modell-basierte oder semantische Methoden	57
3.5.2	Ontologie and Wissensgraph Embeddings	58
3.5.2.1	Grundlegende Konzepte und Anwendungen	58
3.5.2.2	Embeddings und Interoperabilität	60
3.6	Gegenwärtige Softwaretools und Nachschlagverzeichnisse	61
3.6.1	LOOM	61
3.6.2	LogMap	61
3.6.3	ALIN	62
3.6.4	Evaluationstools (OOPS!)	63
3.6.5	Reasoner (HerMiT)	63
3.7	Aktuelle Ansätze und offene Probleme des Ontologie Matchings	64
3.8	Allgemeiner Nutzen des Ontologie Matchings	65
4	Anwendungsfälle von Ontologie Matching im Bereich der IIoT	67
4.1	Einleitendes Beispiel	68
4.2	Plattformbasierte Ansätze	69
4.2.1	oneM2M	70
4.2.2	VICINITY Projekt	73
4.2.3	INTER-IoT Projekt	75
4.2.4	Zusammenfassung und weitere plattformbasierte Ansätze	78
4.3	Weitere Ansätze mithilfe von analytischem Ontologie Matching	81
4.3.1	Alligator und SemCPS	81
4.3.2	Ergänzungen von Standards um Ontologien	83
4.3.3	Anwendung zur Anlagenüberwachung	84
4.3.4	MAPSOM Framework von Jirkovský	86
4.3.5	Zusammenfassung der analytischen Ansätze	89
4.4	Ansätze mithilfe von Wissensgraph Embeddings	89
4.5	Matching von IIoT-Standards: Der I40 Wissensgraph	92

5	Aktuelle Herausforderungen und Kritik	96
5.1	Aktuelle Herausforderungen	96
5.2	Kritik	98
5.3	Zukünftige Handlungsfelder	100
6	Schlussbetrachtung	102
6.1	Zusammenfassung und Interpretation der Ergebnisse	102
6.2	Limitationen	105
6.3	Schlusswort	106
	Literaturverzeichnis	107
	Anhang: Übersicht der besprochenen Anwendungsfälle	122
	Eidesstattliche Erklärung	125

Darstellungsverzeichnis

Darstellung 1: Aufbau eines CPS.....	18
Darstellung 2: Komponenten einer industriellen Anwendung im Kontext der IIoT.....	19
Darstellung 3: Horizontale und vertikale Integration in einem Produktionssystem.....	20
Darstellung 4: Beispiel für heterogene Standards in einem Entwicklungs- und Fertigungsprozess	20
Darstellung 5: Ebenen der Interoperabilität.....	22
Darstellung 6: Informelle und formale Ontologien.....	25
Darstellung 7: Beispielhafte Darstellung einer Ontologie.....	28
Darstellung 8: Typen von Ontologien.....	30
Darstellung 9: Architektur der Common Core Ontologien.....	31
Darstellung 10: Architektur eines Wissensgraphen.....	32
Darstellung 11: Wissensgraph oder visuelle Darstellung einer Ontologie.....	33
Darstellung 12: Die Gene Ontology.....	35
Darstellung 13: Domänenübergreifende Linked Open Data Cloud.....	37
Darstellung 14: Bing-Suche nach Ouagadougou.....	38
Darstellung 15: Überblick über die wichtigsten Konzepte und Relationen der SAREF Ontologie.....	40
Darstellung 16: Beispiel 1 zur Darstellung semantischer Heterogenität.....	42
Darstellung 17: Beispiel 2 zur Darstellung semantischer Heterogenität.....	42
Darstellung 18: Beispiel 3 zur Darstellung semantischer Heterogenität.....	43
Darstellung 19: Arten von Heterogenität.....	44
Darstellung 20: Begriffsbestimmung Ontologie Matching und verwandter Begriffe.....	48
Darstellung 21: Ermittlung eines Alignments.....	49
Darstellung 22: Prozessschritte des Ontologie Matchings.....	49
Darstellung 23: Einfache Datenintegration mithilfe des Alignments.....	50
Darstellung 24: Übersetzungsprozess mittels Mediator.....	51
Darstellung 25: Alignment Architekturen.....	52
Darstellung 26: Klassifizierung analytischer Matching Methoden.....	55
Darstellung 27: Visualisierung der Konzepte (A,B) und Relation (r) als Vektoren im \mathbb{R}^2	58
Darstellung 28: Methoden zur Ermittlung von Embeddings.....	59
Darstellung 29: Beispiel eines Prozesses zur Ermittlung von Alignments mithilfe von Embeddings...	60

Darstellung 30: Vorgehensweise des LogMap-Algorithmus.....	62
Darstellung 31: Vorgehensweise von ALIN.....	63
Darstellung 32: Ergebnisse des Conference tracks 2022 der OAEI.....	63
Darstellung 33: Wissen aus Alignments ziehen	66
Darstellung 34: Layer der Cloud Manufacturing Plattform	68
Darstellung 35: Verbindung des oneM2M Service Layers mit CPS und Anwendungen	70
Darstellung 36: Alignment zwischen SAREF und der oneM2M Base Ontology	71
Darstellung 37: Anwendungsbeispiel der oneM2M Base Ontology im Smart Home-Bereich.....	72
Darstellung 38: Überblick über das Netzwerk der VICINITY-Ontologien	74
Darstellung 39: Aufbau des Alignments der GOIoT mit domänenspezifischen Ontologien	77
Darstellung 40: Open Semantic Framework (OSF) von Siemens	79
Darstellung 41: Aufbau des Alligator-Frameworks	81
Darstellung 42: Alternative Wissensgraphen aus SemCPS	83
Darstellung 43: Mergen von Ontogien aus dem Produkt- und Produktionsbereich.....	84
Darstellung 44: Anwendung zur Überwachung der Produktionsanlagen	85
Darstellung 45: Zuweisung gültiger und ungültiger Mappings im MAPSOM Framework	87
Darstellung 46: Alignment der COCI, DUL und SSN Ontologien mithilfe des MAPSOM Frameworks	87
Darstellung 47: Prozess zum Ermitteln der Alignments mithilfe von BERT_INT	91
Darstellung 48: Prozess zum Ermitteln der Alignments mithilfe von BERT_INT	91
Darstellung 49: Standards Ontologie (STO).....	93
Darstellung 50: Zuordnung von Standards zu den Referenzarchitekturmodellen RAMI4.0 und IIRA.	95
Darstellung 51: Aufteilung der Patente nach Jahren und Unternehmen	104
Darstellung 52: Link auf Beispiel zu misslungener Kommunikation	106
Darstellung 53: Übersicht der in Kapitel 4 besprochenen Ansätze	124

Abkürzungsverzeichnis

CMfg	Cloud Manufacturing
CPS	Cyber-physical System
ERP	Enterprise Resource Planning
IIoT	Industrial Internet of Things
IoT	Internet of Things
IPSM	Inter Platform Semantic Mediator
JSON	JavaScript Object Notation
MES	Management Execution System
NLP	Natural Language Processing
OAEI	Ontology Alignment Evaluation Initiative
OWL	Web Ontology Language
RDF	Resource Description Framework
SAREF	Smart Applications Reference Ontology
SCADA	Supervisory Control and Data Acquisition
SOSA	Sensor, Observation, Sample and Actuator Ontology
SPARQL	SPARQL Protocol and RDF Query Language
SSN	Semantic Sensor Network Ontology
WoT	Web of Things

1 Einleitung

Integrating all relevant information and automating as many production steps as possible is the central goal of the I40 vision. Instead of envisioning one monolithic system or database, a decentral semantic integration is pursued, i.e., the formal description and linking of all relevant assets and data sources based on an aligned set of ontologies and their data.

(Grangel-González u.a. 2018, S. 131)

1.1 Problemstellung

Das Konzept der Industrial Internet of Things (IIoT), auch als Industrie 4.0 oder I40 bezeichnet, umschließt Begriffe wie Predictive Maintenance, Artificial Intelligence, Machine Learning, Cyber-physical Systems (CPS) oder Smart Factories. All diesen Begriffen gemeinsam geht die Vorstellung voraus, dass Maschinen nicht nur mit Menschen, sondern auch Maschinen untereinander eigenverantwortlich interagieren können.

Folglich ist eine Kernvoraussetzung für die Umsetzung der IIoT eine funktionierende Kommunikation zwischen den Akteuren. Diese soll sowohl horizontal (z.B. von Produktionsmaschine zu Produktionsmaschine) als auch vertikal (z.B. von Produktionsmaschine zu MES zu ERP¹) erfolgen, was zu einigen Herausforderungen in Bezug auf das Gelingen der Kommunikation führt. (Hildebrandt u.a. 2017, S. 1)

Eine Möglichkeit, dies erfolgreich zu bewerkstelligen, ist es, Standards zu definieren. So verweist die „Normungsroadmap Industrie 4.0“ der DIN und DKE (2023, S. 103–105) auf eine weitreichende Zahl an Standards und Standardinitiativen für verschiedenste Bereiche. Ein allumfassender Ansatz hat sich bisher noch nicht herausdestilliert und wird sich mit großer Sicherheit auch nicht.²

Folglich benötigt das Konzept der IIoT aus derzeitiger Sicht unterschiedliche Standards, was zu einer Fülle an Kommunikationsschnittstellen von einem Standard auf den nächsten sowie von einem Akteur auf den nächsten führt. Jede dieser Schnittstellen birgt die Gefahr, dass der Austausch von Daten fehlschlägt. Dabei können Kommunikationsprobleme nicht nur auf syntaktischer Natur, sondern auch auf semantischer auftreten. Ein Beispiel hierfür wäre, wenn unterschiedliche Standards ein und dieselbe Entität unterschiedlich benennen.

Um die Kommunikation und damit die Interoperabilität zwischen Maschinen im Kontext der IIoT zu gewährleisten, steht dabei insbesondere die semantische Interoperabilität im Fokus:

¹ MES steht für Management Executive System, ERP für Enterprise Resource Planning. Eine Beschreibung findet sich weiter unten auf Seite 20 in Darstellung 3.

² Rebstock, Fegel und Paulheim (2008, S. 2) nennen für die Fülle an unterschiedlichen Standards drei Gründe: Neue Technologien, veränderte Geschäftsprozesse und die globalisierte Wirtschaft. Nach den Autoren gibt es zwar immer wieder Bestrebungen, Standards zu vereinheitlichen und den „universalen Superstandard“ zu erstellen. Aufgrund der Anzahl an unterschiedlichen Standards und aufgrund der dynamischen Entwicklung dieser, scheiterten diese Initiativen für gewöhnlich.

The challenge in electronic business transactions is to achieve semantic interoperability, i.e., to avoid mismatch and misunderstanding in integrated business processes, taking into account the parallel usage of standards.

(Rebstock; Fengel; Paulheim 2008, S. 4)

Mit anderen Worten: Semantische Interoperabilität ist einer der Schlüssel zur Umsetzung der Ideen des Industrial Internet of Things. (Schröder; Willner; Diedrich 2021, S. 33)

Schon in der menschlichen Kommunikation kommt es oft zu inhaltlicher Ambiguität. Nicht umsonst spricht man auch von „aneinander vorbeireden“. Bei Maschinen ist dieses Problem in der Regel noch viel stärker ausgeprägt, da sie die Nachricht ohne Kontext erhalten und gerade der Kontext zur korrekten Entschlüsselung einer Nachricht benötigt wird.³

Ein Ansatz, diese semantische Interoperabilität zu gewährleisten, ist die Verwendung von Ontologien und das Bilden von Referenzen zwischen den Ontologien, das sogenannte Ontologie Matching. Während Ontologien das Wissen über die Begriffe und deren Kontext abbilden, ermöglicht das Ontologie Matching das Übersetzen der Begriffe von einer Ontologie in die andere. (Rebstock; Fengel; Paulheim 2008, S. 4; Teslya; Ryabchikov 2019, S. 139)

Der Vorteil, den dieser Ansatz verspricht, ist somit nichts anderes als die semantische Referenz unterschiedlicher Standards und in weiterer Folge die semantische Interoperabilität zwischen beliebigen Akteuren im Bereich der IIoT. Abgeleitet von dieser Problemstellung ergeben sich folgende Forschungsfragen:

FF1: Wie unterstützen Ontologien und das Ontologie Matching die semantische Interoperabilität?⁴

FF2: Welche aktuellen Methoden, Anwendungen und Problemstellungen gibt es im Bereich der IIoT zur Unterstützung semantischer Interoperabilität?

FF3: In welchen Branchen und Domänen wurden diese Methoden zur Unterstützung semantischer Interoperabilität in den letzten Jahren angewendet?

Eng verbunden mit der semantischen Interoperabilität ist der Begriff des Semantic Web, dessen Ziel es ist, Kommunikation zwischen Menschen und Maschinen als auch zwischen Maschinen untereinander über das Internet zu ermöglichen. (Berners-Lee 2015) Im Kern betrachtet diese Arbeit somit nur einen Ausschnitt des Semantic Web, kann damit aber auch auf den derzeitigen Forschungs- und Entwicklungsstand zurückgreifen. Im Zuge dieser Arbeit werden immer wieder Begriffe und Technologien des Semantic Web verwendet. Aufgrund des eingeschränkten Umfangs werden diese meist nur kurz erklärt. Für das Erlangen eines tieferen

³ Ein eindrückliches Beispiel, warum der Kontext auch für menschliche Kommunikation unumgänglich ist, liefert eine Fernsehwerbung aus dem Jahr 1999. („Fruit joy ad 1999“ 2018) Ein QR-Code mit dem Link auf das Video findet sich auf Seite 106.

⁴ Der Schwerpunkt dieser Arbeit liegt auf dem Ontologie Matching, da dies nachzeitigem Stand der Forschung der vielversprechendste Weg zu sein scheint, um semantischer Interoperabilität zu ermöglichen und insbesondere im naturwissenschaftlichen Kontext erfolgreich angewendet wird (siehe Abschnitt 3.8).

Verständnisses über diese Begriffe sei an dieser Stelle auf Szilagyí und Wira (2016) oder direkt auf das World Wide Web Consortium (W3C 2021, 2022) verwiesen.

1.2 Aufbau der Arbeit

Diese Arbeit ist wie folgt gegliedert:

Kapitel 1 beschäftigt sich mit der Problemstellung dieser Arbeit und der Methodik, mit der die oben erwähnten Forschungsfragen beantwortet werden sollen. Zusätzlich wird auch auf die intendierte Zielgruppe verwiesen, um etwaiges Vorwissen, das für die Lektüre vorausgesetzt wird, zu erwähnen und zu erklären, warum welche Schwerpunkte gesetzt wurden.

Kapitel 2 geht auf die grundlegenden Begriffe dieser Arbeit ein und arbeitet die in der Einleitung besprochenen Problemfelder genauer aus. Abschnitt 2.1. beschäftigt sich hierbei mit den wichtigsten Konzepten der IIoT sowie mit semantischer Interoperabilität und versucht die Bedeutung letzterer für die IIoT herauszustreichen. In Abschnitt 2.2. wird der Begriff der Ontologie vorgestellt. Dabei wird einerseits auf die Definition des Begriffs eingegangen, als auch unterschiedliche Arten und Beispiele von Ontologien vorgestellt, mit dem Ziel, den allgemeinen Nutzen von Ontologien zu beschreiben.

Dem Verknüpfen von unterschiedlichen Ontologien zur Ermöglichung semantischer Interoperabilität – dem sogenannten Ontologie Matching – ist **Kapitel 3** gewidmet. Ausgehend von den in Kapitel 2 besprochenen Konzepten und Problemfeldern werden grundlegende Begriffe, Zielsetzungen und Methoden des Ontologie Matching vorgestellt, um das Problem der semantischen Interoperabilität allgemein lösen zu können und damit Forschungsfrage FF1 beantworten zu können.

Im Anschluss daran werden in **Kapitel 4** konkrete Arbeiten und Projekte aus dem IIoT-Umfeld besprochen, die sich der in Kapitel 3 beschriebenen Methoden bedienen, um semantische Interoperabilität zu ermöglichen. Dieses Kapitel bildet den Kern zur Beantwortung der in Abschnitt 1.1. erwähnten Forschungsfragen FF2 und FF3.

Abschnitt 5 gibt einen Überblick über die aktuellen Herausforderungen und ungelösten Probleme. Darüber hinaus werden auch Forschungsschwerpunkte erwähnt, die aus Sicht des Autors zukünftig besonders relevant sind, um die der Arbeit zugrunde gelegten Probleme nachhaltig zu lösen.

Zuletzt werden in **Kapitel 6** die Forschungsfragen zusammenfassend beantwortet, um darauf aufbauend eine Interpretation der Ergebnisse für industrielle Anwender zu liefern.

Damit kann der Aufbau und das Ergebnis dieser Arbeit in folgendem Satz zusammengefasst werden:

Das Problem der semantischen Interoperabilität im Kontext der IIoT (Abschnitt 2.1) kann mithilfe von Ontologien (Abschnitt 2.2), die miteinander verbunden werden (Kapitel 3) gelöst werden, wie theoretische und praktische Beispiele (Kapitel 4) zeigen, wobei es hier noch einige offene Fragestellungen und Probleme (Kapitel 5) gibt und eine generische Anwendung noch auf sich warten lässt (Kapitel 6).

1.3 Methodik

Die Forschungsfragen sollen durch eine theoretisch-konzeptionelle Arbeit beantwortet werden. Dabei sollen auf Basis bestehender Literatur, die aus unterschiedlichen Forschungsgebieten stammt, neue Erkenntnisse gewonnen werden. (Goldenstein; Hunoldt; Walgenbach 2018, S. 85)

Zur Identifikation der relevanten Literatur werden in einem ersten Schritt die Methoden der Schlagwortsuche sowie der konzentrischen Kreise verbunden. Hierbei wird eine Einstiegsquelle gewählt und dessen Literaturverzeichnis nach relevanten Beiträgen untersucht. Diese werden anschließend nicht nur inhaltlich ausgewertet, sondern es wird wiederum auch deren Inhaltsverzeichnis analysiert. (Kache u.a. 2015, S. 8–9)

Während die Schlagwortsuche sich eignet, um Literatur zu identifizieren, die noch nicht zitiert wurde, hilft die Methode der konzentrischen Kreise, wichtige Beiträge und Akteure im jeweiligen Themenbereich zu bestimmen. Darüber hinaus ermöglicht diese Methode das Identifizieren weiterer Schlagworte. Dies ist insbesondere für das in dieser Arbeit gewählte Thema von Vorteil, da es für die im Titel genannten Begriffe „Semantische Interoperabilität“, „IIoT“ und „Ontologie“ eine Vielzahl an Synonymen oder verwandten Begriffen gibt.⁵ Um die Suchergebnisse der Schlagwortsuche weiter zu spezifizieren, werden zusätzlich Boolesche Operatoren verwendet.⁶ (Goldenstein; Hunoldt; Walgenbach 2018, S. 60, 65–66; Schlosser u.a. 2006)

Zur Suche der relevanten Literatur greift der Autor auf die ihm zur Verfügung stehenden Datenbanken zu. Dies betrifft dabei insbesondere die Online Literaturabfrage „OLAV“ („Olav – Online Literaturabfrage der FH Vorarlberg“ 2022) und die dadurch digital oder physisch in der Bibliothek der Fachhochschule Vorarlberg vorhandenen Literaturexemplare. Zusätzlich werden weitere Datenbanken (ProQuest, SpringerLink, ScienceDirect, ResearchGate, WISO, IEEE, Deutsche Gesellschaft für Informatik) und Suchmasken wie GoogleScholar oder SemanticScholar verwendet.

Im zweiten Schritt erfolgt auf Basis der identifizierten Literatur eine Selektion, um die relevantesten Werke zu bestimmen. Dabei werden einerseits die Inhalte durch Lesen der Abstracts, Forschungsfragen und Darstellungen bewertet. Andererseits werden Verfasser, Verlag, Erscheinungsjahr sowie Zitierindices berücksichtigt. (Brink 2013, S. 118–123)

Die am Ende dieses Selektionsprozesses ermittelten Quellen dienen als Basis der Gliederung dieser Arbeit und zur Beantwortung der Forschungsfragen.

⁵ Beispiele für Synonyme oder verwandte Begriffe zu „Ontologie“ wären „Knowledge graph“ oder „Schema“, anstelle von „Ontologie Matching“ wird auch von „ontology mapping“, „extension“, „alignment“, „integration“ oder „inclusion“ gesprochen. Im Bereich IIoT gibt es verwandte Begriffe wie „IoT“, „Industry 4.0“, „Cyber-physical (Production) Systems“, „Digital Twin“, „Verwaltungsschale“ und deren Abkürzungen sowie Übersetzungen in andere Sprachen.

⁶ Ein Beispiel wäre der Ausdruck *"interoperability" AND ("IIoT" OR "Industry 4.0") AND ("ontology" OR "knowledge graph") AND ("mapping" OR "matching")* samt der Einschränkung auf Artikel seit 2019, der bei GoogleScholar dennoch 3.330 Einträge liefert. (Stand 18.02.2023)

1.4 Zielgruppe und Bezug zum Studiengang

Diese Arbeit richtet sich an Personen, die mit den Grundkonzepten der IIoT bzw. der Industrie 4.0 vertraut sind, jedoch noch keine vertiefenden Kenntnisse in Ontologien und der Problematik der semantischen Interoperabilität im Kontext der IIoT haben. Dementsprechend liegt der Fokus dieser Arbeit auf der Vorstellung des Begriffes Ontologie sowie des Ontologie Matchings, um anschließend aktuelle Anwendungen zur und Herausforderungen in der Ermöglichung semantischer Interoperabilität im Kontext der IIoT zu präsentieren. Begriffe aus der IIoT werden nur in dem Rahmen eingeführt, in dem diese für das weitere Verständnis von Relevanz sind. Zur Vertiefung der notwendigen Kenntnisse im Bereich der IIoT sei hierbei auf die Werke von Veneri und Capasso (2018), Heidel (2019) oder Pistorius (2020) verwiesen.

Des Weiteren erwähnt diese Arbeit Konzepte und Technologien des Semantic Web, wie beispielsweise das RDF- oder OWL-Format. Auch hier werden die Begriffe nur kurz eingeführt. Auf detailliertere Erklärungen wird an gegebener Stelle verwiesen. Eine erste Einführung bieten neben den in Abschnitt 1.1 erwähnten Quellen auch die Videos von Sack (2020).

Zuletzt werden im Zuge dieser Arbeit wichtige Begriffe mithilfe mathematischer Symbole formal definiert. Eine Wiederholung, der meist aus dem Schulunterricht bekannten mengentheoretischen Symbole, findet sich beispielsweise bei Ebbinghaus (2021) in Kapitel 1 und 2.

Auch wenn zugegebenermaßen nicht unbedeutende Anforderungen an das Vorwissen der Leserschaft gestellt werden, so ist der Bezug zum Studiengang des Business Process Managements naheliegend. Im Kern dieser Arbeit geht es darum, Geschäftsprozesse zwischen maschinellen, aber auch menschlichen Akteuren zu ermöglichen und die Grundlage für eine Steuerung dieser zu schaffen. Damit bespricht diese Arbeit wichtige Aspekte zur digitalen Transformation von Geschäftsprozessen und baut dabei direkt auf dem durch den Studiengang vermittelten Wissen auf.

2 Grundlegende Konzepte

2.1 IloT, CPS und die Bedeutung der semantischen Interoperabilität

2.1.1 IloT

Die Idee hinter dem Begriff Industrial Internet of Things (IloT) ist es, die Konzepte des Internet of Things auf die industrielle Anwendung, aber auch auf ganze Gebäude, Städte oder Lieferketten zu übertragen. Neben IloT werden parallel weitere Begriffe wie Industry 4.0, Industrial Internet oder Internet of Production synonym verwendet. (Boyes et al., 2018; Wortmann & Flüchter, 2015) Nach Schröder, Willner und Diedrich (2021, S. 34) wird dabei ein Netzwerk von dezentral organisierten Systemen über das Internet gebildet, die miteinander interagieren und autonom und intelligent handeln können, um ihre eigenen oder gemeinsamen Zielsetzungen zu erreichen. Die Hypothese ist, dass dieser dezentrale Ansatz die Komplexität großer Systeme besser beherrschbar macht und es so ermöglicht, schneller, flexibler und auf die individuellen Anforderungen der Kunden maßgeschneiderte Produkte und Dienstleistungen zu erstellen. (Teslya; Ryabchikov 2019, S. 138–139) Als Konsequenz ist es erforderlich, dass die Akteure, also die „Things“ im Akronym IloT, miteinander kommunizieren können, was wiederum dazu führt, dass die bisher meist rein physischen (Produktions-)Systeme um ein informationsverarbeitendes System ergänzt werden müssen. Dies hat zur Geburt des Begriffes Cyber-physical Systems geführt.

2.1.2 Cyber-physical Systems

Cyber-physical systems (CPS) are physical and engineered systems whose operations are monitored, coordinated, controlled and integrated by a computing and communication core.

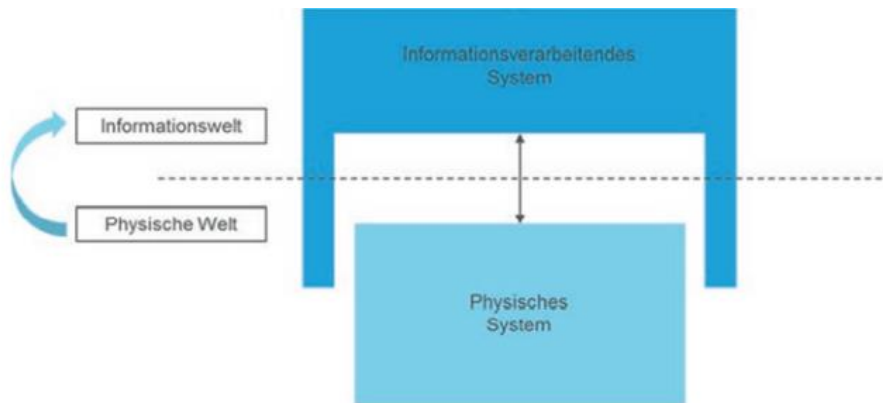
(Rajkumar u.a. 2010, S. 1)

Ein CPS⁷ verbindet somit mechanische Komponenten (Maschine) mit digitalen Komponenten (Software, Informationstechnologie), wie in Darstellung 1 abgebildet. Dadurch ist es möglich, CPS über das Internet zu steuern oder die von den CPS erhobenen Daten in Echtzeit zu analysieren.

CPS bestehen nach Silva und Kaminski (2016, S. 4) im Idealfall aus einem Computer, einem Kontroller, Aktoren und Sensoren, einem Verarbeitungsprozess sowie Kommunikationsschnittstellen für Mensch-Maschinen-Interaktionen sowie für das Internet. Die Idee der CPS ist eng verbunden mit weiteren Konzepten im Bereich der IloT, wie Big Data (=Analyse der Daten von CPS), Digital Twin (digitales Abbild des physischen Systems, das mit diesem bidirektional kommunizieren kann), Predictive Maintenance (Daten des CPS werden ausgewertet, um notwendigen Wartungsbedarf einzuschätzen und um damit Kosten zu reduzieren und die

⁷ Im Produktionsbereich auch als Cyper-physical Production Systems (CPPS) bezeichnet.

operative Lebenszeit zu verlängern) sowie autonomen Systemen (CPS wertet selbständig Daten aus und trifft eigenständige Entscheidungen).



Darstellung 1: Aufbau eines CPS

Beschreibung: Das CPS verbindet die physische Welt mit der digitalen Informationswelt in einem System, um damit die Digitalisierung der Produktion zu ermöglichen

Quelle: Schröder, Willner, Diedrich (2021, S. 34)

Der Begriff CPS wurde erstmals 2006 von Helen Gill erwähnt. (Lee; Seshia 2016, S. 5) Die Idee, physische Komponenten mit Computern zu verbinden, begann jedoch schon kurz nach dem Zweiten Weltkrieg und wurde seither kontinuierlich weiterentwickelt, wie Kim und Kumar in ihrem Artikel ausführen. (Kim; Kumar 2012, S. 1288–1289)

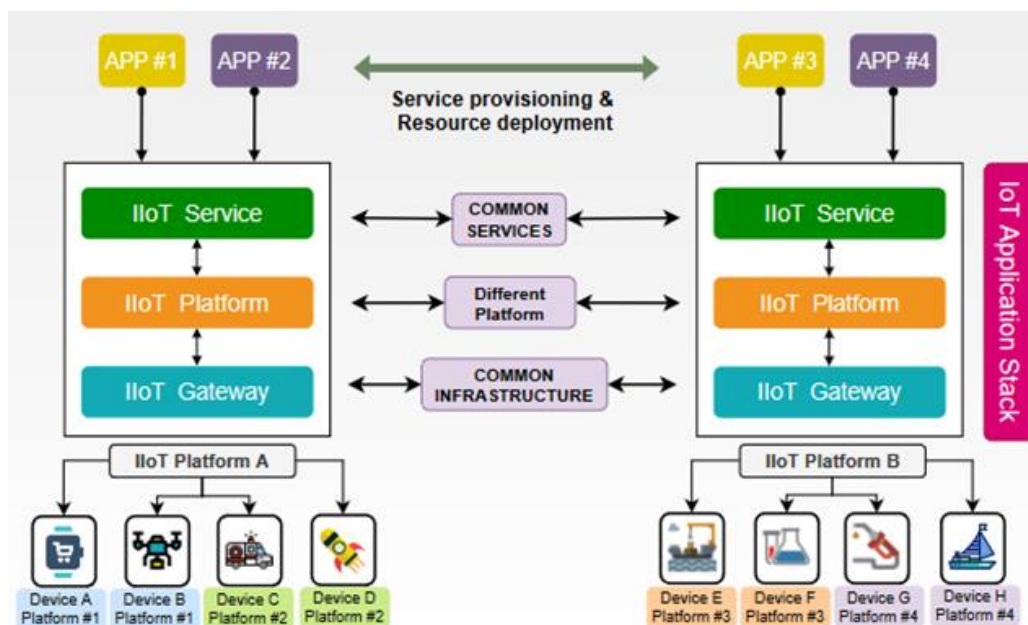
Somit versucht der Begriff CPS im Grunde ein (mittlerweile) weitgehend triviales Konzept zu beschreiben: Ein an das Internet angeschlossener Computer mit spezifischer Peripherie. Diese spezifischen Peripherien können beispielsweise Bohr- und Fräselemente (CNC-Fräse), Mix- und Kochelemente (Produkte wie die Küchenmaschine Thermomix) oder ein ganzes Auto (autonom fahrendes Auto) sein. Während im Endkundenbereich meist „smart“-Begriffe wie „smart device“ oder „smart home“ verwendet werden, ist CPS ein Synonym, dessen Besonderheit es ist, dass dieses nicht von IT-, sondern von Industrieunternehmen geprägt wurde. Mit anderen Worten: CPS ist nach Ansicht des Autors ein Begriff, der ein allgegenwärtiges Phänomen – nämlich die Digitalisierung – beschreibt, für das es in unterschiedlichen Domänen auch unterschiedliche Begriffe gibt. Anders formuliert: CPS ist kein spezifisches Phänomen der IIoT.

2.1.3 Interoperabilität als grundlegende Herausforderung der IIoT

Im Bereich der Produktion können CPS verwendet werden, um Informationen über die Produktionslinie zu sammeln, diese auszuwerten und auf Basis der Auswertungen den physischen Prozess zu steuern. Dies kann einerseits auf lokaler Ebene geschehen, beispielsweise wenn eine CNC-Maschine die Position des Bohrkopfes ermittelt, überprüft, ob diese korrekt ist und anschließend entweder die Position anpasst oder zum nächsten Prozessschritt übergeht, wodurch der Bohrer aktiviert wird. Andererseits kann dies auch auf höherer Ebene

passieren, in dem mehrere CPS in ein großes System integriert und ganze Produktionslinien gesteuert werden. Hierbei spricht man von smart factory⁸ beziehungsweise auch von digital factory, wenn mehrere smart factories verbunden werden. (Mabkhot u.a. 2018, S. 4–5)

Konzepte wie jenes der IIoT bedienen sich dieser Idee, mit dem Ziel, die Produktivität zu steigern oder neue Geschäftsmodelle zu entwickeln. Voraussetzung dafür ist die Integration der CPS in ein Gesamtsystem. Für dieses Gesamtsystem werden neben CPS weitere Komponenten wie Plattformen und Services betrachtet, wie in Darstellung 2 dargestellt ist. Industrielle Anwendungen bestehen dabei aus verschiedenen Komponenten, die miteinander kommunizieren müssen. Auf unterster Ebene finden sich die CPS, die über Plattformen miteinander oder mit Services interagieren.



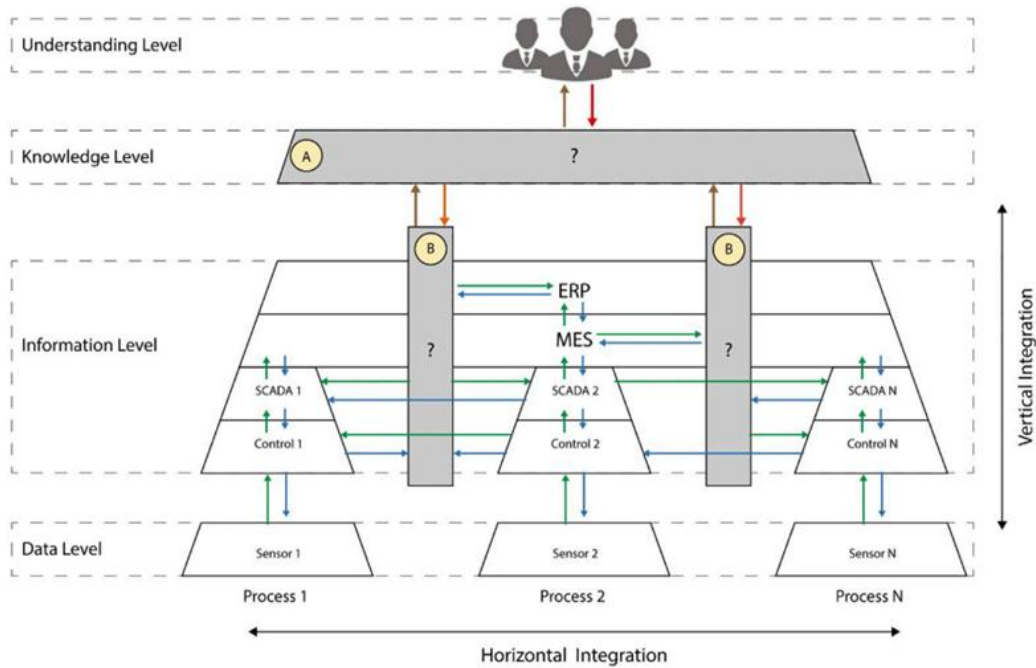
Darstellung 2: Komponenten einer industriellen Anwendung im Kontext der IIoT

Quelle: Hazra u.a. (2023, S. 10)

Da diese Komponenten in der Regel sehr heterogen sind, ist die Integration zu einem funktionierenden Gesamtgebilde eine der großen Hürden in der Umsetzung der IIoT. (Jirkovský 2017, S. 13–14, 17; Teslya; Ryabchikov 2019, S. 139) Diese Integration muss dabei sowohl horizontal, sprich entlang der Wertschöpfungskette, als auch vertikal entlang der Organisationsstruktur vom Shopfloor bis zum Management bzw. den jeweiligen Informationssystemen

⁸ Eine „smart factory“ ist nach Radziwon u.a. (2014, S. 1188) „a manufacturing solution that provides such flexible and adaptive production processes that will solve problems arising on a production facility with dynamic and rapidly changing boundary conditions in a world of increasing complexity.“ Synonym verwendet werden Begriffe wie „cloud manufacturing“, „smart production“ oder „digital manufacturing“. (Nieuwenhuize 2016, S. 9)

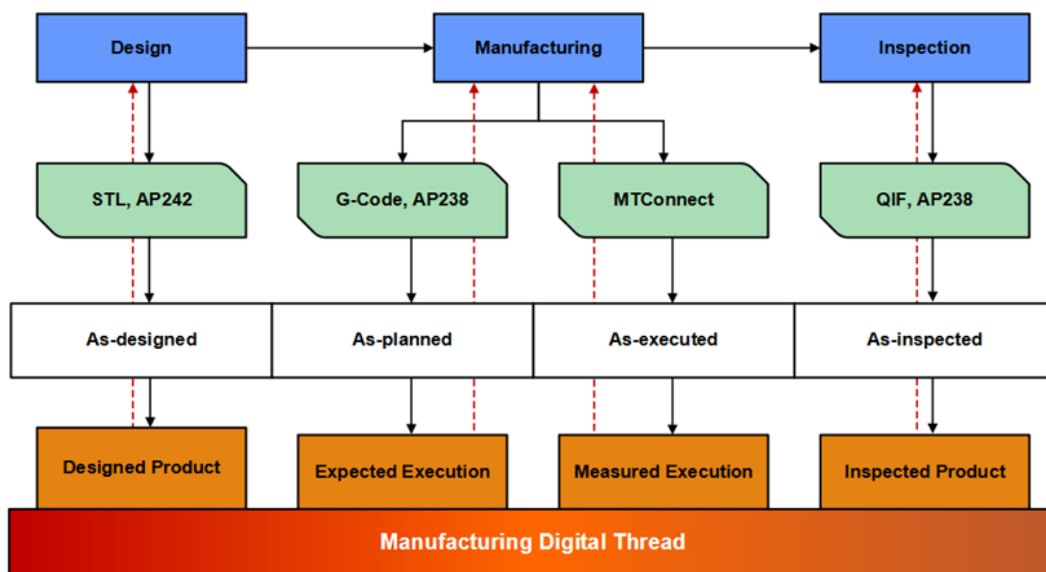
(SCADA, MES, ERP) erfolgen (siehe Darstellung 3). (Plattform Industrie 4.0 2015, S. 19,26; Sony; Naik 2020, S. 2)



Darstellung 3: Horizontale und vertikale Integration in einem Produktionssystem.

Quelle: Pereira, Szejka und Canciglieri (2021, S. 424)

Wie komplex allein die horizontale Integration sein kann, zeigt das Beispiel eines Entwicklungs- und Fertigungsprozesses von Lu, Xu und Wang (2020, S. 22–23), bei dem vom Design über die Produktion bis zur Qualitätskontrolle unterschiedliche Standards (STL, AP242, G-Code, AP238, MTConnect, QIF, AP238) verwendet werden.



Darstellung 4: Beispiel für heterogene Standards in einem Entwicklungs- und Fertigungsprozess

Quelle: Lu, Xu und Wang (2020, S. 22)

Im Folgenden betrachten wir die möglichen Ebenen einer Integration von Systemen, um das Verständnis für die Begriffe der semantischen Integration zu schärfen. Hierzu definieren wir zuerst den Begriff Interoperabilität.

2.1.4 Einordnung semantischer Interoperabilität

Ide und Pustejovsky (2010, S. 2) definieren Interoperabilität als Maß für die Fähigkeit unterschiedlicher Objekte, Systeme oder Organisationen zusammenzuarbeiten, um ein gemeinsames Ziel zu erreichen. Dabei kann Interoperabilität zwischen digitalen Systemen in die drei Ebenen Application (Anwendung), Middleware und Network (Netzwerk) unterteilt werden. Um Informationen austauschen zu können, müssen alle Ebenen betrachtet werden.⁹ (Hazra u.a. 2023, S. 2)

Jede dieser Ebenen unterteilt sich in weitere Subebenen (siehe Darstellung 5). So wird die Netzwerk-Ebene in der Regel durch das Open Systems Interconnection (OSI) Modell realisiert. Für die Middleware werden meist Abstraktionsebenen eingeführt, für die es Standards wie OPC-UA¹⁰ und oneM2M¹¹ gibt. Während die Netzwerk-Ebene die technische Interoperabilität abdeckt, bildet die Middleware die Voraussetzung für die syntaktische sowie semantische Interoperabilität, die in der obersten Ebene, der Anwendung, zentral sind. (Schröder; Willner; Diedrich 2021, S. 35) Syntaktische Interoperabilität wird beispielsweise erschwert, wenn unterschiedliche Formate wie OPC UA, oneM2M oder MTConnect verwendet werden. (Jirkovský 2017, S. 70)

Semantik bezeichnet die Bedeutung eines Begriffes. Umgangssprachlich beschrieben sind zwei Begriffe semantisch ident, wenn sie dasselbe „meinen“, d.h. sie zeigen auf dasselbe reale Objekt oder gedankliche Konstrukt (Konzept). Darüber hinaus beschreibt Semantik auch die Beziehung zwischen den Konzepten bzw. Objekten und damit den Kontext. (Gardner 2005, S. 1003–1004) Semantische Interoperabilität ist folglich das Maß für die Fähigkeit mehrerer Akteure, sich gegenseitig zu verstehen.

Im Bereich der IIoT wäre Interoperabilität theoretisch leicht zu erreichen, wenn die unterschiedlichen Akteure auf jeder Ebene dieselben Protokolle und Standards verwenden würden. Aus diesem Grund liegt ein großer Schwerpunkt der Initiativen in der IIoT auf dem Erstellen und Harmonisieren von Referenzarchitekturen¹² und Standards.¹³ Aktuell scheint eine auf allen Ebenen homogene Kommunikation zwischen CPS die Ausnahme zu sein. Viel

⁹ Es gibt noch weitere Unterteilungen von Interoperabilität. So differenzieren Cimmino u.a. (2019, S. 99–100) Interoperabilität in semantische, syntaktische und technische. Andere Modelle wie das „levels of conceptual interoperability model“ (LCIM) von Tolk und Muguira (2003) oder jenes von Noura, Atiquazzaman und Gaedke (2019) haben noch weitere Stufen. Allen gemein ist, dass die semantische Interoperabilität eine dieser Stufen ist.

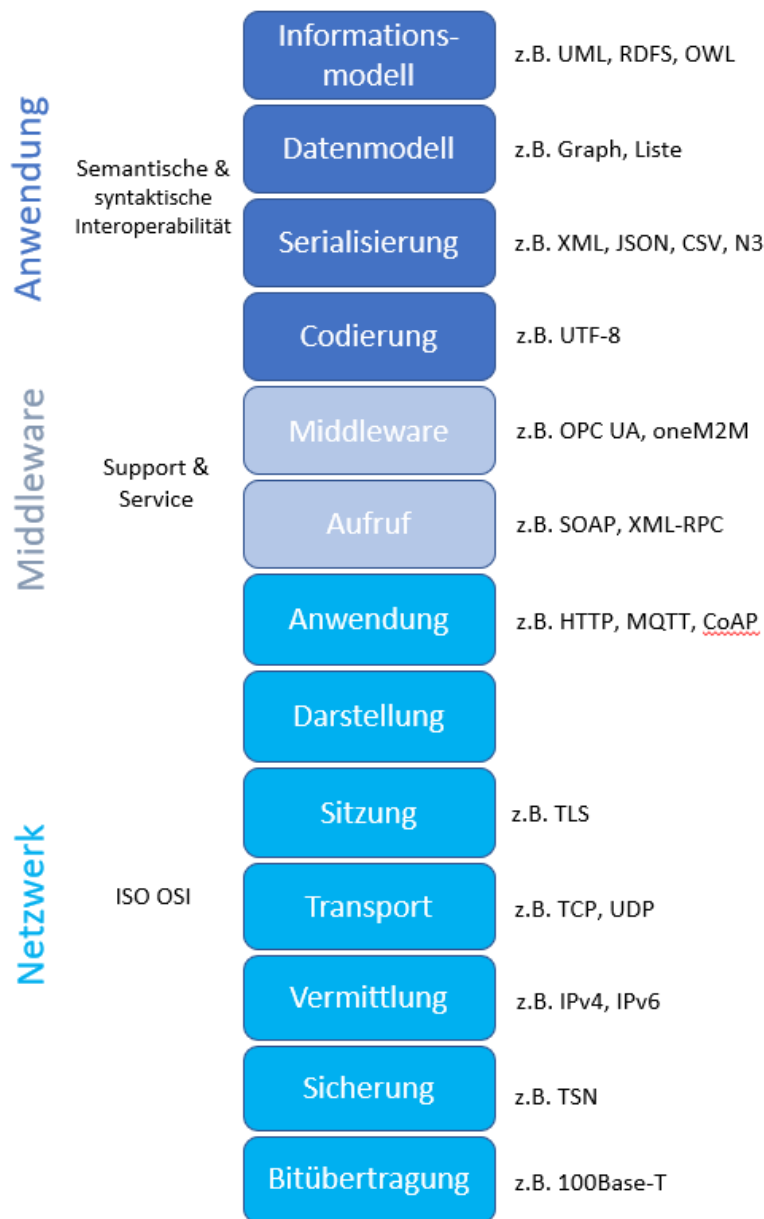
¹⁰ <https://opcfoundation.org/about/opc-technologies/opc-ua> (Zugriff am: 30.05.2023)

¹¹ <https://www.onem2m.org/using-onem2m/what-is-onem2m> (Zugriff am: 30.05.2023)

¹² Ein Beispiel für ein Referenzarchitekturmodell ist RAMI4.0. (Adolphs u.a. 2015; Heide 2019)

¹³ Einen aktuellen Überblick über die Vielzahl an Organisationen, die in diesem Feld tätig sind, bieten Hazra u.a. (2023, S. 22–25) sowie die Normungsroadmap Industrie 4.0 der DIN und DKE (2023, S. 103–105).

wahrscheinlicher ist das Vorliegen einer Heterogenität.¹⁴ (Vila u.a. 2022, S. 3; Hazra u.a. 2023, S. 2–3)



Darstellung 5: Ebenen der Interoperabilität

Quelle: In Anlehnung an Schröder, Willner und Diedrich (2021, S. 34, 35)

Der Fokus dieser Arbeit liegt auf der semantischen Interoperabilität. Genauer gesagt: Wie kann Wissen – gemeint ist die Bedeutung der Begriffe samt dem Kontext, in dem sie sich befinden – zwischen mehreren CPS übertragen werden? Dafür bedarf es zum einen einer Lösung, Wissen maschinenlesbar abzubilden und zum anderen Heterogenität in

¹⁴ Hazra u.a. (2023, S. 2) verwenden sogar den Ausdruck „extreme heterogeneity“, um zu beschreiben, dass Interoperabilitätsprobleme im Bereich der IIoT meist auf allen drei Ebenen (Network, Middleware, Application) vorkommen.

Wissensrepräsentationen unterschiedlicher CPS aufzulösen, da Heterogenität Interoperabilität verhindert.¹⁵ Lösungsansätze hierfür bedienen sich meist des Konzeptes der Ontologie, (Fensel u.a. 2020, S. 31) welches im folgenden Kapitel genauer beschrieben wird. Die möglichen Arten von Heterogenität und wie damit umgegangen werden kann, beleuchtet Kapitel 3.

¹⁵ Hierbei ist anzumerken, dass Heterogenität auch Vorteile hat. So sind maßgeschneiderte und damit untereinander heterogene Systeme in der Regel effizienter beim Lösen ihrer (lokalen) Probleme. (Jirkovský 2017, S. 40) Problematisch wird die Heterogenität erst, wenn Interoperabilität benötigt wird, was – wie oben erläutert – eine Kernanforderung in der IIoT ist.

2.2 Ontologien

“The ontologies are the pillar component when implementing an interoperable environment, and specifically to implement the semantic interoperability. The ontologies allow to define terms and relationships in a problem domain without ambiguity, and allow to establish a consensus between different communities about the meaning and the use of data.”

(Cimmino u.a. 2019, S. 121)

2.2.1 Versuch einer Definition

Der Begriff „Ontologie“ (griechisch ὄν (seiend) und λόγος (Lehre), also „Lehre des Seins“) stammt aus der Philosophie und bezeichnet nach Aristoteles die Lehre über die von der Natur gegebenen Eigenschaften und Beziehungen von Dingen. Der Begriff wurde in den 1960er-Jahren von der Computerwissenschaft aufgenommen, da er sich insbesondere im Bereich der künstlichen Intelligenz als nützlich erwies. Aus der Beschreibung der Realität folgt zwangsläufig, dass eine Ontologie auch das vorhandene Wissen über diese beinhaltet. Aus diesem Grund erfuh die Ontologie hier nicht nur ein zweites Leben, sondern auch eine Bedeutungsveränderung¹⁶. (Neuhaus 2018, S. 5, 14) Demnach war eine Ontologie nun eine formale, explizite Spezifikation einer geteilten Konzeptualisierung.¹⁷ (Studer; Benjamins; Fensel 1998, S. 184) „Formal“ bedeutet, dass die Ontologie in einer formalen Sprache abgebildet wird. Hierauf soll in Abschnitt 2.2.2 eingegangen werden.

Während in der Fachliteratur meist auf diese Nominaldefinition verwiesen wird,¹⁸ ist sie keineswegs allgemein akzeptiert. Die Kritik betrifft einerseits die Verwendung des Begriffs „Konzeptualisierung“ zur Beschreibung einer Ontologie, da dieser für sich selbst schon einer genaueren Definition bedarf, was wiederum aufgrund unterschiedlicher Tendenzen in der Literatur nicht trivial ist. (Neuhaus 2018, S. 3) Andererseits argumentieren Guarino und Giaretta (1995, S. 2,7), dass die Definition aufgrund des Wortes „Spezifikation“ zu eng gefasst wurde und schlagen deshalb die alternative Definition „*a logical theory which gives an explicit, partial account of a conceptualization*“ vor.¹⁹

In diesem Zusammenhang erwähnen Leser und Neumann (2007, S. 274) die Unterscheidung zwischen informellen und formalen Ontologien. Neben unstrukturierten Termlisten und Thesauri sind auch Kataloge Varianten von informellen Ontologien. So gibt etwa ein

¹⁶ Im Englischen erfuh der Begriff auch eine lexikalische Anpassung: Während der philosophische Begriff *Ontology* nach wie vor großgeschrieben wird, wird für die Computerwissenschaften *ontology* verwendet.

