



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna | Austria



Diplomarbeit

Textuelle Analyse von Jobausschreibungen zur Bewertung der Kompetenzentwicklung in der industriellen Instandhaltung

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines
Diplom-Ingenieurs (Dipl.-Ing.), eingereicht an der TU Wien,
Fakultät für Maschinenwesen und Betriebswissenschaften, von

David Lichtenschopf, BSc.

Mat.Nr.: 01326607



unter der Leitung von

Priv.-Doz. Dr.-Ing. Fazel Ansari, MSc.

(E330 Institut für Managementwissenschaften, Bereich: Smart and Knowledge Based Maintenance)

Univ.Ass. Dipl.-Ing. Steffen Nixdorf, BSc.

(E330 Institut für Managementwissenschaften, Bereich: Smart and Knowledge Based Maintenance)

Wien, im März 2022

David Lichtenschopf

Ich nehme zur Kenntnis, dass ich zur Drucklegung dieser Arbeit nur mit Bewilligung der Prüfungskommission berechtigt bin.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass die vorliegende Arbeit nach den anerkannten Grundsätzen für wissenschaftliche Abhandlungen von mir selbstständig erstellt wurde. Alle verwendeten Hilfsmittel, insbesondere die zugrunde gelegte Literatur, sind in dieser Arbeit genannt und aufgelistet. Die aus den Quellen wörtlich entnommenen Stellen, sind als solche kenntlich gemacht.

Das Thema dieser Arbeit wurde von mir bisher weder im In- noch Ausland einer Beurteilerin/einem Beurteiler zur Begutachtung in irgendeiner Form als Prüfungsarbeit vorgelegt. Diese Arbeit stimmt mit der von den Begutachterinnen/Begutachtern beurteilten Arbeit überein.

Ich nehme zur Kenntnis, dass die vorgelegte Arbeit mit geeigneten und dem derzeitigen Stand der Technik entsprechenden Mitteln (Plagiat-Erkennungssoftware) elektronisch-technisch überprüft wird. Dies stellt einerseits sicher, dass bei der Erstellung der vorgelegten Arbeit die hohen Qualitätsvorgaben im Rahmen der geltenden Regeln zur Sicherung guter wissenschaftlicher Praxis „Code of Conduct“ an der TU Wien eingehalten wurden. Zum anderen werden durch einen Abgleich mit anderen studentischen Abschlussarbeiten Verletzungen meines persönlichen Urheberrechts vermieden.

Wien, im März 2022

David Lichtenschopf

Danksagung

Mein größter Dank gilt Alexandra, Christine, Herbert, Gertrude, Franz, Stefanie, Barbara, Matthias, Valentin, Charlotte und Dorothea, welche mich ausnahmslos in jeder noch so schwierigen Zeit meines Studiums und Lebens begleitet haben und hoffentlich auch noch lange begleiten werden.

Meiner Godi Christine für die Unterstützung und Korrekturlesen.

Mein Dank gilt auch den Mitgliedern des gesamten ASA Kulturvereins, welche mich bei allen Höhen und Tiefen unterstützt haben.

Ein Dank geht an meine Betreuer dieser Arbeit Dipl.-Ing. Nixdorf und Prof. Ansari für die operative Betreuung und Dr. Hartner-Tiefenthaler für das kurzfristige Einspringen als Zweitprüferin.

Weiters geht mein Dank an Prof. Linhardt, welcher mich auf meinem akademischen Weg begleitet hat und mir immer wieder neue Perspektiven aufgezeigt hat.

Des Weiteren möchte ich mich bei allen Kommilitonen_innen, die ich im Laufe des Studiums kennen und schätzen gelernt habe, bedanken.

Kurzfassung

Der Arbeitsmarkt ist im stetigen Wandel. Im produzierenden Gewerbe wird seit 2011 der Begriff Industrie 4.0 häufig als Schlagwort verwendet und steht für die Digitalisierung und Automatisierung von Produktionen. Diese Veränderung hat zur Folge, dass sich auch das Anforderungsprofil der Kompetenzen von Beschäftigten in diesen Betrieben verändert hat und weiter verändern wird.

Das Ziel dieser Arbeit ist es, ein Konzept zu entwickeln und zu validieren, mit welchem ein nachfrageorientierter Kompetenzkatalog von Beschäftigten in der Instandhaltung im produzierenden Gewerbe erstellt werden kann. Mit der Erstellung von mehreren zeitlich getrennten Auswertungen sollen Veränderungen der nachgefragten Kompetenzen in der Instandhaltung abgeleitet werden können.

Als Grundlage für diesen Kompetenzkatalog werden online verfügbare Stellenanzeigen als Datenbasis und der ESCO Standard als Basis für spezifische Kompetenzen verwendet. Die Auswertung erfolgt durch einen Abgleich der ESCO Kompetenzen mit jenen der Stellenanzeigen unter der Verwendung von Text Mining Methoden. Es werden dafür zwei verschiedene Ansätze verfolgt. Umgesetzte Ansätze sind die Auswertung der Anzahl von einzelnen Kompetenzen und die Auswertung eines Topic Modells. Der Ansatz bezüglich der Anzahl soll einen generellen Kompetenzkatalog für die Instandhaltung und zeitliche Trends erkennen lassen. Der Ansatz bezüglich des Topic Modells soll eine Zuordnung zu spezifischen Stellen ermöglichen.

Die größte Herausforderung des Ansatzes besteht darin, dass die Kompetenzen aus den online Stellenanzeigen automatisiert erkannt werden können, da die verwendeten Begriffe in den Stellenbeschreibungen zumeist nicht direkt mit dem ESCO Standard übereinstimmen. Aus diesem Grund ist eine ontologiebasierende Synonymerkennung umgesetzt.

Mit dem Ansatz der Wortanzahl ist es möglich, einen Kompetenzkatalog zu erstellen und zeitliche Veränderungen zu beobachten, jedoch ist dies nur auf der gesamten Datenbasis möglich und lässt keine Auswertung von Kompetenzen zu spezifischeren Stellen (z.B. IT, elektrische mechanische Instandhaltung) zu. Die Evaluierung zeigt, dass 24% der manuell gefundenen Kompetenzen auch automatisiert gefunden werden. Durch eine Eingrenzung auf eine spezifische Stelle und einer manuell erstellten Erweiterungsliste anhand von Kompetenzen aus 15 Stellenanzeigen wird die Evaluierung auf 76% gehoben.

Für die Zuordnung von Kompetenzen zu spezifischen Stellen wird der Ansatz des Topic Modelings verfolgt, jedoch werden keine interpretierbaren spezifischen Stellen mit dieser Methode erkannt.

Abstract

The job market is in steady change. Since 2011 Industry 4.0 is one of the major keywords used in production industry and standing for digitalization and automatization of production. This change has ongoing influence on the requirement profile of various jobs.

The objective of this thesis is to develop and validate a concept, which evaluates the demand-oriented competence portfolio of maintenance staff in a manufacturing environment. Furthermore, trends in the competence portfolio should be derivable.

As bases for the competence portfolio and trends recognition online job advertisements are used as data base and the ESCO standard is used as base for specific competences. The evaluation process compares the included competences from the job advertisements to the ESCO standard with the use of text mining methods. Two different approaches are realized. One approach is based on the count of competences and the second approach is a topic model. With the count of competences approach a general competence portfolio for maintenance stuff and trends regarding the competence alterations should be found and with the topic modeling approach further clustering of competences to a specific job should be realized.

The major obstacle for the count approach is that competences mentioned in the job advertisements need to be found automatically, but usually the wording of these competences differs from the wording used in the ESCO standard. For this reason, an ontology-based synonym recognition is realized.

With the count approach it is possible to construct a competence portfolio and observe trends within the data, but this is only possible on the general level of the database and there is no possibility to evaluate the competences of specific levels of jobs (e.g. IT, electric, mechanical related maintenance jobs). Furthermore, the synonym recognition has several limitations. In the realization there is no possibility to recognize contextual synonyms and the ontology-based synonym recognition also has its limitations. The evaluation of the results shows, that 24% of the manually found competences are also found with the matching algorithm. Considering only a specific job and adding a manually developed extensional competence list based on competences from 15 job advertisements the evaluation results increased to 76%.

The clustering of the competences to specific jobs with the topic modeling approach, with which clusters the corpus to different topics, did not deliver interpretable results and therefore no further clustering was possible.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Ausgangssituation	1
1.2	Problemstellung (Forschungsfragen).....	2
1.3	Problemrelevanz.....	3
1.4	Ansatz zur Lösung der Problemstellung	4
1.5	Vorgehensweise	5
1.6	Aufbau und Strukturierung dieser Arbeit.....	6
2	Theoretischer Hintergrund	8
2.1	Überblick Industrieentwicklung und Instandhaltung.....	8
2.2	Kompetenzenmanagement	13
2.3	Text Mining.....	30
3	Stand der Forschung.....	45
3.1	Existierende Lösungsansätze für die Problemstellung	46
3.2	Ähnliche Lösungsansätze für die Problemstellung	50
3.3	Stand der Praxis	50
4	Design und Umsetzung.....	52
4.1	Vom Text Mining Prozess zum abstrakten Konzept	52
4.2	Erweiterungen auf Basis der ersten Evaluierung.....	61
5	Auswertung und Resultate	64
5.1	Auswertung der angewendeten Methoden	64
5.2	Resultate in Bezug auf die Forschungsfragen.....	79
5.3	Möglicher Nutzen in Bezug auf die Problemstellung	82
6	Diskussion und Ausblick	83
6.1	Diskussion der Ergebnisse	83
6.2	Einschränkungen der Ansätze und Ergebnisse	84
6.3	Weiterentwicklungsmöglichkeiten.....	86
7	Anhang.....	88
8	Literaturverzeichnis	101
9	Abbildungsverzeichnis	113
10	Tabellenverzeichnis	114

11	Formelverzeichnis	115
12	Abkürzungsverzeichnis	116

1 Einleitung

In der Einleitung wird die Ausgangssituation (1.1), die Problemstellung (1.2), die Relevanz des Problems (1.3), die Methodik (1.4), die Vorgehensweise (1.5) und der weitere Aufbau (1.6) dieser Arbeit beschrieben.

1.1 Ausgangssituation

Der stetig zunehmende Wettbewerbsdruck in der Produktionstechnik stellt Firmen des Maschinen- und Anlagenbaus immer wieder vor neue Herausforderungen, um auch in Zukunft wettbewerbsfähig bleiben zu können. Gleichzeitig wird die industrielle Produktion immer dynamischer und komplexer.¹ Große Hoffnung wird dabei in die zunehmende Digitalisierung von Industriebetrieben gesetzt.

Die zunehmende Digitalisierung von Industriebetrieben, oft auch in Schlagwortform als Industrie 4.0 bezeichnet, hat zum Ziel die Produktion schneller, effektiver, flexibler und mit erhöhter Qualität, bei gleichzeitiger Senkung der Kosten, bewerkstelligen zu können.² Um diese Ziele erreichen zu können, bedarf es verschiedener struktureller und technologischer Veränderungen. Diese Veränderungen haben großen Einfluss auf Beschäftigte in den Betrieben und deren Kompetenzen, welche als ein zentrales Element in der Vision Industrie 4.0 angesehen werden.³ Die Kompetenzen betreffen unterschiedliche Bereiche in den Unternehmen und damit einhergehend unterschiedliche Beschäftigte. Die Kompetenzen erstrecken sich vom technischen Kontext (z.B. Umgang mit CAE Systemen, Programmierfähigkeiten, Simulationsfähigkeiten) bis zum nicht technischen Kontext (z.B. Kreativität, Teamwork und lebenslanges Lernen).⁴

Ein Baustein, um die obigen Ziele erreichen zu können, ist die Umgestaltung der Instandhaltungsstrategien, sodass die Gesamtverfügbarkeit der Produktionsanlagen maximiert werden kann. Die Gesamtverfügbarkeit kann mit Hilfe der vorausschauenden Instandhaltungsstrategie erhöht werden.⁵

Diese Veränderung hat zunehmend Einfluss auf die benötigten Kompetenzen der Beschäftigten in der Instandhaltung.⁶ Krason et al. hebt in diesem Artikel folgende zukünftige Kompetenzen hervor (1) IT-Kompetenz, um auftretende Probleme an die Spezialisten kommunizieren zu können (2) Umgang mit 3D Simulationen (3) Arbeiten

¹ Vgl. Kagermann et al. (2013), S. 19.

² Vgl. Bauernhansl et al. (2016), S. 8.

³ Vgl. Kagermann et al. (2013), S. 27.

⁴ Vgl. Maisiri et al. (2019), S. 98.

⁵ Vgl. David Stark (2010), S. 19.

⁶ Vgl. Krason et al. (2018), S.52.

mit zusätzlichen Informationsquellen (Augmented Reality) (4) Umgang und Analyse von großen Datenmengen (5) kommunikative und sozialen Kompetenzen.⁷

Unternehmen stellen die Nachfrageseite des Jobmarktes dar und fragen nach dezidierten Kompetenzen in Jobausschreibungen. Diese Jobbeschreibungen sollten nicht nur die derzeitigen benötigten Kompetenzen von Beschäftigten umfassen, sondern auch auf die langfristige Strategie eines Unternehmens ausgerichtet sein und somit zukünftig benötigte Kompetenzen enthalten.⁸ Die Herausforderung ist, zu erkennen welche Kompetenzen auf dem Vormarsch sind und zumindest mittelfristig einen wichtigen Stellenwert für das Unternehmen einnehmen werden.

Gleichzeitig steht auch die berufliche Erst- und Weiterbildung vor der Herausforderung, ihre Curricula an die Anforderungen der Unternehmen anzupassen, um deren Absolventen und Absolventinnen eine gute Basis an Wissen, Kompetenzen und Fähigkeiten vermitteln zu können, welche zukünftig von Unternehmen nachgefragt werden.⁹

1.2 Problemstellung (Forschungsfragen)

Wie bereits in Kapitel 1.1 erwähnt, hat die Umsetzung der Vision Industrie 4.0 großen Einfluss auf die KSAOs (*Knowledge, Skills, Abilities and other Characteristics* = Wissen, Fertigkeiten, Fähigkeiten und weitere Charakteristiken) von Beschäftigten in den unterschiedlichsten Bereichen eines produzierenden Unternehmens und damit wiederum Einfluss auf den Erfolg des gesamten Unternehmens und wiederum Einfluss auf die gesamte Volkswirtschaft.