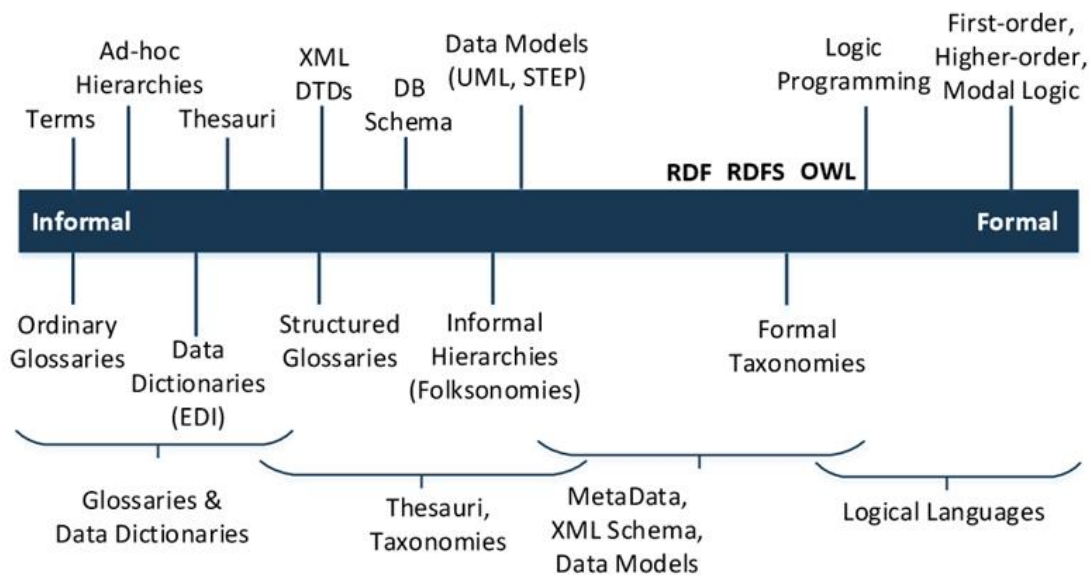
¹⁷ Original: „formal, explicit specification of a shared conceptualization“. Diese Definition ist eine Erweiterung jener von Gruber (1993), der Ontologie als „an explicit specification of a conceptualization“ definierte.

¹⁸ Wie zum Beispiel bei Leser und Naumann (2007, S. 272) sowie bei Schröder, Willner und Diedrich (2021, S. 38).

¹⁹ Guarino und Giaretta geben dazu passend auch eine Definition von Konzeptualisierung an: „*an intensional semantic structure which encodes the implicit rules constraining the structure of a piece of reality*“ (Guarino; Giaretta 1995, S. 6)

Firmenkatalog bereits Auskunft über die wichtigsten Entitäten (Produkte, Produktfamilien und deren Eigenschaften) und deren Beziehung zueinander. Folglich ist auch jedes Sachbuch eine informelle Ontologie.²⁰ Formale Ontologien wiederum haben eine formale Spezifikation, die es erlaubt, aus nicht explizit angegebenen Informationen mithilfe von Regelwerk neue Informationen zu ermitteln oder Integritätsbedingungen²¹ zu formulieren. Insbesondere sind formale Ontologien maschinenlesbar, weshalb sie ein wichtiges Werkzeug für die IIoT sind und damit im Fokus dieser Arbeit stehen. Aus diesem Grund ist im Weiteren mit dem Begriff „Ontologie“ implizit immer eine formale Ontologie gemeint.

Darstellung 6 fasst die unterschiedlichen Arten von Ontologien sowie deren Zugehörigkeit zu eher informellen oder formalen Ontologien zusammen.



Darstellung 6: Informelle und formale Ontologien

Quelle: Grangel-González (2018, S. 21) in Anlehnung an Guarino, Oberle und Staab (2009, S. 20, 21)

Die Unterscheidung in formale und informelle Ontologien löst jedoch nicht das Problem mit der Definition von Ontologien. Betrachten wir deshalb die Problematik mit dem Begriff „Konzeptualisierung“ genauer. Die Idee hinter einer Ontologie und jedes anderen Modells ist es, einen abstrahierten und vereinfachten Ausschnitt aus der Welt, eine Miniworld, zu beschreiben. (Gruber 1993, S. 2) Die Kernfrage lautet in Anlehnung an Willard Van Orman Quine: „Was gibt es?“ (Quine 1948, S. 1) Welche Objekte, welche Beziehungen und welche Eigenschaften? In einer Ontologie werden diese mit Symbolen (Namen) gekennzeichnet. Die

²⁰ Damit sind Standards ebenfalls Ontologien, da auch hier die Bedeutung der Begriffe beschrieben wird. Meist sind Standards jedoch keine formalen Ontologien, was mitunter ein Problem darstellt. In Abschnitt 4.3.2 werden Beispiele besprochen, in denen Standards um formale Ontologien ergänzt werden, um das Problem der semantischen Interoperabilität zu lösen.

²¹ Integritätsbedingungen sind überprüfbare Regeln, die beispielsweise Wertebereiche von Attributen einschränken oder Relationen einfordern (z.B.: „Jeder Vater muss eine „ist-Elternteil-von“-Relation haben). (Leser; Naumann 2007, S. 20)

Symbole zeigen einerseits auf ein Konzept unserer Vorstellung (die sogenannte Intension) als auch auf eine Menge realer Objekte (Extension), die durch dieses Konzept zusammengefasst und beschrieben werden.²² (Leser; Naumann 2007, S. 74–75)

Das Symbol „Apfel“ verweist beispielsweise auf ein Konzept in unserer Vorstellung (ein rundliches Kernobst mit Stiel in der Größe einer Faust) als auch auf die Menge aller real existierenden Äpfel. Die Frage, ob Konzepte wie zum Beispiel „Apfel“ oder abstraktere wie „Wille“ und „Zeit“ existieren oder rein gedankliche Konstruktionen sind,²³ beschäftigt die Philosophie schon seit Platons Ideenlehre und hatte im Mittelalter mit dem Universalienstreit einen Höhepunkt, der bis heute nicht endgültig beigelegt ist.

So definierte der Philosoph Willard Van Orman Quine (1909-2000) Klassen zuerst rein extensional, wodurch diese rein aus der Summe der beinhalteten Gegenstände und deren Eigenschaften definiert werden. Im Zuge des sogenannten „ontologischen Debakels“, das durch Erkenntnisse in der Quantenphysik ausgelöst wurde, kam Quine jedoch von dieser strikten Definition ab und ließ auch abstrakte Konzepte zu. (Quine 1976, S. 499–503; Fulda 1988, S. 76–78)

Während die Frage, ob und welche Konzepte existieren, in der Philosophie nach wie vor diskutiert wird, ist der Ansatz in den Computerwissenschaften – analog zu jenem der Informatik, Mathematik und Physik – ein durchaus pragmatischer. Genauso wie in der Mathematik Definitionen aufgestellt werden, in der Informatik bzw. der objektorientierten Programmierung Klassen²⁴ gebildet werden, oder in der Physik mit imaginären Zahlen gerechnet wird, ohne dass geklärt oder gar diskutiert wird, ob dies tatsächlich existiert, so werden auch in Ontologien abstrakte Konzepte definiert und verwendet. Dass diese Einstellung nicht nur Vorteile mit sich bringt, zeigt das Beispiel der double articulation, die in Abschnitt 3.1.2 besprochen wird.

Eine exakte und durchgängig anerkannte Nominaldefinition des Begriffs Ontologie scheint somit bis jetzt nicht gefunden worden zu sein.²⁵ Es lassen sich jedoch grundlegende

²² Die Frage, woher die Symbole ihre Bedeutung erhalten, also mit ihren Konzepten verbunden sind, wird in der Literatur als „symbol grounding problem“ bezeichnet und ist ein grundlegendes Problem der Kognitions- und Computerwissenschaften, auf das in dieser Arbeit jedoch nicht weiter eingegangen wird. Ausführungen hierzu findet sich bspw. bei Harnad (1990) und Cubek u.a. (2015).

²³ Da die Mathematik voll von abstrakten Konzepten ist (z.B. Stetigkeit, geometrische Figuren, imaginäre Zahlen, ...) führt die Frage, ob solche Konzepte existieren oder nur erfunden sind zwangsläufig zur Frage, ob Mathematik entdeckt oder erfunden wird. So hielt der Mathematiker Leopold Kronecker (1823-1891) beispielsweise folgendes fest: „Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.“ Davon abgeleitet bildete sich auch in der Mathematik mit dem Konstruktivismus eine Richtung, die nur Existenzaussagen zuließ, die sich mit einer endlichen Zahl an Schritten von existenten Objekten ableiten lassen, was die Welt der Mathematik und der Naturwissenschaften, wie wir sie heute kennen, stark einschränken würde, da Konzepte wie jene aus der Infinitesimalrechnung (Differenzieren, Integrieren) oder die imaginären Zahlen nicht existent und damit nicht untersuchungsrelevant wären. (Strick 2016) Dieser Konflikt wurde auch in Robert Musils (1880-1942) „Die Verwirrungen des Zöglings Törleß“ aufgegriffen, in dem er dem verzweifelten Törleß, der am Sinn der Verwendung der imaginären Zahlen zweifelt, wenn es sie doch gar nicht gibt, ausrichten lässt: „Ich glaube, wenn man allzu gewissenhaft wäre, so gäbe es keine Mathematik.“ (Musil 1906)

²⁴ Auch wenn der Klassen-Begriff in der objektorientierten Programmierung mit jenem der Ontologie verwandt ist, so gibt es hier markante Unterschiede, auf welche Rebstock, Fengel und Paulheim (2008, S. 99) genauer eingehen.

²⁵ Nicht anders ist der Versuch von Neuhaus zu interpretieren, Ontologie über eine eigens geschaffene formale Theorie zu definieren. (Neuhaus 2018)

Eigenschaften und Komponenten von Ontologien benennen. Eine Ontologie besteht nach Corcho und Gómez (2000, S. 81–84) aus folgenden Komponenten:

- Konzepte: hierunter fallen sowohl konkrete als auch abstrakte, elementare oder zusammengesetzte, reale oder fiktive Konzepte wie „Mensch“, „Zeit“, „Arbeit“
- Taxonomie: eine hierarchische Struktur, über die Vererbungsmechanismen angewendet werden können (das Konzept Mutter ist eine Spezialisierung des Konzepts Frau.)
- Axiome: Aussagen, die allgemeingültigen Charakter haben. Sie dienen dazu, Informationen bzw. Teile von Ontologien auf ihre Korrektheit zu prüfen bzw. neues Wissen abzuleiten. („Jede Frau, die ein Kind hat, ist eine Mutter.“)
- Instanzen/Individuen/Literale: Elemente eines Konzeptes. („Eva“ ist eine Instanz des Konzeptes „Frau“.)
- Relationen: gerichtete oder ungerichtete Verbindungen zwischen Konzepten oder Instanzen. („x ist Mutter von y.“)

Ausgehend davon gibt Jirkovský (2017, S. 22–23) eine formale Definition einer Ontologie an. Danach ist eine Ontologie O eine Struktur, definiert durch:

$$O := (C, \leq_C, R, \sigma_R, \leq_R, A, \sigma_A, T)$$

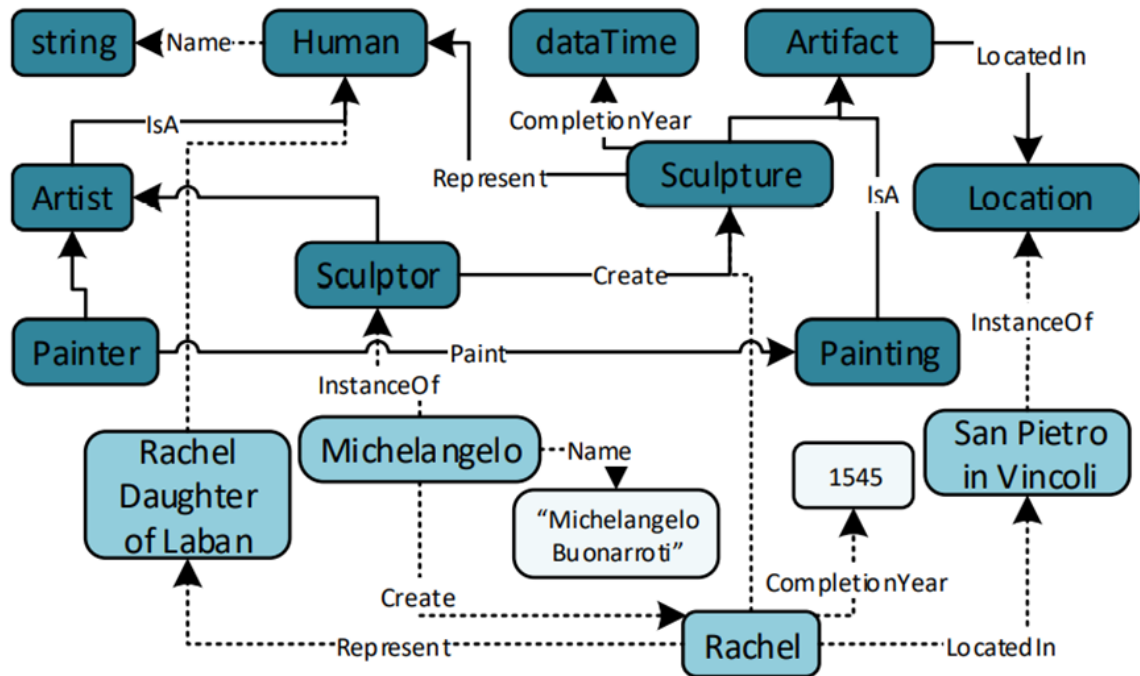
Wobei

- C die Konzepte und Instanzen, R die Relationen zwischen den Konzepten, A die Attribute und T die Datentypen sind
- \leq_C eine hierarchische (algebraische) Struktur auf C mit einem Wurzelknoten $root_C$, genannt Konzepthierarchie (auch als Taxonomie bezeichnet), ist
- $\sigma_R: R \rightarrow C^+$ die Relationssignatur ist, die mehrere Konzepte verbindet
- \leq_R eine hierarchische Struktur auf R ist, genannt Relationshierarchie
- $\sigma_A: A \rightarrow C \times T$ eine Relation ist, eine sogenannte Attributsignatur, die ein Attribut einem Konzept und einem Datentyp zuweist.

Mit anderen Worten: Eine Ontologie ist eine Struktur, die aus Konzepten, Relationen, Attributen und Datentypen besteht, wobei sowohl die Konzepte als auch die Relationen hierarchisch geordnet werden können, die Relationen die Konzepte verbinden und die Attribute den Konzepten samt einem Datentyp zugeordnet werden. Darstellung 7 gibt ein Beispiel für eine solche Ontologie.

Die in Darstellung 7 abgebildete Ontologie besteht aus Konzepten (Human, Artist, Painter, Sculptor, Artifact, Painting, Sculpture, Location), aus Relationen (Represent, LocatedIn, Create, Paint) und aus Attributen (Name, CompletionYear). Die Konzepte haben eine hierarchische Struktur. So sind Sculptor sowie Painter eine Spezialisierung des Konzepts Artist, das wiederum das Konzept Human spezialisiert.

Die Relationen verbinden mehrere (meist 2) Konzepte. So ist die Relation Create eine gerichtete Relation von Sculptor (Bildhauer) zu Sculpture (Skulptur).



Darstellung 7: Beispielhafte Darstellung einer Ontologie

Quelle: Jirkovský (2017, S. 23)

Der Nutzen einer Ontologie liegt dabei nicht nur in der Abbildung von explizitem Wissen, sondern auch darin, aus explizitem Wissen mittels Inferenz implizites Wissen zu entdecken. Ein einfaches Beispiel, um neues Wissen per Inferenz zu identifizieren, wäre die Antwort auf die Frage, wie viele Skulpturen sich in San Pietro in Vincoli befinden.²⁶

Des Weiteren zeigt Darstellung 7 auch den Unterschied zwischen einer Ontologie und einem Datenschema. Eine Ontologie beinhaltet meist nicht nur die abstrakte Ebene der Klassen und Strukturen, auch als T-Box (Terminologische Box) bezeichnet, sondern auch die Instanzen, die sogenannte A-Box (Assertional Box). (Osman; Ben Yahia; Diallo 2021, S. 39) Im obigen Beispiel gehören die dunkelblau gehaltenen Entitäten zur T-Box, die hellblau hinterlegten Entitäten zur A-Box.

Zuletzt kann anhand des obigen Beispiels ein weiterer Begriff beschrieben werden: Der „Kontext“ eines Begriffs ist die Umgebung dieses Begriffs innerhalb der Ontologie (sprich, die durch Relationen verbundenen Konzepte und Instanzen). Der Renaissance-Künstler Michelangelo steht mit anderen Konzepten in Relation, als die Comicfigur der Turtles mit gleichem Namen und befindet sich somit in einem anderen Kontext.

²⁶ Vorausgesetzt, die in Darstellung 7 abgebildete Ontologie ist vollständig, müsste ein Inferenzalgorithmus alle „LocatedIn“-Relationen zählen, die auf das Konzept „San Pietro in Vincoli“ zeigen. Natürlich könnte man sich auch komplizierte Fragestellungen ausdenken wie „Nenne alle Künstler, die sowohl eine Skulptur als auch ein Gemälde geschaffen haben, das in San Pietro in Vincoli steht.“

2.2.2 Formale Repräsentationen von Ontologien

Um eine Ontologie zu beschreiben, braucht es eine Ontologie-Sprache. Im Laufe der Zeit haben sich hierfür mehrere Typen von Sprachen entwickelt, die sich in ihrer Ausdrucksmächtigkeit (Was kann gesagt werden?) und ihrer Fähigkeit, komplexe Inferenzen zu bilden (Welches neue Wissen kann aus dem Bestehenden mittels logischen Schließens erzeugt werden?) unterscheiden. Dabei stellte sich heraus, dass sich diese beiden Eigenschaften diametral gegenüberstehen und es keine Ontologie-Sprache gibt, die für beides in besonderem Maße geeignet ist. Entweder hat eine Sprache eine hohe Ausdrucksmächtigkeit, oder sie hat eine hohe Inferenzfähigkeit. (Corcho; Gómez-Pérez 2000, S. 80)

Ein Beispiel, wie mit diesem Gegensatz umgegangen wird, ist die Web Ontology Language²⁷ (OWL), von der es in der Version 1.0 drei unterschiedliche Ausprägungen gibt.²⁸ Die volle Ausprägung des OWL-Standards (OWL Full) ist vollständig kompatibel mit RFD²⁹ und verfügt damit über eine hohe Ausdrucksmächtigkeit. Dagegen können Ontologien, die den vollen OWL-Standard nutzen, mitunter nicht entscheidbar sein.³⁰ Deshalb schränkt OWL DL (Web Ontology Language Description Logic) die Ausdrucksmächtigkeit zugunsten der Entscheidbarkeit ein, beispielsweise durch Zuweisung einer Ressource zu genau einem Datentyp (Klasse, Individuum, Eigenschaft). Da auch OWL DL für bestimmte Usecases noch nicht die gewünschten Inferenzfähigkeiten lieferte, wurde zusätzlich noch OWL Lite entwickelt, das im Gegenzug weitere Ausdrucksmächtigkeit einbüßte. (Jirkovský 2017, S. 34–35)

Im Allgemeinen können Ontologie-Sprachen in drei Gruppen unterteilt werden:

- Sprachen aus der Beschreibungslogik, zu denen auch OWL gehört
- Sprachen aus der Prädikatenlogik 1. Ordnung, wie z.B. KIF (Genesereth; Fikes 1992)
- Frame-basierte Sprachen wie Ontolingua (Gruber 1992)

Nähere Details zu den Unterschieden finden sich beispielsweise bei Fensel. (2001, S. 61–72)

Da sich diese Arbeit mit Methoden des Semantic Web zur Lösung des Problems der semantischen Interoperabilität beschäftigt und OWL hier der gängige Standard ist, gehen wir im Weiteren davon aus, dass Ontologien mithilfe einer Beschreibungslogik abgebildet werden.

²⁷ www.w3.org/OWL (Zugriff am: 30.05.2023)

²⁸ Die aktuelle Version OWL 2.0 verfügt ebenfalls über mehrere Subvarianten (OWL EL, OWL QL, OWL RL). („OWL 2 Web Ontology Language Profiles (Second Edition)“ 2012) Diese lassen sich jedoch nicht mehr so einfach hierarchisch beschreiben, weshalb zum besseren Verständnis auf OWL 1.0 zurückgegriffen wurde.

²⁹ Ressource Description Framework ist ein Format des W3C, um Informationen im Web darzustellen. Die Darstellung erfolgt dabei durch ein Tripel, das aus zwei Konzepten (Subjekt und Objekt) und einem Prädikat (Relation) besteht. Diese bilden zusammen einen Graphen, wobei die Konzepte die Knoten und die Relation die Kante ist, die die beiden Knoten verbindet. (Hartig; Champin; Kellogg 2023)

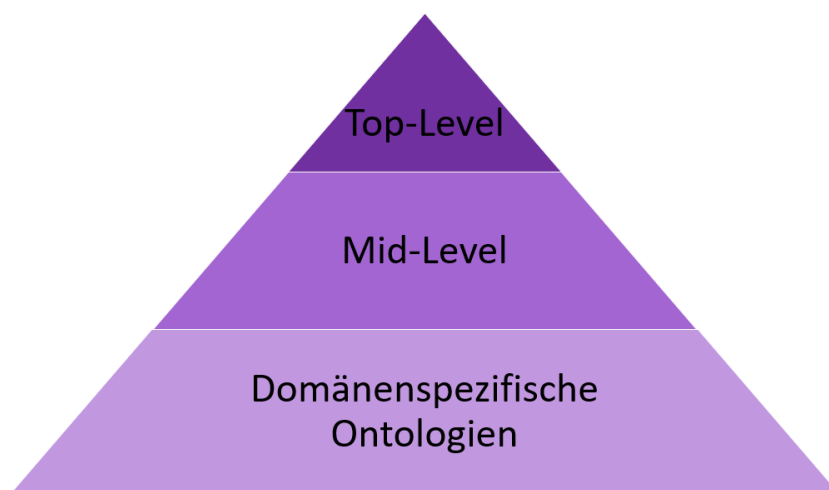
³⁰ Nicht entscheidbar bedeutet, dass Aussagen getroffen werden können, für die es nicht möglich ist zu entscheiden, ob diese richtig oder falsch sind. Dieses ist ein fundamentales Problem, das der österreichische Mathematiker Kurt Gödel in seinen Unvollständigkeitssätzen bewiesen hat und die Grundfesten der Mathematik erschütterte. („Gödelscher Unvollständigkeitssatz“ 2022) Ein weiterer Grund, die Ausdrucksmächtigkeit weiter einzuschränken, kann aber darin liegen, dass ein Problem in polynomialer Zeit gelöst werden kann. Hier sei auf das P-NP-Problem verwiesen. („P-NP-Problem“ 2023)

2.2.3 Typen von Ontologien

Wie in Abschnitt 2.2.1 erwähnt, bilden Ontologien mithilfe von Konzepten, Relationen und weiteren Bausteinen Wissen auf maschinenlesbare Weise ab. Mittlerweile gibt es eine Fülle an unterschiedlichen Ontologien. Diese werden meist abhängig von der Spezifität ihrer Konzepte in drei Gruppen unterteilt: Top-Level, Mid-Level und domänenspezifische Ontologien.

Top-Level Ontologien (auch Upper-Ontologien, Referenz-Ontologien oder „foundation ontology“ genannt) sind eine spezielle Form von Ontologien, die grundlegende Konzepte definieren und zueinander in Relation setzen, um das Fundament für spezifischere Ontologien zu erzeugen. Dies schafft den Vorteil, dass diese grundlegenden Konzepte nicht jedes Mal neu definiert werden müssen (z.B. Definition einer „Klasse“, eines „Prozess“ oder einer „Subklassen-Relation“). (Leser; Naumann 2007, S. 276) Mid-Level Ontologien beinhalten allgemeine Konzepte und Relationen, die für eine oder mehrere spezifische Domänen relevant sind, wie beispielsweise für Chemie oder Materialwissenschaften. Unterhalb dieser Mid-Level Ontologien liegen die domänenspezifischen Ontologien. Erst diese Ebene erlaubt es, Wissen in einem solchen Detaillierungsgrad abzubilden, wie er für konkrete Anwendungen benötigt wird. (Adamovic u.a. 2022, S. 18, 24; Rudnicki 2019, S. 2) Darüber hinaus liegt der Vorteil domänenspezifischer Ontologien darin, dass ihr Verwendungszweck begrenzt ist, was wiederum die Wahrscheinlichkeit für eine Akzeptanz der Ontologien unter den beteiligten Akteuren erhöht. (Leser; Naumann 2007, S. 276)

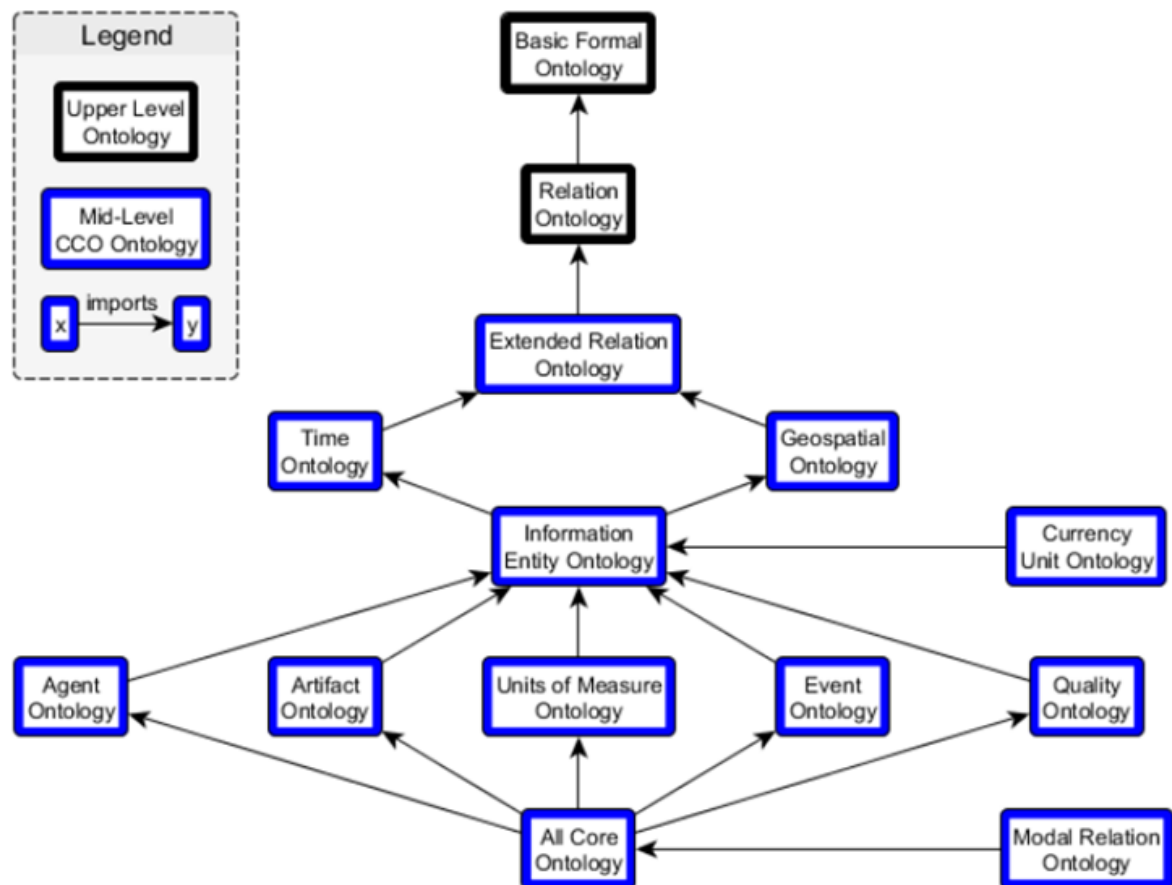
Die Konzepte zwischen Top-Level, Mid-Level und domänenspezifischen Ontologien sind durch Subklassen-Relationen verbunden und werden von oben nach unten immer konkreter. Diese Aufteilung erlaubt dadurch modularen Aufbau von Ontologie-Architekturen. (Lemaignan u.a. 2006, S. 2)



Darstellung 8: Typen von Ontologien

Quelle: In Anlehnung an Adamovic u.a. (2022, S. 5)

Ein Beispiel für eine durchaus komplexe Ontologie-Architektur sind die Common Core Ontologien.³¹ Diese bestehen aus zwei Upper Level Ontologien – die Basic Formal Ontology (BFO) – und die Relation Ontology – und mehreren Mid-Level Ontologien. Interessant ist, dass sich die Ontologien jeweils an Konzepten der anderen Ontologien bedienen, wie in Darstellung 9 ersichtlich ist. Ziel der Common Core Ontologien ist es, ein Basisvokabular zur Verfügung zu stellen, das um domänenspezifische Ontologien erweitert werden kann, um die semantische Interoperabilität innerhalb eines Unternehmens zu ermöglichen. Domänenspezifische Ontologien, die auf den Common Core Ontologien aufbauen, sind beispielsweise die Aircraft Ontology oder die Army Universal Task List Ontology. (CUBRC 2023)



Darstellung 9: Architektur der Common Core Ontologien

Quelle: Rudnicki (2019, S. 23)

Neben den in Darstellung 9 abgebildeten Beispielen für Top-Level Ontologien gehören auch Ontologien wie EMMO, DOLCE, BORO, ISO 15926, OPM und SUMO in diese Kategorie. Beispiele für Domänenspezifische Ontologien sind ScorVoc, MatOnto, SSN, SONA.

Die Unterteilung in Top-Level, Mid-Level und domänenspezifische Ontologien ist in der Literatur jedoch nicht einheitlich. So bezeichnet Charpeney (2019, S. 19) die SSN Ontologie

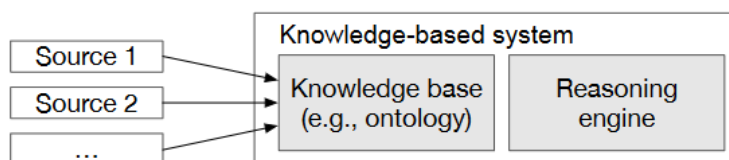
³¹ Wie wichtig Ontologien heutzutage sind, zeigt der Umstand, dass die Regierung der Vereinigten Staaten von Amerika die Rechte an den Common Core Ontologien hält. (Jensen u.a. 2023)

(siehe Abschnitt 2.2.6.2.1) als „upper ontology“, die mit domänenspezifischen Ontologien ergänzt werden muss. Als Beispiel für eine solche domänenspezifische Ontologie nennt der Autor eine, die das Konzept „Temperatur“ beinhaltet, was wiederum ein Konzept ist, das in den Common Core Ontologien in Form der „Units and Measure Ontology“ innerhalb einer Mid-Level Ontologie definiert wird. Analog wird die SAREF Ontologie (siehe Abschnitt 2.2.6.2.2) als domänenspezifisch (Amara; Hemam; Djezzar; Maimor 2022, S. 278) und als unspezifische „core ontology“ (Li u.a. 2019, S. 3) klassifiziert, was als Bezeichnung für Mid-Level Ontologien verstanden werden kann.

Dennoch ist die Unterscheidung aus Sicht des Autors sinnvoll, da sie insbesondere in Hinblick auf die Entwicklung neuer bzw. Verbesserung bestehender Ontologien eine sinnvolle Architektur vorgibt. Nicht zuletzt ist das Teilen gemeinsamer (Meta-)Konzepte ein wichtiger Faktor in der Ermöglichung von Alignments zwischen zwei Ontologien, wie in Abschnitt 3 und 4 genauer besprochen werden wird.

2.2.4 Wissensgraph vs. Ontologie

Der Begriff Wissensgraph (engl. Knowledge Graph) ist eng verwandt mit dem Begriff Ontologie. Während beispielsweise Paulheim und Cimiano (2017, S. 490) sowie Osman u.a. (2021, S. 39) Wissensgraph als graphen-basierte Visualisierung von Wissen³² beschreiben, dessen Konzepte und Instanzen Knoten und dessen Relationen die Kanten sind, verwenden Arenas u.a. (2016, S. 1) Wissensgraph als Bezeichnung für große Ontologien. Ehrlinger und Wöß wiederum sehen Ontologien als Komponenten eines Wissensgraphen, der zusätzlich noch über eine Inferenzmaschine („Reasoning engine“) verfügt, die aus dem bestehenden Wissen neues Wissen erzeugt.³³ Mit letzterer Definition ist auch das Semantic Web ein Wissensgraph, der das gesamte – maschinenlesbare – Wissen im Internet abbildet und dessen Inferenzmaschine die Suchmaschinen wie Google oder Bing sind. (Ehrlinger; Wöß 2016, S. 3–4)



Darstellung 10: Architektur eines Wissensgraphen

Quelle: Ehrlinger und Wöß (2016, S. 3)

³² Die Notation RDF verwendet Tripel (Subjekt, Prädikat, Objekt) die einen Graphen bilden, indem Subjekt und Objekt die Knoten und das Prädikat die Kante darstellen. Osman u.a. (2021, S. 39) beschreiben eine Ontologie wie folgt: „It (Anm.: An Ontology) can be viewed as a labeled directed graph whose nodes are entities, and edges are relations. Nodes are labeled by entity names, and edges are labeled by relation names.“

³³ Ein einfaches Beispiel hierfür wäre, dass aus „Eva ist Mutter von Abel“ und „Eva ist Mutter von Kain“ sowie der Regel „A und B sind Geschwister, wenn sie dieselbe Mutter/denselben Vater haben“ die Aussage abgeleitet wird: „Kain und Abel sind Geschwister.“

Zuletzt definieren Baumgartner u.a. (2023, S. 4) einen Wissensgraphen als Struktur aus drei Mengen

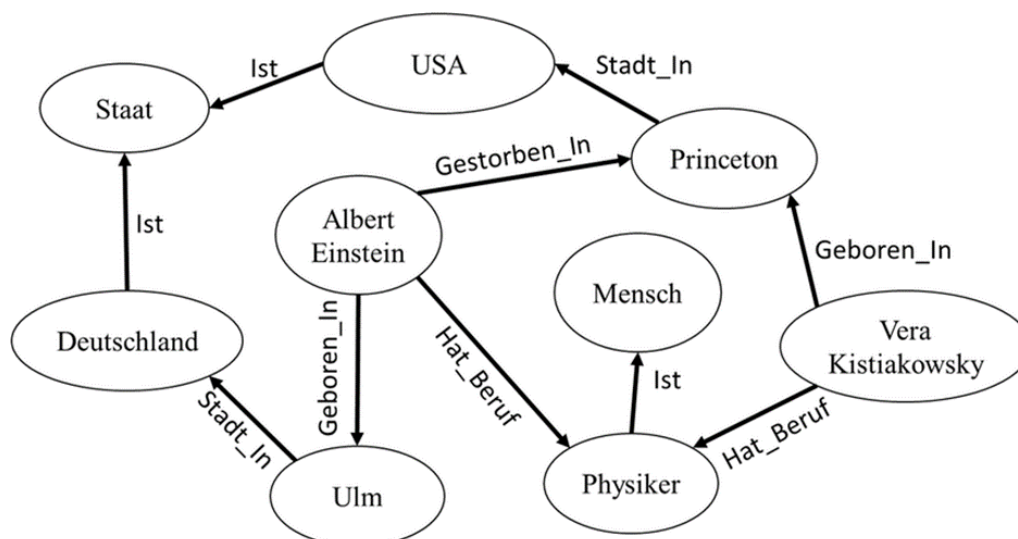
$$G := (V, R, E)$$

wobei V die Menge der Entitäten (Konzepte und Individuen) und R die Menge der Relationen sind. E ist die Menge der (RDF-)Triple $t = (S, P, O)$, mit denen Aussagen wie „Eva ist Mutter von Abel“ abgebildet werden. Hierbei sind S (Subjekt) und O (Objekt) Entitäten (Eva, Abel) und P ist eine Relation („ist Mutter von“). Damit kommt diese Definition jener in Abschnitt 2.2.1 erwähnten formalen Definition einer Ontologie sehr nahe. Bedeutung erlangte der Begriff Wissensgraph vor allem durch Google, das ihn 2012 öffentlichkeitswirksam vorstellte, ohne eine exakte Definition zu liefern. (Heist u.a. 2021, S. 128)

In Abbildung 10 wurde ein weiterer Begriff eingeführt, die Knowledge Base (Wissensbasis). Ehrlinger und Wöß verwenden diesen als Oberbegriff einer Ontologie, ohne ihn genauer zu definieren. Dem gegenüber definiert Jirkovský Knowledge Base als Set aller Instanzen einer Ontologie samt deren Relationen. (Jirkovský 2017, S. 26) Dies kommt somit eher der Definition einer A-Box gleich.

Dass die Unterscheidung zwischen Wissensgraph und Ontologie nicht durchgängig ist, zeigt sich schon daran, dass DBpedia (siehe Abschnitt 2.2.6.1.2) sowohl als Ontologie (Fensel u.a. 2020, S. 7) als auch als Wissensgraph (Heist u.a. 2021, S. 129) bezeichnet wird. Um die babylonische Sprachverwirrung perfekt zu machen, verwenden Javed u.a. (2022, S. 1) den Begriff „ontology graph“.

Aus Sicht des Autors überwiegt in der Literatur jedoch die Verwendung von Wissensgraph als Synonym für bzw. als visuelle Darstellung von Ontologien (siehe Darstellung 7 oder 11). Aus diesem Grund wird im weiteren Verlauf dieser Arbeit der Begriff Ontologie jenem des Wissensgraphen vorgezogen. Ausnahmen bilden Wissensgraph Embeddings (Knowledge Graph Embeddings, siehe Abschnitt 3.5.2), da sich hier in der Literatur die Verwendung des Begriffes Wissensgraph gegenüber jenem der Ontologie durchgesetzt hat. Wissensgraph ist dabei aber als Synonym für Ontologie zu verstehen.



Darstellung 11: Wissensgraph oder visuelle Darstellung einer Ontologie

Quelle: Schönhof, Konertz (2020, S. 41)

2.2.5 Allgemeiner Nutzen von Ontologien innerhalb einer Domäne

Bisher haben wir versucht zu bestimmen, was Ontologien sind, welche Typen von Ontologien es gibt und welche Begriffe damit eng verwandt sind. Bevor wir auf konkrete Beispiele von Ontologien eingehen, die unser tägliches Leben beeinflussen, betrachten wir den allgemeinen Nutzen von Ontologien.

Ontologien geben zuallererst einen Überblick über die wichtigsten Konzepte und deren Beziehungen für jene Domäne, für die sie konzipiert sind. In weiterer Folge bilden sie damit das Domänenwissen ab.

“Given a domain, its ontology forms the heart of any system of knowledge representation for that domain. Without ontologies, or the conceptualizations that underlie knowledge, there cannot be a vocabulary for representing knowledge.”

(Chandrasekaran; Josephson; Benjamins 1999, S. 21)

Ontologien definieren somit das Vokabular einer Domäne samt den Beziehungen der darin befindlichen Begriffe und ermöglichen so (funktionierende) Kommunikationsprozesse und damit den Austausch von Wissen zwischen Expertinnen und Experten sowie zwischen Maschinen. (Leser; Naumann 2007, S. 272, 275; Kumar u.a. 2019, S. 4)

Mit der Idee des Semantic Web kamen Ontologien wieder verstärkt in den Fokus der Forschung und der Industrie. Mittlerweile werden sie in vielen unterschiedlichen Bereichen eingesetzt, wie beispielsweise in Suchmaschinen, beim Natural Language Processing (NLP), in der Robotik oder der industriellen Automation. (Jirkovský 2017, S. 50)

2.2.6 Beispiele für Ontologien bzw. Wissensgraphen und deren Verwendung

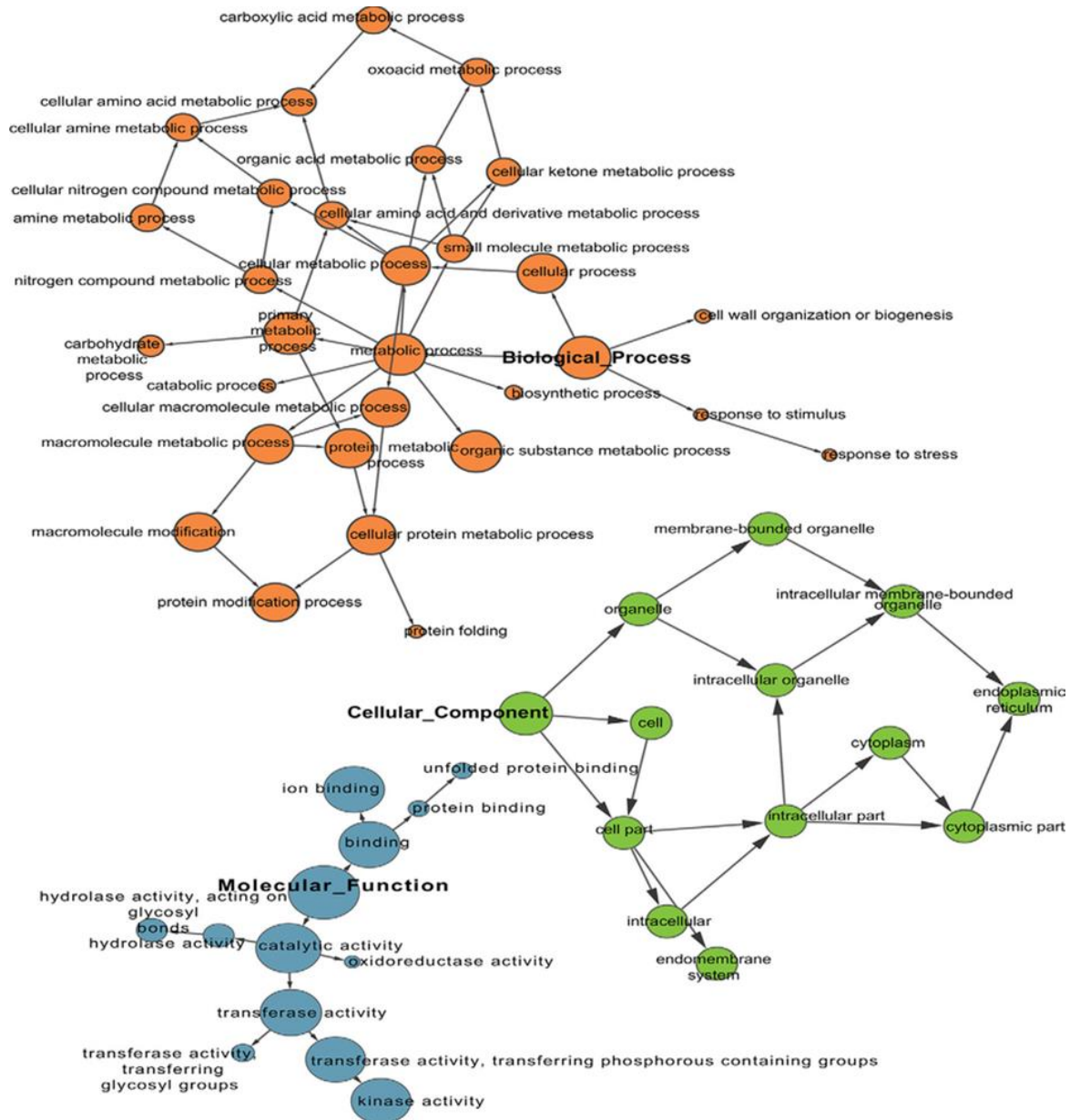
Im folgenden Abschnitt werden exemplarisch einige Ontologien vorgestellt. Der Schwerpunkt liegt dabei weniger auf dem genauen Inhalt der Ontologien als vielmehr auf dem Nutzen und potenziellen Anwendungen dieser Ontologien. Dabei werden nicht nur IloT-spezifische Ontologien besprochen. Seit einigen Jahren gibt es eine Fülle von öffentlich zugänglichen Ontologien, die unser Leben tagtäglich beeinflussen. Aus diesem Grund werden zwei wichtige Vertreter solcher öffentlichen Ontologien besprochen. Begonnen wird jedoch mit einer Ontologie aus dem Bereich der Naturwissenschaften, die einen wichtigen Beitrag zur rapiden Entwicklung innerhalb ihrer Domäne beigetragen hat und heute ein fester Bestandteil in der modernen Forschung in diesem Bereich ist.

2.2.6.1 Ontologien außerhalb des IloT-Kontexts

2.2.6.1.1 Gene-Ontologie

Der menschliche Organismus besitzt um die 20.000 Gene, die für die Bildung von Proteinen verantwortlich sind. (Hartley 2022) Diese Proteine sind auf molekularer Ebene an chemischen

Reaktionen beteiligt, die in Summe unterschiedliche biologische Prozesse steuern. Bei der Erforschung der Funktionen von Genen und Proteinen sind Forschende aus unterschiedlichen Ländern und Wissenschaftsdomänen (Medizin, Biologie, Pharmazie, ...) beteiligt, die mitunter unterschiedliche Begriffe für dieselben Konzepte verwenden. Dadurch war es bis zur Einführung der Gene Ontology³⁴ beinahe unmöglich, alle Forschungsergebnisse, die zu einem bestimmten Gen publiziert wurden, zu verknüpfen. Dies hat sich mit der Einführung der Gene Ontology 2001 wesentlich verbessert. (Leser; Naumann 2007, S. 269–271)



Darstellung 12: Die Gene Ontology

Beschreibung: Ausschnitt der Gene Ontology mit den drei Root Konzepten: zelluläre Komponenten, molekulare Funktionen und biologische Prozesse.

Quelle: Saeed, Baranwal und Khurana (2016, S. 8)

³⁴ <http://geneontology.org> (Zugriff am: 30.05.2023)

Die Gene Ontology umfasst die Bezeichnungen und Beschreibungen zellulärer Komponenten, molekularer Funktionen und biologischer Prozesse (siehe Darstellung 12). Sie ist nach Eigendefinition die weltweit größte Quelle für Informationen über die Funktionen von Genen und deren Produkten (Proteine) über alle Lebewesen und Viren hinweg. Dieses Wissen ist die Grundlage für die Analyse von groß angelegten molekularbiologischen und genetischen Experimenten in der biomedizinischen Forschung. Mit Stand April 2023 besteht sie aus 43.093 Begriffen und 7.442.411 Relationen. Neben der Schaffung einer einheitlichen Terminologie dient die Ontologie dazu, Wissen, das über einen Organismus entdeckt wurde, auf andere Organismen zu übertragen. Dies geschieht beispielsweise über Gensequenzvergleiche. (The Gene Ontology Consortium u.a. 2021, S. 1; The Gene Ontology Consortium 2023) Die Gene Ontology ist damit ein Paradebeispiel dafür, Wissen in menschen- und maschinenlesbarer Form zu speichern, zu verknüpfen, wieder zu finden und zu teilen.

2.2.6.1.2 DBpedia

DBpedia³⁵ ist der zentrale Baustein des Semantic Web bzw. der Linked Open Data Cloud,³⁶ der Verknüpfung von im Internet frei zugänglichen Datensätzen, da DBpedia mit vielen weiteren Datensätzen verbunden ist. Darstellung 13 zeigt die englischsprachige DBpedia im Zentrum der Linked Open Data Cloud. Die weiteren braun hinterlegten DBpedia Einträge betreffen weitere Sprachvarianten. (Fensel u.a. 2020, S. 7)

³⁵ <https://dbpedia.org> (Zugriff am: 30.05.2023)

³⁶ <https://lod-cloud.net> (Zugriff am: 30.05.2023)

15 Hotels in Ouagadougou - Bestpreisgarantie.
www.booking.com/ouagadougou/hotels
 Anzeige Hotels in Ouagadougou reservieren. Schnell und sicher online buchen.
 Typen: Hotels, Ferienwohnungen, Villen, Hostels, Resorts, B&Bs

Luxus hotels Budget hotels
 Jetzt Buchen Keine Reservierungsgebühr
 Für morgen buchen

Villas Ouagadougou - Jetzt online buchen
www.wimdu.de/villa/ouagadougou Jetzt buchen
 Anzeige Finden Sie die besten Angebote in Ouagadougou für Ihren nächsten Urlaub! Schnell und bequem buchen. Reservieren Sie jetzt online!

Ouagadougou burkina faso - Top-Hotels in Ouagadougou
www.tripadvisor.de
 Anzeige Ouagadougou sensationell günstig. Jetzt Urlaub auf Tripadvisor buchen.
 Hotels - ab 19,00 €/Nacht Hotels mit Pool - ab 60,00 €/Nacht

Top 10 Luxury Hotels Ouagadougou - 5 Star Best Hotels Ouagadougou
www.luxuryhotelsguides.com/ouagadougou Book Hotel
 Anzeige Top 10 Best Luxury Hotels Ouagadougou. 5 Star Small Luxury Hotels.
 Types: Hotels, Motels, Hostels, B&Bs, Resorts, Villas

Status	commune urbaine	Region	Centre
Provinz	Kadiogo	Fläche	219,3 km²
Mehr			

Darstellung 14: Bing-Suche nach Ouagadougou

Quelle: <https://www.bing.com/search?q=ouagadougou> (Zugriff am: 01.06.2023)

2.2.6.1.3 Schema.org

Schema.org³⁷ ist eine Ontologie, die die Suche im Web unterstützen soll. Gestartet als Initiative von Suchmaschinenherstellern (Google, Yahoo, Bing, Yandex) dient Schema.org dazu, Webseiten-Inhalte zu markieren, sodass sie von den Suchmaschinen semantisch verstanden werden können. Schema.org beinhaltet beispielsweise Eigenschaften wie „Name“ oder „Adresse“ oder Typen wie „Airport“ oder „Restaurant“. (Fensel u.a. 2020, S. 15–16) Bindet man die Markierungen von Schema.org in die eigene Webseite ein, so kann damit auch die Reihung in der Suche einer Suchmaschine verbessert werden (Stichwort: Search Engine Optimization) oder – ein banales wie hilfreiches Feature – die Öffnungszeiten und geographische Position eines Supermarktes können als Ergebnis in der Suchmaschine angezeigt werden.

Für schema.org ist eine Erweiterung um Begriffe aus dem Bereich der IoT in Arbeit. (Ancic 2022; Brickley 2023) Konkrete Anwendungen hierfür konnte der Autor bisher nicht ausmachen.

2.2.6.2 IIoT-Ontologien

Nun sollen zwei wichtige Ontologien aus dem Bereich der IIoT besprochen werden. Dabei konzentriert sich der Autor auf die Beschreibung zweier Ontologien, die – zumindest in der untersuchten Literatur – häufig erwähnt werden und auch für das Verständnis der in Abschnitt 4 vorgestellten Anwendungen relevant sind. Im Anschluss folgt ein Verweis auf weitere IIoT-Ontologien.

³⁷ <https://schema.org> (Zugriff am: 01.06.2023)

2.2.6.2.1 Semantic Sensor Network (SSN)

Die Semantic Sensor Network Ontologie³⁸ (SSN) des W3C ist gedacht, um unabhängig von der jeweiligen Domäne Sensoren und deren Beobachtungen zu beschreiben. Die ursprüngliche Version von SSN erwies sich als zu "schwergewichtig" für viele Anwendungen und deckte wichtige Konzepte wie z. B. Aktoren nicht ab. Aus diesem Grund wurde in einer Überarbeitung ein Kern-Modul namens SOSA (Sensor, Observation, Sample, Actuator) geschaffen, das (im Sinne der „Leichtgewichtigkeit“³⁹) nur die zentralen Konzepte wie z.B. „Sensor“, „Actuator“ oder „Observation“ beinhaltet.