In verschiedenen Publikationen wird bereits die Veränderung von nötigen KSAOs im Zusammenhang mit der Implementierung von Industrie 4.0 und der sie charakterisierenden Technologien hervorgehoben und untersucht. Beispiele dafür sind die systematische Literaturanalyse von Silva et al. (2019), die Publikation von Krason et al. (2018), die Literaturanalyse gekoppelt mit Fokusgruppeninterviews von Pinzone et al. (2017), die Literaturanalyse mit subjektiver Interpretation von Guzmán et al. (2020) oder die systematische Literaturanalyse von Maisiri et al. (2019). Bei der Analyse dieser Publikationen fällt auf, dass die nötigen KSAOs zumeist generischer Natur sind und nicht spezifisch auf die Instandhaltung abzielen. Außer jene von Krason et al. (2018), auf welche im folgenden Absatz näher eingegangen wird.

Die Umsetzung dieser, in 2.1.3 aufgezeigten Anwendungen, hat Einfluss auf die benötigten Kompetenzen und Fertigkeiten der Beschäftigten in der Instandhaltung.

⁷ Vgl. Krason et al. (2018), S.50f.

⁸ Vgl. Parry (1996), S. 56.

⁹ Vgl. Cedefop (2018), S. 16.

Krason et al. (2018) heben in deren Publikation folgende zukünftige Kompetenzen von Beschäftigten in der Instandhaltung hervor:¹⁰

- Generelle IT-Kompetenz, um auftretende Probleme an spezialisierte Personen kommunizieren zu können
- Umgang mit 3D Simulationen
- Arbeiten mit zusätzlichen Informationsquellen (wie Augmented Reality)
- Umgang und Analyse von großen Datenmengen
- Kommunikative und soziale Kompetenzen

Daraus ergeben sich folgende Fragen, wie setzt sich die nachgefragte KSAOs-Mischung in der industriellen Instandhaltung zusammen und lassen sich zeitliche Veränderungen erkennen?

1.3 Problemrelevanz

Unternehmen stellen die Nachfrageseite des Jobmarktes dar und fragen in Jobausschreibungen nach dezidierten Kompetenzen. Diese Jobausschreibungen enthalten Jobbeschreibungen und sollten nicht nur die derzeit benötigten Kompetenzen von Beschäftigten umfassen, sondern auch auf die langfristige Strategie eines Unternehmens ausgerichtet sein und somit zukünftig benötigte Kompetenzen enthalten.¹¹ Die Herausforderung ist, zu erkennen welche Kompetenzen auf dem Vormarsch sind und zumindest mittelfristig einen wichtigen Stellenwert für das Unternehmen einnehmen werden.

Gleichzeitig stehen auch die Institutionen der beruflichen Erst- und Weiterbildung vor der Herausforderung, ihre Curricula an die Anforderungen der Unternehmen anzupassen, um ihren Absolventen und Absolventinnen eine gute Basis an KSAOs vermitteln zu können, welche zukünftig von Unternehmen nachgefragt werden und den Absolventen und Absolventinnen einen guten Übergang in das weitere berufliche Leben ermöglichen.¹²

Das Problem liegt darin, dass beide dieser Akteure nur einen kleinen Einblick in die nachgefragten KSAOs haben, entweder aus der Sicht eines Unternehmens oder einer Bildungsinstitution. Diese Arbeit soll dabei helfen, zu einem bestimmten Zeitpunkt einen Einblick in die nachgefragten Kompetenzen in der Instandhaltung zu erhalten. Bei der Durchführung von zeitlich verschobenen Auswertungen sollen Trends erkannt werden.

¹⁰ Vgl. Krason et al. (2018), S. 52.

¹¹ Vgl. Parry (1996), S. 56.

¹² Vgl. Cedefop (2018), S. 16.

Die Veränderungen als Folge der Umsetzung der Vision Industrie 4.0 könnten zudem die Diskrepanz zwischen erworbenen und nachgefragten Fertigkeiten (engl. „skill mismatch“) noch weiter verschärfen. Eine OECD Studie aus dem Jahr 2013 stellte fest, dass bereits 21% der Beschäftigten über- und 13% unterqualifiziert für deren Jobs waren.¹³ Diese Studie untersucht unter anderem die Lese-, Rechen- und Problemlösefertigkeit von Erwachsenen im Alter von 16-65 und zeigt auf, dass der gesellschaftliche Mehrwert, zusammengesetzt aus beruflicher und sozialer Natur, wie etwa die Ausübung eines Ehrenamts, eines Individuums mit diesen Fertigkeiten zusammenhängt.¹⁴ Dabei muss berücksichtigt werden, dass die durchschnittliche Ausbildungsdauer zwar zukünftig weiter steigen wird, sich jedoch die Wachstumsraten verlangsamen und daraus folglich wird die Quantität an Fertigkeiten, die in der gesamten Gesellschaft zur Verfügung stehen, langsamer wachsen als in den letzten Jahrzehnten.¹⁵

1.4 Ansatz zur Lösung der Problemstellung

Diese Arbeit untersucht den Informationsgehalt von im Internet verfügbaren Jobausschreibungen auf Online-Jobportalen in Anbetracht der nachgefragten KSAOs im Instandhaltungsbereich mit Hilfe von Text Mining Methoden.

Ziel ist es, ein Konzept zu erstellen mit welchem die KSAOs von Jobausschreibungen im Instandhaltungsbereich statistisch bewertet werden. In Anbetracht einer zusätzlichen zeitlichen Veränderung können Trends der KSAOs aufgezeigt werden. Dieses Konzept wird daraufhin einem Proof-of-Concept unterzogen werden.

Die Vorgehensweise dieser Arbeit orientiert sich an den sieben Richtlinien von Hevner et al. (2004) für die Gestaltung von Forschungsarbeiten im Bereich von Informationssystemen. Die sieben Richtlinien sind folglich aufgelistet und werden in Bezug auf diese Arbeit gebracht:¹⁶

1. *Design as an Artifact*
2. *Problem Relevance*
3. *Design Evaluation*
4. *Research Contribution*
5. *Research Rigor*
6. *Design as a Search Process*
7. *Communication of Research*

¹³ Vgl. OECD (2013), S. 24.

¹⁴ Vgl. ebd., S. 246.

¹⁵ Vgl. Braconier et al. (2014), S. 16.

¹⁶ Vgl. Hevner et al. (2004), S. 83.

1. Das Artefakt der Arbeit ist die Erstellung des Konzeptes und die Evaluierung der Ergebnisse. Das Konzept erstreckt sich von der Datensammlung bis zur Auswertung und wird mit Hilfe eines Pythonskripts umgesetzt.
2. Es werden erweiterte Einblicke in die derzeitig und zukünftig benötigten Kompetenzen im Instandhaltungsbereich für Firmen und berufliche Erst- und Weiterbildungsinstitutionen geboten. Um aus Firmensicht die passenden Beschäftigten einzustellen, vorhandene Beschäftigte umzuschulen oder als Bildungsinstitution den Absolventen und Absolventinnen die nachgefragten KSAOs zu lehren.
3. Es handelt sich um eine observierende Fallstudie, da ausgeschriebene Jobausschreibungen zur statistischen Auswertung herangezogen werden.
4. Diese Arbeit gibt einen Einblick in das Kompetenzprofil von Instandhaltern und Instandhalterinnen auf Basis von aktuellen Jobausschreibungen und zeigt über zeitlich abgegrenzte Durchführungen Trends der Kompetenzen auf.
5. Text Mining Methoden sind in der Organisationsforschung anerkannt und können gut für die Analyse von großen Datenmengen eingesetzt werden. Die Genauigkeit der Ergebnisse ist stark von den gewählten Methoden abhängig.
6. Die Text Mining Methoden werden nach deren Sinnhaftigkeit und Anwendbarkeit ausgewählt und erprobt. Das Vorgehen und die angewandten Methoden werden mehrfach evaluiert und iterativ verbessert, um die Genauigkeit zu erhöhen.
7. Die Ergebnisse dieser Arbeit richten sich an Firmen mit Instandhaltungspersonal und an berufliche Erst- und Weiterbildungsinstitutionen. Sie helfen dabei, diesen Organisationen einen tieferen Einblick in die nachgefragten Kompetenzen zu erhalten.

1.5 Vorgehensweise

Zuerst werden die theoretischen Grundlagen aus der Literatur im Bereich von Kompetenzmanagement in der Instandhaltung, Jobanalyse, Jobprofilanalyse, KSAOs auf Personenebene, Jobtaxonomien und Text Mining recherchiert und zusammengefasst. Dieses Wissen stellt die Ausgangsbasis für den Entwurf des Konzeptes dar. Dieses Konzept wird mit einem Python Script umgesetzt. Das Konzept umfasst mehrere Schritte des Text Mining Prozesses von der Datenerfassung bis zu der Auswertung. Es werden vor allem Text Mining Methoden aus der Literatur verwendet und eine Jobtaxonomie dient als Vergleichsbasis.

Die Daten der Jobausschreibungen im Instandhaltungsbereich werden mit Hilfe von Web Scraping (automatisierte Datenextrahierung von Websites) von Online-

Jobportalen gewonnen und die Rohdaten mit Preprocessing Methoden für die spätere Bewertung aufbereitet.

Die einzelnen KSAOs werden daraufhin nach deren Sinnhaftigkeit zusammengefasst, um ihre Aussagekraft richtig bewerten zu können. Dabei ergeben sich zwei Schwierigkeiten, einerseits die Verwendung von verschiedenen Begriffen für die gleiche Anforderung oder Tätigkeit, wie „drive innovation“ und „innovative solutions“¹⁷ und andererseits die Verwendung desselben Begriffs für unterschiedliche Anforderungen oder Tätigkeiten, so gilt der Begriff Java als Homonym und hat somit mehrere Bedeutungen. Java ist im Englischen ein Synonym für Kaffee und gleichzeitig eine Programmiersprache.¹⁸

Diese Einteilung und Zusammenfassung dient als Grundlage, um die nachgefragten KSAOs statistisch bewerten zu können und somit nachgefragte Trends im Jobmarkt erkennen zu können. Dabei wäre es sinnvoll, die statistische Auswertung an mehreren Zeitpunkten durchzuführen, um diese Trends besser validieren und ihre Veränderung beobachten zu können. Letzteres wird, aus zeitlichen Gründen, nur in einem eingeschränkten Zeitraum in dieser Arbeit möglich sein und deshalb nicht repräsentativ in Bezug auf Zeitraum, Umfang und Detaillierung sein.

1.6 Aufbau und Strukturierung dieser Arbeit

In Abbildung 1 ist Aufbau und Struktur dieser Arbeit prägnant zusammengefasst und wird folglich genauer erläutert.

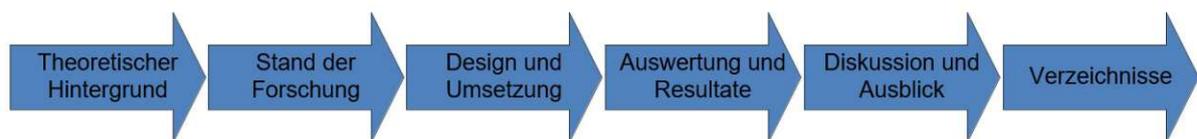


Abbildung 1: Aufbau und Struktur dieser Arbeit¹⁹

Kapitel 2, Theoretischer Hintergrund, fasst alle relevanten Informationen in Bezug auf die Veränderung der Industrie, Kompetenzen und Text Mining in dieser Arbeit zusammen. Detaillierter umfasst dies im kompetenzbezogenen Bereiche KSAOs auf Personenebene in der industriellen Instandhaltung, Kompetenzmanagement in der Instandhaltung, Jobanalyse, Jobausschreibung und Jobtaxonomie, letztere inklusive einiger Standardtaxonomien. In Bezug auf Text Mining werden mehrere Methoden und relevanten Problematiken vorgestellt.

¹⁷ Vgl. Pejic-Bach et al. (2020), S. 421.

¹⁸ Vgl. Zhao et al. (2015), S. 4012.

¹⁹ Eigene Darstellung.

Kapitel 3, Stand der Forschung, beschreibt Lösungsansätze, welche sich mit der gleichen oder einer ähnlichen Problemstellung befassen oder für die verwendete Methodik in dieser Arbeit die Basis bilden und fasst diese zusammen.

Kapitel 4, Design und Umsetzung, beschreibt die angewandten Methoden im Detail inklusive der verwendeten Programmierpakete und Teilen des Python Scripts.

Kapitel 5, Auswertung und Resultate, stellt die gewonnenen Ergebnisse dar und im Anschluss den Bezug zur Problemstellung her.

Kapitel 6, Diskussion und Ausblick, befasst sich mit der kritischen Betrachtung der Methodik und der Ergebnisse. Die Einschränkungen bei diesem Ansatz werden aufgezeigt und mit den möglichen nächsten Schritten dieser Arbeit diskutiert.

Kapitel 7 bis 12 beinhaltet in dieser Reihenfolge Anhang, Literatur-, Abbildungs-, Formelverzeichnis, Tabellen- und Abkürzungsverzeichnis.

2 Theoretischer Hintergrund

2.1 Überblick Industrieentwicklung und Instandhaltung

2.1.1 Evolution der Industrie

Sowohl der Fortschritt als auch die Überlebensfähigkeit und der Wohlstand der Menschheit war seit jeher an ihre Innovationsfähigkeit gekoppelt. Dies zieht sich von der Entdeckung und Nutzung des Feuers, über die Entwicklung des Rades bis hin zum jetzigen Zeitpunkt.

Der Begriff Industrie wird laut DUDEN als ein „Wirtschaftszweig, der die Gesamtheit aller mit der Massenherstellung von Konsum- und Produktionsgütern beschäftigten Fabrikationsbetriebe eines Gebietes umfasst“²⁰ definiert und ist somit stark auf die Herstellung von Massengütern ausgerichtet.

Ende des 18. Jahrhunderts wurde mit der Mechanisierung und der Nutzung von mechanischer Energie (Wasser- und Dampfkraft) die Möglichkeit eröffnet, die Güter mit Hilfe von einfachen mechanischen Produktionsanlagen, anstatt nur mit reiner manueller Arbeit, aufbereiten zu können. Die prägendsten Industrien war dabei die Textilindustrie, das Transportwesen und die Kohleindustrie.²¹ Diese Epoche wird als die erste industrielle Revolution bezeichnet.

Als zweite industrielle Revolution wird die Einführung von arbeitsteiliger Massenproduktion mit gleichzeitigem Nutzen von elektrischer Energie bezeichnet.²² Ein klassisches Beispiel dafür ist die erste Fließbandproduktion bei Ford, welche einen derartigen Preisverfall von Automobilen zur Folge hatte, was es auch Fließbandmitarbeiter_innen von Ford ermöglichte, sich ein Automobil anschaffen zu können.²³

Der Einsatz von Elektronik und Informationstechnologien ab den 1970er Jahren im produzierenden Gewerbe und die damit einhergehende Automatisierung von Produktionsprozessen wird als die dritte industrielle Revolution bezeichnet.²⁴ Dies könnte als erste digitale Ausbaustufe der Industrie verstanden werden.

²⁰ URL: <https://www.duden.de/node/70642/revision/70678>, Stand: 28.08.2020.