Die neue SSN Ontologie integriert alle Konzepte und Relationen von SOSA und erweitert diese um zusätzliche, wie beispielsweise „System“ oder „Latency“. Dabei ist es aber dennoch möglich, SOSA auch alleine zu nutzen, wenn die Erweiterung auf SSN für den Anwendungsfall nicht benötigt wird. (Haller u.a. 2017; Ganzha u.a. 2021, S. 6) In Bezug auf die IIoT ist zu betonen, dass SSN damit wichtige Konzepte für CPS beinhaltet (wie Sensor, Aktor), jedoch keine darüberhinausgehenden Konzepte, wie sie beispielsweise für die Produktion relevant sind (das Konzept „Produkt“ fehlt beispielsweise). Somit ist SSN zwar eine wichtige Ontologie für IIoT-Anwendungen, jedoch nur in Kombination mit anderen nutzenstiftend.

2.2.6.2.2 Smart Appliance Reference (SAREF)

Die Smart Appliance Reference⁴⁰ (SAREF) Ontologie ist ein Standard des Europäischen Instituts für Telekommunikationsnormen (ETSI), der ursprünglich entwickelt wurde, um Anwendungen im smart home Bereich abzubilden. SAREF entwickelte sich im Laufe der Zeit aber zu einer der wichtigsten Ontologien für die Domäne der IIoT. SSN und SAREF haben eine Überlappung, da letztere beispielsweise ebenfalls Konzepte wie „Device“, „Sensor“ oder „Actuator“ beinhaltet (siehe Darstellung 15). (Moreira u.a. 2017, S. 2)

SAREF hat einen modularen Aufbau, der um domänenspezifische Konzepte ergänzt werden kann. Hieraus haben sich unterschiedliche Erweiterungen gebildet. So listet das Linked Open Vocabulary⁴¹ mit Stand 19.05.2023 insgesamt 10 Erweiterungen auf. Hierzu gehören Erweiterungen in Domänen wie Landwirtschaft (SAREF4AGRI), Energiewirtschaft (SAREF4ENER), Bauwirtschaft (SAREF4BLDG) oder industrieller Produktion (SAREF4INMA).

Im Bereich der IIoT gibt es noch eine Vielzahl weiterer Ontologien, wie beispielsweise die Manufacturing's Semantics Ontology (MASON) (Lemaignan u.a. 2006) oder die RAMI Ontology⁴². Eine detailliertere Auflistung bieten Kumar u.a. (2019) und Korecky (2022). Eine Liste öffentlich zugänglicher Ontologien findet sich im Linked Open Vocabulary⁴³. Dass die Entwicklung von Ontologien im Bereich der IIoT noch nicht abgeschlossen ist und dass es

³⁸ <https://www.w3.org/TR/vocab-ssn> (Zugriff am: 19.05.2023)

³⁹ „Leichtgewichtigkeit“ meint einerseits, dass die Anzahl der Konzepte und Relationen im Vergleich zur ursprünglichen Version von SSN geringer ist und andererseits, dass eine erhöhte Performance in Anwendungen realisiert werden kann. (Lelli 2019, S. 11)

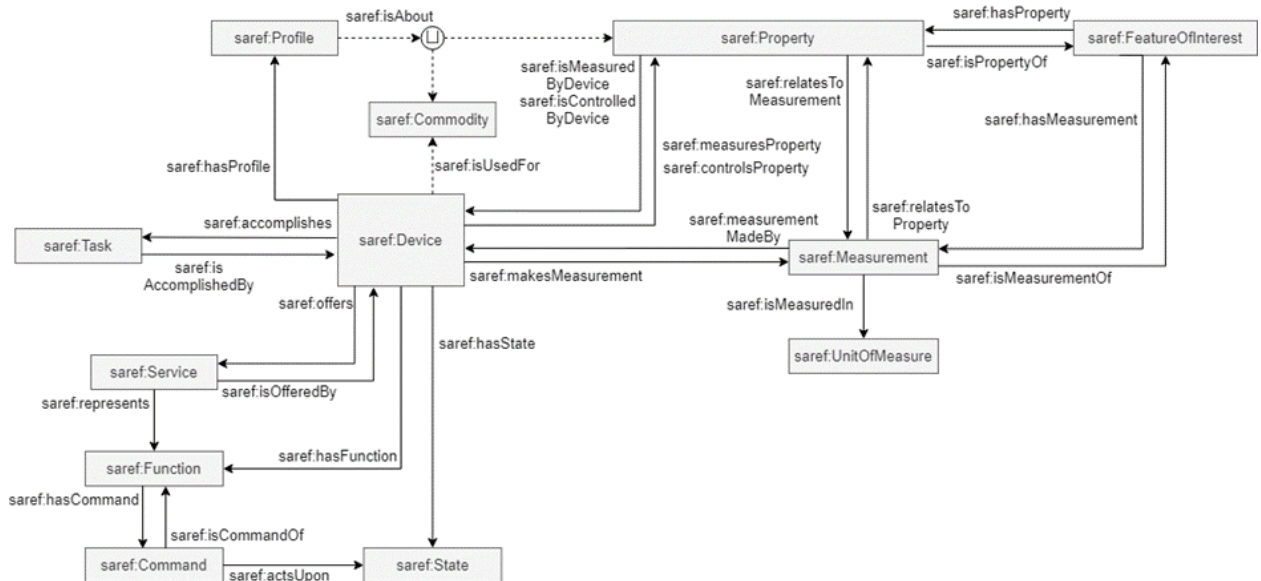
⁴⁰ <https://saref.etsi.org/core/v3.1.1> (Zugriff am: 19.05.2023)

⁴¹ Siehe <https://lov.linkeddata.es/dataset/lov/vocabs?q=SAREF> (Zugriff am: 19.05.2023)

⁴² <https://github.com/i40-Tools/RAMIOntology> (Zugriff am: 19.05.2023)

⁴³ <https://lov.linkeddata.es/dataset/lov> (Zugriff am: 01.04.2023)

insbesondere „die eine IIoT-Ontologie“ noch nicht gibt, zeigen die Beispiele von Vila (2022) und Pereira (2022), die beide unabhängig voneinander neue IIoT-Ontologien entwickeln.



Darstellung 15: Überblick über die wichtigsten Konzepte und Relationen der SAREF Ontologie

Quelle: Daniele u.a. (2020)

Hieraus erwächst jedoch ein fundamentales Problem: Einerseits gibt es viele unterschiedliche domänenspezifische Ontologien, andererseits setzen die horizontale und vertikale Integration, welche Eckpfeiler der IIoT sind, eine Interoperabilität dieser Domänen voraus. Dies ist nicht ohne weiteres möglich, da unterschiedliche Ontologien unterschiedliche Konzepte verwenden oder mitunter dieselben Konzepte mit unterschiedlichen Begriffen. Andere Ontologien wiederum verwenden ähnliche Konzepte, nur ist ihr Detaillierungsgrad oder ihre Perspektive eine andere. Und zuletzt ist die Erstellung einer Ontologie immer mit einem Designprozess verbunden, in dem sich nicht nur das Hintergrundwissen der Entwicklerinnen und Entwickler, sondern auch unterschiedliche Anforderungen spiegeln, sodass unterschiedliche Teams unterschiedliche Designentscheidungen treffen, auch wenn sie an sich dieselbe Domäne beschreiben. (Osman; Ben Yahia; Diallo 2021, S. 38) Folglich können wir davon ausgehen, dass es in Zukunft auch weiterhin unterschiedliche Ontologien benötigen wird.⁴⁴

Um nun aber semantische Interoperabilität zu ermöglichen, benötigt es einen Austausch zwischen den Ontologien, womit wir uns in folgendem Kapitel genauer beschäftigen werden.

⁴⁴ Es gibt Ansätze, wie jenen von Vila (2022), die versuchen, die Interoperabilitätsprobleme mit einer einzigen Ontologie zu lösen. Die Autoren schränken aber selbst ein, dass dies nur dann funktioniert, wenn die unterschiedlichen Akteure diese Ontologie verwenden – was aus Sicht des Autors jedoch eine im Allgemeinen nicht erfüllbare Voraussetzung darstellt. (Vila u.a. 2022, S. 4)

3 Ontologie Matching

Wer fremde Sprachen nicht kennt, weiß nichts von seiner eigenen.

Johann Wolfgang von Goethe (1749 – 1832)⁴⁵

In den obigen Ausführungen haben wir gesehen, dass Ontologien eine Möglichkeit bieten, Wissen formal und maschinenlesbar abzulegen, was sie zu einer wichtigen Technologie für das Semantic Web aber auch für die IIoT macht. Umgekehrt haben wir gezeigt, dass es bis zum heutigen Tag nicht die eine, alles umfassende Ontologie gibt und dass Heterogenität eher die Regel als die Ausnahme ist (siehe Abschnitt 2.1.4). Konsequenterweise braucht es somit für die Gewährleistung der Interoperabilität zwischen Akteuren der IIoT auch eine Interoperabilität zwischen Ontologien. Anders ausgedrückt: Damit CPS, die jeweils unterschiedliche Ontologien verwenden, untereinander kommunizieren können, braucht es eine Verbindung bzw. Übersetzung von einer Ontologie in die andere. Diese Problematik wird u.a. mit dem Begriff Ontologie Matching bezeichnet.

In diesem Kapitel werden das Konzept des Ontologie Matchings sowie unterschiedliche Methoden, verfügbare Softwaretools und aktuelle Herausforderungen besprochen. Begonnen wird dabei mit einer genaueren Betrachtung der Problematik und dem zu erreichenden Ziel:

The goal of matching ontologies is to reduce heterogeneity between them.

(Euzenat; Shvaiko 2013, S. 37)

Dementsprechend wollen wir zuerst die Arten von Heterogenität genauer betrachten.

3.1 Semantische Heterogenität

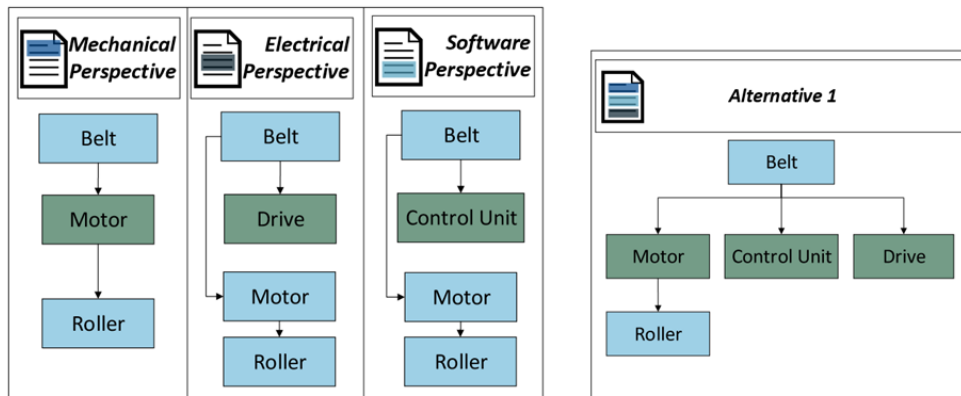
Modelle dienen dazu, unterschiedliche Aspekte der Realität abzubilden. Je nach Zweck, Ziel, Adressaten und Hintergrundwissen der Entwicklerinnen und Entwickler unterscheiden sich Modelle, auch wenn sie an sich denselben Ausschnitt aus der Realität beschreiben. Dies gilt insbesondere auch für Ontologien. Anhand dreier aufeinander aufbauender Beispiele soll das Problem verdeutlicht werden.

3.1.1 Semantische Heterogenität anhand dreier Beispiele

Auf der linken Seite von Abbildung 16 ist der Aufbau eines Förderbandes aus drei unterschiedlichen Perspektiven dargestellt: Aus mechanischer, elektronischer und aus der Software-Perspektive. Je nach Perspektive stehen unterschiedliche Konzepte im Fokus, die sich aber teilweise überschneiden. Im einfachsten Fall (Alternative 1 in Darstellung 16) beinhalten die

⁴⁵ Zitiert nach Goethe (1907)

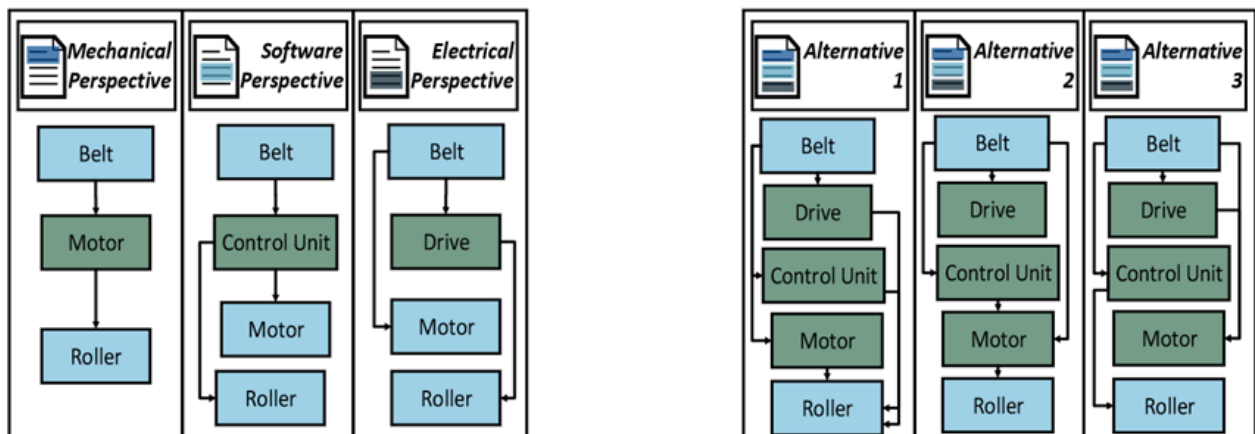
unterschiedlichen Perspektiven nur Teile einer gemeinsamen Ontologie und es ist relativ leicht, die Perspektiven zu einer gemeinsamen Ontologie zusammenzuführen.



Darstellung 16: Beispiel 1 zur Darstellung semantischer Heterogenität

Quelle: Grangel-González (2018, S. 112)

Schwieriger wird es, wenn konzeptionelle und strukturelle Unterschiede wie verschiedene Teilbeziehungen verwendet werden. Wie in Darstellung 17 gezeigt, ist ein Überführen der Perspektiven auf eine gemeinsame Ontologie nicht mehr eindeutig möglich. Stattdessen gibt es drei mögliche Alternativen.



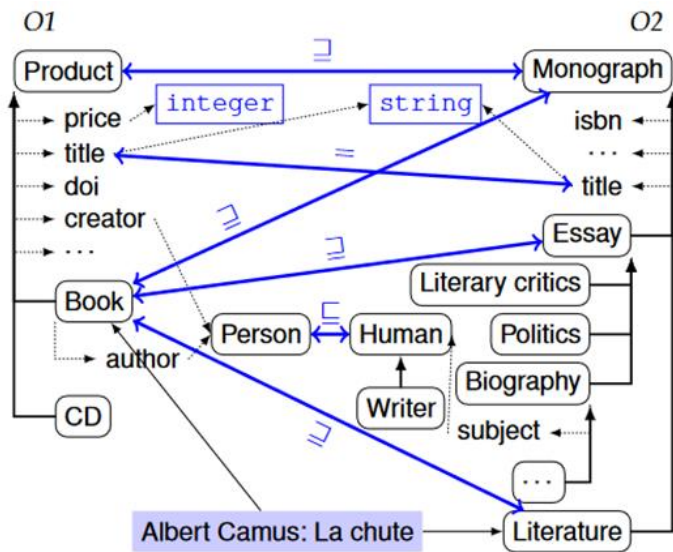
Darstellung 17: Beispiel 2 zur Darstellung semantischer Heterogenität

Quelle: Grangel-González (2018, S. 113)

In Beispiel 3 wird die Komplexität der Unterschiede noch einmal gesteigert (siehe Darstellung 18). In der Ontologie O1 ist „Book“ ein Unterkonzept von „Produkt“ und wird auch nicht weiter separiert. Die Ontologie O2 hat als Konzepte „Monograph“, „Essay“ und „Literature“, welche zumindest in diesem Beispiel durch das Symbol \supseteq als Spezialisierungen⁴⁶ des Konzeptes

⁴⁶ Zur Verdeutlichung der Schreibweise: Aus dem Ausdruck „A ist Spezialisierung von B“, wird in der symbolischen Schreibweise „ $A \supseteq B$ “ oder „ $B \supseteq A$ “.

„Book“ aus O1 gekennzeichnet sind.⁴⁷ Somit unterscheiden sich die beiden Ontologien strukturell, da O2 in dieser Hinsicht granularer ist als O1. Ein Synonym für den in O1 vorkommenden Autor („Author“) findet sich in Ontologie O2 mit dem Konzept „Writer“. Jedoch ist dieser in O1 nur als Attribut von „Book“ beschrieben, in O2 aber als Entität, womit eine weitere konzeptionelle Heterogenität vorliegt.



Darstellung 18: Beispiel 3 zur Darstellung semantischer Heterogenität

Quelle: Euzenat und Shvaiko (2013, S. 2)

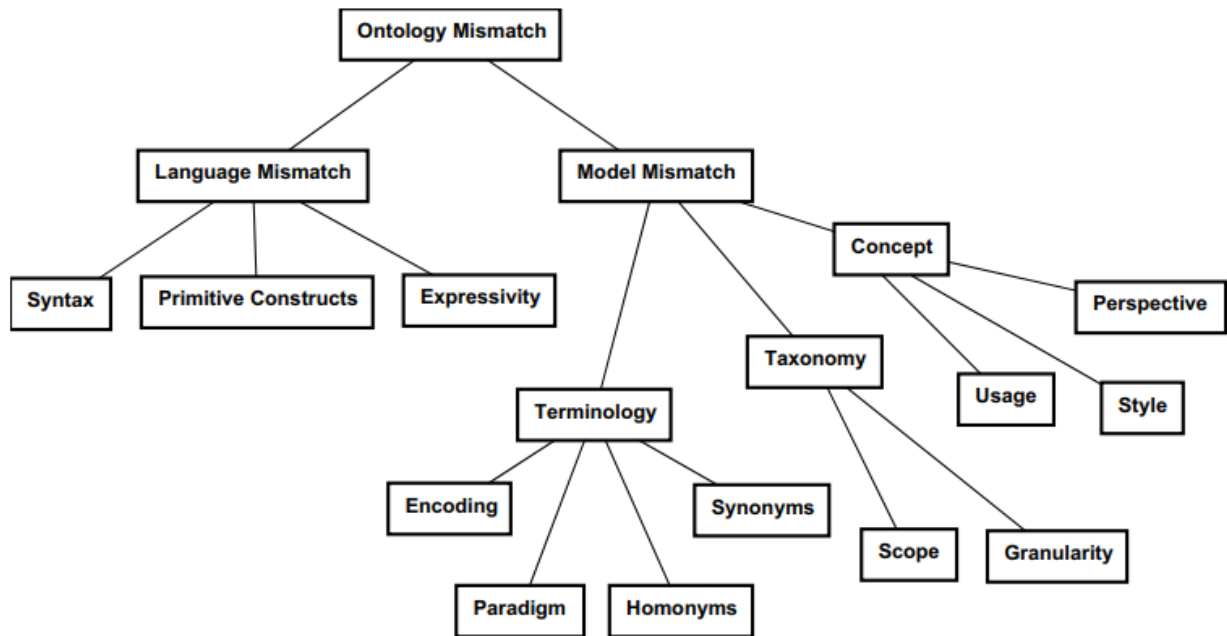
3.1.2 Arten von Heterogenität

Die oben bereits angedeuteten Arten von Heterogenität werden in der Literatur nicht einheitlich klassifiziert.

Leser und Naumann (2007, S. 60–61) gliedern Heterogenität in technische, syntaktische, strukturelle, semantische sowie in eine Datenmodellheterogenität. Euzenat und Shaivko (2013, S. 37–39) unterscheiden wiederum zwischen terminologischer, syntaktischer, konzeptioneller und semiotischer Heterogenität.

Rebstock, Fengel und Paulheim (2008, S. 110) klassifizieren die unterschiedlichen Ursachen für eine Nichtübereinstimmung von Ontologien (Ontology Mismatch) durch weitere Unterkategorien, welche in Abbildung 19 dargestellt sind. Da diese Auflistung am detailliertesten ist, wird sie im Folgenden genauer besprochen. (Rebstock; Fengel; Paulheim 2008, S. 109–112)

⁴⁷ Dass „Literature“ eine Spezialisierung des Konzeptes „Buch“ ist, kann nur in bestimmten Kontexten (z.B. im Kontext einer Buchhandlung) als korrekt angesehen werden. Literatur als schriftlich fixiertes Zeugnis kann auch in anderer Form erscheinen, wie beispielsweise der Stein von Rosetta oder, wenn die poetische bzw. ästhetische Funktion als Bedingung vorausgesetzt wird, in Form von Gedichten in Zeitungen und Zeitschriften. Dies ist ein Problem der sogenannten „double articulation“, welche im nächsten Abschnitt besprochen werden wird.



Darstellung 19: Arten von Heterogenität

Quelle: Rebstock, Fengel und Paulheim (2008, S. 110)

Unter sprachlicher Heterogenität (Language Mismatch) nennen die Autoren drei Kategorien: Die **Syntaktische Heterogenität** („Syntax“), welche beispielsweise auftritt, wenn Ontologien in unterschiedlichen Formaten (XML, OWL, FOL⁴⁸) beschrieben wurden. **Heterogenität auf der Konstruktionsebene** („Primitive Constructs“) entsteht, wenn dieselben Aussagen unterschiedlich ausgedrückt werden. Zum Beispiel kann die Aussage „A und B sind disjunkt“ auch durch „A ist eine Teilmenge von Nicht-B“ beschrieben werden. Und schließlich die **heterogene Ausdrucksmächtigkeit** („Expressivity“) einer Modellierungssprache. Damit ist gemeint, dass bestimmte Ausdrücke nur in einer der zu verbindenden Ontologien möglich sind (siehe Abschnitt 2.2.2).

Terminologische Heterogenität betrifft zum einen unterschiedliche **Kodierungen** von Informationen. Dies umfasst zum Beispiel Maßeinheiten (Grad Celsius vs. Fahrenheit)⁴⁹, Währungen oder zeitliche Angaben. Daneben fallen auch sprachliche Unterscheidungen wie „Buch“ vs. „Book“ in diese Kategorie. Ebenfalls in die terminologische Heterogenität fallen **Synonyme** (für ein Konzept gibt es unterschiedliche Symbole/Ausdrücke, z.B. „Exponent“ und „Hochzahl“) oder **Homonyme** (Ein Symbol/Ausdruck kann unterschiedliche Konzepte bezeichnen, z.B. „Michelangelo“⁵⁰). **Paradigma**, also die Art der Denkweise und Weltanschauung, wird von den Autoren nicht genauer spezifiziert.

⁴⁸ FOL steht für First Order Logic

⁴⁹ Ein eindrückliches, da teures Beispiel für diese Art von Heterogenität liefert der 125 Millionen Dollar teure Absturz des Mars Climate Roboters. („Mars Climate Orbiter: Absturz wegen Leichtsinnsfehler beim Rechnen“ 1999)

⁵⁰ Siehe Seite 28.

Auf Taxonomischer Ebene kann es Unterschiede in der **Granularität** („Granularity“) oder im **Scope** geben. So kann zum Beispiel in der einen Ontologie, welche den Fokus auf die literarische Beschreibung von Werken legt, die Klasse „Werk“ mehrere Unterkonzepte wie „Epik“, „Lyrik“, „Dramatik“ haben, während eine andere Ontologie, welche den Schwerpunkt auf die physischen Merkmale legt, die Klasse „Werk“ in die Unterklassen „Hard Cover“ und „Soft Cover“ unterteilt.⁵¹

Konzeptionelle Heterogenität betrifft meist Unterschiede, die während des Designprozesses der Ontologie entstanden sind und hängen in dieser Hinsicht stark mit taxonomischer Heterogenität zusammen. Unterkategorien sind hierbei der **Gebrauch** („Usage“), der **Stil** („Style“) und die **Perspektive** („Perspective“). Konzeptionelle Heterogenität entsteht beispielsweise, wenn Sachverhalte entweder als Subklassen oder als Eigenschaften modelliert werden. Um beim oben erwähnten Beispiel zu bleiben, kann die erste Ontologie den Fakt, dass es sich um ein Hard oder Soft Cover handelt, mithilfe einer Eigenschaft beschreiben, während es in der zweiten Ontologie mithilfe von Unterkonzepten realisiert wird. Dadurch kommt es zu strukturellen Unterschieden.

Ein besonderer Fall der Heterogenität ist die sogenannte „double articulation“⁵². Diese entsteht, wenn ein Konzept aus zwei Perspektiven betrachtet wird und bei der Verwendung des Konzeptes einschränkende Bedingungen angewendet werden, ohne diese zu kommunizieren. Jarrar und Meersman (2008, S. 10–11) nennen als Beispiel das Konzept „Buch“ und drei unterschiedliche Perspektiven. Zum einen die einer Buchhandlung, für die ein Buch ein „verkaufbares Buch“ (insbesondere mit ISBN) ist. Zum anderen die Perspektive einer Bibliothek, die auch „Masterarbeiten“ als Subkonzept des Konzepts „Buch“ betrachten, was bei der Buchhandlung nicht der Fall ist. Zuletzt erwähnen die Autoren das Beispiel eines Museums, für die „Buch“ gleichzusetzen ist mit „altem Buch“, das in der Regel nicht verkäuflich ist, wodurch die Schnittmenge der Extensionen⁵³ des Konzeptes „Buch“ der Buchhandlung und des Museums leer ist.

Aufgrund der double articulation kann es somit zur Situation kommen, dass bei allen beteiligten Personen Einigkeit über einen Begriff herrscht, jedoch alle etwas anderes darunter verstehen, da sie implizit subjektive Einschränkungen aufgrund ihrer intendierten Anwendung mitbetrachten.

Dies ist nicht nur in der Verwendung von Ontologien problematisch, sondern auch bei deren Erstellung, dem sogenannten ontology engineering. In diesem Kontext sehen Adamovic u.a. die Kernaufgabe von Ontologien nicht darin, die meist eingeschränkten Perspektiven der

⁵¹ Weitere Beispiele für unterschiedliche Granularitäten und Scopes, die bereits in dieser Arbeit erwähnt wurden, sind die unterschiedlichen Klassifizierungen von Heterogenität oder die Ontologien aus Darstellung 18 auf Seite 43.

⁵² Der Begriff scheint nur in englischsprachiger Literatur verwendet zu werden, weshalb der Autor diesen unübersetzt lässt. Der auch in der Linguistik verwendete Begriff „Zweifache Gliederung“, der die Untergliederung in bedeutungstragende (Morpheme) und bedeutungsunterscheidende (Phoneme) Einheiten meint, ist hier nicht passend.

⁵³ Wie in Abschnitt 2.2.1 definiert wurde, meint Extension die Menge aller realen Bücher, die unter dem Konzept „Buch“ zusammengefasst werden.

Anwender (z.B. Ingenieure) in Form eines weiteren Schemas (z.B. Datenbankschema) abzubilden, sondern die allgemeinen Konzepte, auf denen die Schemata fußen.

“In this sense, ontologies should not be used as another data structure but instead as the conceptualisation of domain knowledge.”

(Adamovic u.a. 2022, S. 29)

Die double articulation ist dabei ein wichtiger Faktor, der bei der Verwendung bestehender Ontologien immer mitberücksichtigt werden muss. Insbesondere dann, wenn eine Ontologie in einer Domäne mitverwendet werden soll, für die sie nicht konzipiert wurde.

3.1.3 Heterogenität im Kontext der IIoT

Die oben erwähnten Arten von Heterogenität gelten allgemein für Datenaustauschprozesse und betreffen somit auch den Kontext der IIoT. Grangel-González (2018, S. 26–27) erwähnt hierbei beispielhaft terminologische Heterogenität in etablierten Standards. So verwenden OPC UA und AML die Begriffe „InternalElement“ (AML) und „Object“ (OPC UA) synonym. Beispiele für Homonyme sind in den ISO-Standards 15704 und 10303 zu finden, die beide den Begriff „Ressource“ unterschiedlich definieren. (Grangel-González; Vidal 2021, S. 18–19)

Weitere Heterogenität entsteht dadurch, dass unterschiedliche Referenzarchitekturmodelle ein und denselben Standard in unterschiedlichen Ebenen (Layers) oder Dimensionen verwenden, was zu Homonymen oder taxonomischer Heterogenität führen kann.⁵⁴ In Bezug auf CPS können dabei dieselben Objekte mit unterschiedlichem Detaillierungsgrad modelliert werden. Ein weiteres simples Beispiel, wie Heterogenität im Bereich der IIoT entsteht, ist, wenn unterschiedliche CPS verschiedene Maßeinheiten oder Restriktionen für die Befüllung von Attributen oder Attributwerten verwenden.⁵⁵ Zuletzt besteht auch die Möglichkeit, dass bestimmte Konzepte nur in einem der Standards bzw. CPS vorhanden sind. So verfügte die erste Version von SSN (siehe Abschnitt 2.2.6.2.1) beispielsweise nicht über das Konzept „Aktor“, SAREF oder die neue Version von SSN jedoch sehr wohl.

Nachdem wir die Arten von Heterogenität besprochen haben, beschäftigen wir uns nun damit, wie diese Heterogenität aufgelöst werden kann. Das Beispiel in Darstellung 18 auf Seite 43 gibt bereits die Idee vor: Ziel ist es, Verbindungen zwischen den Objekten der unterschiedlichen Ontologien zu erzeugen, um damit die Ontologien zusammenzuführen und die Heterogenität aufzulösen. Dies kann beispielsweise über Äquivalenzen (Pfeile mit =) oder Spezialisierungs-Relationen (Pfeile mit \sqsubseteq) erfolgen. Die Disziplin, solche Verbindungen und deren Eigenschaften zu entdecken, wird meist mit dem Begriff „Ontologie Matching“ bezeichnet und soll im nächsten Abschnitt genauer besprochen werden.

⁵⁴ Genaueres wird in Abschnitt 4.5 noch besprochen werden.

⁵⁵ Ein konkretes Beispiel wird in Abschnitt 4.2 erwähnt.

3.2 Begriffsbestimmung und formale Definition des Ontologie Matchings

Das Zusammenführen von Ontologien wird mit einer Vielzahl von Begriffen beschrieben. So ist wahlweise von „Matching“, „Mapping“ oder „Alignment“ die Rede.⁵⁶ Im Folgenden wird versucht, die wichtigsten Begriffsvarianten zu beschreiben und auf deren Unterschiede einzugehen, wobei betont werden muss, dass auch diese Unterscheidung laut Osman u.a. (2021, S. 40) in der Literatur nicht immer streng vollzogen wird.⁵⁷

Ontologie Matching ist jener Prozess, in dem Beziehungen zwischen zwei oder mehreren Entitäten aus zwei oder mehreren Ontologien gesucht werden. (Faria u.a. 2019, S. 9)

Ontologie Alignment bezeichnet das Resultat des Matching-Prozesses. Es ist das Set von Mappings zwischen den Ontologien, die während des Matching-Prozesses ermittelt wurden. Ein Mapping ist eine (meist) gerichtete Verbindung von Konzepten einer Ontologie auf Konzepte einer anderen. (Euzenat; Shvaiko 2013, S. 39; Faria u.a. 2019, S. 10) Somit entspricht ein Alignment einem Homomorphismus⁵⁸ zwischen zwei Strukturen. Verwandt mit dem Alignment ist eine Brücken Ontologie (auch Mapping Ontologie oder Articulation genannt), welche die Verbindungen der Entitäten und Relationen zwischen den Ontologien formal abbildet.⁵⁹ (Kalfoglou; Schorlemmer 2003, S. 4; Hodges; García; Ray 2017, S. 30)

Ontologie Merging wiederum beschreibt den Gesamtprozess des Zusammenführens oder Verlinkens der Konzepte zweier Ontologien zu einer neuen kohärenten Ontologie, wohingegen bei **Ontologie Integration** eine Ontologie durch die Einbettung der anderen Ontologien in diese verändert wird. Zuletzt gibt es auch noch die **Ontologie Reconciliation**, welche einen Prozess beschreibt, in dem zwei oder mehrere Ontologien harmonisiert werden. Das bedeutet, dass sie so verändert werden, dass ein Matching ermöglicht bzw. verbessert wird, beide Ontologien aber weiterhin eigenständig bestehen bleiben. (Euzenat; Shvaiko 2013, S. 40; Osman; Ben Yahia; Diallo 2021, S. 44)

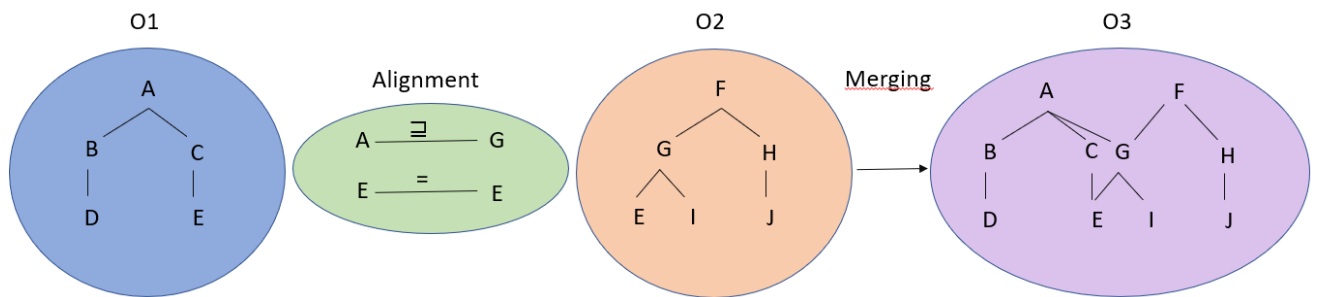
Darstellung 20 soll diese Begriffe noch einmal verdeutlichen. Durch das Ontologie Matching wird aus den Ontologien O1 und O2 ein Alignment ermittelt, das aus zwei Mappings ($A \sqsupseteq G$ und $E = E$) besteht. Aus dem Alignment lässt sich nun mittels Ontologie Merging eine gemeinsame Ontologie O3 erzeugen. Hätte O2 nicht über das Konzept G verfügt, so hätte G mittels Ontologie Reconciliation zuerst zu O2 hinzugefügt werden können, um anschließend das Alignment zu ermitteln.

⁵⁶ Wobei diese Aufzählung keinesfalls vollständig ist.

⁵⁷ Wie erwähnt wurde, enthält dieses Kapitel nur die wichtigsten Begriffe, die für den Inhalt dieser Arbeit von Bedeutung sind. Einen vollständigeren Auflistung samt Beschreibungen liefern Euzenat und Shvaiko (2013).

⁵⁸ Ein Homomorphismus ist eine strukturerhaltende Abbildung. Alignments sind Homomorphismen, da sie die Struktur der Ontologien beibehalten. Dies bedeutet, dass beispielweise zwei Konzepte, die in der einen Ontologie eine Mutter-Tochter-Beziehung haben, auch in der anderen Ontologie eine analoge Beziehung haben, wobei mit „analoger Beziehung“ jene Beziehung gemeint ist, auf die die Mutter-Tochter-Beziehung gemappt wird.

⁵⁹ Ein einfaches Beispiel für ein Konzept aus der Brücken Ontologie wäre ein Oberkonzept, für Konzepte aus den zu matchenden Ontologien. Für ein Konzept „MalerIn“ aus einer Ontologie O1 und ein Konzept „BildhauerIn“ aus Ontologie O2, wäre ein mögliches Oberkonzept der Bridge Ontologie „KünstlerIn“.



Darstellung 20: Begriffsbestimmung Ontologie Matching und verwandter Begriffe

Quelle: In Anlehnung an Osman, Ben Yahia und Diallo (2021, S. 5)

Das Ermitteln des Alignments kann dabei formal als eine Funktion⁶⁰ f gesehen werden, die aus

- einer Menge Ω von Ontologien O_n mit $n \in \mathbb{N}$,
- einem initialem Alignment A , dem sogenannten Anker Alignment, das bereits bekannte Verbindungen der Ontologien beinhaltet, sowie
- Matching-Parametern p , welcher den Matching-Algorithmus beeinflussen (z.B. Gewichte, Schwellenwerte) und
- externen Ressourcen r , wie beispielsweise bekanntes Hintergrundwissen, Top-Level Ontologien oder domänenspezifische Thesauri

ein erweitertes Alignment A' ermittelt. (Euzenat; Shvaiko 2013, S. 41; Osman; Ben Yahia; Diallo 2021, S. 40)

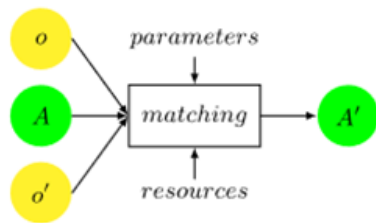
Damit gilt:

$$A' = f(\Omega, A, p, r)$$

Betrachten wir die Matching-Box aus Darstellung 21 genauer, so ist dieses im Allgemeinen ein stark oder vollständig automatisierter Prozess. Er beginnt nach Faria u.a. (2019, S. 12–14) mit dem Einlesen und Aufbereiten der Daten, um diese im Anschluss besser verarbeiten zu können. Diese Aufbereitung kann beispielsweise ein Transformieren von semistrukturierten Daten in strukturierte sein, um anschließend schneller auf diese zugreifen zu können.

Daraufhin erfolgt ein Matching auf lexikalische Äquivalenz – d.h. es wird überprüft, ob idente Begriffe oder Synonyme enthalten sind. Dies ist der erste und meist auch wichtigste Schritt, um Verbindungen zwischen den Ontologien zu ermitteln.

⁶⁰ Erwähnt werden muss, dass ein Alignment selbst im Allgemeinen keine Funktion ist, da ein Konzept beispielsweise auf mehrere andere Konzepte gemappt werden kann, was dem Funktionsbegriff widerspricht. Im Gegensatz dazu kann das Ermitteln des Alignments, wie es im Folgenden beschrieben wird, als Funktion definiert werden.

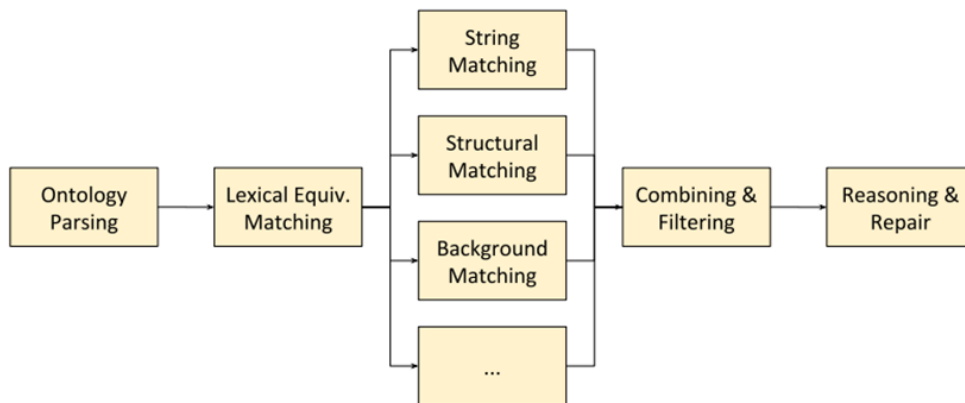


Darstellung 21: Ermittlung eines Alignments

Quelle: Euzenat und Shvaiko (2013, S. 41)

Als nächstes kann eine Fülle von weiteren Matching-Algorithmen angewendet werden, um weitere Verbindungen zu ermitteln (siehe dazu Abschnitt 3.6). Anschließend werden die entdeckten Mappings gesammelt und mithilfe eines Ähnlichkeitscores bewertet. Handelt es sich um ein Ontologie Merging, so kann auch eine Harmonisierung erfolgen, indem beispielsweise zwei idente Konzepte zu einem Konzept vereint werden.

Zuletzt kann noch ein sogenannter Reasoner verwendet werden, der versucht, die logischen Folgen des ermittelten Mappings zu ermitteln, um dadurch den Abgleich besser bewerten zu können. Zusätzlich werden Inkonsistenzen repariert und das erhaltene Mapping mit den Anforderungen der gegebenen Anwendung verglichen.



Darstellung 22: Prozessschritte des Ontologie Matchings

Quelle: Faria u.a. (2019, S. 12)

Jedes Alignment ist ein Set von Mappings. Diese Mappings können als Tupel $\langle c_1, c_2, r, s \rangle$ dargestellt werden, wobei

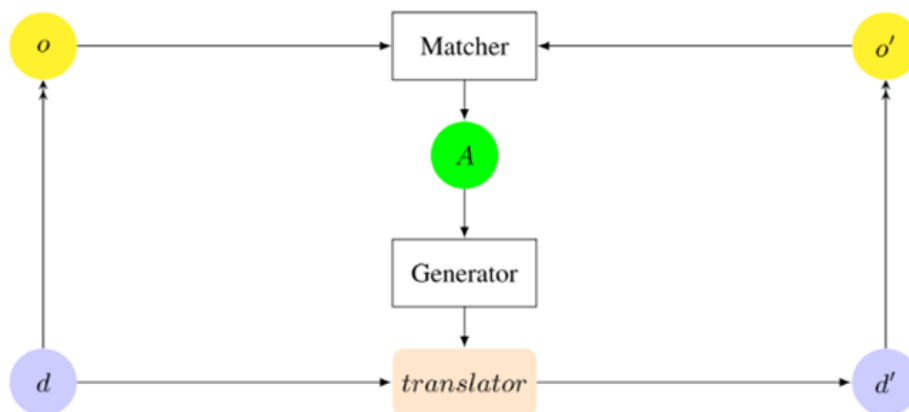
- c_1 und c_2 Entitäten (und damit Konzepte, Instanzen, Eigenschaften, Datentypen, Relationen usw.) sind,
- r eine Mapping-Relation zwischen den Entitäten ist, wie bspw. eine Gleichsetzung oder eine Subklassen-Relation und
- s der Konfidenz-Score des Mappings ist, der zwischen 0 und 1 liegt und damit ein Maß für die Korrektheit der Übereinstimmung angibt. Meist wird dies als Wahrscheinlichkeit interpretiert.

Für den Konfidenz-Score wird in der Regel ein Limit (Threshold) gesetzt, das erreicht werden muss, damit das Mapping nicht verworfen wird.

Das Ziel dieses Abschnittes war es, die grundlegenden Konzepte des Ontologie Matchings zu präsentieren. Bevor wir tiefer in die Methoden des Ontologie Matchings eintauchen, betrachten wir zuerst, wie mithilfe dieses Werkzeuges die semantische Interoperabilität auf Datenebene ermöglicht werden kann.

3.3 Vom Alignment zur erfolgreichen Interoperabilität

Wie im Abschnitt 3.2 beschrieben, ist das Ergebnis des Matching-Prozesses ein Alignment zwischen zwei oder mehreren Ontologien. Dies alleine reicht jedoch noch nicht für eine erfolgreiche Kommunikation zwischen den IIoT Akteuren. Es braucht ein Mapping auf Datenebene.⁶¹ Analog zu einem Wörterbuch, das die Verlinkung zwischen den Begriffen zweier Sprachen angibt, braucht es für eine erfolgreiche Kommunikation zusätzlich eine Übersetzung der Nachrichten. Mit anderen Worten: Aus dem Alignment (Wörterbuch) muss ein Übersetzungsprogramm erstellt werden. Dieses übersetzt eine Nachricht d mithilfe eines Alignments A in eine Nachricht d' , wie in Darstellung 23 abgebildet.

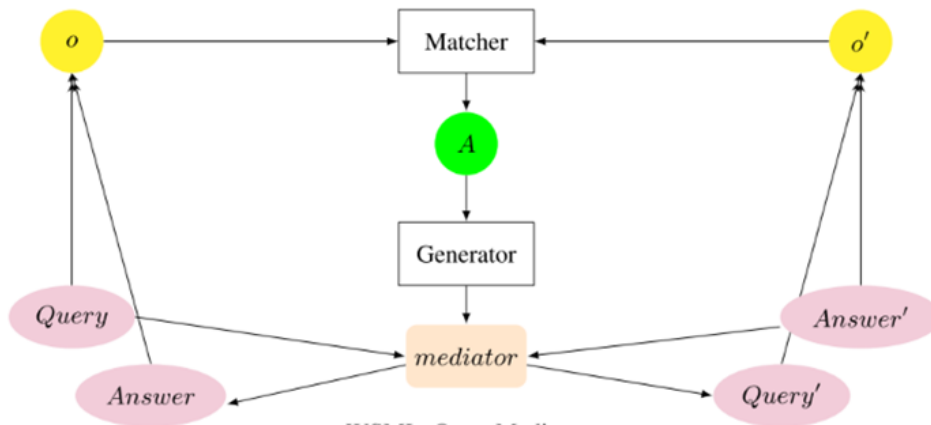


Darstellung 23: Einfache Datenintegration mithilfe des Alignments

Quelle: Euzenat und Shvaiko (2013, S. 385)

Dabei ist zu beachten, dass eine Nachricht nur dann mit Sicherheit korrekt übersetzt werden kann, wenn das Alignment korrekt und vollständig ist. In Darstellung 23 ist zu sehen, dass die Kommunikation nur in eine Richtung erfolgt. Soll eine bidirektionale Kommunikation erfolgen, bei der mitunter auch komplexere Anfragen aus der einen Ontologie durch Antworten aus der anderen Ontologie beantwortet werden sollen, so spricht man von einem Mediator.

⁶¹ Hierbei spricht man auch von semantischer Datenintegration. „Semantic Data Integration is a mechanism which associates different sources of data on the basis of the meaning of data content.“ (Grangel-González u.a. 2018, S. 25)



Darstellung 24: Übersetzungsprozess mittels Mediator

Quelle: Euzenat und Shvaiko (2013, S. 388)

Hierbei steigen jedoch auch die Anforderungen an das Alignment. Während es bei der Übersetzung noch gerichtet (d.h. von o nach o') sein kann, benötigt ein Mediator auch ein Alignment⁶² von o' nach o .

Für die Formulierung der Anfragen kann beispielsweise SPARQL⁶³ verwendet werden. Der Mediator ist dabei ein Stück Software, der die Anfrage übersetzen und die Antwort zurückübersetzen kann.

Nachdem wir nun besprochen haben, wie Ontologie Matching semantische Interoperabilität unterstützt, wollen wir in den folgenden Abschnitten 3.4, 3.5 und 3.6 genauer auf Methoden und Softwaretools zur Ermittlung von Alignments eingehen.

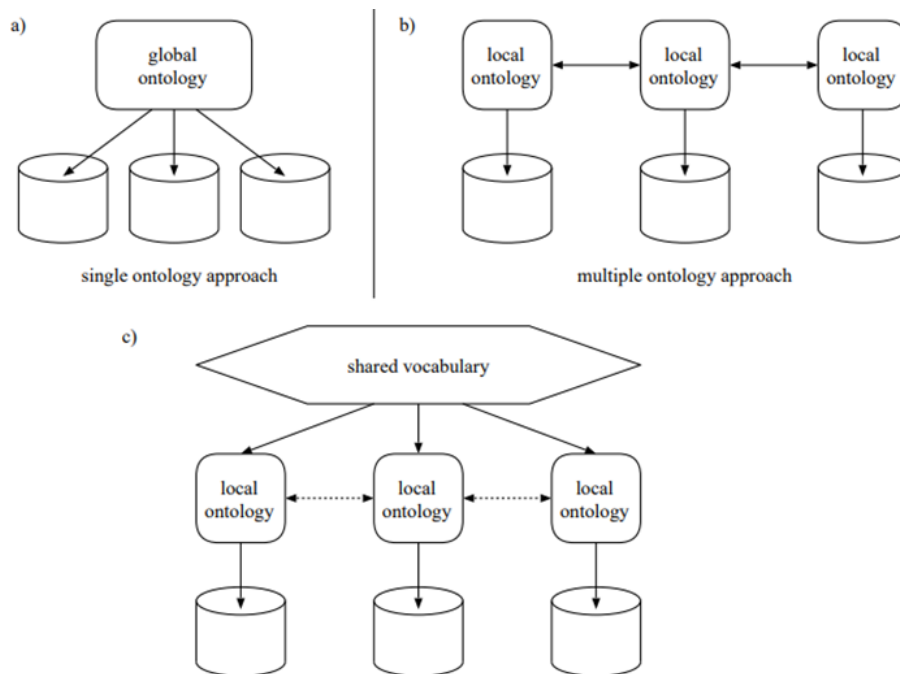
3.4 Alignment Architekturen

Bisher haben wir den Fall betrachtet, dass immer zwei Ontologien direkt miteinander verbunden werden. Im Allgemeinen gibt es aber drei Ansätze, wie mithilfe von Ontologien semantische Interoperabilität ermöglicht werden kann: (Wache u.a. 2001, S. 109–110; Grangel-González u.a. 2018, S. 31)

- a) Globaler Ansatz
- b) Multipler Ansatz
- c) Hybrider Ansatz

⁶² Mathematisch formuliert bedeutet dies, dass ein Isomorphismus zwischen beiden Ontologien benötigt wird.

⁶³ <https://www.w3.org/TR/sparql11-query> (Zugriff am: 08.06.2023)



Darstellung 25: Alignment Architekturen

Quelle: Wache (2001, S. 109)

Der **globale Ansatz** setzt eine globale Ontologie ins Zentrum, mit denen alle Daten aus unterschiedlichen Datenquellen semantisch beschrieben werden können (siehe Darstellung 25). Dies bedeutet, dass die Ontologie alle vorhandenen Konzepte, Relationen und Instanzen beinhalten muss und Änderungen in den Datenstrukturen jedes Mal zu Änderungen in der globalen Ontologie führen können. Dieser Ansatz benötigt kein Ontologie Matching, ist in der Regel aber nur dann möglich, wenn keine semantische Heterogenität vorliegt.

Der **multiple Ansatz** beruht darauf, dass unterschiedliche Bereiche unterschiedliche lokale Ontologien verwenden. Um mit anderen Bereichen zu kommunizieren, benötigt jeder Bereich ein eigenes Alignment zum jeweils anderen Bereich. Dies führt dazu, dass bei Einführung einer weiteren Ontologie mitunter für jede bestehende Ontologie ein Alignment zur neuen erzeugt werden muss.

Der **hybride Ansatz** ist eine Mischung aus den vorigen beiden Methoden. Hierbei wird eine globale Top-Level Ontologie verwendet, die das gemeinsame Informationsmodell auf einer hohen Abstraktionsebene abbildet und zum Matchen mit externen Ontologien dient. Darüber hinaus gibt es lokale, domänenspezifische Ontologien, die entweder über die gemeinsame Top-Level Ontologie gemappt werden können oder über direkte Alignments verfügen, abhängig davon, wie detailliert der Informationsaustausch sein muss. Der Vorteil dieser Architektur ist, dass sie flexibel im Hinblick auf das Hinzufügen neuer Ontologien ist. (Cruz; Parés; Quintela 2021, S. 5–6; Grangel-González 2019, S. 138)

Insbesondere der hybride Ansatz scheint für die Verwendung in IIoT-Plattformen gut geeignet zu sein. Einerseits können die Informationen von den CPS, die jeweils über lokale Ontologien verfügen, an eine IIoT-Plattform übergeben werden, wobei letztere die geteilte Ontologie und damit die wichtigsten Konzepte und Relationen beinhaltet, um die grundlegende

Kommunikation zu externen Systemen gewährleisten zu können. Andererseits erlaubt dieser Ansatz auch den Austausch zwischen den CPS und damit zwischen den lokalen Ontologien. (Hazra u.a. 2023, S. 21, 27) In Abschnitt 4.2 werden Anwendungen aus dem Bereich der IIoT vorgestellt, die einen hybriden Ansatz verwenden.