²¹ Vgl. Hobsbawm (1999), S. 28.

²² Vgl. Kagermann et al. (2013), S. 17.

²³ Vgl. Mokyr und Strotz (1998), S. 9.

²⁴ Vgl. Kagermann et al. (2013), S. 18.

2.1.2 Industrie 4.0 als Vision

Der stetig zunehmende Wettbewerbsdruck in der Produktionstechnik stellt Firmen des Maschinen- und Anlagenbaus immer wieder vor neue Herausforderungen, um auch in Zukunft wettbewerbsfähig bleiben zu können. Gleichzeitig wird die industrielle Produktion immer dynamischer und komplexer.²⁵ Auf der Hannover Messe 2011 wurde erstmals das Schlagwort „Industrie 4.0“ für eine Vision einer vierten industriellen Revolution verwendet. 2016 war das Motto des Weltwirtschaftsforums in Davos „Mastering the Fourth Industrial Revolution“.²⁶ Dies verdeutlicht die zunehmende Wichtigkeit des Themas. Die vierte industrielle Revolution ist die erste industrielle Revolution, die vor ihrer Umsetzung ausgerufen wurde.²⁷

Das allgemeine Ziel in der industriellen Wertschöpfung ist es, die Produktion schneller, effektiver, flexibler und mit erhöhter Qualität, bei gleichzeitiger Senkung der Kosten zu bewerkstelligen. In Bezug auf die Vision Industrie 4.0 werden vor allem die Anforderungen an die Geschwindigkeit, Wirtschaftlichkeit, Flexibilität, Wandlungsfähigkeit und Nachhaltigkeit hervorgehoben.²⁸

Die Vision von Industrie 4.0 will mit drei technologischen Möglichkeiten die Produktionstechnik revolutionieren:²⁹

- Horizontale Vernetzung – Internet der Dinge und Dienste
- Vertikale Integration – Einsatz von Cyber-physischen Systemen
- Echtzeitoptimierung – Verarbeitung von großen Datenmengen und darauf basierende Prognosen

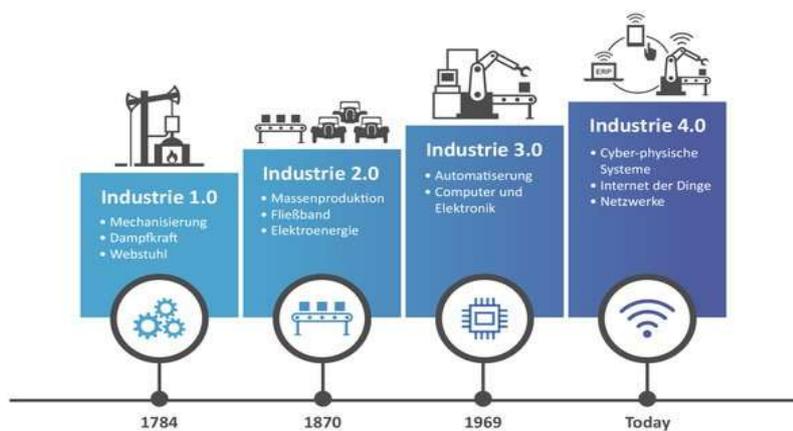


Abbildung 2: Evolution der Industrie³⁰

²⁵ Vgl. Kagermann et al. (2013), S. 19.

²⁶ Vgl. Pfeiffer (2017), S. 107.

²⁷ Vgl. Rojko (2017), S. 88.

²⁸ Vgl. Bauernhansl et al. (2016), S. 8.

²⁹ Vgl. ebd., S. 6.

³⁰ URL: <https://www.btelligent.com/themen/industrie-40/>, Stand: 29.08.2020.

Das Potenzial von Industrie 4.0 erstreckt sich über unterschiedliche Themenfelder:³¹

- Individualisierung der Kundenwünsche
- Flexibilisierung
- Optimierte Entscheidungsfindung
- Ressourcenproduktivität und -effizienz
- Wertschöpfungspotenziale
- Demografie-sensible Arbeitsgestaltung
- Work-Life-Balance
- Wettbewerbsfähigkeit als Hochlohnstandort

Aus einer Firmenumfrage des Unternehmensberaters, Wirtschaftsprüfers und Steuerberaters PwC (PricewaterhouseCoopers) aus 2014 ging hervor, dass Firmen den monetären Mehrwert von Industrie 4.0 Anwendungen im Durchschnitt auf eine mögliche Kostenreduktion von 13,8% und eine mögliche Effizienzsteigerung von 17,9% in den nächsten fünf Jahren schätzten. Des Weiteren geht daraus hervor, dass 2014 ca. 20% der befragten Firmen ihre Schlüsselprozesse bereits digitalisiert haben und dieser Anteil bis 2019 nach eigener Einschätzung auf 85% ausgebaut werden soll.³²

2016 schätzten Bauernhansl et al. die möglichen Kostenreduktionen in unterschiedlichen Unternehmensbereichen ab und kamen dabei zum Schluss, dass beispielsweise die Bestandskosten um 30-40%, die Fertigungskosten um 10-30%, die Komplexitätskosten um 60-70% und die Instandhaltungskosten um 20-30% reduziert werden könnten.³³

Um diese Ziele erreichen zu können, beziehungsweise die Potenziale voll ausschöpfen zu können, bedarf es verschiedener struktureller und technologischer Veränderungen.

Diese Veränderungen haben großen Einfluss auf die Kompetenzen der Beschäftigten in den Unternehmen. Die Beschäftigten werden als ein zentrales Element in der Vision Industrie 4.0 angesehen werden.^{34,35} Die Kompetenzen der Beschäftigten können unter dem Begriff der KSAOs zusammengefasst werden. In Kapitel 2.2.1 wird der Begriff der Kompetenz diskutiert und der Zusammenhang zu KSAOs erläutert.

Verschiedene Aufgaben und Problemstellungen erfordern unterschiedliche KSAOs der Beschäftigten. Die systematische Literaturrecherche von Maisiri et al. aus dem

³¹ Vgl. Kagermann et al. (2013), S. 19.

³² Vgl. Geissbauer et al. (2014), S. 22.

³³ Vgl. Bauernhansl et al. (2016), S. 9.

³⁴ Vgl. Kagermann et al. (2013), S. 27.

³⁵ Vgl. Kaare und Otto (2015), S. 827.

Jahr 2019 analysierte 68 Publikationen, welche auf die Zusammenhänge zwischen Industrie 4.0 und einzelnen Fertigkeiten ausgerichtet waren, fasste diese zusammen und kategorisierte die einzelnen Fertigkeiten. Diese Fertigkeiten erstrecken sich vom technischen Bereich (z.B. Umgang mit CAE Systemen, Programmierfertigkeit, Simulationsfertigkeiten) bis zum nicht technischen (z.B. Kreativität, Teamwork und lebenslanges Lernen).³⁶

2.1.3 Industrie 4.0 in der Instandhaltung

Der Begriff Instandhaltung wird in der DIN 31051 (Ausgabedatum 2019-06) wie folgt definiert, „*Kombination aller technischen und administrativen Maßnahmen sowie Maßnahmen des Managements während des Lebenszyklus eines Objekts, die dem Erhalt oder der Wiederherstellung ihres funktionsfähigen Zustands dient, sodass es die geforderte Funktion erfüllen kann*“³⁷.

Die Ziele der Instandhaltung sind es, die Abnutzungsgeschwindigkeit von Objekten beziehungsweise Anlagen zu verzögern und deren Verfall und mögliche Zerstörung zu verhindern, die Arbeitssicherheit für alle beteiligten Beschäftigten zu gewährleisten und mögliche umweltschädliche Unfälle zu vermeiden.³⁸ Für produzierende und dienstleistende Unternehmen steht dabei oft die Sicherung einer möglichst störungsfreien Nutzung der Objekte und Anlagen im Vordergrund.³⁹ Dieses Ziel soll mit Hilfe von so wenig aufzuwendenden Ressourcen wie möglich erreicht werden, um daraus einen wirtschaftlichen Vorteil erzielen zu können. Aus diesem Grund stellt die von vgl. Bauernhansl et al. (vgl. 2016, S. 9) postulierte mögliche Kostenreduktion von 20-30% einen großen Anreiz für Unternehmen dar, dieses Potenzial voll auszuschöpfen.⁴⁰

Für die Erreichung dieses Ziels werden vier Instandhaltungsstrategien eingesetzt, um für die jeweiligen Objekte und Anlagen je nach deren Anforderungen und damit verbundenen Zielen das anlagenspezifische Optimum erreichen zu können.⁴¹ Die vier Instandhaltungsstrategien sind⁴²:

1. Ausfallsbehebung
2. Zeitgesteuerte periodische Instandhaltung
3. Zustandsorientierte Instandhaltung
4. Vorausschauende Instandhaltung

³⁶ Maisiri et al. (2019), S. 98ff.

³⁷ DIN 31051, S. 4.

³⁸ Vgl. Strunz (2012), S. 2.

³⁹ Vgl. ebd., S. 2.

⁴⁰ Vgl. Bauernhansl et al. (2016), S. 9.

⁴¹ Vgl. Matyas (2002), S. 13.

⁴² Vgl. ebd., S. 13.

Die Ausfallbehebung ist einzige nicht-vorbeugenden Instandhaltungsstrategie, bei welcher ein Fehler erst nach dem Ausfall einer Anlage behoben wird. Diese Strategie wird zumeist bei kleinen unkritischen Anlagen, mit einer geringen Ausfallwahrscheinlichkeit und Redundanz eingesetzt.⁴³ Selbst bei einem Ausfall der Anlage hat dies nur geringe Folgen für das Unternehmen.⁴⁴

Die zeitgesteuerte periodische Instandhaltung wird eingesetzt, wenn die mittlere Zeit zwischen zwei Schäden bekannt ist und nur eine geringe Streuung aufweist.⁴⁵ Gleichzeitig muss die technische Möglichkeit eines vorzeitigen Tausches gegeben sein.

Die zustandsorientierte Instandhaltung wird verwendet, wenn der Zustand messbar ist und dadurch zufällig Ausfälle vermieden werden können.⁴⁶

Die vorausschauende Instandhaltung umfasst die weitere Vermeidung eines aufgetretenen Fehlers. Zumeist wird die Fehlerursache durch konzeptionelle Änderungen vermieden, zum Beispiel durch eine Konstruktionsänderung.⁴⁸



Abbildung 3: Instandhaltungsstrategien und deren Merkmale⁴⁷

Kriterien für Auswahl der passenden Instandhaltungsstrategien sind Verkettung der Anlagen, Redundanz von Anlagen, Gültigkeit von Qualitäts-, Umwelt-, und Sicherheitsnormen, Arbeitszeitenvereinbarungen, Instandsetzungszeiten, Ersatzteilverfügbarkeit, Materialpuffer zwischen den Anlagen und

⁴³ Vgl. ebd., S. 14.

⁴⁴ Vgl. ebd., S. 14.

⁴⁵ Vgl. ebd., S. 14.

⁴⁶ Vgl. ebd., S. 14.

⁴⁷ Eigene Darstellung angelehnt an Moubrey (1996) zitiert nach Matyas (2002), S. 13.

⁴⁸ Vgl. Matyas (2002), S. 15.

Belastungsspitzen durch saisonale Schwankungen am Markt und bei der Rohstoffverfügbarkeit.⁴⁹

Bei den Instandhaltungsstrategien kann zwischen nicht-vorbeugender und vorbeugender Instandhaltung unterschieden werden, wobei unter Ersteres nur die Ausfallbehebung fällt, bei allen anderen Strategien soll ein möglicher Fehler präventiv vermieden werden.

Die mögliche Kostenreduktion von 20-30% in der Instandhaltung könnten durch folgende drei Hauptanwendungen realisiert werden:⁵⁰

- Optimierung der Lagerbestände von Ersatzteilen
- Zustandsorientierte Instandhaltung
- Dynamische Priorisierung

Zum Beispiel kann die Gesamtverfügbarkeit (OEE = Overall Equipment Efficiency), ein gängiger Key Performance Indikator bei Produktionsanlagen, mit Hilfe einer Kombination von datengestützter, zustandsbasierter und vorausschauender Instandhaltungsstrategie erhöht werden.⁵¹

2.2 Kompetenzmanagement

Der Begriff Kompetenz wird vor allem in den Organisations- und Managementwissenschaften verwendet und in der Literatur unterschiedlich definiert. Die Kompetenz im unternehmerischen Sinn kommt sowohl aus dem strategischen Management als auch aus dem Personalmanagement, welche auf unterschiedlichen organisatorischen Ebenen Einfluss hat.⁵² Dies hat zur Folge, dass es keine eindeutige Definition des Begriffes mit klarem Umfang gibt.⁵³

Im strategischen Management liegt seit den 1990er Jahren der Fokus zumeist auf der organisatorischen Kernkompetenz, welche für das Unternehmen aus wettbewerblicher Sicht wichtig ist und dem Unternehmen zukünftiges Wachstum ermöglichen soll.⁵⁴ Kernkompetenzen verbrauchen sich nicht über die Zeit, sondern werden durch das Anwenden gefördert, viel eher gehen Kompetenzen, welche nicht immer wieder angewandt werden, verloren. Das derzeitige Geschäftsmodell von Unternehmen baut auf die Kernkompetenzen auf und gleichzeitig sollen diese neue Geschäftsmöglichkeiten eröffnen, da sie einen wettbewerblichen Vorteil für das

⁴⁹ Vgl. Matyas (2002), S. 14.

⁵⁰ Vgl. Bauernhansl et al. (2016), S. 9.

⁵¹ Vgl. David Stark (2010), S. 19.

⁵² Vgl. Simpson (2002), S. 55.

⁵³ Vgl. Nordhaug und Grønhaug (1994), S. 91.

⁵⁴ Vgl. Prahalad und Hamel (1997), S. 2.

Unternehmen darstellen.⁵⁵ Kernkompetenzen werden laut Prahalad und Hamel (1997) nach drei Hauptmerkmalen charakterisiert. Sie sollen einen Zugang zu einer Vielzahl an Märkten ermöglichen, der Kundschaft einen wahrgenommenen Mehrwert am Endprodukt vermitteln können und schwer von Konkurrenz nachgeahmt werden können.⁵⁶

Der Fokus des Personalmanagements liegt hingegen auf den personenspezifischen Kompetenzen. Diese personenspezifischen Kompetenzen befassen sich mit den fundamentalen persönlichen Charakteristiken einer Person und deren Handlungen im Zusammenhang mit einer Aufgabe und der gegebenen Situation.⁵⁷

In Abbildung 4 werden die Kernkompetenzen in der Organisation zugeordnet und der Zusammenhang zwischen Kernkompetenzen und individuellen Kompetenzen hergestellt.

In dieser Arbeit liegen die Kompetenzen auf Personenebene im Fokus, da diese in Stellenausschreibungen nachgefragt werden. Deshalb wird im nachfolgenden Kapitel 2.2.1 genauer auf die Kompetenzen auf Personenebene eingegangen.

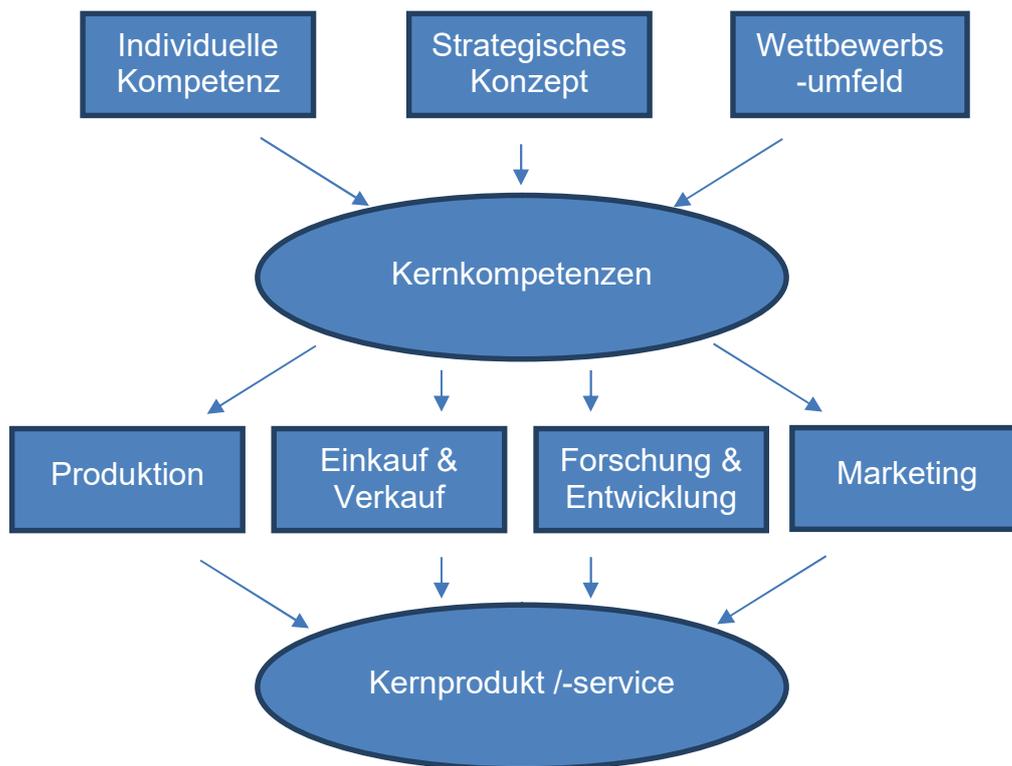


Abbildung 4: Zusammenhang individuelle Kompetenzen und Kernkompetenzen⁵⁸

⁵⁵ Vgl. Prahalad und Hamel (1997), S. 5.

⁵⁶ Vgl. ebd., S. 7.

⁵⁷ Vgl. Bergenhenegouwen (1996), S. 30.

⁵⁸ Eigene Darstellung angelehnt an ebd., S. 30.

2.2.1 Kompetenzen auf Personenebene

Zudem gibt es verschiedene Charakteristiken in Jobbeschreibungen, die nicht eindeutig als Kompetenzen bezeichnet werden können.⁵⁹ Nachfolgend sind einige der gängigsten Definitionen des Begriffs auf der individuellen Ebene angeführt, um einen Eindruck der Vielfältigkeit des Kompetenzbegriffes zu erhalten.

Boyatzis (1982) definiert den Begriff Kompetenz als unterschwellige Eigenschaften eines Individuums, welche in Zusammenhang mit einer effektiv ausgeführten Arbeit steht.⁶⁰

Fleishman et al. (1995) beschreiben die Kompetenz als eine Mischung aus Wissen, Fertigkeiten, Fähigkeit, Motivation, Glauben, Werten und Interessen.⁶¹

Nordhaug (1993) fasst den Begriff Kompetenz als das arbeitsbezogene Wissen, die Fertigkeiten und die Fähigkeit einer Person zusammen. Er räumt aber bereits ein Jahr später in seinem Artikel ein, dass diese Definition zu eng gefasst ist, da sie nur jene Kompetenzen anspricht, welche zu diesem Zeitpunkt verwendet werden.⁶²

Spencer und Spencer (1993) verstehen unter Kompetenz die einem Individuum zugrunde liegenden Charakteristiken, welche in lockerem Zusammenhang mit einer effektiven oder herausragenden beruflichen Leistung stehen. Diese Charakteristiken sind demnach: Motive, Eigenschaften, Selbstkonzept (Werte, Attitüde und Selbstbild), Wissen und Fertigkeiten.⁶³ Sie unterteilen diese Charakteristiken des Weiteren anhand ihrer Erkennbarkeit. Gut erkennbare Charakteristiken sind etwa Wissen und Fertigkeiten, welche auch von außenstehenden Personen relativ schnell erkannt werden können und von Personen durch Training erlangt wurden. Versteckte Charakteristiken sind im Gegensatz dazu tiefer mit der Persönlichkeit verwoben und daher relativ schwierig von Dritten zu erkennen oder weiterzuentwickeln. Dennoch stellen gerade diese versteckten Charakteristiken den kosteneffektivsten Weg bei einer Auswahl für neues Personal dar, da die versteckten Eigenschaften im Nachhinein nur sehr schwer beeinflusst werden können. Als illustrative Darstellung entwickelten sie das Eisberg-Modell, siehe Abbildung 5.⁶⁴

⁵⁹ Vgl. Ellington et al. (2015), S. 1.

⁶⁰ Vgl. Boyatzis (1982), S. 23.

⁶¹ Vgl. Fleishman et al. (1995), S. 553.

⁶² Vgl. Nordhaug und Grønhaug (1994), S. 93.

⁶³ Vgl. Spencer und Spencer (1993), S. 9.

⁶⁴ Vgl. ebd., S. 11.

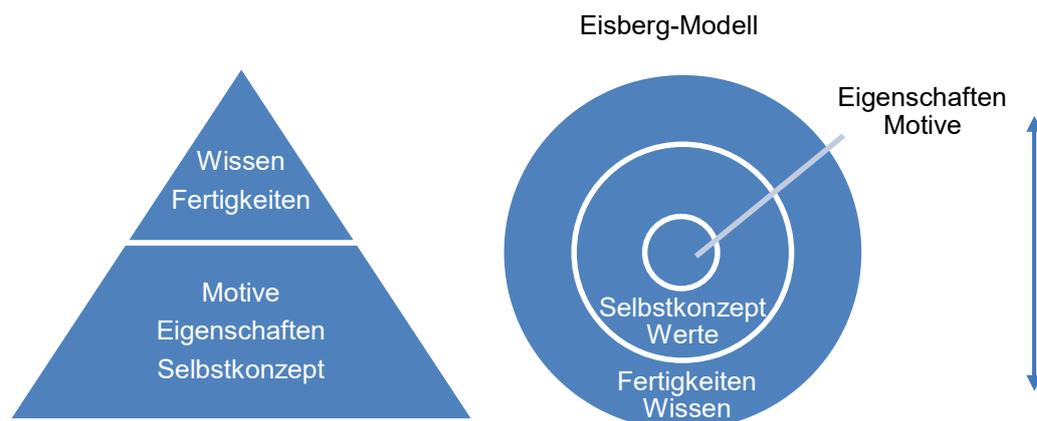


Abbildung 5: Oberflächliche und tieferliegende Eigenschaften⁶⁵

Kompetenzen können auch nach deren Detaillierungsgrad unterteilt werden. Dabei wird unterschieden, ob eine Kompetenz spezifisch auf einen Job ausgerichtet oder die Kompetenz generisch ist und in vielerlei Jobs benötigt wird. Beherrscht man zum Beispiel den Umgang mit einer speziellen Maschine, welche nur in einer Branche oder einem Unternehmen eingesetzt wird, stellt dies eine sehr spezifische Kompetenz dar. Im Gegensatz dazu steht die Kommunikationskompetenz, welche in unterschiedlichen Ausprägungen in jedem Job benötigt wird. In der Personalentwicklung birgt dieser Unterschied die Herausforderung, ob spezifische oder generische Kompetenzen in den Weiterbildungen der Beschäftigten gefördert werden sollten.⁶⁶

Die unterschiedlichen Kompetenzen können in verschiedenen Phasen eines Jobs wichtig sein, um neue Aufgaben gut bewerkstelligen zu können. Vor allem generische Kompetenzen befähigen Beschäftigte, sich selbstständig weiter zu entwickeln und ermöglichen ihnen damit auch einen leichteren Wechsel zwischen verschiedenen Jobs und Unternehmen.⁶⁷

Außerdem ist es wichtig, auf die sprachliche Bedeutung der Kompetenz hinzuweisen. In der englischen Sprache werden umgangssprachlich *Competence* und *Competency* häufig synonym verwendet, wobei es hier wichtig wäre, eine klare Unterscheidung zu treffen. *Competence* ist auf ein Ziel ausgerichtet und die dafür nötigen KSAOs werden im Kontext des Zieles abgeleitet. Es handelt sich hierbei also immer um einen Status. *Competency* wiederum wird als Verhalten verstanden, welches von erfolgreichen Personen abgeleitet wird.⁶⁸

⁶⁵ Eigene Darstellung angelehnt an Spencer und Spencer (1993), S. 11.

⁶⁶ Vgl. Elkin (1990), S. 4.

⁶⁷ Vgl. ebd., S. 4.

⁶⁸ Vgl. Teodorescu (2006), S. 28.

Im deutschen Sprachraum kann man in ähnlicher Weise zwischen Kompetenz und Qualifikation unterscheiden. Qualifikation steht für die Fähigkeit, bestimmte konkrete berufliche Situationen zu meistern und ist somit anwendungsbezogen. Kompetenz hingegen wird holistischer verstanden und umfasst neben spezifischem Wissen auch transversale Fähigkeiten, dennoch wird nach wie vor der Begriff Schlüsselqualifikation als Synonym von Kompetenz verwendet.⁶⁹

2.2.1.1 Wissen

Wissen wird manchmal als konkrete Manifestation aus abstrakter Intelligenz verstanden, jedoch ist es eigentlich ein Ergebnis aus der Interaktion von Intelligenz (Fähigkeit zu Lernen) und Situation (Möglichkeit zum Lernen), hieraus lässt sich ableiten, dass Wissen stark von der sozialen Umgebung beeinflusst wird und nicht rein von der Intelligenz abhängig ist.⁷⁰

Das Wissen kann in explizites und implizites Wissen unterteilt werden. Explizites Wissen ist auf unmittelbare Vermittlung von Theorien und Konzepten gestützt. Implizites Wissen stammt aus der Erfahrung bereits durchgeführter Aufgaben und kann von der Person meist nicht artikuliert werden. Neu errungenes Wissen tritt zuerst als implizites Wissen auf und muss erst in explizites Wissen überführt werden, um dieses Wissen auch anderen Menschen zugänglich machen zu können und somit auch andere davon profitieren können.⁷¹

Des Weiteren kann Wissen, auch wie bei den Kompetenzen erwähnt, anhand des Detaillierungsgrades unterschieden werden. Wissen erstreckt sich von spezifischem Wissen, wie Detailwissen einer bestimmten Maschine und deren Verhaltens, bis zu sehr generischem Wissen, wie ein großer Wortschatz.⁷²

In Jobausschreibungen handelt es sich immer um explizites Wissen nach welchem das ausschreibende Unternehmen nachfragt, da der Firma bewusst sein muss, welches Wissen für eine Stelle benötigt wird. Der Detaillierungsgrad des Wissens ist stark vom jeweiligen Jobprofil abhängig. Das nötige Wissen wird in der Regel mit Hilfe einer Jobanalyse evaluiert, welche in Kapitel 2.2.3 genauer beleuchtet wird.

2.2.1.2 Fertigkeit

Pear (1927) definierte als einer der Ersten den Begriff der Fertigkeit im Arbeitsumfeld und viele der darauffolgenden Definitionen sind an sein Grundgerüst angelehnt.

Das Konzept der Fertigkeit nach Pear (1927) umfasst die Integration von gut angepassten Leistungen und nicht die Zusammenfügung von Gewohnheiten. Die

⁶⁹ Vgl. Winterton et al. (2006), S. 34.

⁷⁰ Vgl. ebd., S. 9.

⁷¹ Vgl. Polanyi (1966), S. 1.

⁷² Vgl. Weinert (1999), S. 23.

Fertigkeit steht in direktem Zusammenhang mit der Quantität und Qualität einer Leistung. Eine Fertigkeit muss erlernt sein, um von einer Person ausgeführt werden zu können. Die bloße Fähigkeit der Person ermöglicht dieser nicht automatisch die Ausführung der Fertigkeit. Ist die Fertigkeit jedoch erlernt so verschmilzt diese mit der Fähigkeit.⁷³

Renold (1928) definiert den Begriff der Fertigkeit etwas umfassender, jegliche für die Industrie sinnvolle Kombination aus mentalen und physischen Qualitäten, welche durch Übung angeeignet wurden.⁷⁴

Im Allgemeinen sind Fertigkeiten nur in einem eingeschränkten Ausmaß übertragbar, das bedeutet, dass das Trainieren von Fertigkeiten nur bei genau diesen Fertigkeiten oder sehr ähnlichen Fertigkeiten zu Verbesserungen der Leistung führt. Dies gilt sowohl für motorische Fertigkeiten⁷⁵ als auch für kognitive Fertigkeiten⁷⁶.

2.2.1.3 Fähigkeit

Fähigkeit stellt eine andauernde generelle Eigenschaft beziehungsweise Tauglichkeit eines Individuums dar, welche durch die Ausübung verschiedener Aufgaben charakterisiert wird.⁷⁷

Zu Fähigkeiten zählen Eigenschaften wie die kognitive, soziale und emotionale Intelligenz einer Person.⁷⁸

2.2.1.4 Andere Merkmale

Unter anderen Merkmalen sind alle persönlichen Eigenschaften und individuellen Charakteristiken zu verstehen, welche einen Einfluss auf die Arbeitsleistung haben.

Nimmt man zum Beispiel die Definition der Kompetenz von Fleishman et al. (1995) als Ausgangspunkt, umfassen die anderen Merkmale Motivation, Glauben, Werte und Interessen.⁷⁹

Nach Spencer und Spencer (1993) wiederum umfassen die anderen Merkmale die Motive, die Eigenschaften und das Selbstkonzept einer Person. Diese Merkmale sind zumeist nicht sofort offen sichtbar an einer Person, sondern spiegeln viel mehr die Persönlichkeit der Person wider.⁸⁰

⁷³ Vgl. Pear (1927) zitiert nach Adams (1987), S. 41f.