3.5 Methoden des Ontologie Matching

Ein Ontologie Matching ist im Grunde ein Übersetzungsprozess von einer Sprache (Ontologie) in eine andere. Dementsprechend kommen beim Ontologie Matching ähnliche oder mitunter dieselben Techniken zum Einsatz, wie bei sprachlichen Übersetzungen.

Diese Techniken können in analytische und Embedding-basierte unterteilt werden. Ein einfaches Beispiel für erstere wäre eine Wort-für-Wort-Übersetzung („Mappen“) von einem Begriff in den anderen durch Erkennen von Synonymen. Analytische Regeln betrachten Eigenschaften der Ontologie und versuchen anhand derer Mappings zu erstellen. Embedding-basierte Ansätze verwenden sogenannte Knowledge Graph Embeddings, die Begriffe und Strukturen einer Ontologie auf Zahlenräume – genauer gesagt Vektorräume – abbilden, wodurch der Übersetzungsprozess zu einem mathematischen bzw. statistischen Problem wird. Diese Methode findet auch Anwendung in Natural Language Processing-Modellen wie Generative Pre-trained Transformers (GPT) und Bidirectional Encoder Representations from Transformers (BERT),⁶⁴ die derzeit für einen Hype im Bereich der künstlichen Intelligenz sorgen. (Javed u.a. 2022, S. 6)

Zentral für die Entwicklung und die Beurteilung bestehender und neuer Matching Methoden ist die Ontology Alignment Evaluation Initiative (OAEI)⁶⁵. Diese veranstaltet jährlich Evaluationsevents zur Bewertung und zum Vergleich bestehender Methoden und Tools⁶⁶. Die Vergleiche erfolgen dabei nach unterschiedlichen Kategorien (z.B. nur T-Box Matching oder inklusive A-Box Matching⁶⁷) und unterschiedlichen Ontologien, wobei sich die Unterschiede sowohl auf die Größe der zu matchenden Ontologien als auch auf deren Domänen beziehen. (Jiménez-Ruiz u.a. 2020, S. 1) Die Ergebnisse der Vergleiche werden auf der Homepage der OAEI veröffentlicht.

Zur Bewertung von Alignments werden dabei im Allgemeinen drei Maße herangezogen: Precision, Recall und F-measure. (Ochieng; Kyanda 2018, S. 4–5)

Precision ist das Verhältnis der korrekt ermittelten Mappings zu der Gesamtanzahl der Mappings.

$$Precision = \frac{\text{Anzahl korrekter Mappings}}{\text{Gesamtanzahl der Mappings}}$$

Damit ist die Precision ein Maß für die Korrektheit des ermittelten Alignments.

⁶⁴ Eine gute Einführung in die Funktionsweisen von BERT und GPT liefern Zhou u.a. (2023)

⁶⁵ <https://oaei.ontologymatching.org> (Zugriff am: 27.05.2023)

⁶⁶ Unter Tool wird im Folgenden Software bezeichnet, die dazu dient, Alignments zu erstellen und hierbei meist mehrere Methoden kombiniert.

⁶⁷ Siehe Seite 28.

Der Recall dagegen ist das Verhältnis von korrekten Mappings zu erwarteten Mappings. Damit ist es ein Maß, um die Vollständigkeit der Mappings abzuschätzen.

$$\text{Recall} = \frac{\text{Anzahl korrekte Mappings}}{\text{Anzahl erwarteter Mappings}}$$

F-measure ist ein Maß, das versucht, Precision und Recall zu einer Kennzahl zu verbinden, um unterschiedliche Alignments besser vergleichen zu können.

$$F - \text{measure} = \frac{2 \cdot \text{Recall} \cdot \text{Precision}}{\text{Precision} + \text{Recall}}$$

Derzeit⁶⁸ hat sich weder eine Standardmethode noch ein Standardtool für das Erstellen von Alignments herauskristallisiert. Dies gilt sowohl für die Vergleichstests der OAEI, wo unterschiedliche Methoden je nach Kategorie unterschiedlich gut abschneiden, als auch in der einschlägigen Literatur, wo eine Fülle von unterschiedlichen Ansätzen verwendet wird. (Adamo-
vic u.a. 2022, S. 29)

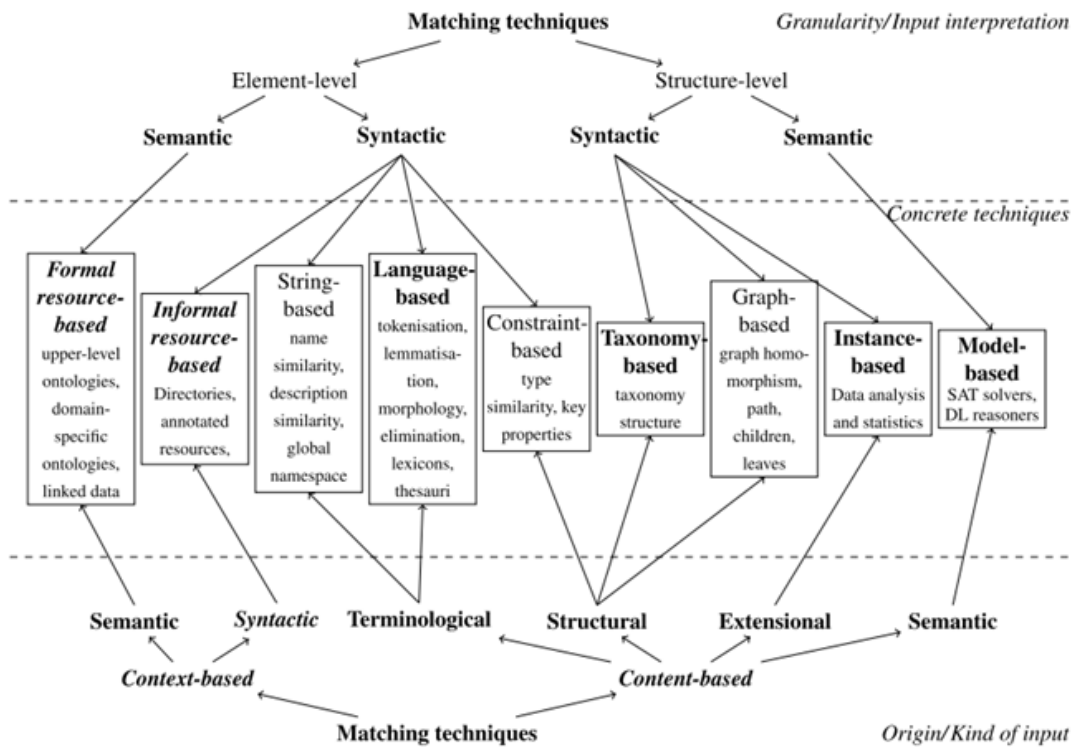
Im Folgenden werden sowohl analytische als auch Embedding-basierte Ansätze besprochen. Dabei liegt der Schwerpunkt auf der Vermittlung der grundlegenden Ideen, die für ein Verständnis für die nachfolgenden Abschnitte relevant sind. Weil die Ermittlung von Alignments in der Regel teil- oder vollautomatisiert von Maschinen durchgeführt wird, werden in Abschnitt 3.6 auch einige Tools exemplarisch vorgestellt.

3.5.1 Analytische Methoden

Euzenat und Shvaiko präsentieren in ihrem 2013 in zweiter Auflage erschienen Werk „Ontology Matching“ auf über 100 Seiten den bis zum heutigen Zeitpunkt umfangreichsten Überblick über unterschiedliche Matching Methoden.

Die Autoren klassifizieren die Methoden dabei nach zwei Dimensionen: nach der Granularität und nach der Art des Inputs. Granularität meint dabei, ob Methoden nur die Entitäten isoliert voneinander betrachten (Element-Level Techniques) oder die strukturelle Information, also die Relationen zwischen den Entitäten, mitberücksichtigen (Structural-Level Techniques). Letztere unterteilen sich wiederum in interne Methoden (content-based), da sie nur den Inhalt der zu vergleichenden Ontologien betrachten, und externe Methoden (context-based), da sie auch externe Ressourcen wie Thesauri oder Top-Level Ontologien zur Ermittlung der Mappings verwenden. (Euzenat; Shvaiko 2013, S. 73-74,76) Beide Klassifizierungsarten spannen einen Baum mit weiteren Unterteilungen auf, wobei sie sich die Blattkonzepte teilen, wie in Darstellung 26 zu sehen ist.

⁶⁸ Stand Juni 2023



Darstellung 26: Klassifizierung analytischer Matching Methoden

Quelle: Euzenat und Shvaiko (2013, S. 77)

Im Folgenden werden diese Blattkonzepte, also die Arten analytische Methoden, kurz vorgestellt.⁶⁹

3.5.1.1 Auf formalen Ressourcen basierte Methoden

Auf formalen Ressourcen basierte Methoden verwenden externe Ontologien⁷⁰, um ein Mapping zu finden. Die externen Ontologien dienen dabei als Mediator, über die eine Verbindung zwischen den zu mappenden Ontologien hergestellt wird. Dabei können sowohl domänen-spezifische als auch Top- und Mid-Level Ontologien verwendet werden. Beispiele wären die Top-Level Ontologien DOLCE und SUMO, aber auch die in Abschnitt 2.2.6.1.2 besprochene DBpedia. Man spricht hierbei auch von der Verwendung von Hintergrundwissen zur Ermittlung von Alignments. Es handelt sich um einen hybriden Ansatz (siehe Abschnitt 3.4).⁷¹

⁶⁹ Wenn nicht explizit angegeben, bezieht sich der Autor hierbei auf Euzenat und Shvaiko (2013, S. 79–82). Aktuellere Beiträge, die inhaltlich jedoch keine neuen Methoden präsentieren, finden sich auch bei Ramar und Gurunathan (2016, S. 678–682)

⁷⁰ Zur Unterscheidung zwischen formalen und informellen Ontologien siehe Abschnitt 2.2.1.

⁷¹ Ein typisches Szenario sähe wie folgt aus: Gegeben seien die Konzepte A (aus Ontologie O1), B (aus Ontologie O2) und C aus einer Mediator-Ontologie O3. Weiß man, dass die Konzepte A und C durch eine Gleichsetzung gemappt werden können als auch die Konzepte B und C, so erhält man durch die Transitivität das Mapping zwischen A und B.

3.5.1.2 Auf informellen Ressourcen basierte Methoden

Im Gegensatz zu der zuvor erwähnten Methode wird hierbei auf informelle Ressourcen zurückgegriffen. Ein Beispiel wäre, wenn zwei Konzepte unterschiedlicher Ontologien mit demselben Bild verbunden sind. Gewöhnlich nutzen diese Methoden Verfahren der Datenanalyse, um aus meist großen Datenmengen Zusammenhänge zu erkennen.

3.5.1.3 Auf Zeichenketten (Strings) basierte Methoden

Bei dieser Art von Methoden werden Zeichenketten einzelner Begriffe miteinander verglichen. Zwei Entitäten gelten dabei als ähnlich, wenn ihre Zeichenketten (deren Namen) ähnlich sind. Hierbei werden Ähnlichkeitsmetriken wie die Jaccard-Distanz, n-gram oder die Livenshtein-Distanz verwendet. Letztere zählt wie viele Ziffern ausgetauscht werden müssen, um von einem Namen auf den anderen zu kommen. Diese Methode eignet sich gut, um idente Konzepte zu identifizieren, die sich in der Schreibweise (z.B. Groß-/Kleinschreibung, mit und ohne Umlaute) unterscheiden. Zur Ermittlung von Synonymen eignet sich diese Methode jedoch nicht. Ebenso versagt sie bei Homonymen.

3.5.1.4 Sprachen-basierte Methoden

Sprachen-basierte Methoden sehen Begriffe als Wörter einer bestimmten natürlichen Sprache an. Um Mappings zu erkennen, wird dabei auf externe Wörterbücher bzw. Thesauri zurückgegriffen,⁷² um beispielsweise Konzepte zu mappen, die in unterschiedlichen Sprachen oder durch Synonyme beschrieben werden. Zusätzlich fallen auch weitere linguistische Verfahren hierunter wie zum Beispiel morphologische Analysen, um Wortstämme zu identifizieren. Ein Beispiel hierfür wäre das Erkennen einer Gemeinsamkeit von „Theorie“ und „theoretischer Beitrag“.

3.5.1.5 Auf Einschränkungen basierende Methoden

Diese Art von Methoden betrachten interne Einschränkungen von Entitäten wie deren Eigenschaften, Wertebereiche oder Datentypen. Ein Mapping könnte zum Beispiel dadurch ermittelt werden, dass die Eigenschaften der betreffenden Konzepte denselben oder einen ähnlichen Wertebereich haben.

3.5.1.6 Graph-basierte Methoden

Bei Graph-basierten Methoden (auch relationale Methoden genannt) werden die Ontologien als Graphen betrachtet (siehe Darstellung 20 auf Seite 48), mit Entitäten als Knoten und Relationen als Kanten. Zwei Entitäten unterschiedlicher Ontologien sind dabei ähnlich, wenn ihre Position in den Graphen ähnlich ist. In weiterer Folge wird zusätzlich angenommen, dass wenn zwei Knoten (Entitäten) ähnlich sind, dann auch deren Umgebung (benachbarte Knoten, Relationen) ähnlich sein muss.

⁷² An diesem Beispiel ist gut zu erkennen, dass die Trennung zwischen den Methoden nicht eindeutig ist. Auch Sprachen-basierte Methoden greifen auf externe formale oder informelle Ressourcen zu. Eine ähnliche Problematik in der Unterscheidung herrscht bei Graph-basierten, Taxonomie-basierten und Instanzen-basierten Methoden.

3.5.1.7 Taxonomie-basierte Methoden

Dies ist ein Spezialfall der Graph-basierten Methoden, bei dem nur die hierarchische Struktur der Konzepte betrachtet wird. Der Ansatz beruht auf der Überlegung, dass Konzepte entlang einer hierarchischen Struktur semantische Ähnlichkeiten aufweisen. Gibt es bereits ein Mapping zwischen zwei Entitäten, so ist es wahrscheinlich, dass es auch ein Mapping für die über- oder untergeordneten Konzepte gibt. Problematisch bei diesem Typ von Methoden sind unterschiedliche Klassifizierungen in den zu mappenden Ontologien.

3.5.1.8 Instanzen-basierte Methoden

Instanzen-basierte Methoden betrachten die Instanzen unterschiedlicher Klassen. Haben diese eine Ähnlichkeit, d.h. sind zum Beispiel viele Instanzen sowohl in der einen als auch in der anderen Klasse, so besteht womöglich ein Mapping zwischen den Klassen. Auch hier kommen meist Datenanalyse- und mathematisch-statistische Verfahren wie die Formale Begriffsanalyse – und hier beispielsweise die Galoisverbindungen – zur Anwendung. (Euzenat; Shvaiko 2013, S. 114)

3.5.1.9 Modell-basierte oder semantische Methoden

Bei dieser Methode wird ausgehend von einem bereits ermittelten Mapping (einem sogenannten Anker) mittels deduktiven Schließens auf weitere Mappings geschlossen. Die Idee dabei ist, dass, wenn zwei Entitäten gleich sind, dann teilen sie sich auch dieselbe Interpretation. Dabei macht man sich zu Nutze, dass Ontologien – wie in Abschnitt 2.2.2 beschrieben – in formalen Sprachen beschrieben werden, die maschinenlesbare, logische Aussagen ermöglichen. Unter diese Art von Methoden fällt auch das Erkennen von Inkonsistenzen bzw. das Reparieren von Ontologien – gemeint ist damit das Erweitern von Entitäten bzw. anpassen von Mappings, um Inkonsistenzen aufzulösen.

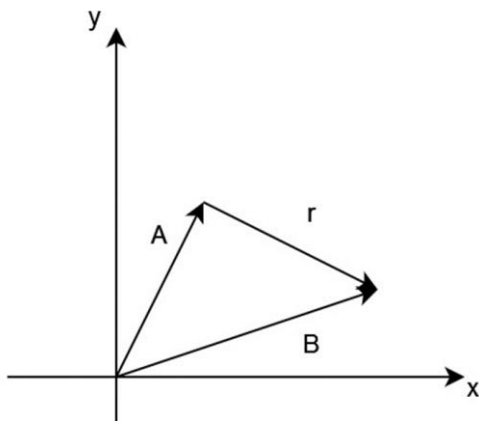
Neben diesen Arten von Methoden richten sich neuere Beiträge zur Ermittlung von Alignments vor allem auf Laufzeitprobleme beim Matching von großen Ontologien. Ochieng und Kyanda (2018) als auch Fensel u.a. (2020, S. 80) erwähnen Methoden zur Partitionierung bzw. Modularisierung von Ontologien, um dadurch die Anzahl der potenziellen Mappings zu verkleinern und somit Rechenaufwand und Rechenzeit zu reduzieren.

3.5.2 Ontologie and Wissensgraph Embeddings

3.5.2.1 Grundlegende Konzepte und Anwendungen

Bisher haben wir Ontologie Matching so betrachtet, als wären es direkte Verbindungen von Konzepten einer Ontologie auf Konzepte einer anderen Ontologie. Einen anderen Ansatz verfolgen sogenannte Wissensgraph Embeddings.

Unter Wissensgraph Embeddings (auch Knowledge Graph Embeddings (KGE)) versteht man strukturerhaltende Abbildungen (Homomorphismen) von Ontologien oder Wissensgraphen in den Vektorraum \mathbb{R}^n , wobei n die Dimension des Vektorraums angibt und meist größer als 100 ist. (Nilsson; Sandin; Delsing 2019, S. 4) Darstellung 27 zeigt exemplarisch zwei Konzepte (A und B) und eine Relation (r), in einem zweidimensionalen Vektorraum. Die Idee dabei ist, dass Konzepte und Instanzen, die semantisch zusammengehören eine ähnliche Repräsentanz im Vektorraum haben.⁷³



Darstellung 27: Visualisierung der Konzepte (A,B) und Relation (r) als Vektoren im \mathbb{R}^2

Quelle: Autor

Die Überführung in einen Vektorraum macht den Weg frei, eine Fülle mathematischer Verfahren, wie jene aus der Vektorrechnung und dem Bereich des Machine Learning, zu verwenden, um neue Erkenntnisse zu gewinnen (i.e. „berechnen“). (Grangel-González 2019, S. 151) Wenn Konzepte und Relationen als Vektoren betrachtet werden, so kann beispielsweise die Behauptung „Peter hat die Tochter Anna“ als Vektoraddition betrachtet werden:

$$\overrightarrow{Peter} + \overrightarrow{hat_Tochter} = \overrightarrow{Anna}$$

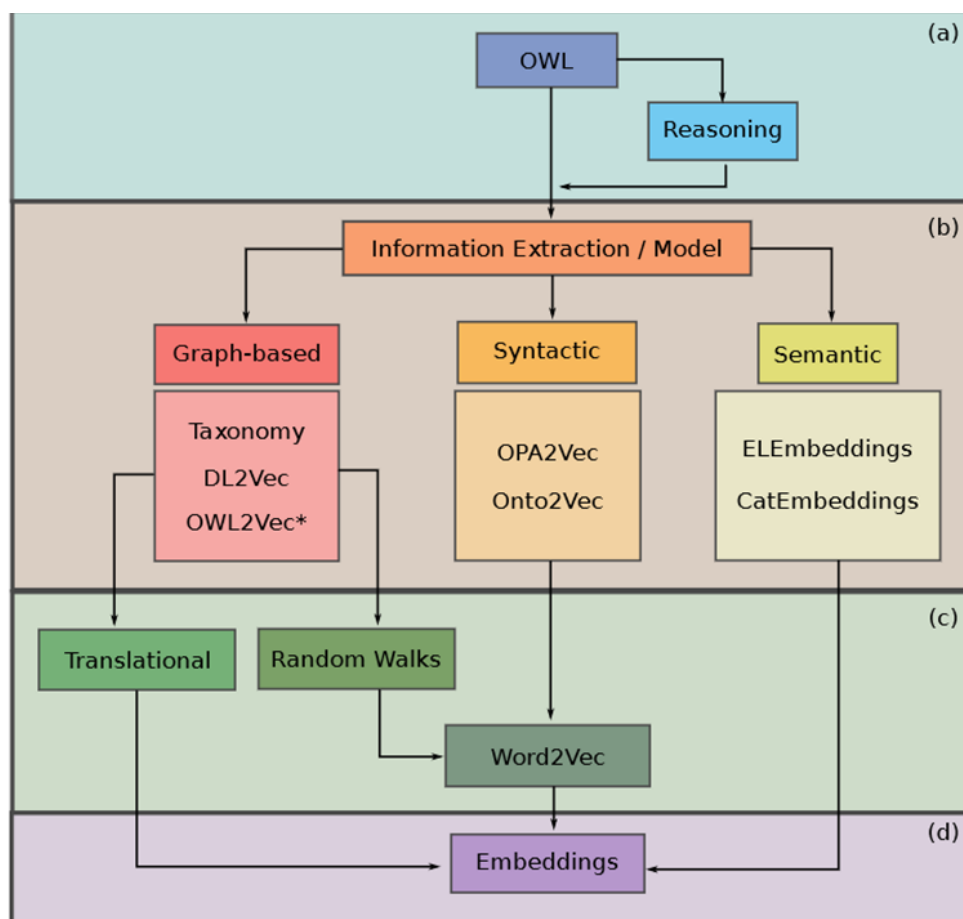
Beziehungweise, falls diese Information nicht explizit in der Ontologie festgehalten wurde, kann neue Information errechnet werden, in dem die Differenz gebildet wird

$$\|\overrightarrow{Peter} + \overrightarrow{hat_Tochter} - \overrightarrow{Anna}\|$$

⁷³ Anders formuliert: Sind zwei Begriffe semantisch nah beieinander, sind auch die entsprechenden Punkte (Ortsvektoren) im Vektorraum nahe beieinander.

und das Ergebnisses als Maß für die Korrektheit dieser Aussage interpretiert wird.⁷⁴ Die Methoden, solche Embeddings durchzuführen, können in drei Kategorien unterteilt werden: (Kulmanov u.a. 2021, S. 7–8)

- Graphen-basierte, bei denen Ontologien als Wissensgraphen interpretiert werden,
- syntaktische, welche Axiome in Ontologien wie Sätze betrachten und versuchen, die syntaktischen Regeln wie zum Beispiel die Häufigkeit des gemeinsamen Auftretens zweier Konzepte im Embedding zu erhalten und
- semantische oder Modell-theoretische, welche die semantische Beziehung direkt erhalten wollen – so wird beispielsweise in der EL-Embedding-Methode die Beziehung „A ist eine Spezialisierung von B“ in den \mathbb{R}^n überführt, sodass A als n-dimensionale Kugel in der n-dimensionalen Kugel von B enthalten ist.



Darstellung 28: Methoden zur Ermittlung von Embeddings

Quelle: Zhapa-Camacho (2022, S. 46)

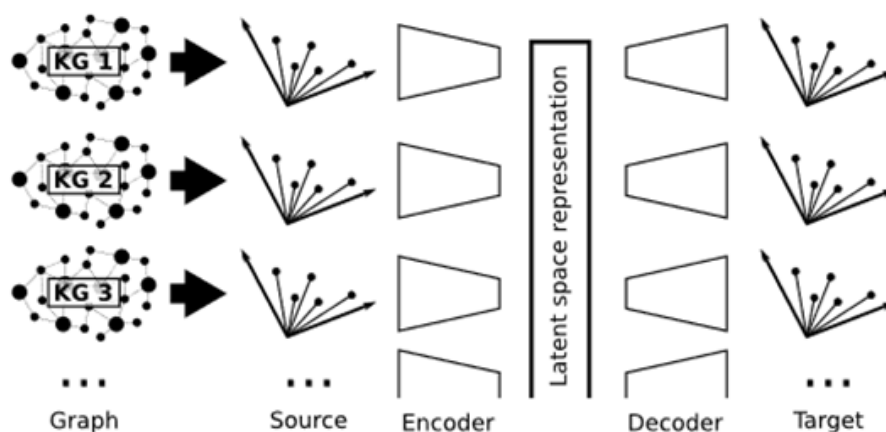
⁷⁴ Je näher das Ergebnis bei 0 ist, desto größer ist die Wahrscheinlichkeit, dass die Aussage korrekt ist. Der Fachbegriff, neue Erkenntnisse aus Ontologien oder Wissensgraphen zu ermitteln, heißt „Link Prediction“. In weiterer Folge ist es auch möglich, mithilfe von Embeddings Ontologien bzw. Wissensgraphen zu vervollständigen („Knowledge Graph Completion“). Die Herausforderung in diesem Bereich ist das Erstellen guter Embeddings. D.h.: auf welche Vektoren sollen die Entitäten abgebildet werden? Hierfür gibt es unterschiedliche Methoden. Eine erste Einführung liefern Lin u.a. (2018), Baumgartner u.a. (2023) und – stark verkürzt – Kamakoti (2020) sowie Liang (2020)

Die Herausforderung all dieser Methoden liegt darin, die Strukturen vollständig zu erhalten. So nennt Zhapa-Camacho als größten Nachteil der Graphen-basierten Methode, dass beim Mapping in der Regel nicht alle Knoten und Kanten erhalten bleiben und dadurch Information verloren geht. (Zhapa-Camacho 2022, S. 15) Darstellung 28 gibt einen Überblick über Methoden zur Ermittlung von Embeddings aus Ontologien.

Konkrete Anwendungen solcher Embeddings sind das Vervollständigen von Wissensgraphen (Knowledge Graph Completion), durch das fehlende Beziehungen ermittelt werden, und die Relationsextraktion, bei der aus Texten Konzepte und Relationen identifiziert werden. Abstraktere Anwendungen, die auf diesen Techniken aufbauen, finden beispielsweise im Natural Language Processing statt, bei dem die Bedeutung eines Satzes ermittelt wird, in dem der Kontext der im Satz verwendeten Begriffe (sprich: Positionen und Umgebungen der Knoten in bekannten Wissensgraphen) bestimmt wird. Technische Umsetzungen sind hierbei u.a. das syntaktische Word2Vec-Modell (Mikolov u.a. 2013), das Sprachmodell BERT (Devlin u.a. 2019) und schließlich auch ChatGPT sowie GPT-4 (Zhou u.a. 2023).

3.5.2.2 Embeddings und Interoperabilität

Wenn nun das Überführen von Ontologien in Vektorräume Vorteile bringt, da es erlaubt, neue Werkzeuge aus dem Bereich der Mathematik im Allgemeinen und des Machine Learning im Speziellen zu verwenden, so liegt es auf der Hand, auch Alignments zwischen den Vektorräumen zu betrachten. Mit anderen Worten: Anstelle des Verknüpfens von Ontologien wird versucht, ihre Embeddings zu matchen. Hierfür gibt es wiederum verschiedene Methoden, die entweder auf Distanzmessungen – falls die Vektorräume dieselben sind – auf Matrix-Projektionen zwischen den Vektorräumen oder auf ein Überführen in einen gemeinsamen Zwischenraum (sogenannter „latent space“) beruhen. (Baumgartner u.a. 2023, S. 4)



Darstellung 29: Beispiel eines Prozesses zur Ermittlung von Alignments mithilfe von Embeddings

Quelle: Baumgartner (2023, S. 2)

Darstellung 29 verdeutlicht die Idee, Wissensgraphen (oder eine Ontologie) zuerst in Vektorräume (Source) zu transformieren und diese dann über den „Latent space“ zu matchen, damit dieselben Entitäten auch dieselbe Repräsentation im resultierenden Vektorraum (Target) haben.

Anwendungen von Embeddings und Embedding-Matching im Bereich IIoT werden in Abschnitt 4.4 vorgestellt. Vertiefende Informationen zum Thema Knowledge Graph Embeddings finden sich neben den bereits genannten Quellen bei Cai u.a. (2018), Jiménez-Ruiz u.a. (2020) und Choudhary u.a. (2021).

3.6 Gegenwärtige Softwaretools und Nachschlagverzeichnisse

Das Erstellen, Pflegen, Validieren und Matchen von Ontologien sind heutzutage Tätigkeiten, die stark durch Softwaretools unterstützt werden. Aufgrund der Fülle an solchen Tools verzichtet der Autor auf eine möglichst vollständige Darstellung und verweist stattdessen auf die Arbeiten von Skjæveland, Slaughter und Kindermann (2022), Grangel- González (2018, S. 31–32) sowie Ochieng und Kyanda (2018). Stattdessen liegt der Fokus auf der Vorstellung einiger weniger Matching-Tools, die in den in Kapitel 4 beschriebenen Anwendungsfällen eingesetzt werden.

3.6.1 LOOM

Der Lexical OWL Ontology Matcher (LOOM) ist ein auf formalen Ressourcen basierter Matching-Algorithmus, der nur die Entitäten, nicht aber die Struktur der Ontologien betrachtet. Er vergleicht die Konzepte zweier Ontologien und mappt diese, falls die Begriffe bzw. deren Synonyme ident sind. Synonyme werden dabei mithilfe externer Thesauri ermittelt (z.B. WordNet⁷⁵ oder GermaNet⁷⁶, siehe Abschnitt 3.5.1.1). Dieser Methode bedienen sich auch die weiter unten beschriebenen Tools. Zusätzlich verfügt LOOM über eine Vergleichsfunktion auf Basis der Zeichenkette, wodurch es auch Mappings erkennt, wenn kleinere syntaktische Unterschiede (z.B. Leerzeichen, Unterstriche, Klammern, Groß-Klein-Schreibung) existieren. (Ghazvinian; Noy; Musen 2009, S. 199) LOOM erzeugt dadurch nur Äquivalenzen zwischen Konzepten, kann jedoch keine komplexeren Relationen (z.B. Subklassen) ermitteln.

3.6.2 LogMap

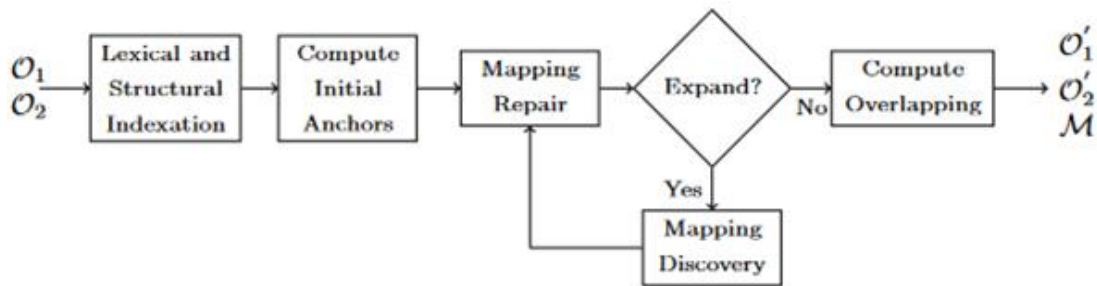
LogMap ist ein Matching System, das neben der Erstellung von Mappings auch Reasoner wie Hermit (siehe 3.6.5) integriert und während des Mappings direkt Inkonsistenzen repariert.

Neben dem lexikalischen Vergleich, der er auch bei LOOM angewandt wird, verfügt LogMap auch über einen strukturellen Vergleich (siehe Abschnitt 3.5.1.6). Dabei geht LogMap iterativ vor. Findet es beispielsweise ein Mapping zwischen zwei Konzepten, so benutzt es dieses als

⁷⁵ <https://wordnet.princeton.edu> (Zugriff am: 26.05.2023)

⁷⁶ <https://uni-tuebingen.de/en/142806> (Zugriff am: 26.05.2023)

Anker und betrachtet die benachbarten Konzepte genauer, da hier mit einer höheren Wahrscheinlichkeit ebenfalls ein Mapping existiert.



Darstellung 30: Vorgehensweise des LogMap-Algorithmus

Quelle: Jiménez-Ruiz und Cuenca Grau (2011, S. 275)

In einem weiteren Schritt wird mittels eines Reasoners überprüft, ob die ermittelten Mappings logische Inkonsistenzen erzeugen, wodurch sie gegebenenfalls wieder fallengelassen werden. (Jiménez-Ruiz; Cuenca Grau 2011, S. 275–276) Darstellung 30 stellt den Prozess graphisch dar.

Die aktuelle Version von LogMap⁷⁷ erlaubt es auch, Mappings mithilfe von Embeddings zu ermitteln. (Chen u.a. 2021, S. 6–8) Damit verbindet LogMap sowohl analytische als auch Embedding-basierte Methoden.

Weitere Softwaretools, die sich ebenfalls lexikalischer als auch struktureller Matcher bedienen, sind AgreementMakerLight (AML)⁷⁸ und BERTMap⁷⁹. Letzteres ist ein auf dem Sprachmodell BERT, und damit auf Embeddings basierendes Tool. (He u.a. 2022, S. 1)

3.6.3 ALIN

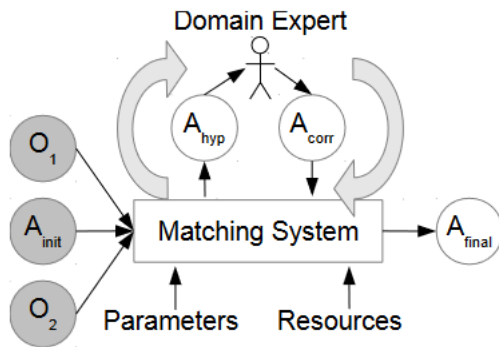
Matching Tools ermitteln in der Regel Mappings, die sie am Ende des Berechnungsprozesses Domänenexperten präsentieren. Diese können die Mappings akzeptieren oder verwerfen, das Expertenfeedback wird jedoch nicht weiter vom Tool verwendet. Mit anderen Worten: Das Expertenwissen hat keinen Einfluss auf das Ermitteln neuer, besserer Mappings. Das Tool ALIN bindet nun dieses Expertenwissen in den Prozess mit ein. Hierfür ermittelt es zuerst potenzielle Mappings. Anschließend präsentiert es diese den Expertinnen und Experten, die diese annehmen oder ablehnen können (siehe Darstellung 31). Die Idee dabei ist, dass immer nur ein Teil der Mappings präsentiert wird, da das Feedback ausreicht, um auf die Gültigkeit anderer Mappings zu schließen. Auf Basis des Feedbacks überarbeitet ALIN die Mappings und präsentiert diese erneut den Expertinnen und Experten, solange, bis alle potenziellen Mappings akzeptiert oder abgelehnt wurden. (Da Silva u.a. 2020, S. 3–5) Ein weiteres Tool, das User-Feedback mitberücksichtigt ist COMA++⁸⁰.

⁷⁷ <https://github.com/ernestojimenezruiz/logmap-matcher> (Zugriff am: 20.05.2023)

⁷⁸ <https://github.com/AgreementMakerLight/AML-Project> (Zugriff am: 26.05.2023)

⁷⁹ <https://github.com/KRR-Oxford/BERTMap> (Zugriff am: 26.05.2023)

⁸⁰ <https://dbs.uni-leipzig.de/Research/coma.html> (Zugriff am: 30.05.2023)



Darstellung 31: Vorgehensweise von ALIN

Quelle: Paulheim, Hertling und Ritze (Paulheim; Hertling; Ritze 2013, S. 36)

Sowohl LogMap als auch ALIN sind in den jährlich stattfindenden Benchmarktests der OAEI regelmäßig unter den besten Algorithmen zu finden. Darstellung 32 zeigt einen Ausschnitt der Ergebnisse aus dem Jahr 2022.

	Precision	Recall	F-measure
LogMap	0.76	0.56	0.64
GraphMatcher	0.75	0.55	0.63
SEBMatcher	0.79	0.48	0.6
ATMatcher	0.69	0.51	0.59
ALIN	0.82	0.44	0.57
edna	0.74	0.45	0.56

Darstellung 32: Ergebnisse des Conference tracks 2022 der OAEI

Quelle: <http://oaei.ontologymatching.org/2022/results/conference> (Zugriff am: 08.06.2023)

3.6.4 Evaluationstools (OOPS!)

Ontologien und in weiterer Folge integrierte Ontologien müssen verständlich, lesbar und konsistent sein. Um diese Eigenschaften zu überprüfen, gibt es Tools, wie den Ontology Pitfall Scanner (OOPS!) (Poveda-Villalón; Gómez-Pérez; Suárez-Figueroa 2014). OOPS! testet dabei nicht nur auf die Vollständigkeit von Metadaten und Inkonsistenzen in Namenskonventionen, sondern auch weitere Kriterien wie Anomalien in Relationen.⁸¹ (Bajaj u.a. 2017, S. 15–16)

3.6.5 Reasoner (Hermit)

Reasoner sind Tools, die Ontologien auf logische Konsistenz prüfen und implizite Relationen zwischen Konzepten ermitteln. Ein Beispiel für ein solches Tool ist der Hermit OWL Reasoner⁸², der in Matchingtools wie LogMap aber auch in Ontologie Editoren wie Protégé⁸³

⁸¹ Ein Überblick über weitere Evaluationstools kann unter <https://www.w3.org/2001/sw/wiki/SWValidators> (Zugriff am: 08.06.2023) gefunden werden.

⁸² <http://www.hermit-reasoner.com> (Zugriff am: 20.05.2023)

⁸³ <https://protege.stanford.edu> (Zugriff am: 08.06.2023)

integriert werden kann. Reasoner sind dabei insbesondere auch für analytische Methoden zum Ermitteln von Mappings (siehe Abschnitt 3.5.1) relevant.

3.7 Aktuelle Ansätze und offene Probleme des Ontologie Matchings

Osman u.a. (2021, S. 59) erwähnen mehrere Herausforderungen und offene Fragen im Bereich des Ontologie Matchings. Zum einen wurde in der bisherigen Fachliteratur das Prinzip der Konservativität weniger betrachtet als jenes der Kohärenz zwischen den Ontologien. Damit gemeint ist, dass für das Ermitteln eines möglichst vollständigen Alignments eher in Kauf genommen wird, die ursprünglichen Ontologien zu erweitern, als diese zu belassen, wie sie sind.⁸⁴ Die Konsequenz ist, dass dadurch mitunter Inkonsistenzen in der ursprünglichen und in der integrierten Ontologie entstehen. Eine automatische Reparatur von Ontologien, die diese Inkohärenzen beseitigt, steht derzeit noch nicht im Fokus der Forschung. Als weiteren noch wenig beachteten Punkt sehen Osman u.a. das Matching von mehr als zwei Ontologien. Der Grund dafür liegt nach Ansicht der Autorin und der Autoren an der erhöhten Komplexität, die dadurch entsteht.

Weiters sehen sie zusätzlichen Forschungsbedarf in der Ermittlung komplexerer Relationen. Der Schwerpunkt liege derzeit hauptsächlich auf der Ermittlung von Äquivalenzen zwischen Konzepten unterschiedlicher Ontologien, wohingegen taxonomische Relationen (wie Subkategorien) oder Disjunktionen (Unvereinbarkeiten) noch von wenigen Tools betrachtet werden. Eine weitere Einschränkung liegt darin, dass die bisherigen Herangehensweisen nur einzelne Domänen betrachtet haben und nicht ohne Weiteres verallgemeinerbar sind.

Konsequenterweise sind auch die Softwaretools, die daraus entstanden sind, nicht für das Matching beliebiger Ontologien geeignet. Dies zeigt sich bis zum heutigen Tag in den Ergebnissen der OAEI, in denen bestimmte Tools jeweils nur in einzelnen Bereichen überdurchschnittliche Ergebnisse erzielen.

Den letzten Punkt, den Osman u.a., aber auch Burns, Cosgrove und Doyle (2019, S. 652) erwähnen, ist der, dass die Skalierbarkeit bei der Verbindung von großen Ontologien meist noch ein Problem ist und hierfür noch wenige Lösungsansätze bestehen. Neben der Größe der Ontologien sind hier aber auch die Anzahl der beteiligten CPS und der von ihnen genutzten Ontologien weitere kritische Faktoren. So ist bis zum heutigen Tag nicht klar, wie die bestehenden Matching Methoden in einer digitalen Fabrik mit tausenden CPS performant umgesetzt werden können.

Euzenat und Shvaiko (2013, S. 1–2) erwähnten bereits vor über 10 Jahren acht Herausforderungen im Bereich des Ontologie Matchings: 1) Matching von großen Ontologien und die Evaluation der Ergebnisse, 2) Verbesserung der Effizienz der Matching Techniken, 3) Einbeziehung von Hintergrundwissen bzw. 4) Einbeziehung von Usern in die Ermittlung von Alignments, 5) Auswahl, Kombination und Tuning von Matching Algorithmen, 6) Darstellung und Erklärung der Matching Ergebnisse, 7) Kollaboratives Matching sowie 8) das Management

⁸⁴ Eine detailliertere Erklärung des Konservativitätsprinzips findet sich u.a. bei Solimando u.a. (2014)

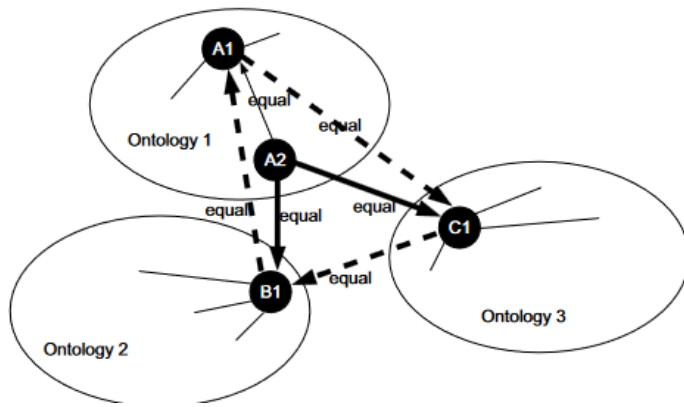
von Alignments, worunter die Autoren das Bereitstellen von Infrastruktur und Support zur Ermittlung sowie zum Abrufen bekannter Alignments verstehen. Während im Bereich der Naturwissenschaften einige dieser Herausforderungen gelöst wurden (beispielsweise durch Alignmentdatenbanken⁸⁵) wird in Abschnitt 4 und 5 gezeigt werden, dass die von Euzenat und Shvaiko vorgebrachten Herausforderungen im Bereich der IIoT nach wie vor ungelöst sind.

3.8 Allgemeiner Nutzen des Ontologie Matchings

Zum Abschluss dieses Kapitels können neben der in dieser Arbeit als Motivation erwähnten Möglichkeit, mithilfe des Ontologie Matchings semantische Interoperabilität heterogener Domänen zu ermöglichen, noch weitere Vorteile aufgezählt werden, die durch das Matching entstehen: (Lambrix; Tan 2006, S. 197; Osman; Ben Yahia; Diallo 2021, S. 38–39)

- Unternehmen können öffentliche Ontologien (wie den in Abschnitt 2.2.6 erwähnten) mit unternehmensspezifischen verbinden, was den Aufwand zur Erstellung einer unternehmensweiten Ontologie verringert. Zusätzlich können Unternehmen so schneller auf neue Anforderungen reagieren.
- Miteinander kooperierende Unternehmen oder Unternehmensbereiche können ihr Wissen in Form von Ontologien teilen und, durch die Verbindung der Ontologien, neues Wissen erzeugen.
- Domänenspezifische Anwendungen, für die es unterschiedliche Ontologien gibt, um diese Domäne aus unterschiedlichen Perspektiven zu beschreiben, können zusammengeführt werden, wodurch eine ganzheitliche Anwendung erzeugt wird.
- Ontologien können in weiterer Folge auch für das Lösen von Problemen verwendet werden. Damit ist nicht nur gemeint, Ontologien als Nachschlagwerke zu verwenden, sondern auch, dass Problemlösewissen abgebildet werden kann oder Antworten automatisiert erzeugt werden können. Ein Beispiel für ersteres beschreiben Ansari u.a. (2018).
- Darüber hinaus können auch die Ontologie Alignments relevantes Wissen beinhalten. So können die darin abgebildeten Beziehungen selbst analysiert und evaluiert werden, wodurch neue Zusammenhänge oder fehlerhafte Daten identifiziert werden. Ein Beispiel hierfür zeigt Darstellung 33. Aus der (durch das Alignment) bekannten Relation „A2 zu B1“ und „A2 zu C1“ folgt die Relation „B1 zu C1“.

⁸⁵ Ein Beispiel ist der Ontology Xref Service, der Auskunft über Mappings zwischen Konzepten unterschiedlicher Ontologien aus dem Bereich der Naturwissenschaften gibt. („EMBL-EBI Ontology Xref Service (OxO)“ 2023) Auch die Gene Ontologie listet Mappings zu weiteren Ontologien (<http://geneontology.org/docs/download-mappings>) (Zugriff am: 15.05.2023))



Darstellung 33: Wissen aus Alignments ziehen

Quelle: Rebstock, Fengel und Paulheim (2008, S. 127)

4 Anwendungsfälle von Ontologie Matching im Bereich der IIoT

Die bisherigen Kapitel haben versucht, die Grundlagen zu legen, um zu verstehen, warum und wie Ontologien und Ontologie Matching bei der Lösung des Problems der semantischen Interoperabilität helfen können. In diesem Kapitel sollen nun Forschungsergebnisse präsentiert werden, die sich mit deren Anwendung im Kontext der IIoT beschäftigen.

Der Aufbau ist dabei wie folgt: In Abschnitt 4.1 wird anhand eines einleitenden praktischen Beispiels der Nutzen von Ontologien und des Ontologie Matchings verdeutlicht. Abschnitt 4.2 beschäftigt sich mit Ansätzen, die IIoT-Plattformen zur Überbrückung des Interoperabilitätsproblems vorsehen. Anschließend werden in Abschnitt 4.3 weitere Beispiele gezeigt, die Ontologie Matching im Kontext der IIoT verwenden, ohne jedoch dezidiert in eine Plattform eingebunden zu sein. Abschnitt 4.4 wiederum beschäftigt sich mit Anwendungen des Wissensgraph Embeddings zur Ermittlung von Alignments. Zuletzt wird in Abschnitt 4.5 noch der I40 Wissensgraph vorgestellt, der ein Alignment von IIoT-Standards darstellt.

Der Fokus liegt dabei auf der Präsentation unterschiedlicher aktueller Lösungsansätze sowie den Herausforderungen und Einschränkungen, mit denen diese behaftet sind. Die gezeigten Umsetzungen erheben keinen Anspruch auf Vollständigkeit – insbesondere wurden nur Arbeiten berücksichtigt, die ab 2017 erschienen sind – noch handelt es sich hier um eine systematische Analyse der vorhandenen Literatur.⁸⁶ Das Ziel des Autors ist es vielmehr, für jeden Abschnitt anhand einiger ausgewählter repräsentativer Beispiele die grundlegenden Lösungsansätze sowie aktuelle Trends zu präsentieren. Die Auswahl der Beispiele erfolgte dabei nach folgenden Kriterien:

- Verwendung von mehreren IoT bzw. IIoT spezifischen Ontologien oder Standards sowie eine Thematisierung von Alignments
- Verwendung der Ontologien zum Zweck der Ermöglichung semantischer Interoperabilität
- Die Dokumentation des Ansatzes sowie die möglichen praktischen Umsetzungen sind nachvollziehbar.
- Die Ansätze unterscheiden sich in relevanten Aspekten und liefern damit unterschiedliche Lösungsmöglichkeiten, um mit dem Problem der semantischen Heterogenität umzugehen.

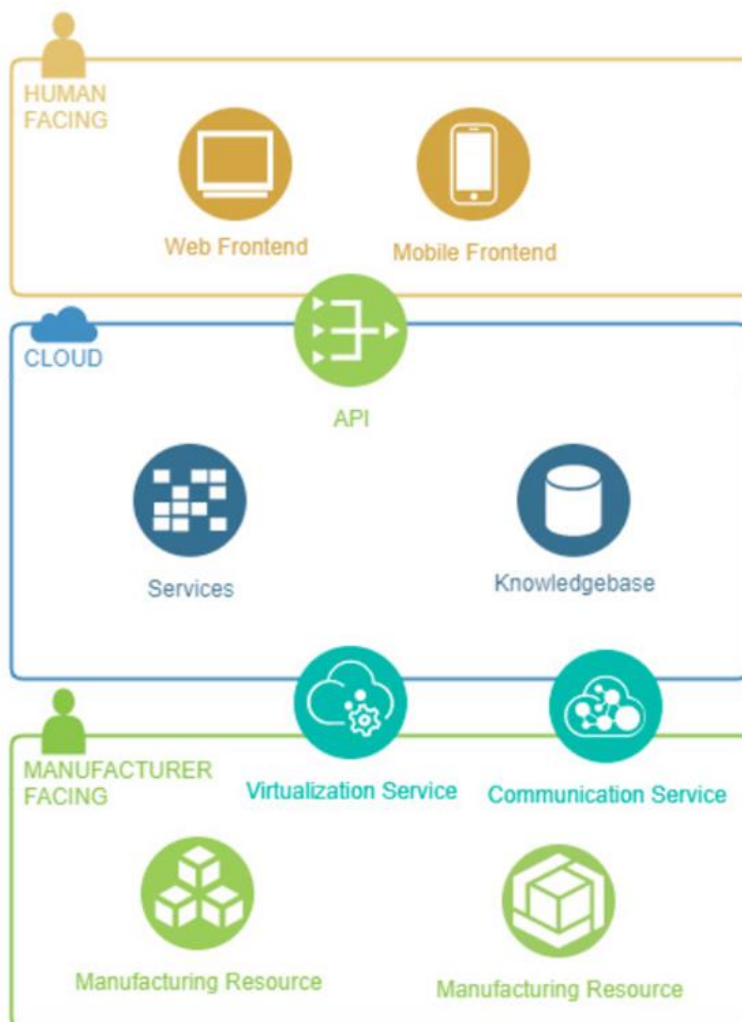
Darstellung 53 im Anhang auf Seite 122-124 gibt einen Überblick über alle in Kapitel 4 besprochenen Ansätze. Wichtig zu betonen ist dabei, dass die bisherigen Umsetzungen sehr individuell sind und nur in wenigen praktischen Erprobungsszenarien umgesetzt wurden. Einen Umsetzungsstandard oder domänenspezifische Best Practice Lösungen scheint es bisher noch nicht zu geben (näheres dazu in Abschnitt 5).

⁸⁶ Eine Suche auf Google Scholar mit der Sucheinschränkung [("Ontology" OR "Knowledge Graph" OR "KG") AND ("alignment" OR "matching") AND ("IIoT" OR "Industry 4.0" OR "I4.0") AND "Interoperability"] liefert allein für den Zeitraum ab 2017 bis zum Tag der Abfrage (27.05.2023) insgesamt 3.200 Einträge. Eine systematische Analyse dieser würde den Rahmen der Arbeit sprengen.