⁷⁴ Vgl. Renold (1928), S. 593.

⁷⁵ Vgl. Cook (1934), S. 762.

⁷⁶ Vgl. Singley und Anderson (1989), S. 35.

⁷⁷ Vgl. Fleishman et al. (1995), S. 553.

⁷⁸ Vgl. Steinmayr (2005), S. 56.

⁷⁹ Vgl. Fleishman et al. (1995), S. 553.

⁸⁰ Vgl. Spencer und Spencer (1993), S. 11.

2.2.2 Kompetenzmanagement in der industriellen Instandhaltung

2.2.2.1 Allgemeines zu Kompetenzmanagement

Jedes Unternehmen möchte konkurrenzfähig, innovativ und effektiv sein. Das Kompetenzmanagement soll zur Erreichung dieser Ziele beitragen, indem die richtige Qualität und Quantität von Kompetenzen zur richtigen Zeit am richtigen Ort zur Verfügung zu stehen.⁸¹ Das Kompetenzmanagement sollte beide Hauptaspekte der Kompetenz, wie in Kapitel 2 erwähnt, umfassen und sowohl auf die strategischen Kernkompetenzen als auch die individuellen Kompetenzen ausgerichtet sein.

Kompetenzmanagement beschäftigt sich auf organisatorischer Ebene mit dem Identifizieren, Messen und Steuern von Kompetenzen, welche die Leistung und den Erfolg in einem Unternehmen beeinflussen.⁸²

In diesem Teil der Arbeit liegt der Fokus auf den Kompetenzen auf Personenebene und nicht auf den strategischen Kernkompetenzen eines Unternehmens. Diese Kompetenzen auf Personenebene werden in Kapitel 2.2.1 beschrieben und umfassen Wissen, Fertigkeiten, Fähigkeiten, Motivation und Selbstkonzept der Person.

Das Kompetenzmanagement besteht aus vier Hauptprozessen:^{83,84}

1. Kompetenzidentifikation
2. Kompetenzbewertung
3. Kompetenzakquise
4. Kompetenzverwendung

Kompetenzidentifikation befasst sich mit der Definition einzelner Kompetenzen und in welchen Bereichen diese benötigt werden. Diese Definitionen sollen sowohl die derzeit verfügbaren Kompetenzen als auch zukünftig wichtige Kompetenzen umfassen und auch strategische Aspekte beinhalten.

Kompetenzbewertung soll die vorhandenen Kompetenzen der Personen quantifizieren und darstellbar machen. Diese Bewertung dient als Grundlage für die weitere Akquise, da sie einen Soll-Ist-Vergleich ermöglicht.

Kompetenzakquise soll die nötigen Maßnahmen planen, um die nötigen Kompetenzen zu erschließen und auszubauen. Es werden in der Regel Weiterbildungsmaßnahmen definiert, welche die Art der Umsetzung (formale / informale Weiterbildung) und den Zeitpunkt umfassen sollen.

⁸¹ Vgl. Heilmann und Heilmann (2011), S. 10.

⁸² Vgl. Baladi (1999), S. 25.

⁸³ Vgl. Corallo et al. (2010), S. 300.

⁸⁴ Vgl. Berio und Harzallah (2005), S. 21f.

Kompetenzverwendung stellt den richtigen Einsatz der vorhandenen Kompetenzen dar. Anhand der Umsetzung des Kompetenzmanagements können die geeigneten Personen einer Aufgabe zugeteilt werden. Diese Aufgaben können dadurch effektiver und effizienter umgesetzt werden und tragen zum Gesamterfolg des Unternehmens bei. Gleichzeitig können so auch Schlüsselmitarbeiter_innen in bestimmten Bereichen identifiziert werden. Des Weiteren sollen bei diesem Prozess auch etwaige Lücken zwischen den nötigen und den vorhandenen Kompetenzen festgestellt werden. Sollten Lücken vorhanden sein, gilt es mit Hilfe von Maßnahmen aus der Kompetenzakquise gegenzusteuern.

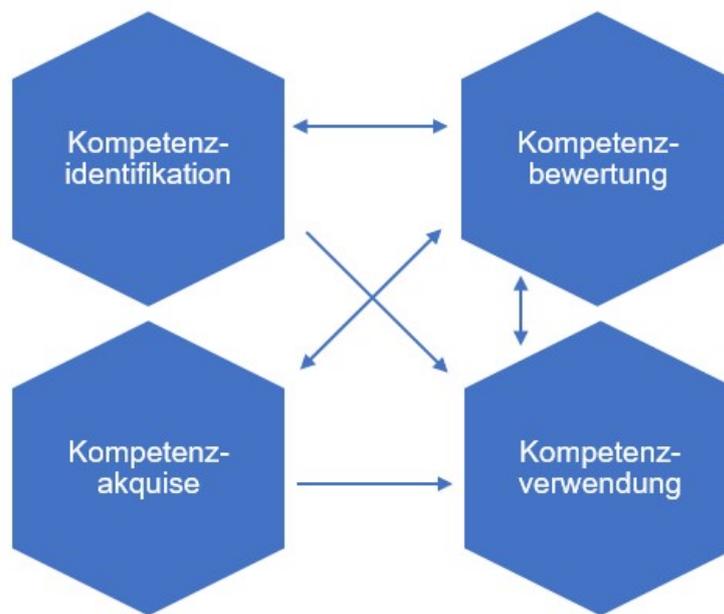


Abbildung 6: Kompetenzmanagement Prozessrahmenmodell⁸⁵

Gut umgesetztes Kompetenzmanagement sollte mit Hilfe von informationstechnologischen Systemen unterstützt werden, um der Organisation offenzulegen wer zu den fachkundigen Personen und Talenten in den unterschiedlichen Bereichen zählt, den vorhandenen Kompetenzbestand leichter zugänglich zu machen und jedem Beschäftigten einen klaren Einblick zu ermöglichen, wie ausgeprägt dessen Kompetenzen sind.⁸⁶ Sind gleichzeitig auch alle Aufgaben im Unternehmen an Kompetenzen orientiert, hilft dies bei der Ermächtigung der Beschäftigten. Dadurch wird ihnen auf klare und strukturierte Weise gezeigt, wie und in welchen Bereichen sie sich für ihren weiteren Karriereweg verbessern müssen, um die von ihnen angestrebten Aufgaben sowohl aus persönlicher als auch aus Unternehmenssicht bewältigen zu können.⁸⁷

⁸⁵ Eigene Darstellung angelehnt an Berio und Harzallah (2005), S. 22.

⁸⁶ Vgl. Corallo et al. (2010), S. 300.

⁸⁷ Vgl. Houtzagers (1999), S. 27.

Als Grundlage für das Kompetenzmanagement dienen zumeist Kompetenzmodelle. Diese gibt es in sehr generischer Form wie das Rahmenmodell der *Employment and Training Administration* (ETA), welches aus fünf Hauptebenen besteht:⁸⁸

1. Managementkompetenzen
2. Spezifisch technische und professionelle Kompetenzen
3. Einfache betriebliche Kompetenzen
4. Wissenschaftliche und akademische Kompetenzen
5. Persönliche effektivitätsorientierte Kompetenzen

Die generischen Kompetenzmodelle werden dann für einzelne Berufsgruppen spezifiziert. Ein Beispiel für die Anwendung des ETA Modells auf den industriellen Instandhaltungsbereich ist von Bohlouli et al. (2012) und umfasst die Integration des ETA Modells in eine bestehende Software im Instandhaltungsbereich, um auf einfache Weise das Kompetenzmanagement direkt in den Instandhaltungsbereich integrieren zu können. Ein anderes Beispiel ist die umfangreiche Arbeit der staatlichen Abteilung für öffentliche Arbeit der Republik Südafrika - (Secretary of the Department of Transport and Regional Development) (2017), welche auf die Instandhaltung in der Bauwirtschaft ausgerichtet ist und alle Bereiche der Instandhaltung von der Planung bis zur Kontrolle umfasst und auch einen expliziten Kompetenzkatalog enthält.

2.2.2.2 Relevanz des Kompetenzmanagements in der industriellen Instandhaltung

1993 zeigte eine Studie von 122 Instandhaltungsvorfällen mit menschlichem Zutun auf, dass diese Vorfälle zu 30% auf Grund von mangelhaftem Einbau, zu 56% auf Grund von Unterlassungen, zu 8% auf Grund falscher Teile und zu 6% aus anderen Gründen ausgelöst wurden.⁸⁹ Diese Fehler führen zu einem Wettbewerbsnachteil für das Unternehmen oder in schlimmeren Fällen auch zu geschädigten Personen.

1983 sank der Öldruck an allen drei Motoren des Flugzeugs (L-1011) aufgrund von fehlenden O-Ringen. Eine nachfolgende Untersuchung stellte fest, dass die Ursachen dafür schlecht durchgeführte Wartung und ein mangelhafter Beschaffungsprozess waren.⁹⁰

⁸⁸ Vgl. Employment and Training Administration, S. 1.

⁸⁹ Vgl. Secretary of the Department of Transport and Regional Development, S. 2.

⁹⁰ Vgl. Dhillon (2002), S. 125 zitiert nach Tripp (1999).

Dieses Beispiel und die Studie über menschliche Fehler in der Instandhaltung zeigen deren weitreichende Folgen auf. Diese Fehler können aus unterschiedlichsten Gründen auftreten:⁹¹

- Schlechte Arbeitsaufteilung
- Schlechtes Equipmentdesign
- Schlecht aufbereitete Instandhaltungsprozeduren
- Komplexe Instandhaltungsaufgaben
- Schlechte Arbeitsumgebung (Temperatur, Luftfeuchtigkeit, etc.)
- Ermüdetes Instandhaltungspersonal
- Veraltete Instandhaltungsanleitungen
- Unzureichende Ausbildung und Erfahrung

Diese Punkte verdeutlichen die Wichtigkeit eines guten Kompetenzmanagements in der Instandhaltung, um menschliche Fehler reduzieren zu können.

2.2.2.3 Umsetzung von Kompetenzmanagement in der industriellen Instandhaltung

Da es im Bereich des Kompetenzmanagements in der industriellen Instandhaltung kaum Fallstudien gibt, wird hier jene von Heilmann und Heilmann (2011) genau beleuchtet.

Heilmann und Heilmann (2011) führten 2007 29 Interviews in einem finnischen Unternehmen in der Forstindustrie durch. Alle der Befragten waren in leitenden Positionen, 22 der Interviewten arbeiteten in der Instandhaltung und sieben in der Produktion. Sie gingen der Frage nach, wie das Kompetenzmanagement im Unternehmen umgesetzt ist und wie das Unternehmen diese Kompetenzen sicherstellen beziehungsweise weiterentwickeln kann. Bei diesen Interviews stellten sie fest, dass das Unternehmen neue Beschäftigte immer zuerst als temporäre Beschäftigte einstellt und nur jene mit den adäquaten Kompetenzen eine unbefristete Anstellung erhalten. Bei einem Abgang von einem langjährigen Beschäftigten wird eine ausreichende Übergangszeit eingeplant, um einen reibungslosen Übergang zwischen altem und neuem Beschäftigten zu ermöglichen. Dies ist vor allem bei Abgängen in den Ruhestand sehr gut möglich. Die Grundlagen werden in der Grundausbildung der neuen Beschäftigten in berufsbezogenen Lehrgängen beigebracht, die in einen theoretischen und praktischen Teil untergliedert sind. Das implizite Wissen wird mit Hilfe von Mentoringprogrammen von den erfahreneren Beschäftigten an die neueren weitergegeben. Sollte es dennoch bei permanenten Beschäftigten zu mangelnden Kompetenzen kommen, sollen diese durch geplante Trainingseinheiten beseitigt werden. Spezialisierungen von einzelnen Personen

⁹¹ Vgl. Dhillon (2002), S. 128 zitiert nach Dhillon (1986).

sollen sich an deren persönlichen Interessen orientieren, um die intrinsische Motivation der Beschäftigten zu fördern und somit auch den Lernprozess möglichst effizient zu gestalten. Um Instandhaltung erfolgreich umsetzen zu können, ist das Wissen über das Equipment und die Prozesse entscheidend. Die Vorarbeiter_innen müssen sowohl über die nötigen technischen als auch planerischen Kompetenzen verfügen, um ihre Beschäftigten möglichst effizient einsetzen zu können. Alle Beschäftigten in der Instandhaltung sollen ein grundlegendes Verständnis für den jeweiligen Produktionsprozess haben und gut mit den Beschäftigten der Produktion kooperieren. Wenn im Einzelfall notwendig sollen sachkundige Personen aus anderen Produktionsstätten wegen Problemen hinzugezogen werden. Die Kompetenzen aller Beschäftigten sollen gelegentlich validiert werden, um einerseits mit Hilfe von Informationstechnologie dem gesamten Unternehmen Zugriff auf die derzeitigen Kompetenzen aller Beschäftigten zu ermöglichen und andererseits einen für jeden Beschäftigten individuellen Weiterbildungsplan erstellen zu können. Bei der Erstellung des Weiterentwicklungsplans sollte jedoch die Motivation und nicht die bereits vorhandenen Kompetenzen eines Beschäftigten im Vordergrund stehen.

2.2.3 Jobanalyse

Eine Jobanalyse ist ein systematischer und zielgerichteter Prozess, welcher jegliche arbeitsbezogenen Aspekte einer Jobstelle untersucht, sammelt und analysiert. Diese Aspekte umfassen die arbeitsbezogenen Aktivitäten, die Eigenschaften des Ausführenden und das Arbeitsumfeld.⁹² Die Jobanalyse versucht vier Fragen zu erläutern:⁹³

1. Welche physischen und mentalen Aufgaben müssen bewerkstelligt werden?
z.B.: Fräsen, Drehen, Materialtransport, Planen, ...
2. Wie werden diese Aufgaben erledigt?
Der Fokus liegt dabei auf den Methoden, mit welchen die Aufgaben ausgeführt werden.
z.B.: Netzplan, spezifische Computerprogramme, Transport mit Hilfe von Kran oder Gabelhubwagen.
3. Warum wird diese Aufgabe ausgeführt?
Soll den Zweck und die Zuständigkeiten der Aufgabe klären.
z.B.: Instandhalten einer Maschine, sodass diese im Produktionsprozess genutzt werden kann.

⁹² Vgl. Sackett et al. (2012), S. 61.

⁹³ Vgl. United States Small Business Administration Office of Management (1980), S. 3.

4. Welche Qualifikationen werden für die Ausführung der Aufgabe benötigt?
Soll alle KSAOs, welche für die Erfüllung einer Aufgabe nötig sind, zusammenfassen.
z.B.: Aufgabenspezifisches Wissen, Fertigkeiten, Fähigkeiten und andere Charakteristiken.