4.1 Einleitendes Beispiel

Anhand eines Beispiels, das bei Drexel und Hoch (2022) beschrieben wird, soll der Nutzen von Ontologien und deren Anwendung im Kontext der IIoT gezeigt werden.

In ihrer Arbeit beschreiben die Autoren eine Plattform für Cloud Manufacturing⁸⁷ (CMfg). Diese ist in drei Layer unterteilt: „Customer“, „Cloud Provider“ und „Manufacturer“. Die Idee hinter CMfg ist, dass Kunden beispielsweise über einen Webshop eine Bestellung für ein Produkt tätigen können, über den Cloud Provider werden die dafür benötigten Produktionsressourcen und Capabilities ermittelt und im untersten Layer beginnt die Herstellung. Zusätzlich werden von unten nach oben weitere Informationen zum gerade laufenden Prozess oder zu freien Kapazitäten übermittelt (siehe Darstellung 34).



Darstellung 34: Layer der Cloud Manufacturing Plattform

Quelle: Drexel und Hoch (2022, S. 2952)

⁸⁷ CMfg versucht, analog zur Idee des Cloud Computing, bei dem Arbeitsschritte auf unterschiedliche Rechner aufgeteilt werden, Produktionsschritte auf unterschiedliche Maschinen, Systeme oder Unternehmen aufzuteilen. Die Idee dabei ist, „manufacturing capabilities as a service“ anzubieten. (Drexel; Hoch 2022, S. 2951)

In jedem dieser Bereiche wird unterschiedliches Wissen benötigt. Während im Manufacturer-Layer Wissen über den Produktionsprozess der Produkte wichtig ist, benötigt der Cloud Provider-Layer Wissen über Arbeitspläne und Produktionsfortschritte. Drexel und Hoch unterteilen das Wissen dabei in drei Bereiche: Produkt, Produktion und Prozesse. Um dieses Wissen maschinenlesbar abzulegen, erstellen die Autoren für jeden dieser Bereiche eine Ontologie. Zusätzlich wird eine Top-Level Ontologie hinzugefügt, die abstrakte Konzepte und deren Relationen beschreibt, wie beispielsweise Parameter, Datentyp, Eigenschaft. (Drexel; Hoch 2022, S. 2952)

Die drei Ontologien (Produkt, Produktion, Prozess) sind über gemeinsame Konzepte miteinander verbunden. Ein gemeinsames Konzept, das in allen Ontologien vorkommt, ist „Produkt“. Während dieses in der Produkt-Ontologie noch genauer beschrieben wird (z.B. konfigurierbares und konfiguriertes Produkt), liegt der Fokus in der Produktions-Ontologie mehr auf den Ressourcen, die für das Produkt benötigt werden. Analog dazu sind die Ontologien auch jeweils für spezielle Layer relevant. Die Produkt-Ontologie ist für den Customer-Layer (Human Facing Layer) relevant, da der Kunde für die Auswahl des Produktes viel über dessen Eigenschaften wissen will. Umgekehrt ist für den Cloud Provider-Layer das Wissen über den Produktionsprozess und damit die Prozess-Ontologie relevant, während für den Produktions-Layer das Produkt aus Sicht der einzelnen Komponenten und den zu deren Herstellung relevanten Ressourcen bedeutend ist (Produktions-Ontologie).

Um den nun oben geschilderten Cloud Manufacturing-Usecase abbilden zu können, muss es zwischen den einzelnen Layern eine Interaktion geben. Dies bedeutet aber auch, dass es zwischen den Ontologien ein Alignment geben muss. Im gegebenen Beispiel ist dies vergleichsweise einfach, da die Ontologien so entworfen wurden, dass sie über das Konzept „Produkt“ direkt verbunden werden können. Allerdings muss dies nicht immer der Fall sein.

Eine Alternative wäre es gewesen, eine gemeinsame Ontologie für alle drei Layer zu wählen. Die Autoren nennen jedoch Vorteile, die durch die Verwendung von drei unterschiedlichen Ontologien entstehen. So bietet dies die Möglichkeit, einzelne Ontologien zu erweitern, ohne das Gesamtsystem zu beeinträchtigen. Hierbei ist insbesondere die Top-Level Ontology relevant, da sie die möglichen Erweiterungen einschränkt und damit die Kopplung zwischen den Ontologien unterstützt. (Drexel; Hoch 2022, S. 2953–2954)

4.2 Plattformbasierte Ansätze

IIoT-Plattformen fungieren als Middleware zwischen CPS und Anwendungen (siehe Darstellung 2 auf Seite 19). Hierbei gibt es eine Fülle an unterschiedlichen Plattform-Ausprägungen⁸⁸ mit unterschiedlichen Schwerpunktsetzungen, wie beispielsweise domänenübergreifende, föderierte Plattformen (z.B. INTER-IoT, IoT EPI, BIG IoT), domänenspezifische Plattformen (Hitachi Lumada) oder Plattformen großer Konzerne (Microsoft Azure, Google Cloud, IBM Watson IoT, Amazon IoT, Siemens MindSphere, Bosch IoT Suite). (Arnold u.a. 2022, S. 929;

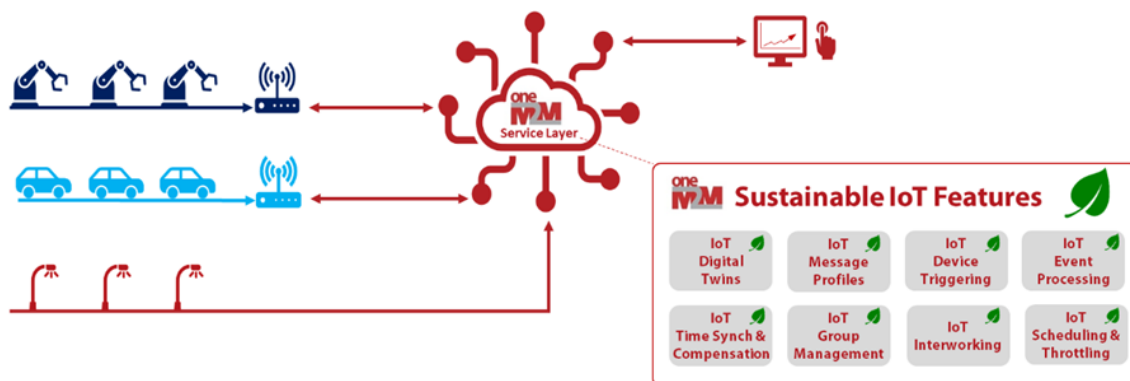
⁸⁸ Arnold u.a. (2022, S. 940–941) listen allein 78 verschiedene Anbieter von Plattformen auf, die in insgesamt 5 Kategorien eingeteilt werden.

Hazra u.a. 2023, S. 15–16) Allen gemein ist, dass sie CPS und Anwendungen miteinander verbinden und damit die heterogene IIoT-Landschaft orchestrieren, was sie zu einem zentralen Element in der Bewerkstellung der Interoperabilität macht. Semantische Interoperabilität ist dabei nur ein Aspekt. Aufgrund des Fokus dieser Arbeit wird jedoch nur dieser besprochen.

Ziel des folgenden Abschnittes ist es, aktuelle Plattform-zentrierte Methoden in der Bewerkstellung semantischer Interoperabilität zu untersuchen. Da es dem Autor nur eingeschränkt oder gar nicht möglich war, zu eruieren, auf welche Art und Weise die oben genannten kommerziellen Plattformen Ontologien verwenden,⁸⁹ beschränken wir uns im Folgenden auf eine Auswahl von drei Plattformen, für die passende Publikationen gefunden werden konnten. Dabei sei erwähnt, dass die vorgestellten Plattformen für den Bereich IoT vorgesehen sind. Deren Ansätze sind jedoch aus Sicht des Autors generisch genug, um auch auf den Bereich der IIoT übertragen zu werden.⁹⁰

4.2.1 oneM2M

Ein Beispiel für ein Alignment mithilfe einer Top-Level Ontologie wird bei der oneM2M Plattform verwendet. oneM2M ist eine Initiative mehrerer Standardisierungsorganisationen (u.a. European Telecommunication Standards Institut (ETSI)) und Industrieunternehmen, die das Ziel verfolgt, Interoperabilität zwischen IoT-Akteuren durch die Entwicklung technischer Spezifikationen zu ermöglichen. (oneM2M 2023a) Im Zentrum der oneM2M Architektur steht dabei eine als „Service Layer“ bezeichnete Middleware (siehe Darstellung 35).⁹¹



Darstellung 35: Verbindung des oneM2M Service Layers mit CPS und Anwendungen

Quelle: oneM2M (2022, S. 4)

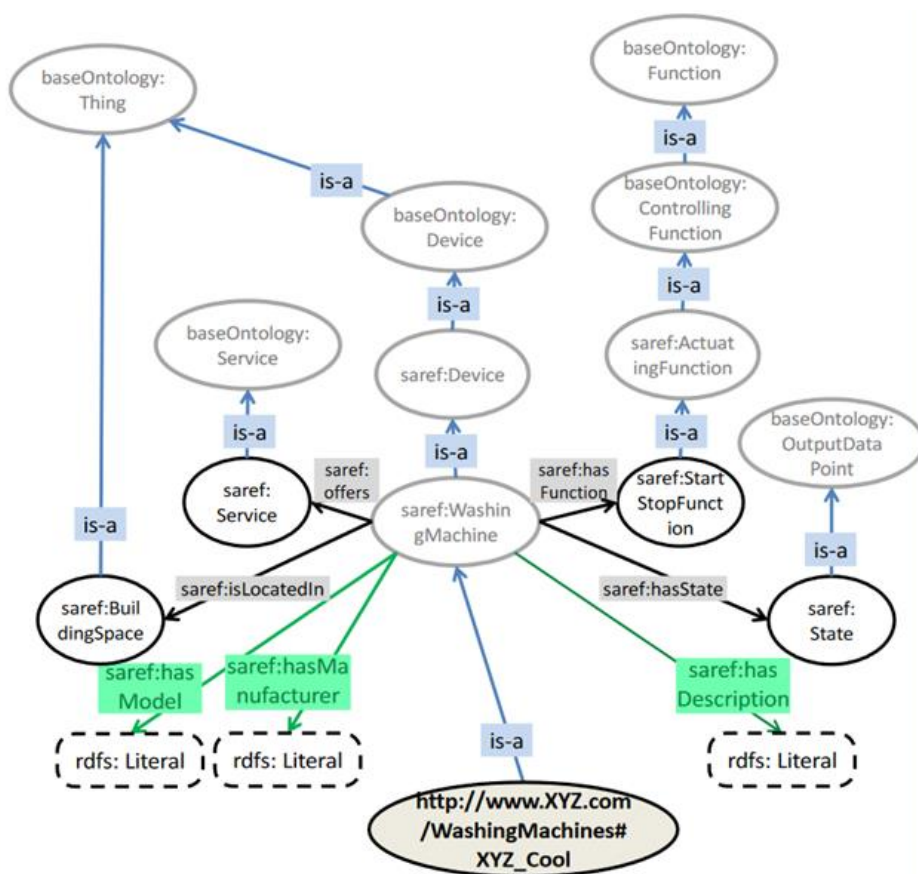
⁸⁹ Ausnahme ist die Bosch IoT Suite, die in Abschnitt 4.2.4 erwähnt wird, wobei die Informationen sehr eingeschränkt sind. Für Siemens MindSphere können Ontologien definiert und mit dem Semantic Data Interconnect (SDI) gemappt werden, jedoch scheint es keine zentrale Ontologie zu geben. (Boland 2020; Bamunuarachchi u.a. 2021, S. 6) Ähnliches gilt für Microsoft Azure, wobei hier zumindest domänenspezifische Ontologien zur Verfügung stehen. (Anderson; Barnstedt; Weinstock 2023)

⁹⁰ Auch Hazra u.a. (2023, S. 14) und Arnold u.a. (2022, S. 931) betrachten für ihre IIoT-Untersuchungen Plattformen, die initial für IoT vorgesehen sind.

⁹¹ Es ist zu erwähnen, dass oneM2M Interoperabilität auf unterschiedlichen Ebenen betrachtet (siehe (oneM2M 2023b)). Im Folgenden wird nur der Aspekt der semantischen Interoperabilität betrachtet.

Für den Service Layer stellt oneM2M die sogenannte Base Ontology zur Verfügung. (oneM2M 2019a) Während die einzelnen CPS unterschiedliche domänenspezifische Ontologien verwenden können, muss jede dieser Ontologien nur kompatibel mit der Base Ontology sein. Letztere ist eine domänenunspezifische Top-Level Ontologie, die ein minimales Set an grundlegenden abstrakten Konzepten (u.a. „Device“, „Operation“ oder „Function“) und Relationen (u.a. Subclass-Relation) beinhaltet, die mittels Äquivalenzen oder Subklassen-Relationen auf die Konzepte der CPS-Ontologien gemappt werden, damit diese in die oneM2M Plattform integriert werden können. Die aus diesem Alignment entstandene Ontologie wird als Common Ontology bezeichnet. (oneM2M 2019b) In einer IoT-Anwendung sollen dadurch die unterschiedlichen Ressourcen einheitlich dargestellt und verwaltet werden können. (Cavalieri 2021, S. 13; Dickerson u.a. 2021, S. 129)

Als Beispiel nennt die Spezifikation der Base Ontology ein Alignment mit SAREF⁹², welches in Darstellung 36 schemenhaft dargestellt ist.



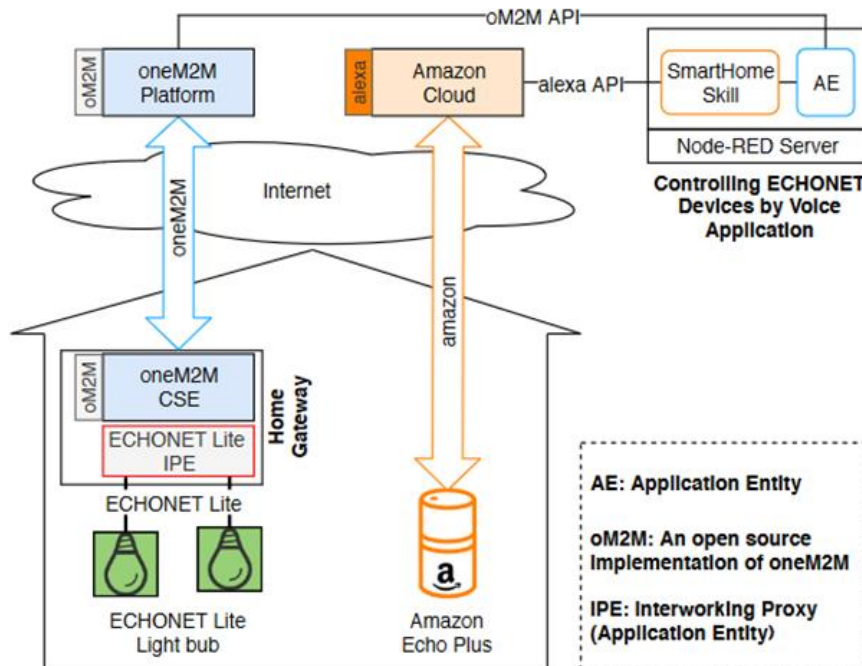
Darstellung 36: Alignment zwischen SAREF und der oneM2M Base Ontology

Quelle: oneM2M (2019a, S. 72)

Anwendungsbeispiele, die auf der Common Ontology beruhen, sind bei Pham, Lim und Tan (2019) für den Smart Home-Bereich sowie bei Cavalieri (2021) für den Smart Grid-Bereich zu finden.

⁹² Die SAREF Ontologie wurde in Abschnitt 2.2.6.2.2 vorgestellt.

Pham, Lim und Tan (2019, S. 785–787) mappen die ECHONET Ontology, eine Ontologie zur Beschreibung der Smart Home-Domäne, mit der oneM2M Base Ontology, um smarte Glühbirnen mit der oneM2M Plattform zu verbinden, die wiederum mit Amazon Alexa verknüpft ist. Dadurch wird ermöglicht, Leuchtmittel sprachgesteuert ein- und auszuschalten.



Darstellung 37: Anwendungsbeispiel der oneM2M Base Ontology im Smart Home-Bereich

Quelle: Pham, Lim und Tan (2019, S. 787)

Cavalieri (2021) verknüpft die oneM2M Base Ontologie zusammen mit SAREF4ENER, einer Erweiterung von SAREF mit Begriffen aus der Energiewirtschaft, (Daniele u.a. 2016) um damit Energieproduzenten (Kraftwerke) mit maschinellen Energiekonsumenten zu verbinden.

Weil die Base Ontologie nur sehr abstrakte Konzepte beinhaltet, ist ein Mapping mit anderen Ontologien meist einfach umzusetzen (siehe Darstellung 36). Dies macht in weiterer Folge den Ansatz sehr flexibel, da er sich, wie die Beispiele oben zeigen, für verschiedene Domänen eignet. Unterschiedliche Ressourcen (z.B. Glühbirnen, Kraftwerke, Kühlschränke, smarte Lautsprecher) und deren Services können in die oneM2M Plattform integriert werden.

Nachteilig ist hingegen aus Sicht des Autors, dass diese Integrationstiefe gering ist. Im Gegensatz zum in Abschnitt 3.4 vorgestellten hybriden Ansatz ist in der OneM2M Variante kein direktes Mapping zwischen den domänenspezifischen Ontologien vorgesehen. Da nun die Ontologien nur auf die abstrakten Konzepte der Base Ontologie gemappt werden, ist auch nur auf dieser Ebene ein Mapping zwischen den domänenspezifischen Ontologien möglich.

Diese Behauptung soll an einem Beispiel verdeutlicht werden: Zwei Kühlschränke verwenden unterschiedliche Ontologien. Dabei misst der eine Kühlschrank die Temperatur in Grad Celsius und der andere in Kelvin. Beide haben den Service „Kühlen“, der mit dem Service-Konzept der Base Ontologie gemappt wird. Dadurch kann in der oneM2M Plattform der Service

„Kühlen“ aktiviert werden, jedoch findet hier nicht das Mapping der Temperaturmaße statt. Dies müsste zusätzlich über ein Mapping der unterschiedlichen Kühlschranksontologien erfolgen.⁹³ Aus Sicht des Autors geben die oneM2M Spezifikationen derzeit keine schlüssige Antwort, wie mit dieser Problematik auf Ebene der Ontologien umgegangen werden soll. Eine Lösungsmöglichkeit wäre, die benötigten Konzepte (Temperatur, Einheiten) mit in die zentrale Ontologie mitaufzunehmen. Diesen Ansatz verfolgt das VICINITY Projekt.

4.2.2 VICINITY Projekt

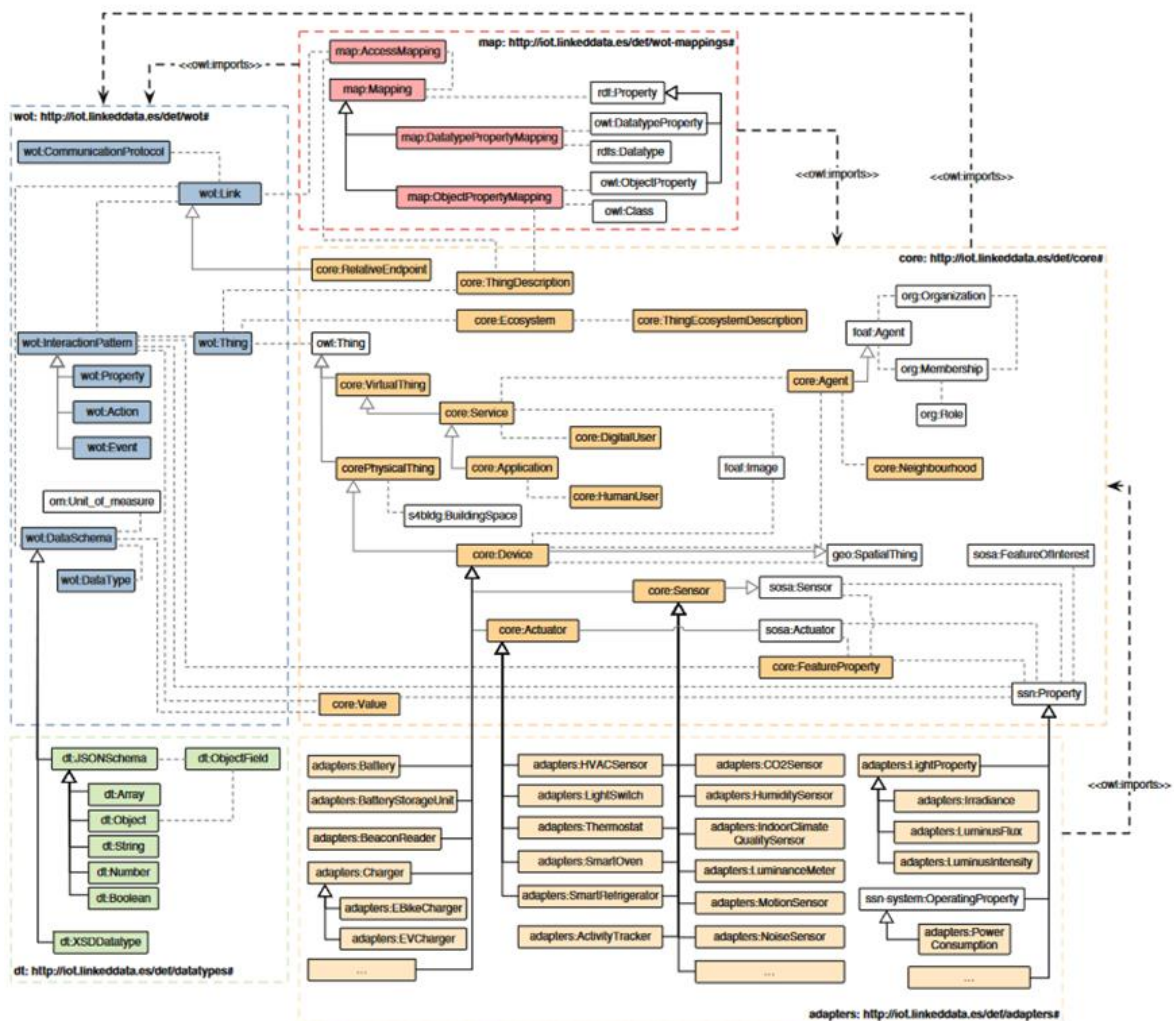
Das von der Europäischen Union finanzierte VICINITY Projekt hat zum Ziel, eine Cloud basierte Plattform zur Verfügung zu stellen, um die Interoperabilität zwischen smarten Objekten und Menschen in einer sogenannten virtuellen Nachbarschaft zu ermöglichen. Unter der Idee „interoperability as a service“ fokussiert das Projekt insbesondere auf die Ermöglichung der semantischen Interoperabilität. (VICINITY 2016; Amara; Hemam; Djezzar; Maimor 2022, S. 139) VICINITY kombiniert hierfür fünf Ontologien, um neben domänenspezifischen Konzepten aus der IoT auch domänenunspezifische Konzepte wie Maßeinheiten und auch Konzepte zur Unterstützung der Interoperabilität abzubilden. Die Ontologien basieren wiederum auf bereits bestehenden Ontologien wie SSN/SOSA und SAREF4BLDG (siehe Abschnitt 2.2.6.2) und lassen sich wie folgt beschreiben: (Ontology Engineering Group 2020a)

- **Web of Things (WoT) Ontologie** modelliert allgemeine Konzepte der Web of Things Domäne und wurde von der W3C WoT working group aufgenommen, um die Web of Things Ontologie bzw. Thing Description zu entwickeln. (Cimmino u.a. 2019, S. 112; Kaebisch u.a. 2020) Zur Ontologie gehören Konzepte wie „Action“, „Event“, „Thing“.
- **VICINITY Core Ontologie** beinhaltet grundlegende Begriffe, die benötigt werden, um Daten zwischen Akteuren über die IoT-Plattform auszutauschen. Sie beinhaltet Konzepte wie „Agent“, „Application“, „Contract“, „Human_user“, „Sensor“ und „Service“. Darüber hinaus verbindet die Core Ontologie die weiteren Ontologien (siehe Darstellung 38). Sie ist vergleichbar mit der oneM2M Base Ontology.
- **Adapters Ontologie** beinhaltet Konzepte und Eigenschaften von Geräten („Devices“), die in einer VICINITY Nachbarschaft verwendet werden können, wie beispielsweise Sensoren („WindSensor“) oder Aktoren („Light switch“).
- **Datatypes Ontologie** listet die relevantesten Datentypen wie „array“, „number“ bzw. deren Instanzen wie „float“ oder „string“ auf.
- **WoT Mappings Ontologie** ist eine spezifische Ontologie, deren Zweck es ist, Begriffe zu definieren, die für das Transformieren von RDF- auf JSON-Format⁹⁴ benötigt wird.

⁹³ Hierbei muss erwähnt werden, dass für das Umrechnen der Einheiten Ontologien alleine nicht ausreichen. Mit dem Alignment der Ontologien kann nur die Relation zwischen den Einheitsmaßen gekennzeichnet werden. Zusätzlich benötigt es (Umrechnungs-)Regeln, welche beispielsweise mithilfe der Semantic Web Rule Language (SWRL, <https://www.w3.org/Submission/SWRL> (Zugriff am: 09.06.2023)) abgebildet werden können. (Hatzivasilis u.a. 2019, S. 4)

⁹⁴ JSON steht für JavaScript Object Notation (<https://www.json.org/json-en.html> (Zugriff am: 09.06.2023)).

Sie dient somit in erster Linie für die syntaktische Interoperabilität. Konzepte dieser Ontologie sind beispielsweise „datatype“, „Mapping“ und „object_property“. Die VICINITY Plattform benötigt diese Ontologie, da CPS über eine REST⁹⁵-Schnittstelle und mithilfe von JSON-Dateien mit der Plattform kommunizieren, die Informationen innerhalb der Plattform aber in RDF verarbeitet werden. (Cimmino u.a. 2019, S. 111–112)



Darstellung 38: Überblick über das Netzwerk der VICINITY-Ontologien

Quelle: Cimmino u.a. (2019, S. 107)

Wie in Darstellung 38 ersichtlich, gibt es ein Alignment, um die fünf Ontologien miteinander zu verbinden. Darüber hinaus ist an den weiß hinterlegten Konzepten zu sehen, dass VICINITY auch Konzepte weiterer Ontologien wie SSN/SOSA, SAREF4BLDG (siehe 2.2.6.2), GEO⁹⁶ oder FOAF⁹⁷ hinzuzieht.

⁹⁵ REST steht für „Representational State Transfer“ und ist ein Stil in der Softwarearchitektur, der insbesondere zur Erstellung von Webservices verwendet wird. Einführende Erläuterungen finden sich beispielsweise unter <https://restfulapi.net> (Zugriff am: 23.06.2023).

⁹⁶ <https://www.w3.org/2005/Incubator/geo/XGR-geo-ont> (Zugriff am: 09.06.2023)

⁹⁷ <http://xmlns.com/foaf/0.1> (Zugriff am: 09.06.2023)

Der Unterschied zu oneM2M liegt bei VICINITY nun insbesondere darin, dass der Zusammenschluss der fünf Ontologien eine tiefere Integration der CPS erlaubt, da detailliertere Mappings mit den domänenspezifischen Ontologien ermöglicht werden. Das in Abschnitt 4.2.1 erwähnte Problem mit den Kühlschränken, die unterschiedliche Einheiten für die Temperatur verwenden, kann in diesem Beispiel direkt über die VICINITY Plattform gelöst werden. (VICINITY 2019)

Ein möglicher Nachteil am VICINITY-Ansatz könnte darin bestehen, dass aufgrund der verwendeten Ontologien die Flexibilität in Bezug auf die Anwendungsbereiche geringer ist. VICINITY wurde insbesondere für den Smart Home-Bereich entwickelt, worauf insbesondere Konzepte der Adapters Ontologie hinweisen (siehe Darstellung 38). Dies schränkt die Anwendungsmöglichkeiten für IIoT-Bereiche auf den ersten Blick ein.

Allerdings muss dies nicht zwangsläufig der Fall sein, da analog zum einleitenden Beispiel in Abschnitt 4.1 jeweils eine der Ontologien – beispielsweise die Adapters Ontologie – erweitert oder verändert werden könnte, ohne dass die anderen Ontologien davon betroffen sind. So könnte die Adapters Ontologie um weitere, domänenspezifische Konzepte und Instanzen ergänzt werden, ohne dass dies Einfluss auf die anderen Ontologien hat.

Darüber hinaus wurde die Interoperabilität zu anderen Ontologien wie oneM2M oder SAREF im VICINITY Projekt explizit berücksichtigt.⁹⁸ (Ontology Engineering Group 2020b; Fernández-Izquierdo; García-Castro 2019, S. 577–578)

4.2.3 INTER-IoT Projekt

Während oneM2M und VICINITY den Fokus darauf legen, CPS über eine zentrale Plattform zu verbinden, geht INTER-IoT⁹⁹ einen Schritt weiter. Ziel dieses ebenfalls aus Mitteln der Europäischen Union finanzierten Projektes ist es, ein Framework zu erstellen, das es erlaubt, Interoperabilität zwischen heterogenen IoT-Plattformen zu ermöglichen.¹⁰⁰ Damit setzt INTER-IoT eine Ebene über oneM2M und VICINITY an und betrachtet zudem explizit auch Ontologie Alignments mit domänenspezifischen Ontologien. (Moreira u.a. 2017, S. 2)

Wesentlich für die Umsetzung der semantischen Interoperabilität im INTER-IoT-Framework ist der sogenannte "Inter Platform Semantic Mediator" (IPSM). Dieser ermöglicht Alignments zwischen unterschiedlichen Ontologien mithilfe einer zentralen Ontologie, genannt „Generic Ontology of IoT Platforms“ (GOIoTP) bzw. deren Erweiterung GOIoTPex¹⁰¹. (Fortino u.a.

⁹⁸ Wobei diese Überprüfung nicht mittels Ontologie Matching sondern über das Softwaretool Themis (<https://themis.linkeddata.es> (Zugriff am: 09.06.2023)) umgesetzt wurde. Themis überprüft dabei, ob Ontologien die ihnen zu Grunde liegenden Anforderungen erfüllen. (Fernández-Izquierdo; García-Castro 2019)

⁹⁹ <https://inter-iot.eu/> bzw. <https://inter-iot.readthedocs.io/en/latest/#main-concepts> (Zugriff jeweils am: 09.06.2023)

¹⁰⁰ Im Projekt wurden die Domänen Logistik und Gesundheitswesen genauer betrachtet, wobei der Ansatz domänenunspezifisch ist. (Moreira u.a. 2017, S. 2)

¹⁰¹ Siehe <https://inter-iot.github.io/ontology> (Zugriff am: 09.06.2023). Auch wenn GOIoTP und GOIoTPex im Zuge von INTER-IoT entwickelt wurden, betonen die Verantwortlichen, dass auch eine

2018, S. 218; Ganzha u.a. 2021, S. 9) Somit ähnelt der Ansatz jenem von oneM2M in solcher Weise, dass beliebige Ontologien mit einer zentralen Ontologie verbunden werden. Andererseits besteht die GOloTP aus mehreren Modulen (Device, User, Service, Platform, Observation&Actuation, Location, Unit), um den Bereich der IoT möglichst ganzheitlich abzudecken, womit INTER-IoT wiederum eine Verwandtschaft zu VICINITY aufweist, den Ansatz aber beispielsweise um das Platform-Modul erweitert. Dieses Modul ist geeignet, eine Plattform oder Middleware samt deren Ontologie zu modellieren und in weiterer Folge das Wissen, ob zwei oder mehrere Plattformen miteinander kompatibel sind, abzubilden. Ebenfalls analog zu VICINITY bauen die Module von GOloTP auf bestehenden Ontologien wie SSN/SOSA oder GeoSPQRQL¹⁰² auf. (Ganzha u.a. 2021, S. 10)

Wie sieht nun die konkrete Umsetzung aus? Um semantische Interoperabilität zwischen zwei Plattformen oder CPS¹⁰³ zu ermöglichen, die dabei unterschiedliche Ontologien (O1 und O2) verwenden, wird zuerst ein Alignment zwischen O1 und GOloTP bzw. zwischen O2 und GOloTP erstellt. Anschließend tritt die zentrale Ontologie GOloTP als Mediator auf und ermöglicht somit das Alignment zwischen O1 und O2.

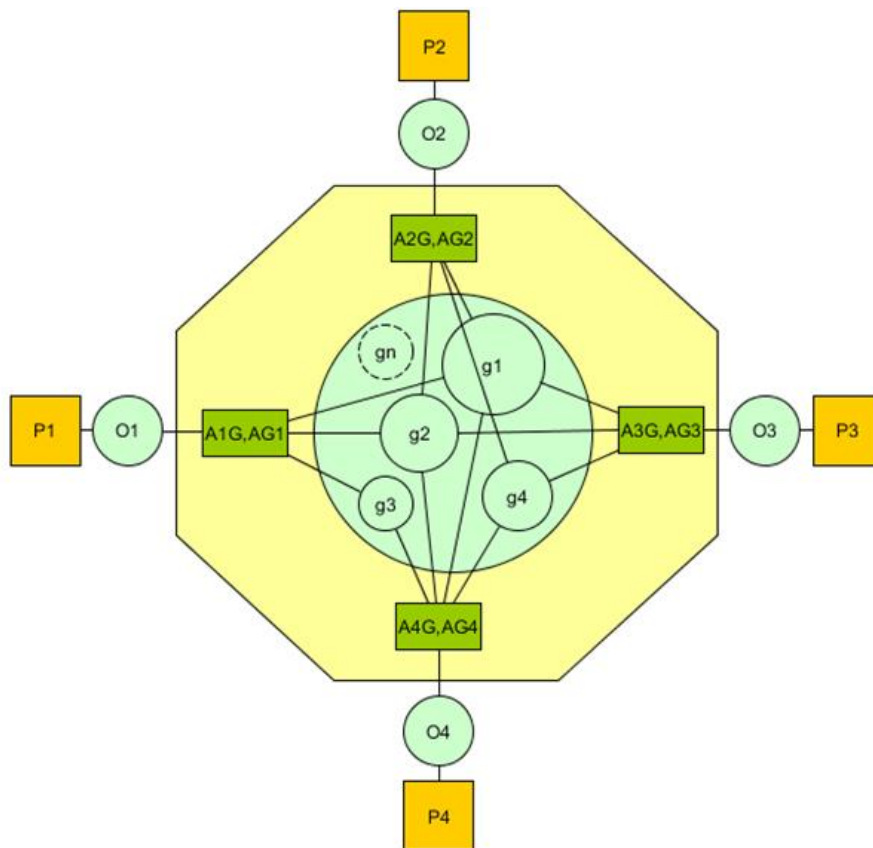
Darstellung 39 zeigt die Idee dieses Alignmentprozesses für vier domänenspezifischen Ontologien (O1-O4). Alle vier Ontologien werden mit der im Zentrum befindlichen GOloTP gematcht. Über die GOloTP sind diese Ontologien nun auch miteinander verbunden und die semantische Interoperabilität zwischen den Plattformen (P1-P4) wird ermöglicht. Der große Vorteil einer solchen Architektur ist, dass die Anzahl der benötigten Alignments nur linear mit der Anzahl der unterschiedlichen Ontologien steigt und somit linear skaliert: Für jede neue Ontologie braucht es ein weiteres bidirektionales Alignment. Demgegenüber würde ein direktes Alignment zwischen den externen Ontologien dafür sorgen, dass die Anzahl der benötigten Alignments annähernd quadratisch zur Anzahl der beteiligten Ontologien zunimmt.¹⁰⁴ Ein weiterer Vorteil von GOloTP ist die Aufteilung der Ontologie in unterschiedliche Ontologiemodule. Dadurch müssen für ein Alignment nicht immer alle Konzepte der GOloTP Ontologie berücksichtigt werden, sondern nur jene Module, die relevant für das Mapping sind, wie durch Darstellung 39 verdeutlicht wird. (Fortino u.a. 2018, S. 218–219; INTER-IoT Projekt 2019)

andere zentrale Ontologie verwendet werden könnte, vorausgesetzt, diese ist geeignet, ein Alignment zwischen den externen Ontologien zu ermöglichen. (Ganzha u.a. 2021, S. 18)

¹⁰² <https://opengeospatial.github.io/ogc-geosparql/geosparql11/index.html> (Zugriff am: 09.06.2023)

¹⁰³ Auch wenn der Ansatz für die Interoperabilität von Plattformen entwickelt wurde, spricht aus Sicht des Autors nichts dagegen, diesen auf die Interoperabilität von CPS zu übertragen.

¹⁰⁴ Beispiel zur Verdeutlichung: Während man beim INTER-IoT-Ansatz bei 4 Ontologien genau 4 bidirektionale Alignments benötigt (jede Ontologie wird mit GOloTP verbunden), wären es bei einem direkten Alignment zwischen jeder der 4 Ontologien insgesamt 6 Alignments (die erste (O1) hat 3 Mappings (zu O2, O3, O4), die zweite (O2) braucht noch 2 (zu O3, O4) und die dritte (O3) braucht noch eines zur letzten (O4) damit jede Ontologie mit jeder verbunden ist). Bei 5 Ontologien würde man bei INTER-IoT 5 und bei einer direkten Anbindung der Ontologien 10 Alignments benötigen. Mathematisch betrachtet beträgt die Zahl für die Alignments im Fall „Jede Ontologie mappt mit jeder anderen“ bei n Ontologien $(n - 1) + (n - 2) + \dots + 1 = \frac{n \cdot (n-1)}{2} = \frac{n^2 - n}{2} = o(n^2)$ Alignments. Denselben kombinatorischen Vorteil liefert jede Form von Middlewares oder Plattformen. Bei der Umformung der obigen Summe wurde die Gauß'sche Summenformel verwendet (siehe https://de.wikipedia.org/wiki/Gau%C3%9Fsche_Summenformel (Zugriff: 01.07.2023)).



Darstellung 39: Aufbau des Alignments der GOloTP mit domänenspezifischen Ontologien

Beschreibung: Im Zentrum des IPMS (gelbes Achteck) befindet sich die GOloTP (türkiser Kreis) mit den Modulen (g1, g2, usw.). Die Plattformen (P1, P2, P3, P4) verfügen über Ontologien (O1, O2, O3, O4). Für jede der vier Ontologien wird ein gerichtetes Alignment zur GOloTP (A1G, A2G, usw.) und ein gerichtetes Alignment zurück (AG1, AG2, usw.) ermittelt. Über das Alignment zur GOloTP sind die Plattform-Ontologien miteinander verbunden.

Quelle: INTER-IoT Project (2019)

Zur Ermittlung der Alignments verweist das INTER-IoT Projekt auf bestehende Tools wie Log-Map oder ALIN (siehe Abschnitt 3.6). Die ermittelten Alignments werden dabei im IPMS verwaltet. (Ganzha u.a. 2021, S. 11; Singh u.a. 2022, S. 181)

Praktische Umsetzungen erfuhr das INTER-IoT Projektes im Gesundheitswesen (Ibáñez-Sánchez u.a. 2021), im Bereich der Hafenlogistik (Giménez u.a. 2021) oder im Bereich des Wassermanagements (Singh u.a. 2022).

Das Besondere am INTER-IoT Projekt ist, dass es nicht nur analog zu oneM2M und VICINTY auf einer zentralen Ontologie aufsetzt, mit der die domänenspezifischen Ontologien verbunden werden, sondern auch den Bedarf an Alignments und das Ermitteln dieser explizit thematisiert und hierfür auf bestehende Lösungen zurückgreift. Folglich ist der Ansatz generisch und ermöglicht es, unterschiedliche Domänen miteinander zu verbinden – vorausgesetzt, die Alignments sind korrekt.

4.2.4 Zusammenfassung und weitere plattformbasierte Ansätze

Aus den drei präsentierten Projekten lässt sich ein gemeinsamer Trend zum hybriden Ansatz (siehe Abschnitt 3.4) unter Verwendung einer Hub and Spoke-Architektur¹⁰⁵ erkennen. Zuerst wird eine zentrale Top-Level Ontologie definiert, die die wichtigsten abstrakten Konzepte beinhaltet. An diese können entweder beliebige domänenspezifische Ontologien gemappt werden (oneM2M) oder die Plattform gibt weitere Mid-Level Ontologien oder Ontologie-Module vor (VICINITY, INTER-IoT). Ein typisches Beispiel für ein solches Modul ist eine Ontologie zu Maßeinheiten. Dadurch können externe Ontologien auf einer tieferen Ebene verbunden werden.

Szmeja u.a. (2022, S. 662) erwähnen hierbei folgende Anforderungen an Top- und Mid-Level Ontologien:

1. Sie müssen alle Begriffe, mit denen im Ökosystem kommuniziert wird, beinhalten.
2. Sie müssen spezifisch genug sein, dass Abfragen oder Schlussfolgerungen möglich sind.
3. Sie müssen über einen modularen Aufbau verfügen, der es erlaubt, dass Module unabhängig verwaltet und versioniert werden können.

Die Vorteile dieser Methode können dabei wie folgt zusammengefasst werden: (Adamovic u.a. 2022, S. 19)

1. Anstelle der Erzeugung neuer Ontologien können bestehende verwendet werden.
2. Es wird ermöglicht, unterschiedliche nutzerspezifische Sichten auf dieselben Probleme abzubilden, indem unterschiedliche Ontologien verwendet werden. Z.B. kann ein Problem aus physikalischer und aus chemischer Sicht betrachtet werden.
3. Dadurch wird die Wahl einer passenden Ontologie erleichtert, die der eigenen Sprache der Nutzer entspricht. Gleichzeitig kann mittels Alignments eine Angleichung an andere Repräsentationen erfolgen, wodurch das Risiko von Konflikten sinken soll.
4. Es besteht die Möglichkeit des Austausches einzelner Module, ohne dass zwangsläufig das Gesamtgebilde verändert werden muss.

Kritisch anzumerken ist insbesondere für Punkt zwei und drei, dass das Problem der double articulation (siehe Abschnitt 3.1.2) in keinem der vom Autor analysierten Beiträge angesprochen wird. Gerade bei der Verknüpfung mehrerer Domänen ist aus Sicht des Autors mit Homonymen und insbesondere mit double articulation zu rechnen.

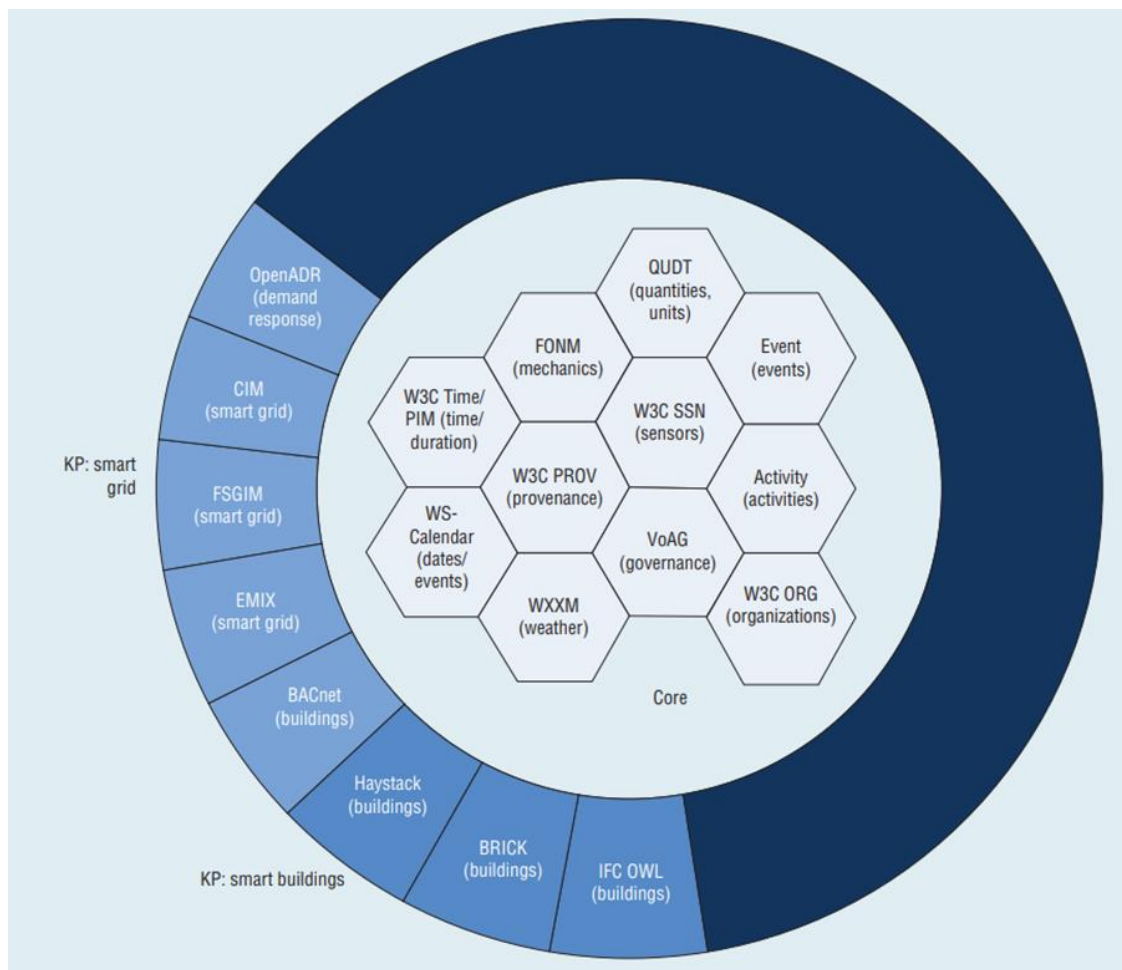
Das oben angesprochene Muster der Hub and Spoke-Architektur zeigt sich im Übrigen auch bei der Erstellung neuer Ontologien. Neben der in Abschnitt 2.2.3 erwähnten Common Core Ontologie, die aus zwei Top-Level und mehreren Mid-Level Ontologien besteht, veröffentlichte die Industrial Ontologies Foundry (IOF), an der u.a. auch das National Institut of Standards

¹⁰⁵ Die Hub and Spoke- (bzw. Nabe und Speiche-) Architektur beschreibt ein Prinzip, bei dem zwei Elemente immer über ein zentrales Element verbunden sind. Dieses Architekturschema wird auch in anderen Bereichen, wie z.B. dem Data Warehousing, verwendet. Darstellung 39 verdeutlicht diese Idee.

and Technology (NIST) beteiligt ist, im Februar 2023 die IOF Core Ontologie. Dies ist eine Top-Level Ontologie, die zentrale Begriffe der noch in Ausarbeitung befindlichen domänen-spezifischen Ontologien, wie beispielsweise der Maintenance Ontology oder der Supply Chain Reference Ontology (SCRO), beinhaltet. Ziel der IOF ist dabei, ein Portfolio an Ontologien zu entwickeln, die den Bereich der digitalen Produktion vollständig abbilden werden. (IOF 2023a, 2023b)

Auch einige kommerzielle Plattformentwickler scheinen diesen Ansatz zu verfolgen. So verwendet Bosch für seine IoT Suite die sogenannte Bosch I4.0 Core Ontology, eine Top-Level Ontologie auf Basis von IEC 62264. (OntoCommons 2020)

Siemens wiederum stellte das Open Semantic Framework (OSF) vor, das mehrere sogenannte Core Ontologien (u.a. SSN) mit domänenspezifischen Ontologien verbindet.¹⁰⁶ (Mayer u.a. 2017) Darstellung 40 zeigt ein Beispiel für die Zusammenstellung des OSF mit den Core Ontologien im Zentrum und weiteren domänenspezifischen Ontologien (genannt „Knowledge Packs“ (KP)) in der Peripherie, die je nach Anwendungsfall unterschiedlich sein können.



Darstellung 40: Open Semantic Framework (OSF) von Siemens

Quelle: Mayer u.a. (2017, S. 98)

¹⁰⁶ Der Autor fand keinen Hinweis darauf, dass die OSF in Siemens IIoT-Plattform MindSphere verwendet wird.

Parallel dazu versucht auch das EU Horizon2020 Projekt „Ontology-Driven Data Documentation for Industry Commons“¹⁰⁷ (OntoCommons) bestehende industriespezifische Ontologien mithilfe von Top- und Mid-Level Ontologien zu verbinden, um ein Ökosystem an Ontologien („Ontology Commons EcoSystem“ (OCES)) zu erstellen. (Adamovic u.a. 2022, S. 4) Dieses Projekt hat sich darüber hinaus zum Ziel gesetzt, Alignments zwischen IIoT-Ontologien zu verwalten. Die bisherigen Ergebnisse sind jedoch noch überschaubar, da die Mappings nur auf Äquivalenzen (identische URI) oder mithilfe des Softwaretools LOOM ermittelt wurden, welches nur lexikalische Ähnlichkeiten prüft, wie in Abschnitt 3.6.1 erwähnt wurde.¹⁰⁸

Neben den hier erwähnten gibt es noch weitere dokumentierte plattform-basierte Projekte, die sich ebenfalls mit semantischer Interoperabilität im Kontext der IIoT beschäftigen. Zu nennen sind beispielsweise BIG IoT oder symbloTe. Ein Überblick über diese findet sich bei Fortino u.a. (2018, S. 204–206) oder Amara u.a. (2022, S. 139–140)

So ambitioniert und vielversprechend die oben gezeigten Ansätze auch erscheinen mögen – mehr als Ansätze, die in einzelnen Versuchsanordnungen praktisch umgesetzt wurden, sind es zum jetzigen Stand noch nicht. Dies hat mehrere Gründe, wobei in Bezug auf die semantische Interoperabilität vor allem zwei Aspekte herausstechen:

- Jeder der gezeigten Ansätze verwendet andere Ontologien. Insbesondere ist nicht einmal die Top-Level Ontologie dieselbe.
- Jeder Ansatz benötigt Ontologie Alignments, thematisiert deren Ermittlung jedoch entweder so gut wie gar nicht oder verweist auf bestehende Lösungen, ohne auf mögliche Probleme, die bei der Ermittlung von Alignments entstehen können, einzugehen.

Der erste Punkt kann dabei nach Ansicht des Autors als Symptom für die Vielschichtigkeit hinter dem Begriff IIoT bzw. als Symptom für die Fülle an unterschiedlichen Szenarien, die unter dem Begriff IIoT subsumiert werden, gesehen werden. Der zweite Punkt ist dagegen auf das Design und den Fokus der einzelnen Projekte zurückzuführen, da hier mehr die Präsentation einer Anwendung als die Auseinandersetzung mit kleineren Details, wie fehlenden Mappings, im Vordergrund steht.

Aus diesem Grund sollen im nächsten Abschnitt auch Arbeiten besprochen werden, die sich dezidiert mit dem Ermitteln von Alignments zwischen IIoT-Ontologien beschäftigen.