Eine gut umgesetzte Jobanalyse hat Einfluss in folgenden Punkten:⁹⁴

1. **Jobbeschreibung** dient als Zusammenfassung der wichtigsten Eckpunkte, wie Aufgaben und Verpflichtungen, eines Jobs.
2. **Jobklassifizierung** clustert verschiedene Jobs in Übergruppen, welche oft an Verpflichtungen und Autoritäten orientiert sind und hilft bei der Zuordnung zu Vergütungsschemen und Personenauswahl.
3. **Jobevaluation** soll für den Betrieb den richtigen Wert der Arbeiten festlegen, um eine marktkonforme Bezahlung zu gewährleisten und attraktiv für Mitarbeiter zu sein.
4. **Job-, Team- und System(re)design**: Jobdesign fasst einzelne Aufgaben zusammen, welche einen Job charakterisieren. Teamdesign fasst einzelne Aufgaben oder Aufgabenbereiche für ein Team zusammen. Systemdesign überschneidet sich mit dem Teamdesign und umfasst des Weiteren die Einteilung des nötigen Equipments und einzelner Personen zu spezifischen Aufgaben. Es wird als Design bezeichnet, wenn der Job neu geschaffen wird und Redesign, wenn ein vorhandener Job umgestaltet wird. Das Redesign umfasst alle Veränderungen der Zusammensetzung der Aufgaben und kann zum Ziel haben, dass die Arbeit effizienter, produktiver, sicherer oder für den Beschäftigten motivierender gestaltet wird.
5. **Anforderungen an das Humankapital**, welche für das Ausführen eines Jobs nötig sind. Diese werden oft in Form von KSAOs zusammengefasst. Clifford (1994) wies bereits auf die Sinnhaftigkeit der Erstellung von organisationsspezifischen Kompetenzkatalogen hin, welche hilfreich bei Entscheidungen von Weiterbildungsmaßnahmen sein können.⁹⁵
6. **Leistungsbewertung und Leistungsmanagement**: Die Leistungsbewertung befasst sich mit dem Vergleich zwischen erwarteter und erbrachter Leistung eines Individuums und dient oft als Grundlage für Gehaltserhöhungen und

⁹⁴ Vgl. Morgeson et al. (2019), S. 3 ff.

⁹⁵ Vgl. Clifford (1994), S. 335.

Feedback an die Beschäftigten. Das Leistungsmanagement dient der Leistungssteuerung von Individuen oder Teams.

7. **Weiterbildungsmaßnahmen** können mit Hilfe der Jobanalyse speziell auf den einzelnen Beschäftigten und dessen Aufgaben zugeschnitten werden. Dies erfolgt häufig mit einem Abgleich zwischen den bereits vorhandenen KSAOs eines Beschäftigten und den für die Aufgaben nötigen.
8. **Arbeitermobilität:** Wenn Beschäftigte und Job zusammenpassen, hat dies meist sowohl für das Unternehmen als auch für den Beschäftigten Vorteile, wie intrinsische Motivation und daraus folglich ein erhöhter Einsatz des Beschäftigten. Diese Jobwechsel können sowohl auf horizontaler als auch auf vertikaler Ebene stattfinden. Manche Unternehmen haben bereits mögliche Karrierewege für ihre Beschäftigten vorgesehen, um deren Weiterentwicklung bestmöglich zu fördern und die Beschäftigten damit sukzessive auf die nächsten Aufgaben vorzubereiten.
9. **Personalplanung:** Für Unternehmen ist es wichtig, dass die richtigen Beschäftigten in den Jobs arbeiten, um dabei ein gewisses Maß an Planungssicherheit zu haben. Für den Fall, dass ein Beschäftigter in einen anderen Job wechselt, sollte ein möglichst reibungsloser Übergang mit dessen Nachfolge möglich sein. Dafür ist es nötig, die vorhanden KSAOs im Unternehmen zu kennen.
10. **Effizienz:** Anhand der Jobanalyse werden die Abläufe genau untersucht und ermöglichen dadurch eine Vereinfachung der Aufgaben. Auf individueller Ebene wären Beispiele dafür (1) das Reduzieren der physischen Bewegungen, (2) Arbeitshilfen, wie Checklisten oder (3) das Verwenden von praktikableren Werkzeugen. Auf Teamebene ist die Vermeidung von Redundanzen der KSAOs sinnvoll und eine Arbeitsaufteilung anhand der Spezialisierungen der Beschäftigten.
11. **Sicherheit und Gesundheit:** Wie bei der Effizienz kann durch die genaue Untersuchung der Abläufe auch die Sicherheit und die Gesundheit der Beschäftigten beeinflusst werden. Dies kann durch Unterstützung mit zusätzlichen Hilfsmitteln erfolgen, um einen Arbeitsablauf sicherer, ergonomischer oder psychisch weniger anspruchsvoll zu gestalten.
12. **Gesetzliche Rahmenbedingungen:** Verschiedene Gesetze befassen sich mit arbeitsrelevanten Thematiken wie der Anstellung, Sicherheit und Gesundheit,

Anstellverfahren, Weiterbildung, Vergütung oder Kündigungen von Beschäftigten. Die Jobanalyse hilft Unternehmen dabei, die einzelnen Jobs und die gesetzlichen Rahmenbedingungen leichter in Einklang zu bringen.

Diese Liste ist primär auf betriebliche Aspekte ausgerichtet und kann deshalb nicht als vollständig angesehen werden. Weitere Punkte, auf die nicht näher eingegangen wird, wären zum Beispiel die Work-Life-Balance oder makroökonomische Aspekte wie die Berufsorientierung. Die Berufsorientierung soll arbeitssuchenden Menschen dabei helfen, die ihren Interessen und Fähigkeiten entsprechenden Jobs zu kennen und ihnen somit auch Branchenwechsellmöglichkeiten zu eröffnen.⁹⁶

Es gibt eine Vielzahl von Methoden und Vorgehensweisen für die Erstellung von Jobanalysen. Auf die Einzelnen wird in dieser Arbeit nicht im Detail eingegangen. Für Beschreibungen auf einzelne Methoden sei hier auf das Buch „Methods, Research, and Applications for Human Resource Management“ von Morgeson et al. (2019) verwiesen. Die Methoden lassen sich jedoch in zwei Hauptgruppen unterteilen, sie sind entweder aufgabenorientiert oder beschäftigtenorientiert. Aufgabenorientierte Methoden sind darauf ausgerichtet was und wie eine Person die Aufgaben ausführt. Beschäftigtenorientierte Methoden sind auf die Charakteristiken des Individuums ausgerichtet, welche für eine erfolgreiche Durchführung nötig sind. Des Weiteren gibt es noch hybride Methoden, die Elemente von beiden Hauptgruppen aufweisen.⁹⁷

Die Expertenumfrage von Levine et al. (1983) untersuchte verschiedene Methoden und bewertet diese in elf Kategorien nach deren jeweiligen Effektivität. Die verwendeten Kategorien waren, bis auf die Zusammenfassung von Effizienz und Sicherheit und Gesundheit, gleich wie die obigen zwölf Punkte des Einflusses der Jobanalyse.⁹⁸

Clifford (1994) wies bereits auf die Kosteneffektivität der Jobanalyse hin.⁹⁹ Einige Jahre später wurde der positive Einfluss sowohl auf die jobbezogene Leistung von Safdar et al. (2010) nachgewiesen.¹⁰⁰ Suthar et al. (2014) wiesen den Einfluss auf die Leistung der gesamten Organisation statistisch nach.¹⁰¹

2.2.4 Jobausschreibung

Mit Hilfe von Jobausschreibungen, oft auch als Stellenausschreibungen bezeichnet, gibt ein Unternehmen bekannt, dass es auf der Suche nach einem neuen

⁹⁶ Vgl. Morgeson et al. (2019), S. 6.

⁹⁷ Vgl. Rogelberg (2017), S. 2.

⁹⁸ Vgl. Levine et al. (1983), S. 342.

⁹⁹ Vgl. Clifford (1994), S. 335.

¹⁰⁰ Vgl. Safdar et al. (2010), S. 32.

¹⁰¹ Vgl. Suthar et al. (2014), S. 178.

Beschäftigten für einen Job ist. Gut 60% aller Stellenausschreibungen führen zu einer Einstellung.¹⁰²

Jobausschreibungen können entweder unternehmensintern oder -extern erfolgen. Der Grundaufbau von Jobausschreibungen ist jedoch zumeist gleich, oft erfolgen auch beide Vorgehensweisen parallel. Interne Jobausschreibungen bieten jedoch den Vorteil, dass die Beschäftigten einerseits schon bekannt sind und bereits sowohl über unternehmensspezifisches Wissen als auch Netzwerke verfügen.¹⁰³

Eine Jobausschreibung soll einerseits die richtigen Bewerbenden ansprechen und diese überzeugen, sich zu bewerben und weiters kann es auch als Marketinginstrument für das Unternehmen dienen.¹⁰⁴

Der Grundaufbau einer Jobausschreibung sollte folgende Punkte enthalten:^{105,106,107}

1. Unternehmensprofil
2. Stellenbezeichnung
3. Tätigkeitsgebiet / Jobbeschreibung
4. Anforderungen
5. Angaben zum Bewerbungsverfahren

Das Unternehmensprofil soll das Unternehmen kurz vorstellen und alle wichtigen Informationen, wie etwa Branche, Tätigkeitsfeld, Größe, Marktposition und Unternehmenskultur/-werte kurz zusammenfassen.¹⁰⁸ Die formale Stellenbezeichnung soll enthalten sein, Beispiele dafür sind „Elektrische Instandhaltung“ oder „Konstrukteur“. Die Jobbeschreibung wird zumeist aus der Jobanalyse (siehe Kapitel 2.2.3) abgeleitet und enthält alle wichtigen Eckpunkte der Stelle. Das Anforderungsprofil kann sowohl formale Eigenschaften wie Ausbildungsabschlüsse oder Zertifizierungen enthalten, als auch formlose wie Kommunikationsstärke oder ähnliche generische Eigenschaften enthalten. Diese Anforderungen können noch in Muss- und Kann-Anforderungen unterteilt werden, wobei es zu beachten gilt, dass die essenziellen Punkte enthalten sind, es aber auf keinen Fall zu viele enthält und dadurch potenziell passende Bewerbende abgeschreckt werden.¹⁰⁹ Die Angaben zum Bewerbungsverfahren sollen dem Betrachtenden die folgenden Schritte, Termine und Kontaktdaten ersichtlich machen.

¹⁰² Vgl. Reinders und Thönißen (2020).

¹⁰³ Vgl. ebd.

¹⁰⁴ Vgl. Lorenz und Rohrschneider (2009), S. 31.

¹⁰⁵ Vgl. Wilk (2011), S. 97.

¹⁰⁶ Vgl. Reinders und Thönißen (2020).

¹⁰⁷ Vgl. Lorenz und Rohrschneider (2009), S. 31 ff.

¹⁰⁸ Vgl. ebd., S. 33.

¹⁰⁹ Vgl. ebd., S. 32.

Auch wenn der Grundaufbau bei den meisten Jobausschreibungen gleich ist, handelt es sich dennoch in der Regel um unstrukturierten Text, der keinerlei formalen Anforderungen genügt. Dies ist eine Folge aus der Verwendung der Ausschreibung als Marketinginstrument und der Notwendigkeit, dass Unternehmen spezifisch auf sich aufmerksam machen müssen, um auf die richtigen Bewerbenden attraktiv zu wirken. Aus diesem unstrukturierten Text ergibt sich die Schwierigkeit, verschiedene Jobausschreibungen bezüglich deren Anforderungen und KSAOs vergleichen zu können. Dies birgt zwei unterschiedliche Fehlerpotenziale, einerseits die Verwendung von verschiedenen Begriffen für die gleiche Anforderung oder Tätigkeit, wie „drive innovation“ und „innovative solutions“¹¹⁰ und andererseits die Verwendung des gleichen Begriffs für unterschiedliche Anforderungen oder Tätigkeiten. Abhilfe für das Verwenden von verschiedenen Begriffen für eine Bedeutung sollen Jobtaxonomien schaffen, welche zum Ziel haben, die verwendeten Begriffe zu vereinheitlichen. Dies gelingt derzeit jedoch nur mit mäßigem Erfolg. Als Beispiel für die Verwendung eines Begriffs mit unterschiedlichen Bedeutungen dient das Wort „Java“. Im englischen Sprachraum ist es ein Synonym für Kaffee und gleichzeitig eine Programmiersprache.¹¹¹ Java ist ein Homonym. Dieses kann, wenn es sich in unterschiedlichen Kontexten im Korpus befindet, nicht unterschieden werden.

2.2.5 Jobtaxonomie

Mit Hilfe von Taxonomien versuchen Initiativen wie das amerikanische „*Occupational Information Network*“ (O*NET)¹¹², das europäische „*European Skills, Competences, Qualifications and Occupations*“ (ESCO)¹¹³ oder auch das österreichische Berufsinformationssystem (BIS)¹¹⁴ die verwendeten Begriffe bestimmter KSAOs zu vereinheitlichen. Weiters zu einzelnen Jobs zuzuordnen und damit spezifische Kompetenzkataloge abzuleiten. Diese Kompetenzkataloge clustern unterschiedliche KSAOs zu einem Job und sollen dabei helfen, dass ein einheitliches Bild eines Jobs entsteht. Dies hilft Firmen bei deren Ausschreibungen, da sie eine externe Referenz haben und Jobsuchenden, da sie Jobs anhand ihrer KSAOs finden können. Falls ein bestimmter Job angestrebt wird, können in den Taxonomien die dafür nötigen KSAOs gefunden werden und sie können daraus ihre persönlichen Weiterbildungsmaßnahmen ableiten.

Die Taxonomien, welche immer auf Expertenmeinungen basieren, beinhalten unterschiedliche Ebenen und Ausbaustufen, welche in Tabelle 1 zusammengefasst sind. ESCO und O*NET stellen die beiden umfänglichsten Standards dar, weshalb auf diese folglich genauer eingegangen wird.

¹¹⁰ Vgl. Pejic-Bach et al. (2020), S. 421.

¹¹¹ Vgl. Zhao et al. (2015), S. 4012.

¹¹² Vgl. National Research Council (2010).

¹¹³ Vgl. Directorate-General for Employment, Social Affairs and Inclusion (2017).

¹¹⁴ Vgl. Arbeitsmarktservice.