¹⁰⁷ Siehe <https://www.ontocommons.eu> (Zugriff am: 09.06.2023). Das Projekt war zum Zeitpunkt der Erstellung dieser Arbeit noch nicht abgeschlossen. Die Ziele und ein relativ aktueller Stand sind bei Adamovic (2022) bzw. auf der angegebenen Projekt-Seite zu finden.

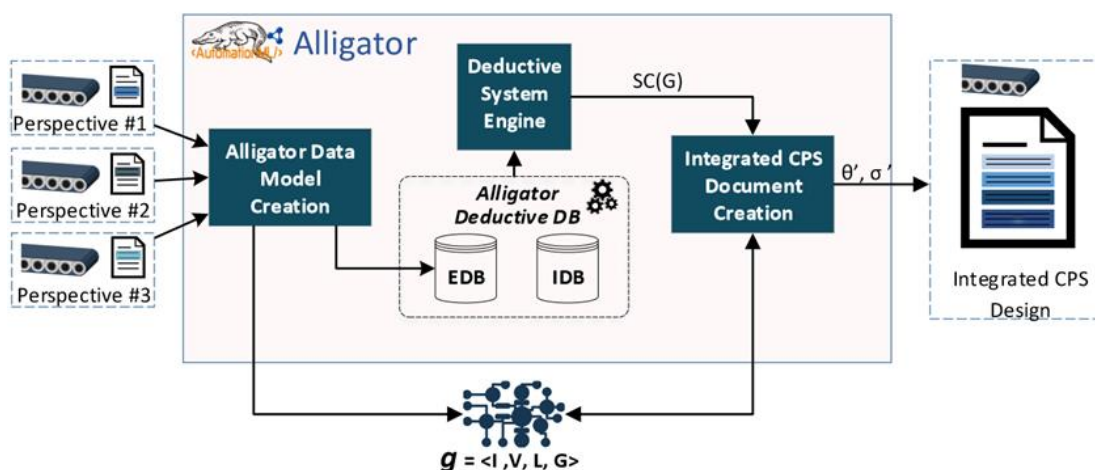
¹⁰⁸ Der aktuelle Stand ist auf <http://industryportal.enit.fr/landscape> bzw. <http://industryportal.enit.fr/mappings> einsehbar. (Zugriff jeweils am: 09.06.2023)

4.3 Weitere Ansätze mithilfe von analytischem Ontologie Matching

4.3.1 Alligator und SemCPS

Grangel-González (2019, S. 109–126) beschreibt einen Ansatz, dessen Ziel es ist, aus AutomationML-Dokumenten¹⁰⁹, die ein CPS jeweils aus einer unterschiedlichen Perspektive¹¹⁰ beschreiben, ein gemeinsames Dokument zu erstellen. Hierbei sollen semantische Konflikte, die durch das Verbinden der Perspektiven entstehen, aufgelöst werden. Die Umsetzung erfolgt über mehrere Prozessschritte und verwendet unterschiedliche Technologien.

Zuerst werden die AutomationML-Dokumente in Wissensgraphen überführt, wobei hierbei AMLO (Kovalenko u.a. 2018), eine Ontologie für AutomationML sowie Krestor¹¹¹, ein Framework, um XML-Dateien in das RDF-Format zu transformieren, verwendet werden.



Darstellung 41: Aufbau des Alligator-Frameworks

Beschreibung: Die Architektur des Alligator-Frameworks: Die AutomationML-Dokumente des CPS werden in einen Wissensgraphen (g) und als Datalog-Prädikate (EDB) in eine deduktive Datenbank überführt. Aus zuvor definierten Regeln (IDB) werden Mappings und Heterogenitäts-Konflikte ermittelt und versucht, ein konsolidiertes AutomationML-Dokument zu erzeugen.

Quelle: Grangel-González (2019, S. 118)

Im nächsten Schritt wird versucht, diese unterschiedlichen Wissensgraphen in einen konsolidierten zu überführen. Hierzu müssen Mappings erstellt sowie Konflikte, die beim Mappen auftreten können, identifiziert und aufgelöst werden. Technisch erfolgt dies zuerst mit zweierlei

¹⁰⁹ <https://www.automationml.org> (Zugriff am: 09.06.2023). AutomationML ist ein XML-basiertes Datenformat zur Modellierung, Speicherung und zum Austausch von Engineering-Daten (z.B. Anlagendaten) und ist insbesondere für den Kontext der IIoT entwickelt worden. (AutomationML 2023)

¹¹⁰ Grangel-González verwendet hierbei den Begriff „Perspektive“ (siehe auch Darstellung 17 in Abschnitt 3.1.1). Das Konzept kann aus Sicht des Autors aber auch auf unterschiedliche Ontologien (und Standards) überführt werden. (Grangel-González u.a. 2018, S. 130)

¹¹¹ <https://github.com/EIS-Bonn/krestor> (Zugriff am: 09.06.2023)

Hilfsmitteln: Neben der Verwendung von eCl@ss¹¹² als Top-Level Ontologie, um syntaktische und terminologische Heterogenität zu identifizieren, kommt eine deduktive Datenbank zum Einsatz. Dies ist eine Erweiterung einer relationalen Datenbank um Regeln, die dabei helfen, Mappings zu ermitteln und Konflikte zu entdecken. Die Regeln¹¹³ werden dabei mithilfe von Datalog beschrieben, einer Ableitung der Programmiersprache Prolog.

Das Resultat dieses Ansatzes ist einerseits ein (mitunter unvollständiges) Alignment und andererseits ein Set an Konflikten, die nicht eindeutig aufgelöst werden können. Das Framework für diesen Ansatz wurde von Grangel-González mit dem Namen „Alligator“ bezeichnet und ist in Darstellung 41 beschrieben.

Da Alligator auf deterministischem Schließen beruht und damit nur – aus Sicht des Algorithmus – sichere Mappings und Konflikte ermittelt, für letztere aber keine Lösungsvorschläge liefert, entwickelte Grangel-González einen weiteren Ansatz, genannt SemCPS, der auf der Probabilistic-Soft-Logic beruht. Für die ermittelten Konflikte werden hierbei mögliche Alternativen von Wissensgraphen erzeugt. Die Kanten sind dabei mit einem Gewicht $w \in [0,1]$ versehen, welches das Maß für die Unsicherheit angibt.¹¹⁴ Die in Beispiel 2 auf Seite 42 (Darstellung 17) gezeigten alternativen Wissensgraphen können nun mit ihren Wahrscheinlichkeitsfaktoren beschrieben werden (siehe Darstellung 42). Zur Ermittlung der Gewichte kommen auch hier wieder vorab definierte Regeln zur Anwendung.¹¹⁵ Der Vorteil dieser Methode ist somit, dass sie auch mit Unsicherheit umgehen kann und direkt Lösungsvorschläge für die oben erwähnten Konflikte liefert.

Im Anschluss an die Ermittlung der alternativen Graphen werden alle Kanten verworfen, die einen zuvor definierten Grenzwert unterschreiten. Das Ergebnis ist dann die Vereinigung aller verbliebenen Knoten und Kanten und damit der integrierte Wissensgraph.

Zuletzt vergleicht Grangel-González die Resultate von SemCPS und Alligator mit jenen anderer Ansätze, namentlich EDOAL/Alignment API¹¹⁶ und SILK¹¹⁷. Hierbei erzielte SemCPS fast durchgehend bessere Ergebnisse, vorausgesetzt, es war kein konfliktfreies Mappen möglich und damit eine Unsicherheit gegeben.

¹¹² http://www.ebusiness-unibw.org/ontologies/eclass/5.1.4/eclass_514en.html (Zugriff am: 09.06.2023)

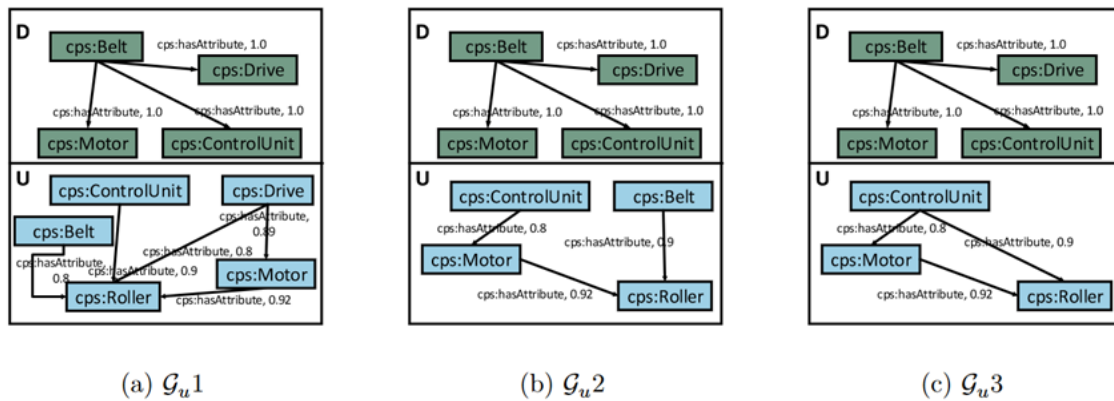
¹¹³ Ein von Grangel-González (2019, S. 121) erwähntes Regelbeispiel wäre „Falls die Konzepte X und Y beide das Attribut „eClassIRDI“ haben und die Werte dieselben sind, dann haben die Konzepte dieselbe „eClassIRDI“. Formal: „hasAttributeName(X, "eClassIRDI") \wedge hasAttributeName(Y, "eClassIRDI") \wedge hasAttributeValue(X,Z) \wedge hasAttributeValue(Y,Z) \Rightarrow sameEClassIRDI(X,Y)“

¹¹⁴ Die aus Fußnote 113 bekannte Regel wird dabei um einen Wahrscheinlichkeitsfaktor von 0.9 (90%) ergänzt: „hasAttributeName(X, "eClassIRDI") \wedge hasAttributeName(Y, "eClassIRDI") \wedge hasAttributeValue(X,Z) \wedge hasAttributeValue(Y,Z) \Rightarrow sameEClassIRDI(X,Y) | 0.9“ (Grangel-González 2019, S. 121)

¹¹⁵ Dies ist aus Sicht des Autors zugleich ein Schwachpunkt dieser Methode, da die Regeln als auch die zu setzenden Gewichte für eine Regel vom User selbst definiert werden müssen und somit subjektiv-gewichtet sein können.

¹¹⁶ <https://moex.gitlabpages.inria.fr/alignapi> (Zugriff am: 10.06.2023)

¹¹⁷ <http://silkframework.org> (Zugriff am: 10.06.2023)



Darstellung 42: Alternative Wissensgraphen aus SemCPS

Beschreibung: SemCPS ermittelt alternative Wissensgraphen (a, b, c), die jeweils aus einem mittels Alligator deterministisch bestimmten Teilgraphen (D) und einem mit Unsicherheit behafteten Teilgraphen (U) bestehen. Bei letzterem ist jede Relation mit einer Wahrscheinlichkeit für deren Korrektheit versehen.

Quelle: Grangel-González (2019, S. 126)

Einschränkend muss jedoch erwähnt werden, dass die Rohdaten – die AutomationML-Dokumente, welche integriert werden sollten – automatisch generiert wurden, da passende Benchmark-Daten in diesem Bereich nicht existieren. Somit sind die Ergebnisse nicht ohne Weiteres verallgemeinerbar. (Grangel-González u.a. 2018, S. 122–123)

4.3.2 Ergänzungen von Standards um Ontologien

Im obigen Abschnitt wurde erwähnt, dass der AutomationML Standard um eine Ontologie (AMLO) ergänzt wurde. Dieses Muster, nämlich aus einem bestehenden Standard eine Ontologie abzuleiten, findet sich auch an anderer Stelle, wie folgende zwei Beispiele zeigen sollen.

Rubí und Gondim (2020) erstellen aus einem bestehenden Standard (openEHR¹¹⁸) zuerst eine Ontologie, um diese anschließend mit der SSN Ontologie (siehe Abschnitt 2.2.6.2.1) zu verbinden. Dies bildet die Voraussetzung für das spätere Erstellen einer Plattform für eHealth-Anwendungen.

Analog dazu leiten Hodges und Garcia (2017, S. 33) zuerst eine Ontologie aus dem EDDL Standard¹¹⁹ ab und mergen diese anschließend mit der QUDT¹²⁰ Ontologie, um damit einen digitalen Zwilling zu erzeugen. Ziel ist es, den Status des CPS (in diesem Fall ein Druckmessgerät¹²¹) zu ermitteln, die Software upzudaten oder Simulationen durchzuführen. Dieses

¹¹⁸ https://specifications.openehr.org/releases/BASE/Release-1.0.3/architecture_overview.html (Zugriff am: 30.05.2023)

¹¹⁹ Electronic Device Description Language (<https://www.fieldcommgroup.org/integration-technologies/eddl#4257225834-2034166732>) (Zugriff am: 10.06.2023))

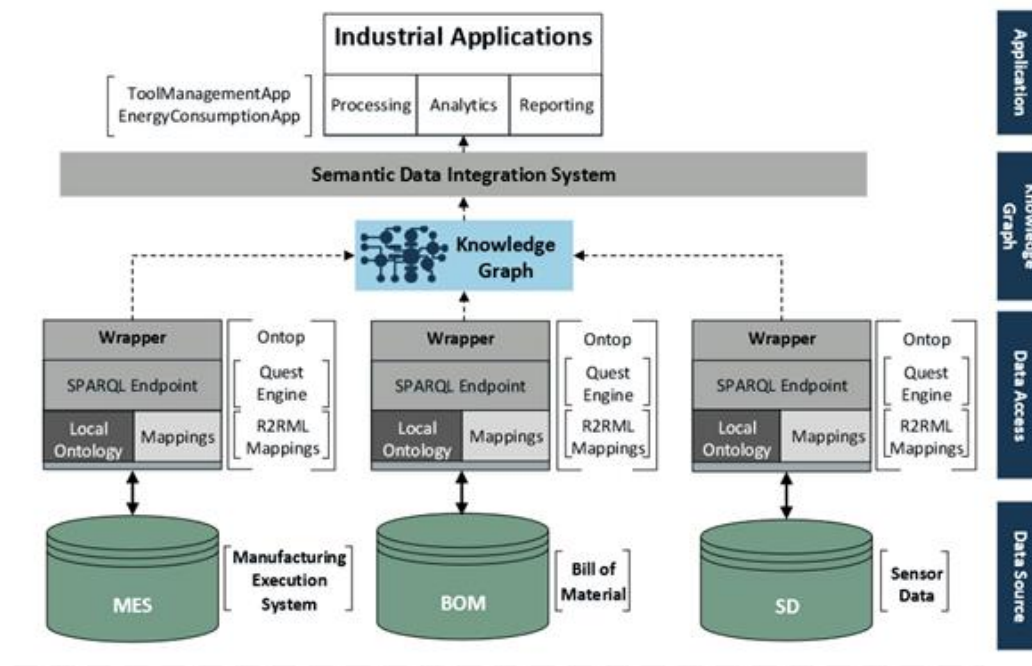
¹²⁰ Quantities, Units, Dimensions and Datatypes (<https://www.qudt.org>) (Zugriff am: 10.06.2023))

¹²¹ <https://www.siemens.com/global/en/products/automation/process-instrumentation/pressure-measurement.html> (Zugriff am: 10.06.2023)

Beispiel zeigt nicht nur, wie aus bestehenden Ontologien neue, für den Einsatzzweck optimierte Ontologien erzeugt werden können. Die Hodges und Garcia verweisen auch auf den Nutzen, Standards mithilfe von Ontologien zu repräsentieren und damit insbesondere auch die Interoperabilität von Standards zu fördern. Gerade das Fehlen von dezidierten Ontologien zu IIoT-Standards und Referenzarchitekturmodellen wie RAMI4.0 (Adolphs u.a. 2015; Heidel 2019) oder IIRA (Lin u.a. 2022) beeinträchtigt die Interoperabilität. (Grangel-González u.a. 2018, S. 34; Pereira; Szejka; Cancigliari Jr 2022, S. 4)

4.3.3 Anwendung zur Anlagenüberwachung

Ein weiteres Beispiel für einen hybriden Ansatz liefern Petersen u.a. (2017). Ziel ist es, eine Anwendung zu erstellen, die auf Daten des Manufacturing Execution Systems (MES), Sensor Daten (SD) der Maschinen und Stücklisten (BOM) bzw. Arbeitspläne zugreifen kann. Im Zentrum der Betrachtung stehen Produktionsmaschinen, die mit unterschiedlichen Werkzeugen ausgestattet und damit für unterschiedliche Fertigungsvarianten verwendet werden können.



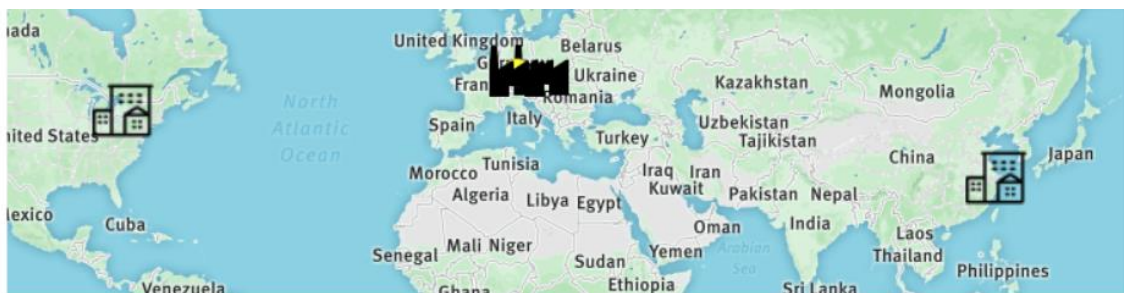
Darstellung 43: Mergen von Ontologien aus dem Produkt- und Produktionsbereich

Quelle: Grangel-González (2019, S. 138) in Anlehnung an Petersen u.a. (2017, S. 7)

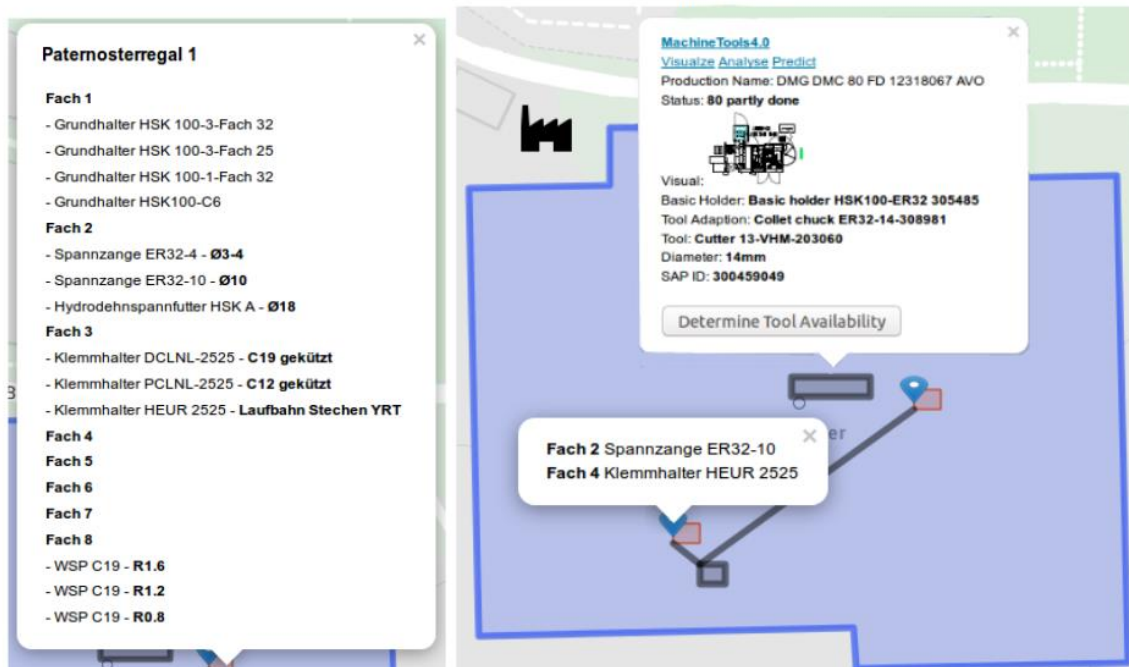
Darstellung 43 zeigt, wie aus den jeweiligen Ontologien eine gemeinsame Ontologie („Knowledge Graph“) erzeugt werden soll. Hierbei werden Konzepte unterschiedlicher Ontologien

(RAMI¹²², VIVO¹²³, NeoGeo¹²⁴, FOAF¹²⁵, SSN¹²⁶) zusammen mit unternehmensspezifischen Konzepten gemappt. Die gemeinsame Top-Level Ontologie beinhaltet dabei nur abstrakte Konzepte und dient als Mediator zwischen den darunter liegenden Ontologien einerseits als auch zu weiteren externen Ontologien andererseits, womit eine starke Ähnlichkeit zu den in Abschnitt 4.2 vorgestellten Ansätzen besteht.

Im Artikel beschreiben die Autoren weiters zwei konkrete Anwendungen, die mit ihrem Ansatz realisiert werden sollen. Die erste Anwendung (siehe Darstellung 44) gibt die aktuelle Konfiguration, den Standort und die Beschaffenheit der installierten Werkzeuge wieder. Damit sollen Maschinenführer schneller den Status einer Maschine überprüfen können, damit in weiterer Folge die Rüstzeiten reduziert werden können.



(a) Global view of the company sites



Darstellung 44: Anwendung zur Überwachung der Produktionsanlagen

Quelle: Petersen u.a. (2017, S. 9)

¹²² <https://github.com/i40-Tools/RAMIOntology> (Zugriff am: 10.06.2023)

¹²³ <https://github.com/vivo-ontologies/vivo-ontology> (Zugriff am: 23.06.2023)

¹²⁴ <https://www.w3.org/2005/Incubator/geo/XGR-geo-ont-20071023> (Zugriff am: 23.06.2023)

¹²⁵ <http://xmlns.com/foaf/0.1> (Zugriff am: 23.06.2023)

¹²⁶ Siehe Abschnitt 2.2.6.2.1.

Anwendung zwei verfolgt das Ziel, mithilfe der in der Maschine verarbeiteten Sensoren die Stromkosten pro Maschine bzw. pro produziertem Stück zu ermitteln.

Wie bereits in den bisherigen Ausführungen erwähnt, weisen auch hier die Autoren darauf hin, dass der Vorteil einer hybriden Architektur darin besteht, dass die darunterliegenden Ontologien flexibel erweiterbar sind und die damit verbundenen Datenstrukturen unterschiedlich sein können, ohne dass es Auswirkungen auf die jeweils anderen Ontologien hat. Die Top-Level Ontologie wiederum benötigt nur jene Konzepte, die für ein Alignment relevant sind. (Petersen u.a. 2017, S. 7)

4.3.4 MAPSOM Framework von Jirkovský

Jirkovský verbindet drei Ontologien (DUL¹²⁷, SSN¹²⁸ und COCI¹²⁹) um Entitäten und Prozesse in einem Wasserkraftwerk abzubilden. Das Matchen und Mergen der Ontologien erfolgt dabei nach dem selbst entwickelten MAPSOM Framework. Hierbei werden die Entitäten unterschiedlicher Ontologien zuerst mithilfe einer Ähnlichkeitsfunktion verglichen. Das Framework verwendet dabei lexikalische Distanzen wie die Levenshtein-Distanz oder das n-gram-Maß (siehe Abschnitt 3.5.1.3). (Jirkovský 2017, S. 94)

Aufbauend auf diesen Distanzmessungen werden Vektoren gebildet. Dabei geben die Koordinaten des Vektors die unterschiedlichen Distanzen von je zwei miteinander verglichenen Entitäten an.¹³⁰

Die Vektoren werden anschließend in zwei Bereiche unterteilt, um Vektoren, die ein gültiges Mapping angeben, von jenen, die kein gültiges Mapping angeben, zu unterscheiden. Dabei wird nach einer initialen Unterteilung eine sogenannte self-organized map (SOM) (Kohonen 1990), eine besondere Form eines zweidimensionalen künstlichen neuronalen Netzes¹³¹ trainiert, dessen Aufgabe es ist, die Vektoren in einen positiven und einen negativen Bereich zu Clustern. Mit anderen Worten: Das neuronale Netz soll anhand der Distanzen zwischen den Entitäten ermitteln, welche Entitäten gemappt werden sollen, und welche nicht. Ein Vektor liegt damit entweder in einem positiven Bereich (i.e. gültiges Mapping) oder in einem negativen Bereich (ungültiges Mapping).

Um das neuronale Netz zu trainieren, werden die zwischenzeitlichen Cluster durch menschliche Akteure bewertet. Darstellung 45 zeigt exemplarisch das Vorgehen bei der Ermittlung positiver und negativer Bereiche. Bei jedem Iterationsschritt entscheidet der User nur für jene

¹²⁷ http://ontologydesignpatterns.org/wiki/Ontology:DOLCE%2BDnS_Ultralite (Zugriff am: 09.06.2023)

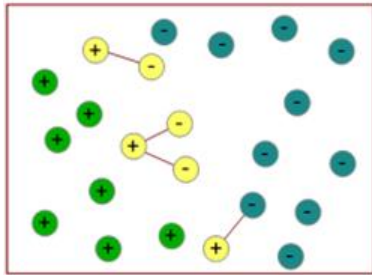
¹²⁸ Siehe Abschnitt 2.2.6.2.1.

¹²⁹ COCI steht für "Cyber-Physical System Ontology for Components Integration" und ist eine von Jirkovský selbst erstellte Ontologie. (Jirkovský 2017, S. 76–81)

¹³⁰ Beispiel: Würden die beiden Konzepte A und B mit jeweils zwei Ähnlichkeitsmaßen x und y verglichen und wäre das Ergebnis des Ähnlichkeitsmaßes x der Wert 0.5 und jenes für Maß y der Wert 0.3, so ergibt sich der Vektor $\vec{v}_{AB} = \begin{pmatrix} 0.5 \\ 0.3 \end{pmatrix}$

¹³¹ Eine erste Einführung in die Grundprinzipien von künstlichen neuronalen Netzwerken findet sich beispielsweise in <https://towardsdatascience.com/introduction-to-artificial-neural-networks-5036081137bb> (Zugriff am: 23.06.2023).

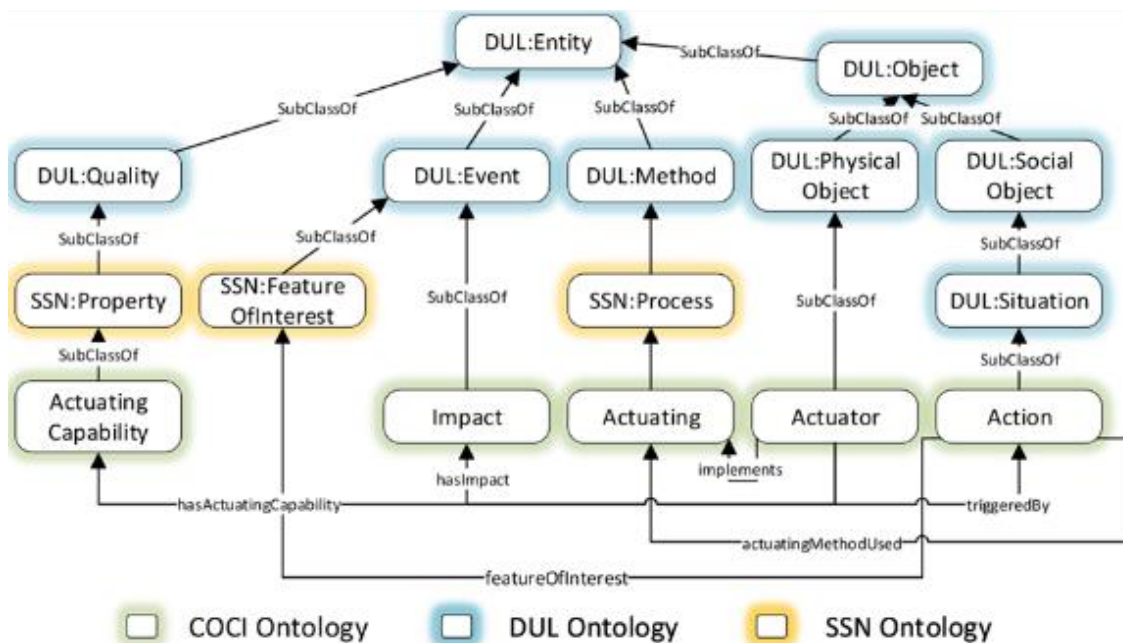
Vektoren, die nahe beieinander liegen, aber unterschiedliche Vorzeichen haben, ob diese korrekt oder falsch klassifiziert wurden. Anhand der dadurch gelabelten Vektoren soll das neuronale Netz lernen, im nächsten Schritt eine bessere Vorhersage zu erstellen. Ziel ist es, dass der User dadurch nur eine geringe Anzahl an potenziellen Mappings bewerten muss und dennoch ein akzeptables Alignment erhält. (Jirkovský 2017, S. 90–93)



Darstellung 45: Zuweisung gültiger und ungültiger Mappings im MAPSOM Framework

Quelle: Jirkovský (2017, S. 93)

Schließlich visualisiert Darstellung 46 das Ergebnis des Alignments der drei Ontologien COCI, DUL und SSN.



Darstellung 46: Alignment der COCI, DUL und SSN Ontologien mithilfe des MAPSOM Frameworks

Quelle: Jirkovský (2017, S. 81)

Interessant ist, dass in der Dokumentation zur aktuellen Version von SSN ebenfalls ein Alignment zwischen DUL und SSN/SOSA beschrieben wird.¹³² Das Ergebnis von Jirkovský weicht jedoch von diesem leicht ab. Da sowohl Jirkovskýs Arbeit als auch die aktuelle SSN-

¹³² https://www.w3.org/TR/vocab-ssn/#DUL_Alignment (Zugriff am: 19.05.2023)

Dokumentation im Jahr 2017 veröffentlicht wurden und Jirkovský letztere nicht in seiner Arbeit erwähnte, liegt der Verdacht nahe, dass er nicht absichtlich ein anderes Alignment verwendet hat, sondern die Diskrepanz auf die unterschiedliche Methodik zur Ermittlung der Alignments zurückzuführen ist.

Einen beinahe identen Ansatz wählen Xue u.a. (2021), ohne dabei aber auf Jirkovský zu verweisen. Auch sie beschreiben eine Methode, Alignments aus einer Kombination von Ähnlichkeitsmaßen und der Verwendung von neuronalen Netzen zu ermitteln. In Ihrem Beispiel betrachten sie drei unterschiedliche Sensor-Ontologien (SSN, MMI Device¹³³ und CSIRO¹³⁴). Zuerst ermitteln sie unterschiedliche Ähnlichkeitsmaße¹³⁵ zwischen den Entitäten der Ontologien. Somit erhalten sie zu je zwei Entitäten aus unterschiedlichen Ontologien eine Liste von Werten, bei der jeder Wert ein Ähnlichkeitsmaß zwischen den beiden Entitäten beschreibt.

Ausgehend von dieser Liste von Werten wird ein neuronales Netz trainiert, welches ermitteln soll, ob zwei Entitäten gemappt werden können oder nicht. Aus Sicht der Autoren bringt dies den Vorteil, dass das oft zeitraubende manuelle Mappen automatisiert werden kann. (Xue u.a. 2021, S. 45) Dieses Argument ist jedoch aus zweierlei Gründen fraglich. Zum einen liefern die Ähnlichkeitsmetriken selbst bereits eine gute Selektion von möglichen Mappings. Warum ein neuronales Netz bessere Ergebnisse liefert als eine Filterung bzw. Sortierung der Ähnlichkeitswerte, wird weder von Xue u.a. noch von Jirkovský begründet. Zum anderen ist es fraglich, ob ein automatisches Ermitteln von Alignments tatsächlich große Zeit- und Aufwandersparnisse bringt. Alignments zwischen zwei Ontologien müssen im Allgemeinen nur einmal ermittelt werden¹³⁶ und im Gegensatz zu naturwissenschaftlichen Ontologien sind IIoT-Ontologien meist überschaubar von ihrer Größe¹³⁷.

Während die Methode von Jirkovský als Input für das neuronale Netz Vektoren verwenden, deren Werte auf Basis von lexikalischen Ähnlichkeitsmaßen ermittelt werden, ist der Ansatz von Xue u.a. flexibler, da er auch strukturelle Zusammenhänge betrachtet. Beiden gemeinsam ist jedoch, dass ihre Methoden eine Mischung aus klassischer analytischer Ermittlung von Alignments und dem Wissensgraph Embedding samt der Verwendung von Machine Learning Algorithmen darstellen. Der Unterschied zum tatsächlichen Wissensgraph Embedding liegt jedoch darin, dass bei letzterem nicht nur Änderungsmaße in den Vektorraum überführt werden, sondern die Konzepte und Relationen der Ontologie. Darüber hinaus werden für das Erstellen des Embeddings bereits neuronale Netze verwendet, nicht erst zur Analyse der Vektoren. Anwendungen aus dem Wissensgraph Embedding werden in Abschnitt 4.4 vorgestellt.

¹³³ Marine Metadata Interoperability (MMI) Device ontology. (Rueda u.a. 2010)

¹³⁴ Commonwealth Scientific and Industrial Research Organisation (CSIRO) sensor ontology (Neuhaus; Compton 2009)

¹³⁵ Angewendete Ähnlichkeitsmaße sind a) lexikalische (Levenshtein-Ähnlichkeit, n-gram-Ähnlichkeit) und b) strukturelle (numerische Hierarchie-Ähnlichkeit). (Xue u.a. 2021, S. 46–47)

¹³⁶ Einschränkend muss hierbei natürlich erwähnt werden, dass das Alignment bei Änderungen einer der Ontologien womöglich angepasst werden muss.

¹³⁷ siehe dazu auch Abschnitt 4.4.

4.3.5 Zusammenfassung der analytischen Ansätze

Fasst man die gezeigten analytischen Ansätze zusammen, so zeigt sich folgendes Muster:

- Es gibt einen Trend, bestehende Standards um Ontologien zu ergänzen und diese anschließend mit weiteren Ontologien zu mergen, um semantische Interoperabilität zu ermöglichen.
- Hierdurch soll eine für den jeweiligen Anwendungsfall optimierte Ontologie entstehen.
- Bereits das Matchen von kleineren Ontologien wird mit einem immensen technischen Aufwand erzeugt, ohne dass dieser von den Autorinnen und Autoren explizit gerechtfertigt wird. Beispiele hierfür sind das Alligator, SemCPS und MAPSOM Framework.
- Auch in dieser Kategorie gibt es einen Trend zu einer hybriden Ontologie-Architektur – analog zu den plattformbasierten Ansätzen aus Abschnitt 4.2. In diesem Zusammenhang weisen Petersen u.a. (2017, S. 7) darauf hin, dass der Vorteil einer hybriden Architektur darin besteht, dass die darunterliegenden Ontologien flexibel erweiterbar sind, ohne Auswirkungen auf die anderen Ontologien. Die Top-Level Ontologie wiederum benötigt nur jene Konzepte, die für ein Alignment und damit für den Austausch zwischen den Ontologien relevant sind.

Nach den analytischen Ansätzen betrachten wir im Folgenden solche, die Methoden des Wissensgraph Embedding verwenden.

4.4 Ansätze mithilfe von Wissensgraph Embeddings

Nilsson, Sandin und Delsing (2019) stellen ein Konzept zur Ermöglichung von Interoperabilität zwischen CPS vor, in dem sie den Alignmentprozess als mathematischen Übersetzungsprozess betrachten. Hierbei werden Nachrichten unter Zuhilfenahme der jeweiligen Wissensgraphen unterschiedlicher CPS in Embeddings überführt und anschließend mit Methoden des Machine Learning bzw. Natural Language Processing von einem Vektorraum in den anderen übersetzt. Die Autoren erwähnen zwar eine Vielzahl an Technologien, wie das Sprachmodell BERT oder das Relational Graph Convolutional Network (R-GCN) (Schlichtkrull u.a. 2018), die beim Übersetzungsprozess helfen könnten, gehen jedoch nicht konkret darauf ein, wie die Embeddings ineinander überführt werden können.

Javed u.a. (2022) schließen an die Arbeit von Nilsson, Sandin und Delsing an und verwenden BERT-INT zur Ermittlung der Alignments. BERT-INT ist eine Kombination aus BERT und einem Interaktions-Modell zur Ermittlung der Mappings zwischen den Entitäten. (Tang u.a. 2021, S. 3175) An einem praktischen Beispiel, bei dem ein Temperatursensor mit einem Aktor kommunizieren soll, erstellen die Autoren ein Alignment zwischen den IIoT-Ontologien SSN und SOSA (siehe Abschnitt 2.2.6.2.1). Ihr Vorgehen ist dabei wie folgt: (Javed u.a. 2022, S. 9–17)

- Ausgangspunkt sind die Wissensgraphen der beiden Ontologien SSN und SOSA sowie ein Datenset von 255 Sensoren, das in weiterer Folge dazu dienen soll, die

Modelle zu trainieren. Die Wissensgraphen sind dabei in sogenannte RDF-Triple unterteilt. Jedes Triple besteht aus einem Subjekt, einer Relation und einem Objekt.

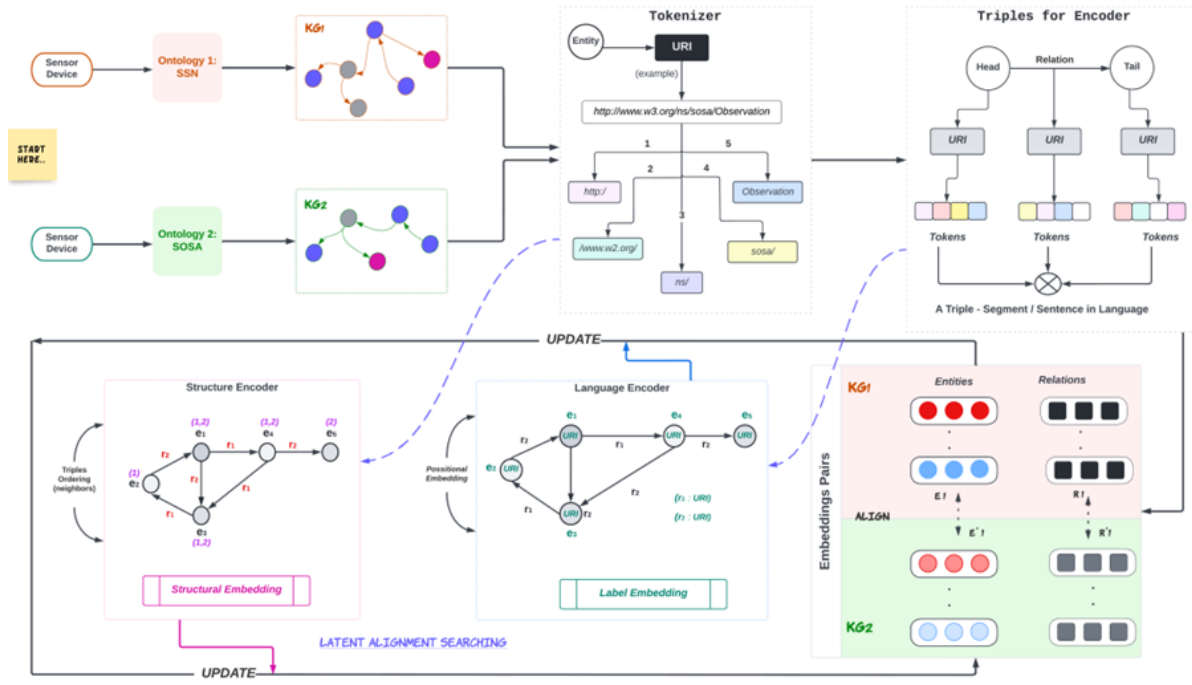
- Die Datensets werden als nächstes in die beiden Ontologien eingelesen, um Instanzen dieser Ontologien zu erzeugen. Mit anderen Worten werden die Daten nun einerseits durch die SSN- und andererseits durch die SOSA-Ontologie beschrieben.
- Jede Entität in einer der Ontologien wird durch ihren Identifier, die sogenannte URI (Uniform Resource Identifier), bezeichnet. Ein Beispiel für das Konzept „Observation“ aus der SOSA-Ontologie wäre <http://www.w3.org/ns/sosa/Observation>. Wie leicht zu sehen ist, ist der Aufbau der URI schematisch. Somit kann die URI mithilfe des sogenannten Tokenizer in ihre Segmente aufgespalten werden.
- Anschließend werden mithilfe des BERT-INT Modells die Embeddings ermittelt, wobei einerseits ein Embedding für die Entitäten (Konzepte, Instanzen, Relationen) – das sogenannte Language Encoding – als auch für die Triple – das sogenannte Structural Encoding – generiert wird.
- Die Idee ist nun, dass jeweils zwei Konzepte, die aus unterschiedlichen Wissensgraphen stammen, auf denselben oder einen ähnlichen Vektor abgebildet werden. Das bedeutet: Für ein Konzept A aus dem SSN-Wissensgraphen wird überprüft, welche Konzepte B aus dem SOSA-Wissensgraphen auf einen ähnlichen¹³⁸ Vektor abgebildet wurden. Darüber hinaus fließt zusätzlich die strukturelle Information mit ein, indem überprüft wird, ob benachbarte Konzepte von A auf ähnliche Vektoren abgebildet wurden, wie benachbarte Konzepte von B.
- Das Ermitteln des bestmöglichen Embeddings erfolgt dabei iterativ. Nach jedem Durchlauf wird anhand einer Verlustfunktion (loss function) ein Score ermittelt, mit dem ein neuronales Netz trainiert wird, um schließlich das bestmögliche Alignment zu erhalten.¹³⁹

Darstellung 47 visualisiert den oben beschriebenen Prozess.

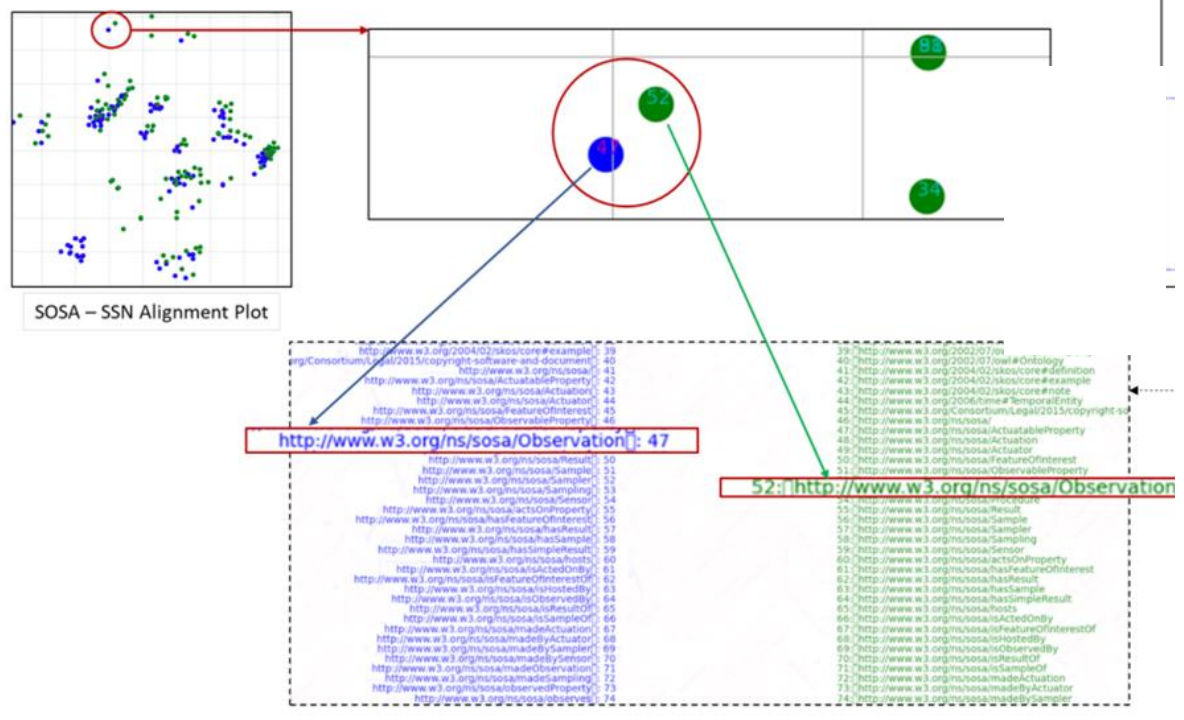
Nach Ermittlung des Embeddings können nun die Distanzen zwischen den Entitäten der unterschiedlichen Ontologien betrachtet werden. Sind die Distanzen gering, kann von einem Mapping ausgegangen werden, wie Darstellung 48 anhand des Konzeptes „Observation“ verdeutlichen soll.

¹³⁸ Im Falle von BERT_INT wird hierbei die Kosinus-Ähnlichkeit (d.h. das normierte Skalarprodukt der Vektoren) ermittelt. (Tang u.a. 2021, S. 3177)

¹³⁹ Wobei dieses Trainieren sowohl überwacht (supervised) als auch unüberwacht (unsupervised) erfolgen kann. (Javed u.a. 2022, S. 13–14)



Darstellung 47: Prozess zum Ermitteln der Alignments mithilfe von BERT_INT
 Quelle: Javed u.a. (2022, S. 9)



Darstellung 48: Prozess zum Ermitteln der Alignments mithilfe von BERT_INT
 Beschreibung: Die Graphik links oben zeigt die zweidimensionale Visualisierung der Embeddings von SSN (grün) und SOSA (blau). Die Graphik rechts oben zeigt eine Vergrößerung des Ausschnittes, auf die die Observation-Konzepte der beiden Ontologien (Graphik unten) abgebildet wurden. Da diese nahe beieinander liegen, ermittelte die Methode ein Mapping zwischen diesen Konzepten.

Quelle: Javed u.a. (2022, S. 17)

Kritisch anzumerken ist bei dieser Arbeit vor allem der Umstand, dass SOSA und SSN keine unabhängigen Ontologien sind. Genauer gesagt inkludiert SSN SOSA, wie in 2.2.6.2.1 bereits erwähnt wurde. Somit ist automatisch jedes Konzept, das in SOSA enthalten ist, auch in SSN enthalten. Dies wirft die Frage auf, warum sich der Ansatz von Javed u.a. überhaupt eines Embeddings bedient hat, wenn eine einfache Gleichsetzung mittels lexikalischen Vergleichs ein 100% korrektes Ergebnis für alle Konzepte, die in SOSA enthalten sind, geliefert hätte. Leider gehen die Autorin und die Autoren nicht auf dieses Problem ein.

Ein weiteres Beispiel für das Alignment von Ontologien mithilfe von Wissensgraph Embeddings zeigen Tayur und Suchithra. (2022) Dabei verwenden sie ein sogenanntes Generative Adversarial Network (GAN) – einer Kombination aus zwei neuronalen Netzen, wobei das eine Netz (genannt „Diskriminator“) die Ergebnisse des anderen („Generator“) bewertet.¹⁴⁰ Das Ziel von Tayur und Suchithra ist es, ein Alignment der IIoT-Ontologien SAREF, SSN und der oneM2M Base Ontologie (siehe 2.2.6.2 und 4.2.1) zu ermitteln. Das grundlegende Prinzip ist dasselbe wie jenes von Javed u.a. Unterschiede finden sich vor allem bei der Verwendung der jeweiligen Algorithmen als auch in der Verwendung von unterschiedlichen Informationen, die im Wissensgraph gespeichert sind.

4.5 Matching von IIoT-Standards: Der I40 Wissensgraph

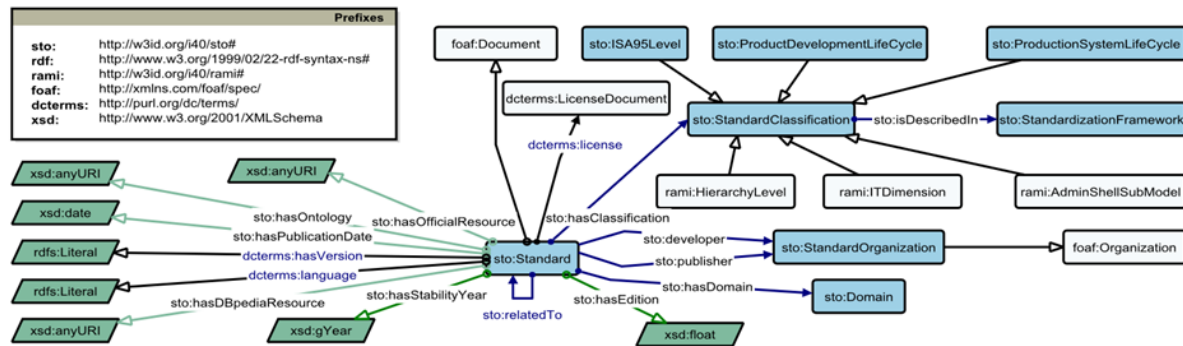
Wie bereits in der Einleitung und in Abschnitt 2.1.4 erwähnt, spielen Standards und Referenzarchitekturmodelle wie RAMI4.0 (Adolphs u.a. 2015; Heidel 2019) und IIRA (Lin u.a. 2022) eine bedeutende Rolle im Ermöglichen von Interoperabilität. Jedoch sind Standards in der Regel nicht als formale Ontologien und damit als maschinenlesbare Dokumente ausgelegt. Wie aber bereits in den Abschnitten 4.3.1 und 4.3.2 besprochen wurde, gibt es Tendenzen, Ontologien zu bestehenden Standards zu entwerfen, um damit semantische Interoperabilitätsprobleme zu lösen. Der logische nächste Schritt wäre es, für die Fülle an IIoT relevanten Standards, Ontologien zu entwickeln und Alignments zwischen diesen zu erzeugen, um damit eine Landkarte mit semantisch interoperablen Standards zu erzeugen. (Grangel-González; Vidal 2021, S. 17)

Einen Schritt in diese Richtung geht der I40KG¹⁴¹ (Industry 4.0 Knowledge Graph) des Fraunhofer-Instituts für Intelligente Analyse- und Informationssysteme (IAIS). Der I40KG verbindet dabei mehr als 220 Standards und 10 Referenzarchitekturmodelle miteinander. Darüber hinaus kategorisiert er die Standards abhängig von den Referenzarchitekturmodellen und codiert diese in einem umfassenden Wissensgraphen. Dadurch ermöglicht er in weiterer Folge, neue, implizit vorhandene Relationen zwischen Standards zu ermitteln, Interoperabilitätskonflikte zu lösen sowie die Relevanz von Standards (in Bezug auf die Anzahl der Verknüpfungen zu anderen Standards) zu beurteilen. (Rivas u.a. 2021, S. 344–345; Grangel-González 2019, S. 147–148)

¹⁴⁰ GANs erlangten in den letzten Jahren insbesondere durch die Verwendung zur Erstellung von Deepfakes Bekanntheit. (Wood 2020)

¹⁴¹ <https://i40-tools.github.io/I40KG/docs/index.html> (Zugriff am 06.05.2023)

Kern des I40KG ist die Standards Ontologie (STO), die als Top-Level Ontologie zur Beschreibung der Standards als auch deren Relationen zueinander dient. Dabei verwendet die STO Konzepte aus bestehenden Ontologien wie der FOAF Ontologie, um Akteure miteinander zu verbinden, oder der RAMI Ontologie¹⁴², um Standards mit den Konzepten aus dem RAMI4.0 Referenzarchitekturmodell zu verknüpfen. (Grangel-González; Vidal 2021, S. 20) Darstellung 49 zeigt einen Ausschnitt der STO.



Darstellung 49: Standards Ontologie (STO)

Beschreibung: Zentrale Konzepte der Standards Ontologie (STO), mit farblicher Trennung zwischen STO eigenen Konzepten (blaue Rechtecke) und Konzepten (weiße Rechtecke) und Eigenschaften (grüne Rauten) aus anderen Ontologien.

Quelle: Grangel-González (2021, S. 20)

Darüber hinaus besitzt der I40KG eine Verbindung zu DBpedia als externe formale Ressource (siehe Abschnitte 2.2.6.1.2 und 3.5.1.1), wodurch die Standards mit weiteren Informationen aus DBpedia verknüpft werden. Grangel-González und Vidal konnten dabei zeigen, dass 96% der durch den I40KG entdeckten Relationen zwischen Standards korrekt waren und die Verbindungen zwischen den Standards – durch Ausnutzung der Transitivität der Relationen – um 80% erhöht werden konnten. (Grangel-González; Vidal 2021, S. 17)

Einschränkend muss erwähnt werden, dass der I40KG nur Relationen zwischen den Standards, nicht aber zwischen deren Konzepten erstellt. Beispielsweise zeigt der I40KG an, dass OPC UA und IEC 62541 Synonyme für denselben Standard sind oder welche Standards in derselben RAMI4.0-Ebene verwendet werden. (Grangel-González; Vidal 2021, S. 21) Der I40KG benutzt somit nur Metadaten der Standards, ohne auf ihre Inhalte (Begriffe, Definitionen) genauer einzugehen. Um auch die Konzepte zwischen Standards zu mappen, wäre es wichtig, dass die Standards in maschinenlesbarer Form, vorzugsweise als Ontologien, vorliegen würden. Dies ist wie oben erwähnt nicht für alle Standards gegeben.

Eine Lösung für dieses Problem schlagen Melluso, Grangel-González und Fantoni (2022) vor, in dem sie die als PDF publizierten Standards, die im I40KG erwähnt werden, zuerst in ein XML-Format überführen und anschließend mithilfe strukturbasierter Regeln Begriffe und Definitionen extrahieren. Im nächsten Schritt werden mittels des bereits oben in Abschnitt 3.5.2.1 bzw. 4.4 erwähnten Sprachmodells BERT Ähnlichkeiten zwischen den in Standards

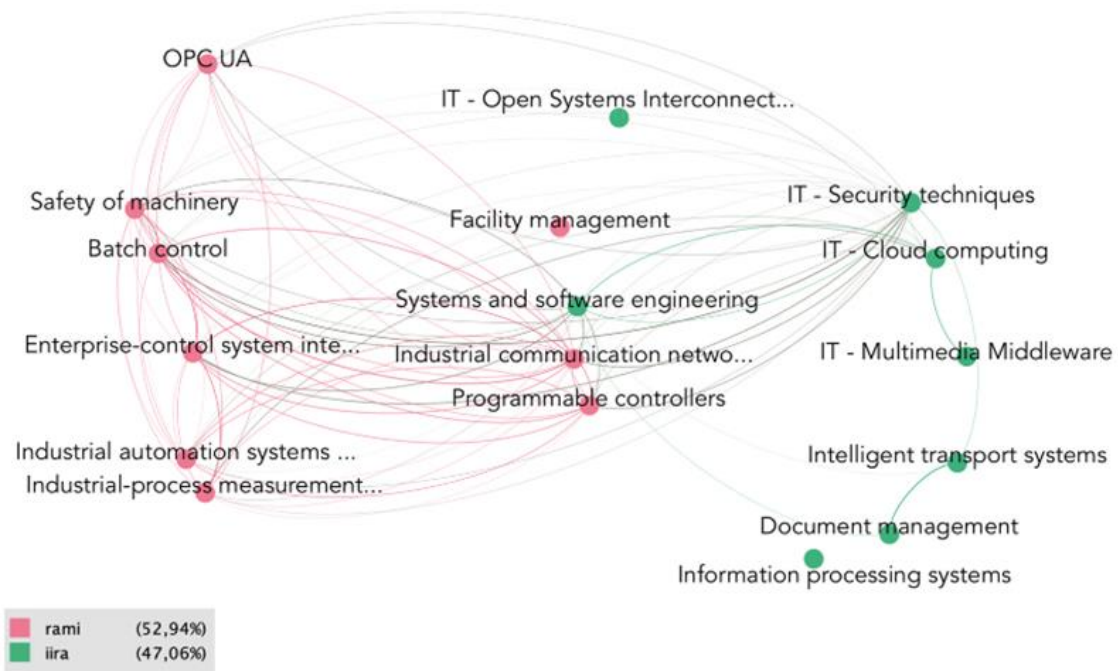
¹⁴² <https://github.com/i40-Tools/RAMIOntology> (Zugriff am: 10.06.2023)

verwendeten Begriffen berechnet. Dabei wird jeder Begriff des einen Standards mit jedem anderen Begriff aus den anderen Standards verglichen. Mit den Definitionen der Begriffe wird analog vorgegangen. Um anschließend die Ähnlichkeit zwischen je zwei Standards zu ermitteln, bilden die Autoren das arithmetische Mittel aller zuvor ermittelten Ähnlichkeiten der Begriffe. Dieser Schritt wird nicht weiter begründet und ist aus Sicht des Autors stark zu hinterfragen. Vorausgesetzt, das mittels BERT ermittelte Embedding ist korrekt, liefert dieser Schritt nur eine Aussage darüber, wie nahe die Begriffe und Definitionen der unterschiedlichen Standards (im n-dimensionalen Vektorraum) im Durchschnitt beieinander liegen.¹⁴³ Es sagt jedoch nichts darüber aus, ob es ein Alignment zwischen den Standards gibt und konsequenterweise, ob die Standards semantische Interoperabilität zulassen.¹⁴⁴

Ausgehend von dieser Ähnlichkeitsanalyse erstellten die Autoren eine Darstellung, die die Nähe zwischen den Standards visualisiert (siehe Darstellung 50). Hierbei ist nach Aussage der Autoren erkennbar, dass die von RAMI4.0 verwendeten Standards semantisch weit von jenen entfernt sind, die mit IIRA verknüpft sind und nur wenige Standards von beiden referenziert werden. Dies zeige, dass die unterschiedlichen Referenzarchitekturmodelle einen anderen Ausschnitt der Realität abzubilden versuchen.

¹⁴³ Selbst diese Information ist nur bedingt aussagekräftig. Wie man bereits aus dem Schulunterricht weiß, ist das arithmetische Mittel anfällig für Ausreißer. D.h. eine sehr schlechte Note sorgt dafür, dass die Durchschnittsnote stark verändert wird, auch wenn die restlichen Noten gut sind. Übertragen auf den vorliegenden Kontext kann man folgendes Beispiel konstruieren: Gegeben seien drei Standards A, B und C, wobei B Teil von A ist, d.h. A beinhaltet alle Begriffe und Definitionen von B und zusätzlich noch weitere. C umgekehrt ist sehr ähnlich zu B. Wenn die Begriffe in A, die nicht in B sind, nun weit entfernt (im Sinne des Embeddings) von den anderen Begriffen von A bzw. B sind, so führt dies dazu, dass bei der oben beschriebenen Methode eine stärkere Verknüpfung zwischen B und C als zwischen A und C ermittelt wird, obwohl A mindestens genauso viele Beziehungen zu C hat wie B.

¹⁴⁴ Nur weil Begriffe nahe beieinander liegen, bedeutet dies nicht, dass es ein sinnvolles Mapping zwischen diesen geben muss. In Anlehnung an das Beispiel der vorigen Fußnote würde sogar herauskommen, dass B eher ein Alignment zu C aufweist als A, da B „näher“ an C ist als A. Da B aber Teil von A ist, gibt es zwangsläufig mindestens genauso viele Mappings zwischen A und C wie zwischen B und C.



Darstellung 50: Zuordnung von Standards zu den Referenzarchitekturmodellen RAMI4.0 und IIRA.

Beschreibung: Jeder Knoten entspricht einem Standard. Die Kanten zeigen die Verbindungen zwischen den Standards, wobei nur Verbindungen angezeigt werden, die einen Wert höher als 0.4 haben. Je weiter die Knoten entfernt sind, desto geringer ist ihre semantische Ähnlichkeit. Die Farbe der Knoten gibt Auskunft, ob die Standards RAMI4.0 (rot) oder IIRA (grün) zuzuordnen sind.

Quelle: Melluso; Grangel-González und Fantoni (2022, S. 10)

Auch wenn der I40KG und die Erweiterung von Melluso, Grangel-González und Fantoni noch einige Einschränkungen aufweisen,¹⁴⁵ so ist zumindest die Intention, die Beziehungen der relevanten IIoT-Standards systematisiert zu analysieren, ein wichtiger erster Schritt, der in weiterer Folge zu Anpassungen bzw. Angleichungen von Standards führen könnte, um semantische Interoperabilität und damit eines der wichtigen Ziele hinter dem Begriff IIoT zu erreichen.

¹⁴⁵ Neben den oben erwähnten sind weitere von den Autoren genannte Einschränkungen, dass 1) die Übertragung von PDF in eine Ontologie fehlerbehaftet sein kann, da beispielsweise Begriffe übersehen werden oder irrelevante Begriffe berücksichtigt werden, 2) der Ansatz von NLP/BERT auch aufgrund der Komplexität der natürlichen Sprache Fehler hervorrufen kann und 3) Melluso, Grangel-González und Fantoni nur 50 Standards betrachtet haben, die nur von zwei Organisatoren (ISO, IEC) stammen, welche für gewöhnlich in der Entwicklung ihrer Standards kooperieren. All dies beeinflusst die Aussagekraft der Ergebnisse. (Melluso; Grangel-González; Fantoni 2022, S. 12)

5 Aktuelle Herausforderungen und Kritik

5.1 Aktuelle Herausforderungen

In Abschnitt 4 wurden mehrere Ansätze beschrieben, um dem Problem der semantischen Interoperabilität im Kontext der IIoT Herr zu werden. Jedoch muss dabei betont werden, dass es nur Ansätze sind, die entweder gar nicht oder nur in einzelnen Pilotprojekten praktische Erprobung fanden. Eine über Pilotprojekte hinausgehende Lösung hat sich bisher noch nicht gezeigt. Dies hat mehrere Gründe, die im Folgenden genau betrachtet werden sollen.

Eines der größten Probleme scheint der Mangel an Ontologien zu gängigen IIoT-Standards zu sein. In diesem Zusammenhang kritisiert Grangel-González (2019, S. 33) weiters, dass in den Standards weder potentielle semantische Heterogenität zu anderen Standards und daraus folgende Konflikte erwähnt werden, noch besprochen wird, wie Ontologien auf Basis des jeweiligen Standards erstellt werden sollen, was zur Folge haben könnte, dass es für denselben Standard mitunter unterschiedliche Ontologien gibt.

Bezüglich des Mangels an Ontologien zu den jeweiligen Standards wurden in den letzten Jahren Aktivitäten gesetzt. So gibt es beispielsweise eine Ontologie für Teile des AutomationML-Standards (Kovalenko u.a. 2018) sowie den I40KG und Methoden, aus Standards automatisiert Ontologien zu ermitteln (siehe Abschnitt 4.5). Auch das OntoCommons-Projekt (siehe Abschnitt 4.2.4) hat sich zum Ziel gesetzt, Standards durch Ontologien zu formalisieren (Adamovic u.a. 2022, S. 19–20). Parallel dazu gibt es auch Initiativen internationaler Normierungsinstitute, Standards ganz allgemein maschinenlesbar zu machen. Zu nennen sind hier beispielsweise das Smart Standards Projekt (CENELEC 2022) zusammen mit Arbeiten der Strategic Group SG 12 der International Electrotechnical Commission (IEC) und der International Organisation of Standards (ISO) (de Ribaupierre u.a. 2021, S. 1, 16; ISO 2022). Auch wenn über deren technische Umsetzungen noch nicht viel bekannt ist, so scheint klar zu sein, dass Ontologien und die Technologien des Semantic Web hier eine große Rolle spielen werden. (DIN; DKE 2023, S. 37, 92–93)

Ein weiteres Problem liegt am Fehlen umfangreicher, öffentlich zugänglicher Alignment-Datenbanken zwischen IIoT-Ontologien. Zwar werden in den in Abschnitt 4 beschriebenen Arbeiten Verbindungen zwischen Ontologien identifiziert, doch werden diese nicht auf Basis von öffentlich zugänglichen und maschinenlesbaren Alignments dokumentiert, wie es beispielsweise in anderen Bereichen der Fall ist.¹⁴⁶ Ein erster Ansatz hierzu dürfte das IndustryPortal¹⁴⁷ des OntoCommons-Projektes (siehe Abschnitt 4.2.4) sein. Diese Websites listet IIoT-Ontologien auf und gibt auch rudimentäre Mappings an.¹⁴⁸ Auch der I40KG und die Erweiterung von

¹⁴⁶ Ein Beispiel aus dem Bereich der Naturwissenschaften ist der Ontology Xref Service, der Auskunft über Mappings zwischen Konzepten unterschiedlicher Ontologien aus dem Bereich der Naturwissenschaften gibt. („EMBL-EBI Ontology Xref Service (OxO)“ 2023)

¹⁴⁷ <http://industryportal.enit.fr> (Zugriff am: 16.06.2023)

¹⁴⁸ Mit Stand Mai 2023 sind die Mappings nur mithilfe identer URI oder mithilfe von LOOM erzeugt worden (siehe Abschnitt 3.6.1). Da das verantwortliche Projekt (<https://www.ontocommons.eu> (Zugriff am: 16.06.2023)) erst Ende 2023 abgeschlossen sein wird, ist dies womöglich nicht der finale Stand.

Melluso, Grangel-González und Fantoni (siehe Abschnitt 4.5) kann als weiterer Schritt in diese Richtung verstanden werden.

Mit den fehlenden Alignments eng verbunden ist der Mangel an Benchmarks für durchgeführte Alignments. Dadurch mangelt es an einer Vergleichbarkeit unterschiedlicher Lösungen, was – insbesondere vor dem Hintergrund der großen Menge an Literatur, die zu diesem Thema produziert wird – erschwert, Fortschritte in der Entwicklung der Alignments im Bereich der IIoT zu entdecken. (Javed u.a. 2022, S. 7,17) Somit zeigt sich, dass die von Euzenat und Shvaiko (2013, S. 1–2) vorgebrachten Herausforderungen im Erstellen von Alignments (siehe Abschnitt 3.7) zumindest im Bereich der IIoT nach wie vor aktuell sind.

Erschwerend zu all diesen genannten Themen kommt noch hinzu, dass viele IIoT-Standards bzw. -Ontologien nicht öffentlich zugänglich sind. (Lelli 2019, S. 11)

Eine weitere Hürde in der Lösung der Interoperabilität ist die Nicht-Existenz einer einheitlichen Begrifflichkeit bzw. einer Meta-Ontologie für den Bereich der Ontologien und des Ontologie Matchings. Wie in den Abschnitten 2.2.4 und 3.2 dargelegt, gibt es viele ähnliche Begriffe mit meist nur leicht unterschiedlicher oder ambiger Bedeutung. Dies führt dazu, dass sich einzelne Forschungsblasen bilden können, die zwar alle dasselbe Problem lösen wollen, jedoch aufgrund unterschiedlicher Begrifflichkeiten nichts voneinander mitbekommen. Mitunter scheint es so, als liefe die Forschung zur Problematik der semantischen Interoperabilität im Bereich der IIoT in dasselbe Dilemma wie die Naturwissenschaften vor einigen Jahrzehnten. Ironischerweise wurde diese Problematik im Bereich der Naturwissenschaften mithilfe von Ontologien gelöst (siehe Abschnitt 2.2.6.1.1).

Bei all diesen Argumenten, die für eine systematische Verwendung von Ontologien und Ontologie Alignments sprechen, ist es aber genauso wichtig, einen Fokus auf das richtige Erstellen von Ontologien und damit auf das Ontologie Engineering zu legen. Li u.a. (2019, S. 5–6) betonen hierbei den Kompromiss zwischen Umfang der Ontologie und ihrer Verwendbarkeit. Je umfangreicher und domänenspezifischer die Konzepte sind, desto stärker wird ihre Verwendbarkeit eingeschränkt. Aus diesem Grund gibt es einen Trend zur Verwendung von Top-Level und Mid-Level Ontologien wie beispielsweise die oneM2M Base Ontologie oder die Core Ontologie aus dem VICINITY Projekt, die nur eine eingeschränkte Anzahl von grundlegenden Konzepten beschreiben und dadurch in unterschiedlichen Kontexten Anwendung finden können.

Diese Ontologien kommen aber nicht ohne die Verknüpfung zu spezifischeren Ontologien wie beispielsweise SSN, SAREF bzw. domänenspezifischen Erweiterungen aus. Li u.a. argumentieren demzufolge für einen stärkeren Fokus auf den modularen Aufbau von Ontologien bzw. Ontologie Integrationen, um die Nutzbarkeit dieser zu erhöhen. Anwender müssten hierbei nur die Grundkonzepte berücksichtigen und jene Teile der Ontologie, die sie für ihre Anwendung benötigen. SSN und SOSA sind bereits Beispiele für solche Ontologien (siehe Abschnitt 2.2.6.2.1). Die in Abschnitt 4.2.24.2 erwähnten Ontologien des VICINITY und des INTER-IoT Projektes sowie das OSF-Framework (Mayer u.a. 2017) (siehe Darstellung 40 auf Seite 79) sind ebenfalls passende Beispiele für einen modularen Aufbau.

Neben der erhöhten Nutzbarkeit hat ein modularer Aufbau nicht zuletzt Vorteile in der Ermittlung von Alignments, wie Jimenez-Ruiz u.a. zeigen. (Jiménez-Ruiz u.a. 2020, S. 7) Umgekehrt

sind Ontologie Alignments selbst eine Methode, Modularität zu ermöglichen. Bisherige Ansätze beruhen meist auf dem Import eines Moduls (oder einer ganzen Ontologie) in eine andere Ontologie. So importiert SSN beispielsweise SOSA. Auch die Common Core Ontologien importieren sich gegenseitig (siehe Darstellung 9 auf Seite 31). Das Problem bei dieser Methode ist einerseits, dass immer das gesamte Modul integriert wird, auch wenn nur einzelne Konzepte benötigt werden, andererseits ist das Importieren eine gerichtete (unidirektionale) Abbildung. Das bedeutet, dass nur jene Ontologie, in die importiert wird, zusätzliches Wissen in Form des importierten Moduls erhält. Das importierte Modul wird dabei nicht um weiteres Wissen angereichert. Demgegenüber sind Ontologie Alignments mitunter bidirektional¹⁴⁹ und die integrierte Ontologie benötigt nur jene Konzepte, für die ein Mapping bzw. eine Abhängigkeit zu gemappten Konzepten besteht. (Haller u.a. 2017)

Nichtsdestotrotz betonen Nilsson und Sandin, dass zwar beinahe alle Ansätze, das Problem der semantischen Interoperabilität zu lösen, über Ontologien führen, bisher jedoch nur Proof of Concepts umgesetzt wurden. Darüber hinaus stellen sie infrage, ob die technischen Lösungen, die Ontologien verwenden, überhaupt skalierbar sind für eine Anwendung in Fabriken mit tausenden CPS. (Nilsson; Sandin 2018, S. 130–131) Auch fünf Jahre nach Nilssons und Sandins Kommentar sind die Ansätze noch nicht über einzelne Leuchtturmprojekte hinausgekommen, wie in Abschnitt 4 gezeigt wurde.

5.2 Kritik

Neben den Problemen, die bei der Lösung der semantischen Interoperabilität nach wie vor bestehen, muss aus Sicht des Autors auch Kritik an den derzeitigen Methoden des Ontologie Matchings angebracht werden. Diese Methoden richten sich weniger auf ein Verstehen der Konzepte als auf abstrakte statistische Verfahren, die schnell eine solche Abstraktionsebene erreichen, dass sie das Ontologie Matching zu einer Blackbox werden lassen, die es selbst Experten schwer oder unmöglich macht, die Rechenwege zu durchblicken. Das Beispiel der Ontologie Embeddings, wie es in Abschnitt 3.5.2 und 4.4 vorgestellt wurde, ist aus heutiger Sicht wohl der Höhepunkt in einem Bereich, der gerade erst im Entstehen ist.¹⁵⁰ Hierbei erfolgt „Verstehen“ auf abstrakten Machine Learning Verfahren, die mit tatsächlichem menschlichem Verstehen meist nur wenig zu tun haben. Die Verwendung dieser Methoden reiht sich nahtlos in den aktuellen Zeitgeist ein, der den Aufstieg sogenannter generativer künstlicher Intelligenz wie ChatGPT oder Bard feiert, ohne potentielle Gefahren dahinter zu sehen, Ergebnisse kritisch zu hinterfragen oder mitunter einfachere, aber derzeit weniger modern erscheinende Methoden vernachlässigt.¹⁵¹ Die Arbeit von Javed u.a. (2022) ist ein eindrückliches Beispiel

¹⁴⁹ Wichtig zu betonen ist, dass dies nicht immer erfüllt sein muss. Im Gegensatz zur Import-Methode besteht hier jedoch zumindest die Möglichkeit dafür.

¹⁵⁰ Nach Aussage von Javed u.a. (2022, S. 16) ist ihre Arbeit die erste, die NLP-Methoden zur Ermittlung von Ontologie Alignments im Kontext der IIoT anwendet.

¹⁵¹ Dem Autor ist es dabei wichtig zu betonen, dass dies nicht als Kritik der Methode des Embeddings bzw. der KI- oder Maschine Learning-Verfahren im Allgemeinen aufgefasst wird, sondern als Kritik an der unreflektierten Verwendung dieser Methoden und der Vernachlässigung der Limitationen, die diese Methoden mit sich führen. Dies wurde auch versucht, in Abschnitt 4.4 zu zeigen.

dafür, wie eine komplexe Methode für eine triviale Aufgabe (nämlich das Matching von SOSA auf SSN) verwendet wurde und sollte aus Sicht des Autors eine Warnung sein, Problemstellungen zuerst gründlich zu überdenken, bevor mit Methoden, die gerade en vogue sind, darauf losgegangen wird.

Erfreulich scheint in diesem Zusammenhang der Ansatz des OntoCommons-Projektes zu sein, das explizit Ontologie-basiertes Schlussfolgern gegenüber generativer KI vorzieht. (Adamovic u.a. 2022, S. 4)

Bei all dieser Kritik muss selbstverständlich erwähnt werden, dass generative KI gerade dabei ist, analytische Methoden in der Ermittlung von Alignments zu überholen und deshalb in der generativen KI mit großer Wahrscheinlichkeit ein Schwerpunkt in der zukünftigen Forschung liegen wird. (He u.a. 2022, S. 5688–5690) Jedoch trifft diese Aussage auf Alignments im Bereich der IIoT mitunter nicht zu. So sehen Javed u.a. (2022, S. 7–8) das Problem von zu geringen Datensätzen im Bereich der IIoT. Sprachmodelle werden in der Regel mithilfe von riesigen Datenmengen trainiert. IIoT-Ontologien sind mitunter sehr klein. So hat die SAREF-Ontologie beispielsweise nur 37 Entitäten, 20 Relationen, sowie 1097 RDF-Triple, was im Vergleich zu Datensätzen natürlicher Sprachen ein Bruchteil ist.¹⁵² Dementsprechend ist es im Kontext der IIoT eine große Herausforderung, die Modelle so zu trainieren, dass sie passable Ergebnisse liefern. Umgekehrt wirft die geringe Größe von IIoT-Ontologien auch die Frage auf, ob es die Methode der Embeddings überhaupt braucht, um eine relativ überschaubare Menge an Mappings zu überprüfen.

Egal, welche Methode angewendet wird, eine Gefahr bleibt schlicht weg darin bestehen, dass Fehler in Mappings nicht auffallen. Mit solchen Fehlern ist allein schon deshalb zu rechnen, da unterschiedliche Ontologien für unterschiedliche Zwecke und Zielgruppen erstellt wurden. Ein weiterer Grund für potenzielle Probleme wurde in Abschnitt 3.1.2 mit der sogenannten „double articulation“ erwähnt.

Eine aus Sicht des Autors viel einfachere Methode, semantische Interoperabilität bzw. die Verbindung von Ontologien zu unterstützen, bei der auch Probleme wie die double articulation verhindert werden könnten, wäre der Austausch der Gremien, die für die unterschiedlichen Standards verantwortlich sind. Würden bereits bei der Entwicklung der Standards Schnittstellen und damit Mappings definiert, bräuchte man eine meist auf Algorithmen und brute force beruhende, nachträgliche Berechnung von Mappings erst gar nicht. Da zusätzlich auch die Expertinnen und Experten direkt beim Erstellen der Mappings beteiligt wären, ist auch die Gefahr geringer einzuschätzen, fehlerhafte Alignments zu erstellen.

Erfreulicherweise gibt es auch hier Initiativen auf unterschiedlichen Ebenen. Zu erwähnen ist beispielsweise, dass sich die Europäische Union der Bedeutung von Standards gewahr geworden ist, was sich in der Standardisation Strategy¹⁵³ oder in Initiativen wie dem Standardisation Booster¹⁵⁴ zeigt. Auch in Deutschland scheint man sich der Problematik bewusst zu

¹⁵² So verfügt beispielsweise der *DBP15K_{FR-EN}*, ein Datensatz aus DBpedia von französisch- und englisch-sprachige begriffen über 66.858 Entitäten, 1.379 Relationen und 192.191 Tripel in englischer Sprache. (Javed u.a. 2022, S. 8)

¹⁵³ https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13099-Standardisation-strategy_en (Zugriff am 06.05.2023)

¹⁵⁴ <https://www.hsbooster.eu> (Zugriff am 06.05.2023)

sein. So erwähnt die Normungsroadmap Industrie 4.0 in ihrer aktuellen Version 5 das Einführen eines Standards für eine gemeinsame I4.0-Sprache, semantische Anreicherungen bestehender Standards zur Verbesserung der Interoperabilität als auch Qualitätskriterien für Ontologien, um beispielsweise Homonyme und Synonyme zu vermeiden bzw. Terminologien zu harmonisieren¹⁵⁵. (DIN; DKE 2023, S. 20, 49–50, 94)

5.3 Zukünftige Handlungsfelder

Aus den in Abschnitt 5.1 und 5.2 besprochenen Problemen können aus Sicht des Autors folgende vier Handlungsfelder abgeleitet werden, auf denen zukünftig der Fokus liegen sollte:

1. Ein Ausbau und eine Harmonisierung maschinenlesbarer Standards zusammen mit einer Datenbank von Alignments zwischen den Ontologien der Standards
2. Ein Berücksichtigen von regel-basierten Alignments in der Architektur von Anwendungen, ähnlich wie es der INTER-IoT-Ansatz (siehe Abschnitt 4.2.3) vorgeschlagen hat, falls die Alignments nicht durch Punkt 1 abgedeckt sind
3. Eine zusätzliche Integration von Embeddings und Machine Learning Algorithmen bzw. NLP (wie beispielsweise dem Sprachmodell BERT), um Alignments zu verbessern bzw. Ontologien zu ergänzen¹⁵⁶, falls dies nicht durch Punkt 1 abgedeckt ist
4. Das Erstellen von Benchmarks für Alignments von IIoT-Ontologien, analog zur OAEI, um Methoden und Tools zu identifizieren, die sich für diesen Bereich eignen

Die Kombination aus regel-basierten Methoden und Embeddings scheint insbesondere bei sich stark ändernden Systemen sinnvoll zu sein, da sich letztere gerade für unvorhergesehene Störungen besser eignen, während erstere die Standardfälle gut abdecken können. (McLean; Walker; Bright 2017, S. 697–698) Der vierte Punkt fordert nichts weniger als das systematische Evaluieren von Alignments sowie der Methoden und Softwaretools, die benötigt werden, diese Alignments zu erstellen. Dies ist insbesondere deshalb relevant, da die bisherigen Ergebnisse der OAEI darauf schließen lassen, dass die Performance von Methoden und Softwaretools stark davon abhängig ist, welche Domänen miteinander verglichen werden. Des Weiteren liegen IIoT-spezifische Ontologien bisher nicht im Fokus der OAEI.

Der wichtigste Punkt bleibt jedoch der erstgenannte. Nur durch die aktive Kommunikation der beteiligten Personen und Institutionen können die in Abschnitt 3.1.2 besprochenen Arten der Heterogenität behoben werden.

Zuletzt wäre es sinnvoll, wenn es zu den einzelnen Projekten auch ein übergeordnetes Programm- und Portfoliomanagement gäbe. Dies ist insbesondere naheliegend, da viele der hier präsentierten Projekte (VICINITY, INTER-IoT, OntoCommons) von der EU finanziert wurden und zeitlich nah aufeinander bzw. überlappend durchgeführt wurden. Dennoch konnte der

¹⁵⁵ Zur Unterstützung der Harmonisierung gibt es bereits Assistenzsysteme wie Harbsafe 2 (<https://www.dke.de/de/arbeitsfelder/cybersecurity/harbsafe2> (Zugriff am 06.05.2023))

¹⁵⁶ Stichwort: Knowledge Graph Completion (siehe auch Fußnote 74 auf Seite 59)

Autor keine Referenzen zwischen den Projekten entdecken, obwohl sich alle mit CPS und semantischer Interoperabilität beschäftigen. Der fehlende Wissenstransfer zwischen den einzelnen Projekten birgt nicht nur die Gefahr, dass sich wichtige Entwicklungen verzögern bzw. Projekte scheitern (mitunter mehrere aus denselben Gründen), sondern ist in Konsequenz auch eine unnötige Verschwendung von Steuergeld. Ein strategisches und ökonomisches Vorgehen wäre wünschenswert für die Zukunft.

6 Schlussbetrachtung

6.1 Zusammenfassung und Interpretation der Ergebnisse

Auf Basis der in dieser Arbeit zusammengetragenen Informationen lassen sich die in 1.1 erwähnten Forschungsfragen wie folgt beantworten:

FF1: Wie unterstützen Ontologien und das Ontologie Matching die semantische Interoperabilität?

Ontologien sind eine Möglichkeit, Wissen abzubilden, das sowohl für Menschen als auch für Maschinen lesbar ist. Demzufolge ist es naheliegend, sie für den Bereich der IloT zu verwenden. Mittels Methoden des Ontologie Matchings (siehe Kapitel 3) kann das Problem der semantischen Heterogenität gelöst und dadurch Interoperabilität ermöglicht werden. Hierzu wurden in Abschnitt 3.1 zuerst Arten semantischer Heterogenität besprochen. Anschließend beschäftigten sich die Abschnitte 3.2 bis 3.8 mit dem Begriff des Ontologie Matchings, grundlegenden Methoden und Softwaretools zur Auflösung semantischer Heterogenität sowie offenen Herausforderungen in diesem Bereich. Dabei wurde nicht nur besprochen, wie Alignments zwischen Ontologien erzeugt werden, sondern auch, wie in weiterer Folge ein erfolgreicher Datenaustausch stattfinden kann (siehe Abschnitt 3.3).

FF2: Welche aktuellen Methoden, Anwendungen und Problemstellungen gibt es im Bereich der IloT zur Unterstützung semantischer Interoperabilität?

Aus den in Kapitel 4 besprochenen Ansätzen scheinen sich vier Trends zu etablieren:

1. Hybride Alignment Architekturen unter Verwendung von Top-Level und Mid-Level Ontologien
2. Modularer Aufbau von Ontologien
3. Erstellen von Ontologien zu Standards und von maschinenlesbaren Standards
4. Verwendung von Wissensgraph Embeddings.

Die hybride Alignment Architektur zeigt sich in einer Reihe von Initiativen, die mittels Top-Level und Mid-Level Ontologien versuchen, domänenspezifische Ontologien zu verbinden.

Damit einher geht der Trend, die Ontologien selbst (z.B. SSN und SOSA) aber auch die Kombination der verwendeten Ontologien modular aufzubauen. Dies bietet nicht nur eine höhere Flexibilität und damit eine gesteigerte Anwendbarkeit, sondern hat auch Vorteile im Ermitteln von Alignments und im Umgang mit Änderungen von einzelnen Modulen. Beispiele hierfür wurden in Abschnitt 4.2 und 4.3 vorgestellt.

Um IloT-Standards miteinander zu verbinden, benötigt es jedoch Ontologien. Hierfür konnten neben Einzelinitiativen, die zu ganz bestimmten Standards Ontologien entwickelt haben

(Abschnitte 4.3.2) auch Initiativen von Wissenschafts- und Normungsorganisationen gefunden werden, deren Ziel es ist, Standards maschinenlesbar zu machen (Abschnitte 4.5 und 5.1).

Der vierte Trend ist die Verwendung von Embeddings und Sprachmodellen wie BERT, der erst kürzlich Einzug in den Bereich der IIoT gehalten hat, bei dem jedoch davon auszugehen ist, dass er in nächster Zeit eine bedeutende Rolle spielen wird (siehe Abschnitte 4.4 und 4.5). Ob er das Potential hat, das Problem der semantischen Interoperabilität nachhaltig und über eine Vielzahl an Bereichen der IIoT zu lösen, muss sich jedoch erst noch zeigen.

Neben diesen Trends konnte in Kapitel 5 gezeigt werden, dass es noch mannigfaltige Herausforderungen gibt, um semantische Interoperabilität zu ermöglichen. Zu nennen sind hierbei neben fehlenden Richtlinien zur Erstellung von Ontologien zu Standards auch ein Mangel an öffentlich zugänglichen Alignments und Benchmarks für Alignments, wie sie beispielsweise in den Naturwissenschaften etabliert sind.

Im Bereich der IIoT stellt sich zusätzlich die Frage, ob die derzeitigen Ressourcen korrekt eingesetzt werden. Der Verdacht liegt nahe, dass der Fokus zu sehr auf der Findung von technischen Lösungen (Stichwort „Embeddings“ und „I40KG“) liegt, als die Ursachen der Probleme semantischer Interoperabilität auf organisatorischer Ebene zu lösen. Als wichtige Handlungsfelder wurden hier die Harmonisierung von Standards und das Erstellen von Ontologien sowie Alignments von den Normungsorganisationen selbst genannt. Mit anderen Worten: Um das Problem der fehlerhaften Kommunikation zwischen Maschinen zu lösen, wäre es am sinnvollsten, wenn zuerst die Menschen bzw. die dahinterstehenden Normungsorganisationen mehr miteinander kommunizieren und die Standards aufeinander abstimmen.

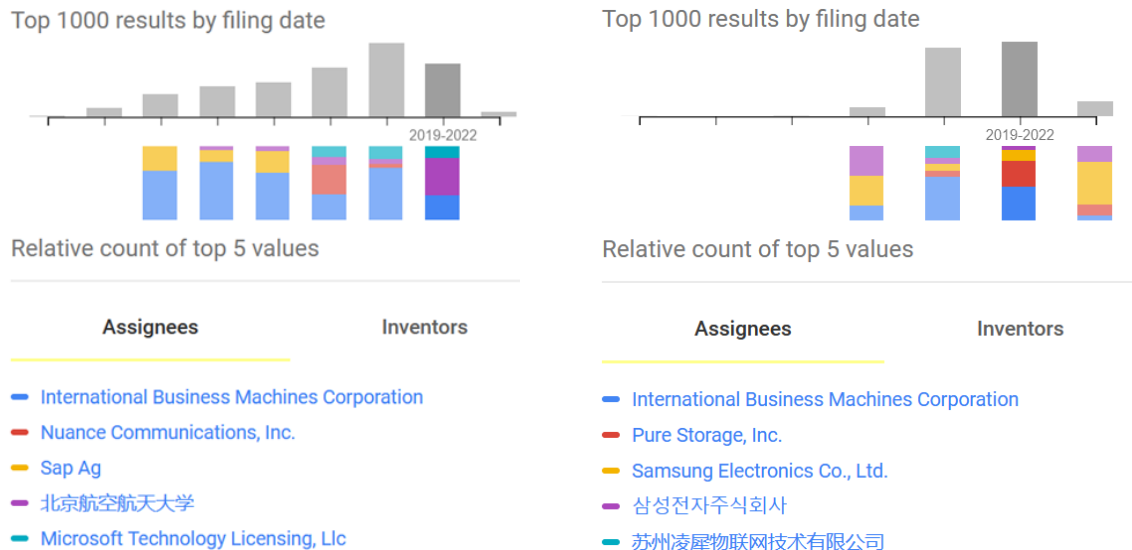
FF3: In welchen Branchen und Domänen wurden Methoden zur Unterstützung semantischer Interoperabilität in den letzten Jahren angewendet?

Die in Kapitel 4 vorgestellten Anwendungsfälle sind im Anhang in Darstellung 53 aufgelistet. Bei der Analyse zeigte sich, dass Ontologie Matchings in unterschiedlichen industriellen Bereichen angewendet werden, vom Gesundheitswesen über die Logistik und den Energiebereich bis zur Produktion. Die dabei präsentierten Lösungsansätze erwähnen keine Gründe, warum deren Methoden nicht auf andere Domänen anwendbar sind. Problematisch dabei scheint vor allem der hohe Aufwand der Umsetzungen – nicht zuletzt auch deshalb, weil in vielen Fällen Ontologien sowie passende Ontologie Alignments erst erstellt werden müssen und nicht nur die semantische Interoperabilität betrachtet wurde, was die Komplexität stark erhöht. Insgesamt kann derzeit kein Best Practice Ansatz identifiziert werden. Viel eher werden in der derzeitigen Forschung unterschiedliche Methoden in den verschiedensten Domänen getestet. Eine Übertragung der Methoden auf weitere Domänen findet gar nicht oder nur in geringem Maße statt. Aus Sicht des Autors ist dies ein starkes Indiz dafür, dass es noch einige Jahre dauern wird, bis es einfach implementierbare und generische Lösungsansätze für das Ermöglichen der (semantischen) Interoperabilität im Bereich der IIoT gibt.

Hazra u.a. (2023, S. 21) sprechen hierbei vom Ziel, „interoperability as a service“ zu ermöglichen. Dies würde es erlauben, neue CPS einfach mittels Plug&Play automatisch in eine digitale Fabrik zu integrieren, genauso, wie man eine Computermaus an einen Computer anschließt. Das System stellt dabei die Interoperabilität automatisch her. Auch wenn dies mit

den in dieser Arbeit besprochenen Ansätzen auf semantischer Ebene durchaus naheliegend erscheint, so gibt es derzeit noch keine Umsetzung, die dies domänenübergreifend ohne menschliches Eingreifen schafft. (Burns; Cosgrove; Doyle 2019, S. 652)

Aus den vorliegenden Ergebnissen kann nun eine Empfehlung für industrielle Unternehmen abgeleitet werden. Die Frage „Gibt es eine bestehende, wirtschaftlich sinnvolle Lösung, um semantische Interoperabilität im Kontext der IIoT im eigenen Unternehmen zu ermöglichen?“ muss derzeit im Allgemeinen klar mit „Nein“ beantwortet werden. Zwar mag es womöglich einige wenige Unternehmen geben, die bestehende Ansätze 1:1 aufgreifen können. Im Allgemeinen sind diese für das Gros an Unternehmen weder generisch genug und damit nicht adaptierbar auf deren individuelle Situation, noch scheinen sie nachhaltig genug zu sein, um auf veränderte Gegebenheiten, wie zusätzliche oder sich verändernde Ontologien in CPS, erfolgreich reagieren zu können. So konnte in dieser Arbeit gezeigt werden, dass es derzeit weder einen Best Practice Ansatz noch einen systematischen Vergleich der bestehenden Ansätze oder eine Methode, den für ein Unternehmen geeignetsten Ansatz zu identifizieren, gibt. Es bleibt somit abzuwarten, ob sich in naher Zukunft die Idee von „interoperability as a service“ verwirklichen lässt. Dass hier mit einem Fortschritt zu rechnen ist, zeigt die Menge an Patenten, die in den letzten Jahren eingereicht wurden (siehe Darstellung 51). Hierbei ist interessant, dass jene Unternehmen, die die meisten Patente halten, wie IBM, Microsoft oder SAP, nach Kenntnis des Autors bisher noch wenig veröffentlicht oder wirtschaftlich (erfolgreich) umgesetzt haben. Dies könnte ein Indiz dafür sein, dass es bisher noch keinem Player gelungen ist, ein erfolgreiches Produkt zusammenzustellen, um die semantische Interoperabilität im Kontext der IIoT zu lösen.



Darstellung 51: Aufteilung der Patente nach Jahren und Unternehmen

Beschreibung: Die linke Graphik zeigt die Aufteilung der Patente für Ontologie Matching, die mithilfe des Suchausdruckes '(("Knowledge Graph" OR "Ontology" OR "Ontologie" OR "Wissensgraph") AND (Alignment OR Mapping OR Matching OR Integration))' gefunden werden konnten. Die rechte Graphik schränkt die Suche weiter auf den Kontext IIoT bzw. IoT ein. Der Suchausdruck lautete dabei '(("IIoT" OR "IoT" OR "Internet of Things" OR "Industry 4.0") AND ("Knowledge Graph" OR "Ontology" OR "Ontologie" OR "Wissensgraph") AND (Alignment OR Mapping OR Matching OR Integration))'.

Quelle: <https://patents.google.com> (Zugriff am: 16.06.2023)

6.2 Limitationen

Die vorliegende Arbeit betrachtete Lösungsansätze, um semantische Interoperabilität im Kontext der IIoT mithilfe von Ontologien zu ermöglichen. Nur angeschnitten wurden dabei weitere Ebenen der Interoperabilität (siehe Abschnitt 2.1.4), was jedoch nicht bedeutet, dass diese weniger relevant für die Umsetzung der IIoT sind. Ganz ausgelassen wurden Ansätze, die semantische Interoperabilität ohne Ontologien zu lösen versuchen.

Unbehandelt blieben in dieser Arbeit auch Aspekte der Daten- und Informationssicherheit sowie des Wissensmanagements. In Bezug auf die Daten- und Informationssicherheit ist es zwar wichtig, dass sich die Systeme untereinander verstehen, ebenso wichtig ist aber auch, dass bei einer Anfrage verstanden wird, ob diese überhaupt an den Anfragesteller beantwortet werden darf. (Miller; Lin 2016, S. 11) Für das Wissensmanagement sind Ontologien eine hilfreiche Technologie, um Wissen abzulegen, zu verknüpfen und – mittels Knowledge Graph Completion – auch zu erzeugen. Analog zur Verwendung der DBpedia Ontologie als maschinenlesbarer Spiegel der Wikipedia, wäre es interessant herauszufinden, welche Potentiale Ontologien zu bestehenden Unternehmenswikis haben.

Die in Abschnitt 5 aufgezählten Ergebnisse beruhen auf der Analyse einer eingeschränkten Menge an Forschungsarbeiten zur semantischen Interoperabilität im Kontext der IIoT. Die hierbei gewählten Selektionskriterien wurden zu Beginn von Abschnitt 4 erwähnt. Es besteht somit die Möglichkeit, dass die erwähnten Ergebnisse eine verzerrte Sicht auf den tatsächlichen Stand der Forschung wiedergeben und keine Allgemeingültigkeit besitzen.¹⁵⁷

¹⁵⁷ Nicht berücksichtigt wurden beispielsweise die „Distributed Ontology, Model, and Specification Language“ (DOL) der OMG (<https://www.omg.org/dol/>) oder die „Expressive and Declarative Ontology Alignment Language“ (EDOAL) (<https://ns.inria.org/edoal/1.0>) (Zugriff jeweils am: 24.06.2023), welches beides Sprachen sind, um Alignments abzubilden.

6.3 Schlusswort

Abschließend ist zu betonen, dass Ontologien und Ontologie Alignments bei Weitem nicht auf den Bereich der IIoT beschränkt sind. Im Kern sind sie Teil jedes Kommunikationsprozesses, sei es, wenn Menschen desselben oder unterschiedlicher Kulturkreise miteinander kommunizieren oder wenn Informationen aus unterschiedlichen Datenbanken miteinander verknüpft werden müssen. Das Besondere ist, dass sich durch das Bewusstwerden, dass Ontologien und mitunter fehlerhafte Ontologie Alignments¹⁵⁸ hier im Spiel sind, ganz neue Sichtweisen, Methoden und Technologien zur Lösung der Kommunikationsprobleme eröffnen.¹⁵⁹ In diesem Sinne wäre es ein nicht ganz ernst zu nehmendes, aber durchaus interessantes Unterfangen, das Kommunikationsproblem des in Abschnitt 1.1 erwähnten Werbespots (der auch im folgenden QR-Code verlinkt ist) aus ontologischer Sicht zu betrachten.



Darstellung 52: Link auf Beispiel zu misslungener Kommunikation

¹⁵⁸ Gemeint ist hier das Phänomen des „Aneinander-vorbei-Redens“.

¹⁵⁹ Dies ist aus Sicht des Autors durchaus vergleichbar mit dem Kennenlernen von Kommunikationsmodellen wie beispielsweise dem 4-Ohren-Modell von Friedemann Schulz von Thun (1995, S. 33).

Literaturverzeichnis

Die Quellenangaben und das Literaturverzeichnis dieser Arbeit wurden mithilfe des Literaturverwaltungsprogrammes Zotero unter Verwendung des Zitierstils der FH Vorarlberg erstellt.

- Adamovic, Nadja u.a. (2022): OntoCommons Roadmap V1. Zenodo. Online im Internet: DOI: 10.5281/ZENODO.7544509 (Zugriff am: 25.04.2023).
- Adolphs, Peter u.a. (2015): „Reference architecture model industrie 4.0 (rami4.0).“ In: ZVEI and VDI, Status report, (2015).
- Amara, Fatima Zahra; Hemam, Mounir; Djeddar, Meriem; Maimor, Moufida (2022): „Semantic Web and Internet of Things: Challenges, Applications and Perspectives.“ In: Journal of ICT Standardization, (2022), S. 261–292.
- Amara, Fatima Zahra; Hemam, Mounir; Djeddar, Meriem; Maimour, Moufida (2022): „Semantic Web Technologies for Internet of Things Semantic Interoperability.“ In: Advances in Information, Communication and Cybersecurity. Herausgegeben von Yassine Maleh u.a. Cham: Springer International Publishing (= Lecture Notes in Networks and Systems), S. 133–143. Online im Internet: DOI: 10.1007/978-3-030-91738-8_13 (Zugriff am: 14.04.2023).
- Ancic, Darko (2022): Welcome to [iot.schema.org](https://github.com/iot-schema-collab/iot-schema-collab). [iot-schema-collab](https://github.com/iot-schema-collab/iot-schema-collab). Online im Internet: URL: <https://github.com/iot-schema-collab/iot-schema-collab> (Zugriff am: 19.05.2023).
- Anderson, Bailey; Barnstedt, Erich; Weinstock, Gil (2023): Adopting DTDL-based industry ontologies - Azure Digital Twins. Online im Internet: URL: <https://learn.microsoft.com/en-us/azure/digital-twins/concepts-ontologies-adopt> (Zugriff am: 26.04.2023).
- Ansari, Fazel u.a. (2018): „A problem-solving ontology for human-centered Cyber-Physical production systems.“ In: CIRP Journal of Manufacturing Science and Technology, 22 (2018).
- Arenas, Marcelo u.a. (2016): Faceted Search Over RDF-Based Knowledge Graphs. Rochester, NY: Online im Internet: DOI: 10.2139/ssrn.3199228 (Zugriff am: 17.03.2023).
- Arnold, Laurin u.a. (2022): „IIoT platforms’ architectural features – a taxonomy and five prevalent archetypes.“ In: Electronic Markets, 32 (2022), 2, S. 927–944. Online im Internet: DOI: 10.1007/s12525-021-00520-0
- AutomationML (2023): What is AutomationML? – AutomationML. Online im Internet: URL: <https://www.automationml.org/about-automationml/automationml/> (Zugriff am: 09.06.2023).
- Bajaj, Garvita u.a. (2017): A study of existing Ontologies in the IoT-domain. arXiv. Online im Internet: URL: <http://arxiv.org/abs/1707.00112> (Zugriff am: 02.12.2022).
- Bamunuarachchi, Dinithi u.a. (2021): „Digital twins supporting efficient digital industrial transformation.“ In: Sensors, 21 (2021), 20, S. 6829.
- Baumgartner, Matthias u.a. (2023): „Towards the Web of Embeddings: Integrating multiple knowledge graph embedding spaces with FedCoder.“ In: Journal of Web Semantics, 75 (2023), S. 100741. Online im Internet: DOI: 10.1016/j.websem.2022.100741

- Berners-Lee, Tim (2015): Semantic Web. Online im Internet: URL: <https://www.w3.org/2000/Talks/0906-xmlweb-tbl/text.htm> (Zugriff am: 26.11.2022).
- Boland, Michael (2020): Semantic Data Interconnect (SDI) – What’s new in 4.0.0 - Mind-Sphere. Online im Internet: URL: <https://blogs.sw.siemens.com/mindsphere/semantic-data-interconnect-sdi-whats-new-in-4-0-0/> (Zugriff am: 26.04.2023).
- Brickley, Dan (2023): IoT and Schema.org: Getting Started. Online im Internet: URL: <https://iot.schema.org/docs/iot-gettingstarted.html> (Zugriff am: 19.05.2023).
- Brink, Alfred (2013): Anfertigung wissenschaftlicher Arbeiten: ein prozessorientierter Leitfaden zur Erstellung von Bachelor-, Master- und Diplomarbeiten. 4., korr.aktualisierte Aufl. 2013. Wiesbaden: Springer Fachmedien Wiesbaden Imprint: Springer Gabler (= Anfertigung wissenschaftlicher Arbeiten).
- Burns, Thomas; Cosgrove, John; Doyle, Frank (2019): „A Review of Interoperability Standards for Industry 4.0.“ In: *Procedia Manufacturing*, 38 (2019), S. 646–653. Online im Internet: DOI: 10.1016/j.promfg.2020.01.083
- Cai, Hongyun; Zheng, Vincent W; Chang, Kevin Chen-Chuan (2018): „A comprehensive survey of graph embedding: Problems, techniques, and applications.“ In: *IEEE transactions on knowledge and data engineering*, 30 (2018), 9, S. 1616–1637.
- Cavalieri, Salvatore (2021): „Semantic Interoperability between IEC 61850 and oneM2M for IoT-Enabled Smart Grids.“ In: *Sensors*, 21 (2021), 7, S. 2571. Online im Internet: DOI: 10.3390/s21072571
- CENELEC (2022): Experts CENELEC. Smart Standards. Online im Internet: URL: <https://experts.cenelec.eu/key-initiatives/smart-standards/> (Zugriff am: 05.05.2023).
- Chandrasekaran, B.; Josephson, J.R.; Benjamins, V.R. (1999): „What are ontologies, and why do we need them?“ In: *IEEE Intelligent Systems and their Applications*, 14 (1999), 1, S. 20–26. Online im Internet: DOI: 10.1109/5254.747902
- Charpenay, Victor (2019): *Semantics for the Web of Things: Modeling the Physical World as a Collection of Things and Reasoning with their Descriptions*.
- Chen, Jiaoyan u.a. (2021): „Augmenting Ontology Alignment by Semantic Embedding and Distant Supervision.“ In: *The Semantic Web*. Herausgegeben von Ruben Verborgh u.a. Cham: Springer International Publishing (= Lecture Notes in Computer Science), S. 392–408. Online im Internet: DOI: 10.1007/978-3-030-77385-4_23 (Zugriff am: 20.05.2023).
- Choudhary, Shivani u.a. (2021): *A Survey of Knowledge Graph Embedding and Their Applications*. arXiv. Online im Internet: URL: <http://arxiv.org/abs/2107.07842> (Zugriff am: 01.06.2023).
- Cimmino, Andrea u.a. (2019): „VICINITY: IoT Semantic Interoperability Based on the Web of Things.“ In: *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. Santorini Island, Greece: IEEE, S. 241–247. Online im Internet: DOI: 10.1109/DCOSS.2019.00061 (Zugriff am: 31.12.2022).
- Corcho, Oscar; Gómez-Pérez, Asunción (2000): *A roadmap to ontology specification languages*. Knowledge Engineering and Knowledge Management Methods, Models, and

Tools: 12th International Conference, EKAW 2000 Juan-les-Pins, France, October 2–6, 2000 Proceedings 12. Springer, S. 80–96.