Der ESCO Standard ist eine Initiative der Europäischen Kommission. Dieser baut auf drei Säulen auf und ist stellen-, kompetenz- und qualifikationsbezogen. Dieser definiert Kompetenzen, ordnet diese zu Übergruppen und spezifischen Stellen zu. ESCO umfasst 2942 verschiedene Stellen. Die Kompetenzen werden in Wissen, Fertigkeiten, Fähigkeiten/Werte und sprachliche Fertigkeiten und Wissen unterschieden.¹¹⁵ Alle diese Kompetenzen sind bereichsunabhängig und in 27 Sprachen verfügbar. Bereichsunabhängigkeit bedeutet in diesem Zusammenhang, dass Stellen aus allen möglichen Arbeitsbereichen (z.B. Militär, Medizin, produzierendes Gewerbe) im Standard enthalten sind. Im Zuge dieser Initiative wurde auch ein eigenes Stellenportal namens EURES geschaffen.¹¹⁶

Der O*NET Standard wurde vom US-amerikanischen Department of Labor / Employment and Training Administration ins Leben gerufen. Der Standard baut auf einzelnen Stellen auf und ordnet diesen Kompetenzen als Wissen, Fertigkeiten oder Fähigkeiten zu. Er umfasst 923 verschiedene Stellen. Der O*NET Standard ist nur in englischer Sprache verfügbar.¹¹⁷

Tabelle 1: Vergleich der kompetenzbasierenden Taxonomien von Experten¹¹⁸

	ESCO	e-CF	O*NET	ISCO-08	ROME
Kompetenzdefinition	✓	✓	✓	x	x
Kompetenztyp	✓	x	✓	x	x
Spezifische Stellen	✓	x	✓	✓	✓
Bereichsunabhängigkeit	✓	x	✓	✓	✓
Mehrsprachigkeit	✓	x	x	x	x

In dieser Arbeit wird der ESCO Standard verwendet, welcher als übergeordneter Europäischer Standard gilt und derzeit rund 13000 Kompetenzen umfasst.¹¹⁹ Dieser Standard ist hierarchisch aufgebaut und frei zugänglich. Dieser Standard wurde gewählt, da die untersuchten Stellenanzeigen aus dem europäischen Raum stammen.

Ein Nachteil dieser Taxonomien besteht darin, dass sie alle auf Expertenmeinungen basieren und daher teuer in der Wartung sind.¹²⁰ Abhilfe für dieses Problem würde eine datengestützte Taxonomie liefern. Beispiele dafür sind die Arbeiten von Zhao et al. (2015), die für das Karriereportal Carrierbuilder mit Hilfe von Text Mining Methoden eine automatische Taxonomie aus Lebensläufen und Stellenanzeigen der Plattform erstellten.¹²¹ Diese Taxonomie ist jedoch nicht für die Allgemeinheit

¹¹⁵ URL: <https://ec.europa.eu/esco/portal/skill>, Stand: 19.11.2021.

¹¹⁶ URL: https://ec.europa.eu/esco/portal/escopedia/Main_Page, Stand: 19.11.2021.

¹¹⁷ URL: <https://www.onetcenter.org/overview.html>, Stand: 19.11.2021.

¹¹⁸ Eigene Darstellung angelehnt an vgl. Khaouja et al. (2021), S. 6.

¹¹⁹ URL: <https://ec.europa.eu/esco/portal>, Stand: 28.08.2021.

¹²⁰ Vgl. Djumalieva und Sleeman (2018), S. 2.

¹²¹ Vgl. Zhao et al. (2015), S. 4012.

zugänglich. Die Arbeit von Djumalieva und Sleeman (2018) ist öffentlich zugänglich und erstellt eine hierarchische Taxonomie basierend auf drei Ebenen aus Jobausschreibungen der Firma Burning Glas Technologies. Als Hauptgruppen dienen berufliche Felder, wie „health and social care“ und „business administration“, die zweite Gruppe richtet sich nach einzelnen Themenbereichen und die unterste Ebene basiert auf einzelnen Fertigkeiten.¹²²

2.3 Text Mining

In diesem Abschnitt der Arbeit werden die verwendeten Methoden des Text Mining genauer beleuchtet. In diesem Abschnitt werden die Fachbegriffe zumeist direkt aus dem Englischen übernommen und sind der Übersichtlichkeit wegen kursiv dargestellt.

Text Mining bezeichnet die zum größten Teil automatisierte Analyse von Textdokumenten, um neues und potenzielles Wissen zu extrahieren und somit bei Entscheidungsfindungen zu unterstützen.^{123,124} Text Mining kommt daher bei der Analyse von großen Textdatenmengen zum Einsatz, bei welchen eine menschliche Analyse entweder kaum bewerkstelligt werden kann oder diese aus wirtschaftlicher Sicht nicht sinnvoll ist. Einsatzpotenziale für Text Mining befinden sich in den unterschiedlichsten Bereichen. Beispiele dafür sind im unternehmerischen Kontext das Kundenbeziehungsmanagement¹²⁵, Competitive Intelligence¹²⁶ und der Zusammenhang zwischen dem Verhalten von Aktien und finanzbezogenen Nachrichten¹²⁷, aber auch im wissenschaftlichen Kontext wie der Organisationsforschung¹²⁸ kann der Einsatz von Text Mining sinnvoll sein.

Das Einschätzen der verschiedenen Datenstrukturen und deren Anteile an der gesamten Datenmenge ist zwar schwierig zu verifizieren, aber schätzungsweise sind 80% aller Daten in einem Unternehmen unstrukturierte Daten und haben jährliche Wachstumsraten von 55-65%.^{129,130}

¹²² Vgl. Djumalieva und Sleeman (2018), S. 10.

¹²³ Vgl. Hippner und Rentzmann (2006), S. 287.

¹²⁴ Vgl. Bohnacker et al. (2002), S. 438.

¹²⁵ Vgl. ebd., S. 437.

¹²⁶ Vgl. Hippner und Rentzmann (2006), S. 287.

¹²⁷ Vgl. Nikfarjam et al. (2010).

¹²⁸ Vgl. Kobayashi et al. (2018b).

¹²⁹ URL: <http://breakthroughanalysis.com/2008/08/01/unstructured-data-and-the-80-percent-rule/>, Stand: 24.10.2020.

¹³⁰ URL: <https://www.datamation.com/big-data/structured-vs-unstructured-data.html>, Stand: 23.10.2020.

Gerade diese unstrukturierten Daten sollen jedoch der Schlüssel sein, um sich einem schnell verändernden Geschäftsumfeld rechtzeitig anpassen zu können und somit einen langfristigen Unternehmenserfolg zu ermöglichen.¹³¹

2.3.1 Bereiche des Text Mining

Die im Prozess eingesetzten Techniken stammen dabei aus den Bereichen des Information Retrieval (IR), Natural Language Processing (NLP), Information Extraction (IE) und der Data Mining (DM) und spielen eine wesentliche Rolle im Prozess des Text Mining.^{132,133}

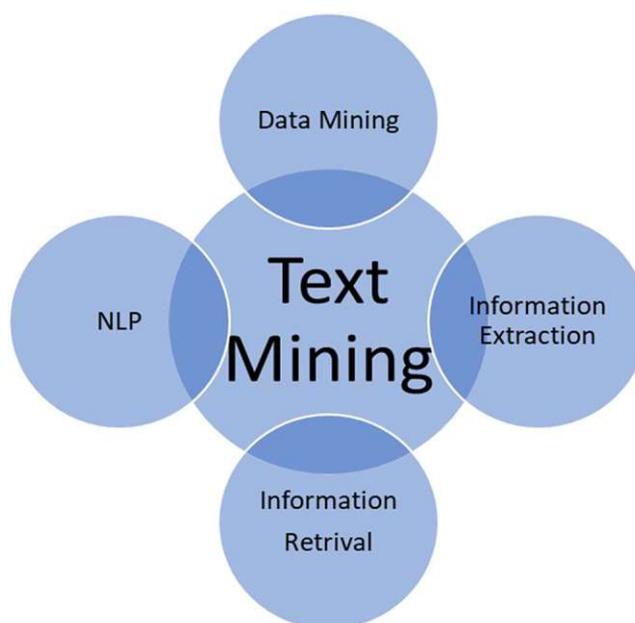


Abbildung 7: Bereiche des Text Mining¹³⁴

Information Retrieval ist eine Systematik, um eine Eingrenzung treffen zu können, welche Dokumente für die Beantwortung einer Problemstellung sinnvoll sind.¹³⁵ Unterschiedliche Filter helfen dabei die Dokumentenmenge zu reduzieren. Beispiele dafür sind Internetsuchmaschinen, welche eine Vielzahl zur Verfügung stehender Dokumente anhand der vorgegebenen Suchwörter eingrenzen.¹³⁶ Eine solche Eingrenzung trägt maßgeblich dazu bei, dass die ressourcenintensiven Data Mining Algorithmen weniger Daten verarbeiten müssen und aus diesem Grund schneller ablaufen können und bessere Ergebnisse erzielen.¹³⁷

¹³¹ Vgl. Hartl und Jacob (2015), S. 1.

¹³² Vgl. Ghosh et al. (2012), S. 223.

¹³³ Vgl. Hippner und Rentzmann (2006), S. 287.

¹³⁴ Eigene Darstellung angelehnt an Kumar und Bhatia (2013), S. 36.

¹³⁵ Vgl. Hippner und Rentzmann (2006), S. 288.

¹³⁶ Vgl. Ghosh et al. (2012), S. 223.

¹³⁷ Vgl. ebd., S. 223.

Natural Language Processing befasst sich mit der Problematik, mit Hilfe eines Computers die menschliche Sprache erfassen und verarbeiten zu können.¹³⁸ NLP dient beim Text Mining dazu, dem System die für die nachfolgende Informationsextrahierungsphase (engl. Information Extraction) nötigen linguistischen Einblicke zu ermöglichen.¹³⁹ Mit Hilfe von NLP kann zum Beispiel die Grammatik eines Textes berücksichtigt werden.

Information Extraction (deutsch Informationsextrahierung) wandelt die unstrukturierten Textdaten in strukturierte Daten um und ist dabei stark von den NLP Daten abhängig.¹⁴⁰ Die nun mehr strukturierten Daten können mit Hilfe von Data Mining Techniken analysiert werden.

Data Mining ist ein Prozess, mit welchem Muster in großen Datenmengen gefunden werden sollen. Das Ziel dabei ist es anhand, dieser Muster neue Erkenntnisse zu gewinnen.¹⁴¹ Im Text Mining Prozess ist es der Informationsextrahierungsphase nachgelagert.

Der Unterschied zwischen Data Mining und Text Mining besteht darin, dass es sich bei Data Mining um strukturierte Daten handelt im Gegensatz zu Text Mining, wobei die unstrukturierten Textdaten erst strukturiert werden müssen.¹⁴² Diese Überführung von unstrukturierten Daten zu strukturierten erfolgt mithilfe von Natural language processing (NLP) und Information Extraction.

Tabelle 2: Unterschiede zwischen Data- und Text Mining¹⁴³

Merkmal	Data Mining	Text Mining
Datenbasis	Strukturiert	Unstrukturiert
Datentyp	Nummern und Dimensionen	Sprachlicher Beitrag
Datenvorverarbeitung (Preprocessing)	Nicht unbedingt notwendig, wenn es sich bereits um einen strukturierten Datentyp handelt, ansonsten können vorbereitende Schritte notwendig sein	Nicht unbedingt notwendig aufgrund des Datentyps, jedoch oft für bestimmte Algorithmen erforderlich
Ziel	Mustererkennung in den Daten	Erschließen von Wissen aus den Daten

Das Ziel des Text Mining, die Generierung von Wissen, baut auf die Algorithmen des Data Mining auf und somit auch auf deren Fähigkeit, Muster zu erkennen. Die

¹³⁸ Vgl. ebd., S. 224.

¹³⁹ Vgl. Ghosh et al. (2012), S. 224.

¹⁴⁰ Vgl. ebd., S. 224.

¹⁴¹ Vgl. ebd., S. 224.

¹⁴² Vgl. Hippner und Rentzmann (2006), S. 287.

¹⁴³ Eigene Darstellung angelehnt an Madreiter (2020), S. 29.

Unterschiede zwischen Data und Text Mining sind in Tabelle 2 zusammengefasst, welche einen kurzen Überblick ermöglicht.

2.3.2 Qualität von textuellen Daten

Das Ziel von Text Mining ist das Erschließen von Wissen, welches wiederum für die Entscheidungsfindung genutzt werden kann. Entscheidungen sind abhängig von der Qualität der Daten, da diese die Grundlage für die Entscheidung bilden.¹⁴⁴ Unzureichende Datenqualität kann alle Unternehmensbereiche betreffen und sowohl in operativen als auch in strategischen Bereichen zu höheren Kosten führen.¹⁴⁵ Das Data Warehouse Institut schätzte 2002 die jährlichen Kosten aufgrund von unzureichender Datenqualität in den USA auf 600 Milliarden USD.¹⁴⁶

Die Algorithmen der Data Mining Methoden, welche für die Mustererkennung eingesetzt werden, sind gut entwickelt und können in der Regel direkt auf Daten angewendet werden.¹⁴⁷ Bei Data Mining bzw. explizit beim Text Mining nimmt die Datenaufbereitung die meiste Zeit in Anspruch.¹⁴⁸ Es gilt zu überprüfen, ob die vorhandenen Daten von ausreichender Qualität für die Wissensgenerierung sind.

Hierbei gilt es jedoch noch zu unterscheiden, aus welcher Quelle die zu analysierenden Daten stammen. Kommen diese aus einem Datenlager (englisch *Data Warehouse*), kann in der Regel davon ausgegangen werden, dass diese bereits von hoher Qualität sind und direkt verwendet werden können. Werden jedoch Daten von Websites herangezogen, kann nur bedingt davon ausgegangen werden, dass die Daten von ausreichender Qualität sind.¹⁴⁹

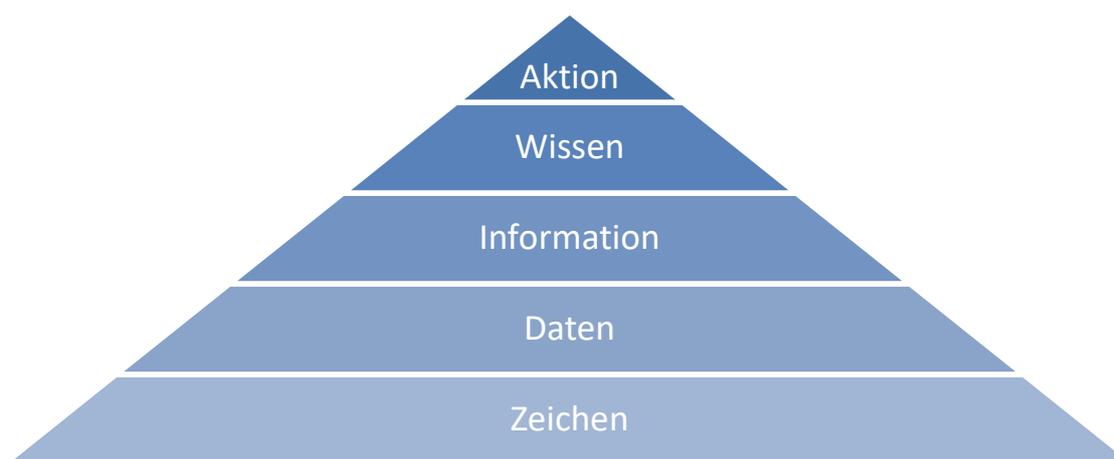


Abbildung 8: Wissenspyramide¹⁵⁰

¹⁴⁴ Vgl. Kiefer (2016), S. 1.