Cruz, Manuel; Parés, Carlos; Quintela, Peregrina (Hrsg.) (2021): Progress in Industrial Mathematics: Success Stories: The Industry and the Academia Points of View. 5. Cham: Springer International Publishing (= SEMA SIMAI Springer Series). Online im Internet: DOI: 10.1007/978-3-030-61844-5 (Zugriff am: 15.10.2022).

Cubek, Richard; Ertel, Wolfgang; Palm, Günther (2015): „A Critical Review on the Symbol Grounding Problem as an Issue of Autonomous Agents.“ In: KI 2015: Advances in Artificial Intelligence. Herausgegeben von Steffen Hölldobler; Rafael Peñaloza; Sebastian Rudolph. Cham: Springer International Publishing (= Lecture Notes in Computer Science), S. 256–263. Online im Internet: DOI: 10.1007/978-3-319-24489-1_21 (Zugriff am: 29.04.2023).

CUBRC (2023): Common Core Ontologies for Data Integratio. Online im Internet: URL: <https://www.cubrc.org/index.php/data-science-and-information-fusion/ontology> (Zugriff am: 18.05.2023).

Da Silva, Jomar u.a. (2020): „A LIN : improving interactive ontology matching by interactively revising mapping suggestions.“ In: The Knowledge Engineering Review, 35 (2020), S. e1. Online im Internet: DOI: 10.1017/S0269888919000249

Daniele, Laura u.a. (2016): „Interoperability for Smart Appliances in the IoT World.“ In: The Semantic Web – ISWC 2016. Herausgegeben von Paul Groth u.a. Cham: Springer International Publishing (= Lecture Notes in Computer Science), S. 21–29. Online im Internet: DOI: 10.1007/978-3-319-46547-0_3 (Zugriff am: 14.04.2023).

Daniele, Laura u.a. (2020): SAREF: the Smart Applications REference ontology. Online im Internet: URL: <https://saref.etsi.org/core/v3.1.1/#introduction> (Zugriff am: 19.05.2023).

Devlin, Jacob u.a. (2019): BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv. Online im Internet: URL: <http://arxiv.org/abs/1810.04805> (Zugriff am: 08.04.2023).

Dickerson, Keith u.a. (2021): „Standards for the IoT.“ In: IoT Platforms, Use Cases, Privacy, and Business Models. Herausgegeben von Carna Zivkovic; Yajuan Guan; Christoph Grimm. Cham: Springer International Publishing, S. 125–147. Online im Internet: DOI: 10.1007/978-3-030-45316-9_6 (Zugriff am: 15.04.2023).

DIN; DKE (2023): Deutsche Normierungsroadmap Industrie 4.0 - Version 5. Online im Internet: URL: <https://www.dke.de/de/arbeitsfelder/industry/die-deutsche-normungs-roadmap-industrie-4-0> (Zugriff am: 05.05.2023).

Drexel, Damian; Hoch, Ralph (2022): „Three-Faceted Manufacturing Knowledge Representation in Cloud Environments.“ In: IFAC-PapersOnLine, 55 (2022), 10, S. 2950–2955. Online im Internet: DOI: 10.1016/j.ifacol.2022.10.180

Ebbinghaus, Heinz-Dieter (2021): Einführung in die Mengenlehre. Berlin, Heidelberg: Springer Berlin Heidelberg. Online im Internet: DOI: 10.1007/978-3-662-63866-8 (Zugriff am: 16.06.2023).

Ehrlinger, Lisa; Wöß, Wolfram (2016): „Towards a definition of knowledge graphs.“ In: SEMANTiCS (Posters, Demos, SuCCESS), 48 (2016), 1–4, S. 2.

- „EMBL-EBI Ontology Xref Service (OxO)“ (2023): EMBL-EBI Ontology Xref Service (OxO). Online im Internet: URL: <https://www.ebi.ac.uk/spot/oxo/index> (Zugriff am: 17.03.2023).
- Euzenat, Jérôme; Shvaiko, Pavel (2013): *Ontology Matching*. Berlin, Heidelberg: Springer Berlin Heidelberg. Online im Internet: DOI: 10.1007/978-3-642-38721-0 (Zugriff am: 30.10.2022).
- Faria, Daniel u.a. (2019): *Ontology Matching in the Biomedical Domain*. Online im Internet: URL: http://disi.unitn.it/~pavel/OM/articles/SWAT4HCLS_Tutorial.pdf (Zugriff am: 04.03.2023).
- Fensel, Dieter (2001): *Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce*. Springer. Online im Internet: URL: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5f9018e9d1738ae312dbf33381df6e5c0527b57c> (Zugriff am: 25.02.2023).
- Fensel, Dieter u.a. (2020): *Knowledge Graphs: Methodology, Tools and Selected Use Cases*. Cham: Springer International Publishing. Online im Internet: DOI: 10.1007/978-3-030-37439-6 (Zugriff am: 24.02.2023).
- Fernández-Izquierdo, Alba; García-Castro, Raúl (2019): Themis: a tool for validating ontologies through requirements. *The 31st International Conference on Software Engineering and Knowledge Engineering*. , S. 573–578. Online im Internet: DOI: 10.18293/SEKE2019-117 (Zugriff am: 21.04.2023).
- Fortino, Giancarlo u.a. (2018): „Towards Multi-layer Interoperability of Heterogeneous IoT Platforms: The INTER-IoT Approach.“ In: *Integration, Interconnection, and Interoperability of IoT Systems*. Herausgegeben von Raffaele Gravina u.a. Cham: Springer International Publishing (= Internet of Things), S. 199–232. Online im Internet: DOI: 10.1007/978-3-319-61300-0_10 (Zugriff am: 02.12.2022).
- „Fruit joy ad 1999“ (2018): Fruit joy ad 1999. Online im Internet: URL: <https://www.youtube.com/watch?v=rHgie5OviYM> (Zugriff am: 26.10.2022).
- Fulda, Hans Friedrich (1988): *Ontologie nach Kant und Hegel*. 17. Klett-Cotta.
- Ganzha, Maria u.a. (2021): „Semantic interoperability.“ In: *Interoperability of Heterogeneous IoT Platforms: A Layered Approach*, (2021), S. 133–165.
- Gardner, Stephen P. (2005): „Ontologies and semantic data integration.“ In: *Drug Discovery Today*, 10 (2005), 14, S. 1001–1007. Online im Internet: DOI: 10.1016/S1359-6446(05)03504-X
- Genesereth, Michael R; Fikes, Richard E (1992): *Knowledge interchange format-version 3.0: reference manual*. Citeseer. Online im Internet: URL: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=b9176f78f6264451396480e12ac34844a5f6b04b> (Zugriff am: 24.06.2023).
- Ghazvinian, Amir; Noy, Natasha; Musen, Mark (2009): *Creating Mappings For Ontologies in Biomedicine: Simple Methods Work*. Annual Symposium proceedings. American Medical Informatics Association, S. 198–202.
- Giménez, Pablo u.a. (2021): „INTER-LogP: INTER-IoT for Smart Port Transportation.“ In: *Interoperability of Heterogeneous IoT Platforms*. Herausgegeben von Carlos E. Palau

- u.a. Cham: Springer International Publishing (= Internet of Things), S. 257–277. Online im Internet: DOI: 10.1007/978-3-030-82446-4_9 (Zugriff am: 22.04.2023).
- Goethe, Johann Wolfgang von (1907): Maximen und Reflexionen. Aphorismen und Aufzeichnungen. Online im Internet: URL: <https://ia800209.us.archive.org/18/items/goethemaximenun00goetgoog/goethemaximenun00goetgoog.pdf> (Zugriff am: 27.04.2023).
- Goldenstein, Jan; Hunoldt, Michael; Walgenbach, Peter (2018): Wissenschaftliche(s) Arbeiten in den Wirtschaftswissenschaften. Wiesbaden: Springer Fachmedien Wiesbaden. Online im Internet: DOI: 10.1007/978-3-658-20345-0 (Zugriff am: 05.11.2022).
- Grangel-González, Irlán u.a. (2018): „Knowledge Graphs for Semantically Integrating Cyber-Physical Systems.“ In: Database and Expert Systems Applications. Cham: Springer International Publishing (= Lecture Notes in Computer Science), S. 184–199. Online im Internet: DOI: 10.1007/978-3-319-98809-2_12
- Grangel-González, Irlán (2019): A Knowledge Graph Based Integration Approach for Industry 4.0. Rheinischen Friedrich-Wilhelms-Universität Bonn. Bonn. Online im Internet: URL: https://www.researchgate.net/profile/Irlan-Grangel-Gonzalez/publication/333532727_A_Knowledge_Graph_Based_Integration_Approach_for_Industry_40/links/5eac262492851cb267694bf5/A-Knowledge-Graph-Based-Integration-Approach-for-Industry-40.pdf (Zugriff am: 04.06.2023).
- Grangel-González, Irlán; Vidal, Maria Esther (2021): Analyzing a knowledge graph of industry 4.0 standards. Companion Proceedings of the Web Conference 2021. , S. 16–25.
- Gruber, Thomas R (1992): Ontolingua: A mechanism to support portable ontologies. Stanford University, Knowledge Systems Laboratory Stanford.
- Gruber, Thomas R (1993): „A translation approach to portable ontology specifications.“ In: Knowledge acquisition, 5 (1993), 2, S. 199–220.
- Guarino, Nicola; Giaretta, Pierdaniele (1995): „Ontologies and knowledge bases.“ In: Towards very large knowledge bases, (1995), S. 1–2.
- Guarino, Nicola; Oberle, Daniel; Staab, Steffen (2009): „What Is an Ontology?“ In: Handbook on Ontologies, (2009), S. 1.
- Haller, Armin u.a. (2017): Semantic Sensor Network Ontology. KNB Data Repository. Online im Internet: DOI: 10.5063/F11C1TTM (Zugriff am: 19.05.2023).
- Harnad, Stevan (1990): „The symbol grounding problem.“ In: Physica D: Nonlinear Phenomena, 42 (1990), 1, S. 335–346. Online im Internet: DOI: 10.1016/0167-2789(90)90087-6
- Hartig, Olaf; Champin, Pierre-Antoine; Kellogg, Gregg (2023): RDF 1.2 Concepts and Abstract Syntax. Online im Internet: URL: <https://www.w3.org/TR/rdf12-concepts/#section-Introduction> (Zugriff am: 30.05.2023).
- Hartley, Gabrielle (2022): The Conversation. The Human Genome Project pieced together only 92% of the DNA – now scientists have finally filled in the remaining 8%. Online im Internet: URL: <http://theconversation.com/the-human-genome-project-pieced-together-only-92-of-the-dna-now-scientists-have-finally-filled-in-the-remaining-8-176138> (Zugriff am: 13.05.2023).

- Hatzivasilis, George u.a. (2019): Secure Semantic Interoperability for IoT Applications with Linked Data. 2019 IEEE Global Communications Conference (GLOBECOM). IEEE, S. 1–6. Online im Internet: DOI: 10.1109/GLOBECOM38437.2019.9013147 (Zugriff am: 24.06.2023).
- Hazra, Abhishek u.a. (2023): „A Comprehensive Survey on Interoperability for IIoT: Taxonomy, Standards, and Future Directions.“ In: ACM Computing Surveys, 55 (2023), 1, S. 1–35. Online im Internet: DOI: 10.1145/3485130
- He, Yuan u.a. (2022): BERTMap: a BERT-based ontology alignment system. Proceedings of the AAAI Conference on Artificial Intelligence. , S. 5684–5691. Online im Internet: URL: <https://ojs.aaai.org/index.php/AAAI/article/view/20510/20269> (Zugriff am: 25.06.2023).
- Heidel, Roland (2019): Industrie 4.0: The reference architecture model RAMI 4.0 and the Industrie 4.0 component. Beuth Verlag GmbH.
- Heist, Nicolas u.a. (2021): „Wissensgraphen im Web: Überblick und Vergleich.“ In: Semantische Datenintelligenz im Einsatz. Herausgegeben von Börteçin Ege; Adrian Paschke. Wiesbaden: Springer Fachmedien Wiesbaden, S. 127–151. Online im Internet: DOI: 10.1007/978-3-658-31938-0_7 (Zugriff am: 03.03.2023).
- Hildebrandt, Constantin u.a. (2017): „Semantic modeling for collaboration and cooperation of systems in the production domain.“ In: 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). Limassol: IEEE, S. 1–8. Online im Internet: DOI: 10.1109/ETFA.2017.8247585 (Zugriff am: 03.03.2023).
- Hodges, Jack; García, Kimberly; Ray, Steven (2017): „Semantic development and integration of standards for adoption and interoperability.“ In: Computer, 50 (2017), 11, S. 26–36.
- Ibáñez-Sánchez, Gema u.a. (2021): „Interoperability Application in e-Health.“ In: Interoperability of Heterogeneous IoT Platforms. Herausgegeben von Carlos E. Palau u.a. Cham: Springer International Publishing (= Internet of Things), S. 231–256. Online im Internet: DOI: 10.1007/978-3-030-82446-4_8 (Zugriff am: 22.04.2023).
- Ide, Nancy; Pustejovsky, James (2010): What does interoperability mean, anyway? Toward an operational definition of interoperability for language technology. Proceedings of the Second International Conference on Global Interoperability for Language Resources. Hong Kong, China: Online im Internet: URL: <https://www.cs.vassar.edu/~ide/papers/ICGL10.pdf> (Zugriff am: 17.03.2023).
- INTER-IoT Projekt (2019): Inter Plattform Semantic Mediator. Online im Internet: URL: <https://inter-iot.readthedocs.io/projects/ipsm/en/latest/> (Zugriff am: 21.04.2023).
- lod-cloud.net (2022): The Linked Open Data Cloud. Online im Internet: URL: <https://lod-cloud.net/> (Zugriff am: 01.06.2023).
- IOF (2023a): IOF Ontologies. Online im Internet: URL: <https://industrialontologies.org/released-ontologies/iof-ontologies-202301-release/> (Zugriff am: 25.04.2023).
- IOF (2023b): Our Mission. Online im Internet: URL: <https://industrialontologies.org/our-mission/> (Zugriff am: 25.04.2023).
- ISO (2022): IEC/ISO SMART. Online im Internet: URL: <https://www.iso.org/smart> (Zugriff am: 06.05.2023).

- Jarrar, Mustafa; Meersman, Robert (2008): „Ontology Engineering – The DOGMA Approach.“ In: *Advances in Web Semantics I*. Herausgegeben von Tharam S. Dillon u.a. Berlin, Heidelberg: Springer Berlin Heidelberg (= Lecture Notes in Computer Science), S. 7–34. Online im Internet: DOI: 10.1007/978-3-540-89784-2_2 (Zugriff am: 01.04.2023).
- Javed, Saleha u.a. (2022): Deep Ontology Alignment with BERT_INT: Improvements and Industrial Internet of Things (IIoT) Case Study. Online im Internet: URL: <https://urn.kb.se/resolve?urn=urn:nbn:se:itu:diva-90194> (Zugriff am: 27.04.2023).
- Jensen, Mark u.a. (2023): GitHub - CommonCoreOntology/CommonCoreOntologies: The Common Core Ontology Repository holds the current released version of the Common Core Ontology suite. Online im Internet: URL: <https://github.com/CommonCoreOntology/CommonCoreOntologies> (Zugriff am: 18.05.2023).
- Jiménez-Ruiz, Ernesto u.a. (2020): Dividing the Ontology Alignment Task with Semantic Embeddings and Logic-based Modules. arXiv. Online im Internet: URL: <http://arxiv.org/abs/2003.05370> (Zugriff am: 26.04.2023).
- Jiménez-Ruiz, Ernesto; Cuenca Grau, Bernardo (2011): „LogMap: Logic-Based and Scalable Ontology Matching.“ In: *The Semantic Web – ISWC 2011*. Herausgegeben von Lora Aroyo u.a. Berlin, Heidelberg: Springer Berlin Heidelberg (= Lecture Notes in Computer Science), S. 273–288. Online im Internet: DOI: 10.1007/978-3-642-25073-6_18 (Zugriff am: 20.05.2023).
- Jirkovský, Václav (2017): Semantic integration in the context of cyber-physical systems. Czech Technical University. Prag. Online im Internet: URL: https://dspace.cvut.cz/bitstream/handle/10467/73622/Disertace_Jirkovsky_2017.pdf?sequence=1 (Zugriff am: 12.12.2022).
- Kache, Martin u.a. (2015): Leitfaden Literaturrecherche. Online im Internet: URL: https://tudresden.de/bu/verkehr/ibb/sft/ressourcen/dateien/lehre/informationen/Leitfaden_Literaturrecherche.pdf?lang=de (Zugriff am: 18.02.2023).
- Kaebisch, Sebastian u.a. (2020): Web of Things (WoT) Thing Description. Online im Internet: URL: <https://www.w3.org/TR/2020/REC-wot-thing-description-20200409/> (Zugriff am: 21.04.2023).
- Kalfoglou, Yannis; Schorlemmer, Marco (2003): „Ontology mapping: the state of the art.“ In: *The Knowledge Engineering Review*, 18 (2003), 1, S. 1–31. Online im Internet: DOI: 10.1017/S0269888903000651
- Kamakoti, Balaji (2020): Medium. Introduction to Knowledge Graph Embedding with DGL-KE. Online im Internet: URL: <https://towardsdatascience.com/introduction-to-knowledge-graph-embedding-with-dgl-ke-77ace6fb60ef> (Zugriff am: 10.03.2023).
- Kim, Kyoung-Dae; Kumar, P. R. (2012): „Cyber–Physical Systems: A Perspective at the Centennial.“ In: *Proceedings of the IEEE*, 100 (2012), Special Centennial Issue, S. 1287–1308. Online im Internet: DOI: 10.1109/JPROC.2012.2189792
- Kohonen, Teuvo (1990): „The self-organizing map.“ In: *Proceedings of the IEEE*, 78 (1990), 9, S. 1464–1480.
- Korecky, Roman (2022): Ontologies in Industry 4.0: Standards, Applications, and Methodologies. Wirtschaftsuniversität Wien. Wien. Online im Internet: URL: <https://www.researchgate.net/profile/Roman->

- Korecky/publication/360270626_Ontologies_in_Industry_40_Standards_Applications_and_Methodologies/links/626e4a31c42af62fe2e49ca2/Ontologies-in-Industry-40-Standards-Applications-and-Methodologies.pdf (Zugriff am: 24.06.2023).
- Kovalenko, Olga u.a. (2018): „AutomationML ontology: modeling cyber-physical systems for industry 4.0.“ In: IOS Press Journal, 1 (2018).
- Kulmanov, Maxat u.a. (2021): „Semantic similarity and machine learning with ontologies.“ In: Briefings in Bioinformatics, 22 (2021), 4, S. bbaa199. Online im Internet: DOI: 10.1093/bib/bbaa199
- Kumar, Veera Ragavan Sampath u.a. (2019): „Ontologies for industry 4.0.“ In: The Knowledge Engineering Review, 34 (2019), S. 14.
- Lambrix, Patrick; Tan, He (2006): „SAMBO—A system for aligning and merging biomedical ontologies.“ In: Journal of Web Semantics, 4 (2006), 3, S. 196–206. Online im Internet: DOI: 10.1016/j.websem.2006.05.003
- Lee, Edward Ashford; Seshia, Sanjit Arunkumar (2016): Introduction to embedded systems: A cyber-physical systems approach. Mit Press. Online im Internet: URL: <https://picture.iczhiku.com/resource/paper/wyKDZEhAzTswsxmm.pdf> (Zugriff am: 20.06.2023).
- Lelli, Francesco (2019): „Interoperability of the Time of Industry 4.0 and the Internet of Things.“ In: Future Internet, 11 (2019), 2, S. 36. Online im Internet: DOI: 10.3390/fi11020036
- Lemaignan, S. u.a. (2006): „MASON: A Proposal For An Ontology Of Manufacturing Domain.“ In: IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06). , S. 195–200. Online im Internet: DOI: 10.1109/DIS.2006.48
- Leser, Ulf; Naumann, Felis (2007): Informationsintegration. Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen. Heidelberg: dpunkt.verlag.
- Li, Wenbin u.a. (2019): „Review of Standard Ontologies for the Web of Things.“ In: 2019 Global IoT Summit (GloTS). Aarhus, Denmark: IEEE, S. 1–6. Online im Internet: DOI: 10.1109/GIOTS.2019.8766377 (Zugriff am: 21.04.2023).
- Liang, Xu (2020): Medium. Summary of Translate Model for Knowledge Graph Embedding. Online im Internet: URL: <https://towardsdatascience.com/summary-of-translate-model-for-knowledge-graph-embedding-29042be64273> (Zugriff am: 11.03.2023).
- Lin, Shi-Wan u.a. (2022): „The Industrial Internet Reference Architecture.“ In: (2022). Online im Internet: URL: <https://www.iiconsortium.org/wp-content/uploads/sites/2/2022/11/IIRA-v1.10.pdf> (Zugriff am: 06.05.2023).
- Lin, Yankai u.a. (2018): Knowledge Representation Learning: A Quantitative Review. arXiv. Online im Internet: URL: <http://arxiv.org/abs/1812.10901> (Zugriff am: 11.03.2023).
- Lu, Yuqian; Xu, Xun; Wang, Lihui (2020): „Smart manufacturing process and system automation – A critical review of the standards and envisioned scenarios.“ In: Journal of Manufacturing Systems, 56 (2020), S. 312–325. Online im Internet: DOI: 10.1016/j.jmsy.2020.06.010
- Mabkhot, Mohammed M. u.a. (2018): „Requirements of the Smart Factory System: A Survey and Perspective.“ In: Machines, 6 (2018), 2, S. 23. Online im Internet: DOI: 10.3390/machines6020023

- Mayer, Simon u.a. (2017): „An Open Semantic Framework for the Industrial Internet of Things.“ In: IEEE Intelligent Systems, 32 (2017), 1, S. 96–101. Online im Internet: DOI: 10.1109/MIS.2017.9
- McLean, Roscoe; Walker, Anthony J.; Bright, Glen (2017): „An artificial neural network driven decision-making system for manufacturing disturbance mitigation in reconfigurable systems.“ In: 2017 13th IEEE International Conference on Control & Automation (ICCA). Ohrid, Macedonia: IEEE, S. 695–700. Online im Internet: DOI: 10.1109/ICCA.2017.8003144 (Zugriff am: 29.04.2023).
- Melluso, Nicola; Grangel-González, Irlan; Fantoni, Gualtiero (2022): „Enhancing Industry 4.0 standards interoperability via knowledge graphs with natural language processing.“ In: Computers in Industry, 140 (2022), S. 103676. Online im Internet: DOI: 10.1016/j.com-pind.2022.103676
- Mikolov, Tomas u.a. (2013): Efficient Estimation of Word Representations in Vector Space. arXiv. Online im Internet: URL: <http://arxiv.org/abs/1301.3781> (Zugriff am: 08.04.2023).
- Miller, Bradford; Lin, Shi-Wan (2016): Interoperability and the IIRA. Online im Internet: URL: https://www.researchgate.net/profile/Bradford-Miller/publication/281971976_Interoperability_and_the_IIRA/links/56002d6a08aec948c4fa67cc/Interoperability-and-the-IIRA.pdf (Zugriff am: 20.05.2023).
- Moreira, João u.a. (2017): Towards IoT platforms' integration: Semantic Translations between W3C SSN and ETSI SAREF. SEMANTICS workshops. Amsterdam: Online im Internet: URL: https://ris.utwente.nl/ws/portalfiles/portal/28202327/Towards_IoT.pdf (Zugriff am: 21.04.2023).
- Musil, Robert (1906): Die Verwirrungen des Zöglings Törleß. Online im Internet: URL: <https://www.projekt-gutenberg.org/musil/toerless/chap004.html> (Zugriff am: 28.04.2023).
- Neuhaus, Fabian (2018): What is an Ontology? Online im Internet: URL: <https://www.semanticscholar.org/paper/What-is-an-Ontology-Neuhaus/a13477eaa1ad804e3adb1a91caef1715f23eeb43> (Zugriff am: 20.12.2022).
- Neuhaus, Holger; Compton, Michael (2009): The semantic sensor network ontology. AGILE workshop on challenges in geospatial data harmonisation. Hannover, Deutschland, S. 1–33.
- Nieuwenhuize, Gerben (2016): Smart manufacturing for Dutch SMEs: Why and how. Master Thesis. Rotterdam School of Management – Erasmus University. Rotterdam. Online im Internet: URL: <https://thesis.eur.nl/pub/34786/Gerben-Nieuwenhuize-419386-Thesis.pdf> (Zugriff am: 24.06.2023).
- Nilsson, Jacob; Sandin, Fredrik (2018): „Semantic Interoperability in Industry 4.0: Survey of Recent Developments and Outlook.“ In: 2018 IEEE 16th International Conference on Industrial Informatics (INDIN). Porto: IEEE, S. 127–132. Online im Internet: DOI: 10.1109/INDIN.2018.8471971 (Zugriff am: 03.11.2022).
- Nilsson, Jacob; Sandin, Fredrik; Delsing, Jerker (2019): „Interoperability and machine-to-machine translation model with mappings to machine learning tasks.“ In: 2019 IEEE 17th International Conference on Industrial Informatics (INDIN). Helsinki, Finland: IEEE, S.

- 284–289. Online im Internet: DOI: 10.1109/INDIN41052.2019.8972085 (Zugriff am: 23.12.2022).
- Noura, Mahda; Atiquzzaman, Mohammed; Gaedke, Martin (2019): „Interoperability in Internet of Things: Taxonomies and Open Challenges.“ In: *Mobile Networks and Applications*, 24 (2019), 3, S. 796–809. Online im Internet: DOI: 10.1007/s11036-018-1089-9
- Ochieng, Peter; Kyanda, Swaib (2018): „Large-Scale Ontology Matching: State-of-the-Art Analysis.“ In: *ACM Computing Surveys*, 51 (2018), 4, S. 75:1-75:35. Online im Internet: DOI: 10.1145/3211871
- „Olav – Online Literaturabfrage der FH Vorarlberg“ (2022): Olav – Online Literaturabfrage der FH Vorarlberg. Online im Internet: URL: <https://www.fhv.at/ueber-die-fh/bibliothek/recherche/olav-online-literaturabfrage-der-fh-vorarlberg/> (Zugriff am: 05.11.2022).
- oneM2M (2019a): Base Ontology. Online im Internet: URL: https://www.onem2m.org/images/pdf/TS-0012-Base_Ontology-V3_7_3.pdf (Zugriff am: 14.04.2023).
- oneM2M (2019b): Ontology based Interworking. Online im Internet: URL: <https://member.onem2m.org/Application/documentapp/downloadLatestRevision/default.aspx?docID=29566> (Zugriff am: 15.04.2023).
- oneM2M (2023a): About Us. Online im Internet: URL: <https://www.onem2m.org/harmonization-m2m> (Zugriff am: 14.04.2023).
- oneM2M (2023b): Functional Architecture. Online im Internet: URL: <https://member.onem2m.org/Application/documentapp/downloadLatestRevision/default.aspx?docID=35753> (Zugriff am: 15.05.2023).
- OntoCommons (2020): Onto Commons Demonstrators. Online im Internet: URL: <https://www.ontocommons.eu/ontocommons-demonstrators#OntoCommons%20Demonstrators> (Zugriff am: 25.04.2023).
- Ontology Engineering Group (2020a): VICINITY Ontologies. Online im Internet: URL: <http://vicinity.iot.linkeddata.es/vicinity/> (Zugriff am: 21.04.2023).
- Ontology Engineering Group (2020b): VICINITY Ontology testing. Online im Internet: URL: <http://vicinity.iot.linkeddata.es/vicinity/conformance.html> (Zugriff am: 21.04.2023).
- Osman, Inès; Ben Yahia, Sadok; Diallo, Gayo (2021): „Ontology Integration: Approaches and Challenging Issues.“ In: *Information Fusion*, 71 (2021), S. 38–63. Online im Internet: DOI: 10.1016/j.inffus.2021.01.007
- „OWL 2 Web Ontology Language Profiles (Second Edition)“ (2012): OWL 2 Web Ontology Language Profiles (Second Edition). Online im Internet: URL: https://www.w3.org/TR/2012/REC-owl2-profiles-20121211/#Computational_Properties (Zugriff am: 12.03.2023).
- Paulheim, Heiko; Cimiano, Philipp (2017): „Knowledge graph refinement: A survey of approaches and evaluation methods.“ In: *Semantic Web*, 8 (2017), 3, S. 489–508. Online im Internet: DOI: 10.3233/SW-160218

- Pereira, Pedro Henrique Morgan (2022): „Interoperability middleware for IIoT gateways based on international standard ontologies and standardized digital representation.“ In: (2022).
- Pereira, Ramon Martinez; Szejka, Anderson Luis; Canciglieri Junior, Osiris (2021): „Towards an information semantic interoperability in smart manufacturing systems: Contributions, limitations and applications.“ In: *International Journal of Computer Integrated Manufacturing*, 34 (2021), 4, S. 422–439.
- Pereira, RM; Szejka, AL; Canciglieri Jr, O (2022): „Ontological Approach to Support the Horizontal and Vertical Information Integration in Smart Manufacturing Systems: An Experimental Case in a Long-Life Packaging Factory.“ In: *Frontiers in Manufacturing Technology*, 2 (2022). Online im Internet: DOI: 10.3389/fmtec.2022.854155 (Zugriff am: 24.06.2023).
- Petersen, Niklas u.a. (2017): „Realizing an RDF-Based Information Model for a Manufacturing Company – A Case Study.“ In: *The Semantic Web – ISWC 2017*. Herausgegeben von Claudia d’Amato u.a. Cham: Springer International Publishing (= Lecture Notes in Computer Science), S. 350–366. Online im Internet: DOI: 10.1007/978-3-319-68204-4_31 (Zugriff am: 25.03.2023).
- Pham, Cu; Lim, Yuto; Tan, Yasuo (2019): „An oneM2M Interworking Proxy Entity for ECHO-NET Lite Protocol.“ In: *2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)*. Osaka, Japan: IEEE, S. 784–787. Online im Internet: DOI: 10.1109/GCCE46687.2019.9015534 (Zugriff am: 15.04.2023).
- Pham, Van Cu u.a. (2019): „An ontology-driven echonet lite adaptation layer for smart homes.“ In: *Journal of Information Processing*, 27 (2019), S. 360–368.
- Pistorius, Johannes (2020): *Industrie 4.0 – Schlüsseltechnologien für die Produktion: Grundlagen • Potenziale • Anwendungen*. Berlin, Heidelberg: Springer Berlin Heidelberg. Online im Internet: DOI: 10.1007/978-3-662-61580-5 (Zugriff am: 16.06.2023).
- Plattform Industrie 4.0 (2015): „Umsetzungsstrategie Industrie 4.0.“ In: (2015). Online im Internet: URL: <https://www.bitkom.org/sites/main/files/file/import/150410-Umsetzungsstrategie-0.pdf> (Zugriff am: 27.04.2023).
- Poveda-Villalón, María; Gómez-Pérez, Asunción; Suárez-Figueroa, Mari Carmen (2014): „Oops!(ontology pitfall scanner!): An on-line tool for ontology evaluation.“ In: *International Journal on Semantic Web and Information Systems (IJSWIS)*, 10 (2014), 2, S. 7–34.
- Quine, Willard Van Orman (1948): „On what there is.“ In: *The review of metaphysics*, (1948), S. 21–38.
- Quine, Willard Van Orman (1976): „Whither Physical Objects?“ In: *Essays in Memory of Imre Lakatos*. Herausgegeben von R. S. Cohen; P. K. Feyerabend; M. W. Wartofsky. Dordrecht: Springer Netherlands (= Boston Studies in the Philosophy of Science), S. 497–504. Online im Internet: DOI: 10.1007/978-94-010-1451-9_29 (Zugriff am: 28.04.2023).
- Radziwon, Agnieszka u.a. (2014): „The smart factory: exploring adaptive and flexible manufacturing solutions.“ In: *Procedia engineering*, 69 (2014), S. 1184–1190.

- Rajkumar, Ragunathan u.a. (2010): „Cyber-physical systems: the next computing revolution.“ In: Proceedings of the 47th Design Automation Conference. Anaheim California: ACM, S. 731–736. Online im Internet: DOI: 10.1145/1837274.1837461 (Zugriff am: 18.02.2023).
- Ramar, Kaladevi; Gurunathan, Geetha (2016): „Technical review on ontology mapping techniques.“ In: Asian Journal of Information Technology, 15 (2016), 4, S. 676–688.
- Rebstock, Michael; Fengel, Janina; Paulheim, Heiko (Hrsg.) (2008): „Ontology Engineering.“ In: Ontologies-Based Business Integration. Berlin, Heidelberg: Springer, S. 97–123. Online im Internet: DOI: 10.1007/978-3-540-75230-1_5 (Zugriff am: 19.10.2022).
- de Ribaupierre, Helene u.a. (2021): „Automatic extraction of requirements expressed in industrial standards: a way towards machine readable standards?“ In: arXiv preprint arXiv:2112.13091, (2021).
- Rivas, Ariam u.a. (2021): „Discover Relations in the Industry 4.0 Standards Via Unsupervised Learning on Knowledge Graph Embeddings.“ In: Journal of Data Intelligence, 2 (2021), 3, S. 336–347. Online im Internet: DOI: 10.26421/JDI2.3-2
- Rubí, Jesús Noel Sáez; Gondim, Paulo Roberto De Lira (2020): „Interoperable Internet of Medical Things platform for e-Health applications.“ In: International Journal of Distributed Sensor Networks, 16 (2020), 1, S. 155014771988959. Online im Internet: DOI: 10.1177/1550147719889591
- Rudnicki, Ron (2019): An Overview of the Common Core Ontologies. Buffalo, NY: CUBRC. Online im Internet: URL: https://www.nist.gov/system/files/documents/2021/10/14/nist-ai-rfi-cubrc_inc_004.pdf (Zugriff am: 25.04.2023).
- Rueda, Carlos u.a. (2010): „The MMI Device Ontology: Enabling Sensor Integration.“ In: American Geophysical Union Fall Meeting – Session, 16 (2010), S. 44–8.
- Sack, Harald (2020): Semantic Web Technologies by Dr. Harald Sack - YouTube. Online im Internet: URL: <https://www.youtube.com/> (Zugriff am: 16.06.2023).
- Saeed, Bushra; Baranwal, Vinay; Khurana, Paramjit (2016): „Identification and Expression Profiling of the Lectin Gene Superfamily in Mulberry.“ In: The Plant Genome, 9 (2016). Online im Internet: DOI: 10.3835/plantgenome2015.10.0107
- Schlichtkrull, Michael u.a. (2018): Modeling relational data with graph convolutional networks. The Semantic Web: 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, Proceedings 15. Springer, S. 593–607.
- Schlosser, Ralf u.a. (2006): „Use of information-seeking strategies for developing systematic reviews and engaging in evidence-based practice: the application of traditional and comprehensive Pearl Growing. A review.“ In: International journal of language & communication disorders / Royal College of Speech & Language Therapists, 41 (2006), S. 567–82. Online im Internet: DOI: 10.1080/13682820600742190
- Schönhof, Raoul; Konertz, Roman (2020): Das technische Phänomen „Künstliche Intelligenz“ im allgemeinen Zivilrecht. Online im Internet: DOI: 10.5771/9783748910503-15
- Schröder, Tizian; Willner, Alexander; Diedrich, Christian (2021): „Semantische Interoperabilität in cyber-physischen Produktionssystemen.“ In: Semantische Datenintelligenz im Einsatz. Springer, S. 33–49.

- Schulz von Thun, Friedemann (1995): *Miteinander reden*. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.
- da Silva, Guilherme Canuto; Kaminski, Paulo Carlos (2016): From Embedded Systems (ES) to Cyber-Physical Systems (CPS): An Analysis of Transitory Stage of Automotive Manufacturing in the Industry 4.0 Scenario. 25th SAE BRASIL International Congress and Display. São Paulo, S. 2016-36–0230. Online im Internet: DOI: 10.4271/2016-36-0230 (Zugriff am: 18.02.2023).
- Singh, Mandeep u.a. (2022): „Data Information Interoperability Model for IoT-enabled Smart Water Networks.“ In: 2022 IEEE 16th International Conference on Semantic Computing (ICSC). Laguna Hills, CA, USA: IEEE, S. 179–186. Online im Internet: DOI: 10.1109/ICSC52841.2022.00038 (Zugriff am: 22.04.2023).
- Skjæveland, Martin G.; Slaughter, Laura Ann; Kindermann, Christian (2022): „OntoCommons D4.3 - Report on Landscape Analysis of Ontology Engineering Tools.“ In: (2022). Online im Internet: DOI: 10.5281/ZENODO.6504670 (Zugriff am: 25.04.2023).
- Solimando, Alessandro; Jiménez-Ruiz, Ernesto; Guerrini, Giovanna (2014): „Detecting and Correcting Conservativity Principle Violations in Ontology-to-Ontology Mappings.“ In: *The Semantic Web – ISWC 2014*. Herausgegeben von Peter Mika u.a. Cham: Springer International Publishing (= Lecture Notes in Computer Science), S. 1–16. Online im Internet: DOI: 10.1007/978-3-319-11915-1_1 (Zugriff am: 26.05.2023).
- Sony, Michael; Naik, Subhash (2020): „Industry 4.0 integration with socio-technical systems theory: A systematic review and proposed theoretical model.“ In: *Technology in Society*, 61 (2020), S. 101248. Online im Internet: DOI: 10.1016/j.techsoc.2020.101248
- Strick, Heinz Klaus (2016): *Der Mathematische Monatskalender: Leopold Kronecker (1823–1891)*. Online im Internet: URL: <https://www.spektrum.de/wissen/leopold-kronecker-1823-1891/1429551> (Zugriff am: 28.04.2023).
- Studer, Rudi; Benjamins, V.Richard; Fensel, Dieter (1998): „Knowledge engineering: Principles and methods.“ In: *Data & Knowledge Engineering*, 25 (1998), 1–2, S. 161–197. Online im Internet: DOI: 10.1016/S0169-023X(97)00056-6
- Szilagyi, Ioan; Wira, Patrice (2016): „Ontologies and Semantic Web for the Internet of Things - a survey.“ In: *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*. , S. 6949–6954. Online im Internet: DOI: 10.1109/IECON.2016.7793744
- Szmeja, Paweł u.a. (2022): „Alignment Format for Semantic Translation.“ In: *Sustainable Technology and Advanced Computing in Electrical Engineering*. Herausgegeben von Vasundhara Mahajan u.a. Singapore: Springer Nature Singapore (= Lecture Notes in Electrical Engineering), S. 651–671. Online im Internet: DOI: 10.1007/978-981-19-4364-5_47 (Zugriff am: 26.04.2023).
- Tang, Xiaobin u.a. (2021): „BERT-INT: a BERT-based interaction model for knowledge graph alignment.“ In: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*. Yokohama, Yokohama, Japan (= IJCAI'20), S. 3174–3180. (Zugriff am: 27.04.2023).
- Tayur, Varun M; Suchithra, R (2022): „Multi-ontology mapping generative adversarial network in internet of things for ontology alignment.“ In: *Internet of Things*, 20 (2022), S. 100616. Online im Internet: DOI: 10.1016/j.iot.2022.100616

- Teslya, Nikolay; Ryabchikov, Igor (2019): „Ontology-Based Semantic Models for Industrial IoT Components Representation.“ In: Proceedings of the Third International Scientific Conference “Intelligent Information Technologies for Industry” (IITI'18). Herausgegeben von Ajith Abraham u.a. Cham: Springer International Publishing (= Advances in Intelligent Systems and Computing), S. 138–147. Online im Internet: DOI: 10.1007/978-3-030-01818-4_14 (Zugriff am: 19.12.2022).
- The Gene Ontology Consortium u.a. (2021): „The Gene Ontology resource: enriching a GOLD mine.“ In: Nucleic Acids Research, 49 (2021), D1, S. D325–D334. Online im Internet: DOI: 10.1093/nar/gkaa1113
- The Gene Ontology Consortium (2023): Gene Ontology Resource. Gene Ontology Resource - Release statistics. Online im Internet: URL: <http://geneontology.org/stats.html> (Zugriff am: 13.05.2023).
- Tolk, Andreas; Muguira, James A (2003): The levels of conceptual interoperability model. Proceedings of the 2003 fall simulation interoperability workshop. Citeseer, S. 1–11.
- Tomaszuk, Dominik; Haudebourg, Timothée (2023): RDF 1.2 Schema. Online im Internet: URL: https://www.w3.org/TR/rdf12-schema/#ch_introduction (Zugriff am: 30.05.2023).
- Veneri, Giacomo; Capasso, Antonio (2018): Hands-On Industrial Internet of Things: Create a Powerful Industrial IoT Infrastructure Using Industry 4. 0. Birmingham, UK: Packt Publishing, Limited. Online im Internet: URL: <http://ebookcentral.proquest.com/lib/vorarlberg/detail.action?docID=5608194> (Zugriff am: 16.06.2023).
- VICINITY (2016): Vicinity | Open virtual neighbourhood network to connect IoT infrastructures and smart objects. Online im Internet: URL: <https://vicinity2020.eu/> (Zugriff am: 21.04.2023).
- VICINITY (2019): 3 Using Semantic Interoperability services. Online im Internet: URL: <https://www.youtube.com/watch?v=xXh9vcLBA-c> (Zugriff am: 21.04.2023).
- Vila, Marc u.a. (2022): „Edge-to-cloud sensing and actuation semantics in the industrial Internet of Things.“ In: Pervasive and Mobile Computing, 87 (2022), S. 101699. Online im Internet: DOI: 10.1016/j.pmcj.2022.101699
- W3C (2021): Data Activity - Building the Web of Data. Online im Internet: URL: <https://www.w3.org/2013/data/> (Zugriff am: 17.06.2023).
- W3C (2022): Semantic Web. Online im Internet: URL: <https://www.w3.org/standards/semanticweb/> (Zugriff am: 17.06.2023).
- Wache, Holger u.a. (2001): Ontology-based integration of information - a survey of existing approaches. Proceedings of the IJCAI-01 Workshop on Ontologies and Information Sharing. Seattle, USA, S. 108–117. Online im Internet: URL: <http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-47/ONTOL2-Proceedings.pdf#page=110> (Zugriff am: 24.06.2023).
- Wood, Thomas (2020): DeepAI. Generative Adversarial Network. Online im Internet: URL: <https://deepai.org/machine-learning-glossary-and-terms/generative-adversarial-network> (Zugriff am: 27.04.2023).
- Xue, Xingsi u.a. (2021): „Artificial Neural Network Based Sensor Ontology Matching Technique.“ In: Companion Proceedings of the Web Conference 2021. New York, NY, USA:

Association for Computing Machinery (= WWW '21), S. 44–51. Online im Internet: DOI: 10.1145/3442442.3451138 (Zugriff am: 11.03.2023).

Zhapa-Camacho, Fernando (2022): Embedding Ontologies Using Category Theory Semantics. King Abdullah University of Science and Technology. Thuwal, Saudi Arabien. Online im Internet: URL: <https://repository.kaust.edu.sa/bitstream/handle/10754/676334/FernandoZhapaThesis.pdf?sequence=1> (Zugriff am: 24.06.2023).

Zhou, Ce u.a. (2023): A Comprehensive Survey on Pretrained Foundation Models: A History from BERT to ChatGPT. arXiv. Online im Internet: URL: <http://arxiv.org/abs/2302.09419> (Zugriff am: 08.04.2023).

„Mars Climate Orbiter: Absturz wegen Leichtsinnsfehler beim Rechnen“ (1999): In: Der Spiegel, 1. Oktober 1999. Online im Internet: URL: <https://www.spiegel.de/wissenschaft/mensch/mars-climate-orbiter-absturz-wegen-leichtsinnsfehler-beim-rechnen-a-44777.html> (Zugriff am: 18.03.2023).

„Gödelscher Unvollständigkeitssatz“ (2022): In: Wikipedia. Online im Internet: URL: https://de.wikipedia.org/w/index.php?title=G%C3%B6delscher_Unvollst%C3%A4ndigkeitssatz&oldid=226580168#Aussage (Zugriff am: 25.02.2023).

„P-NP-Problem“ (2023): In: Wikipedia. Online im Internet: URL: <https://de.wikipedia.org/w/index.php?title=P-NP-Problem&oldid=230259424> (Zugriff am: 12.03.2023).

Anhang: Übersicht der besprochenen Anwendungsfälle

Die folgende Tabelle zeigt die in Kapitel 4 besprochenen Ansätze in der Reihenfolge, in der sie besprochen wurden.

Autoren/Projekt	Kurzbeschreibung	Ansatz	Ontologien	Matching Methode	Anwendung
Drexel und Hoch (2022)	Erstellung einer Cloud Manufacturing-Anwendung	multipl	Selbst erstellte Ontologien (Produktion, Produkt, Prozess)	manuell	Produktion
oneM2M (2019a)	Matching einer Top-Level Ontologie mit domänenspezifischen Ontologien	hybrid	oneM2M Base, SAREF und weitere domänenspezifische Ontologien	unbekannt (vermutlich manuell)	Smart Home/IoT
VICINITY (2016)	Ermöglichen von Interoperabilität in einer virtuellen Nachbarschaft mithilfe einer Kombination von Ontologien	hybrid	WoT/Thing Description, VICINITY Core, Adapters, Datatypes, WoT Mappings	unbekannt (vermutlich manuell)	Smart Home/IoT
INTER-IoT (Fortino u.a. 2018, S. 218)	Mithilfe einer zentralen Ontologie (GOIoT) als Mediator können IIoT-Plattformen kommunizieren	hybrid	GOIoT sowie daran verknüpfte domänenspezifische Ontologien	Bestehende Algorithmen wie LogMap, COMA, Agreement Maker, ALIN, MapOnto Yam++	Transport, Gesundheitswesen, Wassermanagement
Bosch (OntoCommons 2020)	Matching einer Top-Level Ontologie mit domänenspezifischen Ontologien	hybrid	Bosch I4.0 Core Ontology	unbekannt	Produktion
OSF (Mayer u.a. 2017)	Open Semantic Framework von Siemens, das Core Ontologien mit domänenspezifischen Ontologien verbindet.	hybrid	Unterschiedliche Top-, Mid- und domänenspezifische Ontologien wie SSN, QUDT, CIM)	unbekannt	entfällt

OntoCommons (OntoCommons 2020)	Erstellung eines Ökosystems an Ontologien und abgelegten Alignments	hybrid	-	LOOM	Unterschiedliche Bereiche ¹⁶⁰
Grangel-González (2019)	Überführen semantisch heterogener AutomationML-Dokumente in ein gemeinsames Dokument mithilfe von Mappings	global	AMLO, eCI@ss	Alligator, SemCPS	Produktion
Rubí und Gondim (2020)	Integration zweier Ontologien, zur Erstellung einer eHealth-Plattform. Dabei wurde eine Ontologie aus einem Standard abgeleitet.	multipel	openEHR, SSN	unbekannt (vermutlich manuell)	Gesundheitswesen
Hodges und Garcia (2017)	Integration zweier Ontologien, um einen digitalen Zwilling zu erzeugen. Dabei wurde eine Ontologie aus einem Standard abgeleitet.	multipel	EDDL, QUDT	unbekannt (vermutlich manuell)	Produktion
Petersen u.a. (2017)	Verbindung mehrerer Datentöpfe (MES, BOM, Sensoren) um damit Informationen zum Produktionswerk und dessen Inventar zu erhalten.	hybrid	RAMI, VIVO, NeoGeo, FOAF, SSN	unbekannt (vermutlich manuell)	Werkmonitoring
Jirkovský (2017)	Ontologie Integration dreier Ontologien zur Abbildung eines Wasserkraftwerks	multipel	COCI, DUL, SSN	MAPSOM (Verwendung von Ähnlichkeitsdistanzen und einem neuronalem Netz)	Werkmonitoring

¹⁶⁰ Siehe <https://ontocommons.eu/ontocommons-demonstrators> (Zugriff am: 16.06.2023)

Xue u.a. (2021)	Zusammenführen von Sensordaten	multipl	SSN, MMI Device, CSIRO	Ähnlichkeitsmetriken, neuronale Netze, Embeddings	entfällt
Javed u.a. (2022)	Alignment zweier IIoT-Ontologien mithilfe von Methoden des Natural Language Processings	multipl	SSN, SOSA	Embeddings, BERT-INT	entfällt
Tayur und Suchithra. (2022)	Alignment dreier IIoT-Ontologien mithilfe eines Generative Adversarial Network (GAN)	multipl	oneM2M, SAREF, SSN	Embeddings, GAN	entfällt
I40KG (Grangel-González; Vidal 2021; Melluso; Grangel-González; Fantoni 2022)	Verbindungen von IIoT-Standards	Hybrid	Standards Ontologie,	Embeddings (BERT)	entfällt

Darstellung 53: Übersicht der in Kapitel 4 besprochenen Ansätze

Quelle: Autor

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich vorliegende Masterarbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Stellen sind als solche kenntlich gemacht.

Die Arbeit wurde bisher weder in gleicher noch in ähnlicher Form einer anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.



Dornbirn, am 03.07.2023

Raimund Porod