¹⁴⁵ Vgl. Batini et al. (2009), S. 21f.

¹⁴⁶ Vgl. Wayne W. Eckerson (2002), S. 3.

¹⁴⁷ Vgl. Weiss et al. (2010), S. 13.

¹⁴⁸ Vgl. ebd., S. 13.

¹⁴⁹ Vgl. ebd., S. 13.

¹⁵⁰ Eigene Darstellung angelehnt an Fuchs-Kittowski (2001), S. 37.

Abbildung 8 zeigt die Wissenspyramide, welche die Struktur und den Zusammenhang von Zeichen zu Aktionen darstellt. Zeichen stellen die Basis dar. Diese müssen vom Betrachtenden erkannt und gelesen werden, um Daten daraus zu erfassen. Dem Verstehen von Daten folgt eine Interpretation, aus welcher Informationen abgeleitet werden können. Durch das Verbinden von einzelnen Informationen kann Wissen entstehen. Auf Basis dieses Wissens lassen sich Aktionen beziehungsweise Handlungen ableiten und Entscheidungen treffen.¹⁵¹

Eine hohe Datenqualität ermöglicht eine hohe Qualität der daraus gefolgerten Informationen und erleichtert wiederum das Erkennen von Beziehungen zwischen diesen. Eine Erhöhung der Informationsqualität erhöht die Wahrscheinlichkeit, dass aus diesen Informationen Wissen abgeleitet werden kann.¹⁵² Die Aspekte von Daten- und Informationsqualitätsproblemen sind in Tabelle 3 zusammengefasst.

Tabelle 3: Daten- versus Informationsqualitätsprobleme¹⁵³

Probleme der Datenqualität	Probleme der Informationsqualität
Duplikate, Verschiedene Quellen	Unterschiedliche Schlüsse in einer Studie / Analyse
Fehlende Datenbeziehungen	Unklarer Kausalzusammenhang zwischen Effekt und Diagnose
Unlesbare Einträge	Fehlender logischer Fluss in einer Arbeit
Rechtschreibfehler	Uneindeutiger Ausdruck mit grammatikalischen Fehlern
Veraltete Einträge	Veraltete Analysen, die nicht an neue Erkenntnisse oder organisatorischen Veränderungen angepasst wurden
Inkonsistente Datenformate oder Namensgebungen	Inkonsistentes Layout bzw. Navigationsstruktur
Unpassender Speicherort – falscher Speicherort	Unauffindbare Dokumente
Komplizierte Datenbankabfrageprozeduren	Erschwerte Informationsnavigation und -Informationsauswahl
Falsche Datenkennzeichnungen oder Datenformatierung	Unpassende oder unzureichende Kategorisierung
Falsche Dateneinträge aufgrund einer mangelnden Validierung der Quelle	Unzureichende Schlüsse aufgrund mangelnder Evidenz
Nicht gerechtfertigte Manipulation der Daten (Löschen, Verändern)	Manipulation des Entscheidungsprozesses (z.B. Überlasten, Verwirren)

Datenqualitätsprobleme können in der Regel mit Hilfe von Datensäuberungs-, Datenprofilierungs-, Stabilisierungsalgorithmen (z.B. Phonetische Manipulation, Rechtschreibkorrektur), statistischen Prozesskontrollen oder Wörterbuchvergleichen mit automatischen Prozessen gelöst werden, dies ist bei

¹⁵¹ Vgl. Fuchs-Kittowski (2001), S. 37.

¹⁵² Vgl. Eppler (2006), S. 23.

¹⁵³ Eigene Darstellung angelehnt an ebd., S. 28.

Informationsqualitätsproblemen nur eingeschränkt möglich.¹⁵⁴ Einige Algorithmen bzw. Konzepte, mit welchen die Datenqualitätsprobleme bereinigt werden, sind in Kapitel 2.3.3.2 beschrieben.

2.3.3 Text Mining Prozess

In diesem Abschnitt wird auf einzelne Prozessschritte und die dabei verwendeten Methoden des Text Mining genauer eingegangen, wobei deren spezifische Anwendung erst in Kapitel 4 veranschaulicht wird.

Der Prozess des Text Mining lässt sich in folgende Hauptschritte untergliedern, siehe Abbildung 9, und ist dem Data Mining Prozess sehr ähnlich. Die Dokumentenaufbereitung ist jedoch in der Regel aufwendiger als bei Data Mining.¹⁵⁵ Eine Gemeinsamkeit ist, dass Data Mining und Text Mining Methoden auf Algorithmen des maschinellen Lernens aufbauen.¹⁵⁶

Die Begriffe Text und Dokument werden in diesem Zusammenhang synonym verwendet.

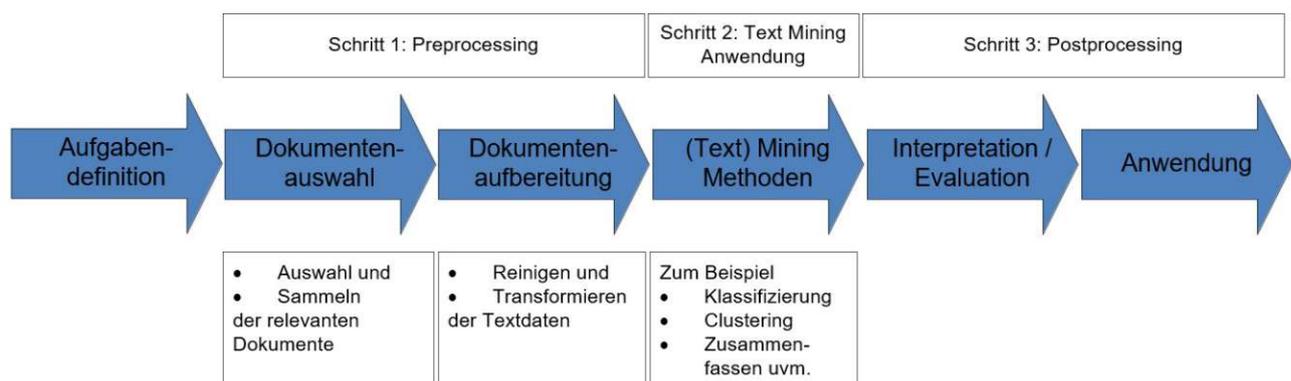


Abbildung 9: Der Text Mining Prozess ^{157,158}

Eine weitere gängige Unterteilung des Text Mining Prozesses ist die Einteilung in *Text Preprocessing*, *Text Mining Operations* und *Postprocessing*. Dabei umfasst das *Text Preprocessing* die Dokumentenauswahl und Dokumentenaufbereitung. *Text Mining Operations* ist die Anwendung von Mining Methoden. Das *Postprocessing* umfasst die Schritte Interpretation / Evaluation und Anwendung.¹⁵⁹

¹⁵⁴ Vgl. ebd., S. 28.

¹⁵⁵ Vgl. Hippner und Rentzmann (2006), S. 288.

¹⁵⁶ Vgl. Weiss et al. (2010), S. 3.

¹⁵⁷ Eigene Darstellung angelehnt an Hippner und Rentzmann (2006), S. 288.

¹⁵⁸ Eigene Darstellung angelehnt an Kobayashi et al. (2018b), S. 739.

¹⁵⁹ Vgl. Zhang et al., S. 681f.

Zusammenfassung der einzelnen Prozessschritte:

1. **Aufgabendefinition:** Festlegung der Problemstellung und Ableiten der Ziele.
2. **Dokumentenauswahl** (englisch Information Retrieval): Eingrenzen und Sammeln der sinnvollen Dokumente, welche für die Erreichung der Ziele relevant sind.
3. **Dokumentenaufbereitung** (englisch Text Transformation) wandelt die unstrukturierten Daten in strukturierte Daten um und kann dabei auch die Grammatik berücksichtigen. Die Techniken stammen dabei sowohl aus dem Bereich des Natural Language Processing als auch der Information Extraction.
4. **Mining Methoden:** Die strukturierten Daten können mit Verfahren aus dem klassischen Data Mining verarbeitet werden. Texte können hiermit zum Beispiel in vorgegebene Kategorien eingeteilt werden (Klassifizierung) oder die Texte können nach deren Ähnlichkeit gruppiert werden, um nur einige Möglichkeiten zu nennen.
5. **Interpretation / Evaluation:** Auswerten und Bewerten der Text Mining Ergebnisse.
6. **Anwendung:** Umsetzung der Ergebnisse in einen Nutzen für die Organisation oder den Forschungsbereich.

2.3.3.1 Auswahl und Sammlung von Dokumenten

Die Quellen, aus welchen die Dokumente gesammelt werden, müssen in kausalem Zusammenhang mit den zu untersuchenden Zielen stehen. Die zu analysierenden Texte können dabei je nach Anwendungsfall aus dem *World Wide Web*, Unternehmensdokumenten (z.B. Memos, Berichte usw.), persönlichen Dokumenten (z.B. E-Mail, SMS, Tagebüchern usw.) oder aus Antworten auf offene Fragen von Umfragen stammen. Für die Analyse mit Mining Methoden muss der Text in digitaler Form verfügbar sein. Handschriftliche Dokumente oder gescannte Dokumente müssen in der Regel erst in eine für den Computer verarbeitbare Form gebracht werden.¹⁶⁰

Der für diese Arbeit relevante Text ist in Stellenausschreibungen, welche in Bezug zu Instandhaltung stehen, enthalten und auf online Jobportalen verfügbar. Es handelt sich dabei bereits um einen digitalen Text.

Bei Stellenausschreibungen von online verfügbaren Jobportalen handelt es sich um Web Textdaten. Diese können entweder mit Anwendungsprogrammierschnittstellen (API – *application programming interface*) oder mit *Web scraping* gesammelt

¹⁶⁰ Vgl. Kobayashi et al. (2018b), S. 735.

werden.¹⁶¹ APIs sind Schnittstellen, welche von den Websites bereits für die Übergabe von Daten für Entwickelnde anderer Anwendungen bereitgestellt werden. Ein Beispiel dafür ist das Nutzen von Logindaten auf unterschiedlichen Websites (z.B. Login über Google). Dies ermöglicht Anwendenden, sich auf Websites von Dritten mit deren Google Account Login anzumelden, was ihnen das Erstellen eines zusätzlichen Accounts erspart. *Web scraping* wiederum wird zum Downloaden von vorgegebenen Daten oder Bereichen einer Website genutzt.¹⁶² Dabei sind standardisierte Dateiformate wie HTML oder XML äußerst nützlich, da diese eine Website in unterschiedliche Bereiche unterteilt.

Es ist wichtig die rechtlichen, ethischen Rahmenbedingungen zu berücksichtigen, welche im Zusammenhang mit Datenzugriff stehen. Gerade im Zusammenhang mit persönlichen Daten gilt es mit Bedacht vorzugehen.¹⁶³ Ein weiterer Aspekt ist die Verwaltung und Verarbeitung von den gesammelten Daten, welcher bei kleineren Projekten zumeist eine untergeordnete Rolle spielt. Bei größeren Projekten ist es jedoch essenziell, sich mit dieser Thematik und den dazugehörigen Problemen, wie Datenspeicherung, -zusammenführung und -steuerung auseinanderzusetzen. Dabei kann der Einsatz von Datenbanksystemen oder *dataware houses* nötig sein. Das Verwenden von Daten, welche bereits in solchen Systemen gespeichert sind, kann aufgrund der bereits erfolgten Datenreinigung und Standardisierung die Erfolgswahrscheinlichkeit der Mining Methoden erhöhen und erleichtert gleichzeitig den Umgang mit den Daten.¹⁶⁴

2.3.3.2 Dokumentenaufbereitung

Das Aufbereiten von Dokumenten besteht aus zwei Teilen, dem Reinigen und dem Transformieren der Daten. Das Reinigen und Transformieren (englisch *text cleaning* and *text transformation*) hat zum Ziel, die unstrukturierten Daten in qualitativ hochwertige, strukturierte Daten überzuführen, welche folglich mit Hilfe von Data Mining Algorithmen ausgewertet werden können.

Textdatenreinigung hat zum Ziel, die Datenqualität (siehe auch Abschnitt 2.3.2) zu erhöhen und damit die Validität der erkannten Muster und Beziehungen zu erhöhen.¹⁶⁵ Dabei wird der Text mit Hilfe von verschiedenen Algorithmen auf eine möglichst einfache Form gebracht und alle Bestandteile mit geringem Informationsgehalt aus dem Text entfernt. Standardreinigungsprozeduren bestehen in der Regel aus folgenden Schritten:¹⁶⁶

¹⁶¹ Vgl. ebd., S. 735.

¹⁶² Vgl. Olston und Najork (2010), S. 176.

¹⁶³ Vgl. van Wel und Royackers (2004), S. 129.

¹⁶⁴ Vgl. Inmon (1996), S. 49f.

¹⁶⁵ Vgl. Kobayashi et al. (2018b), S. 740.

¹⁶⁶ Vgl. ebd., S. 740.

- *Tokenizing / Text segmentation*
- *Stop word removal*
- *Lowercase conversion*
- *Lemmatizing / Stemming*
- *Multiword grouping*

Tokenizing oder auch *Text segmentation* (deutsch Textsegmentierung) umfasst das Untergliedern des Textes in einzelne Textelemente (*tokens*), wie Wörter und Satzzeichen (*types*). Diese Textelemente stellen die Grundlage für die weitere Analyse dar.¹⁶⁷ *Tokens* unterscheiden sich von *types* indem, dass diese öfters in einem Text vorkommen und *types* immer nur ein spezielles Wort oder Satzzeichen charakterisieren, daraus folglich ist die Anzahl an *tokens* viel größer als jene der *types*.¹⁶⁸ Dieser Schritt umfasst des Weiteren das Entfernen von überflüssigen Leerzeichen als auch Formatierungskennzeichnungen. Letztere spielen gerade bei Texten, welche aus dem Web stammen eine bedeutende Rolle, da diese in der Regel viele Formatierungskennzeichnungen von Website Standardformaten (z.B. HTML oder XML) enthalten.

Stop word removal entfernt Bindeworte, Präpositionen und Pronomen (z.B. und, für, der, die, das, ...), da diese in der Regel nur sehr wenig bis keinen zusätzlichen Informationsgehalt besitzen und für die Bedeutung des Textes nur eine untergeordnete Rolle spielen.¹⁶⁹ Mit Hilfe dieser Methode kann die spätere Analyse beschleunigt werden, da die Gesamtdatenmenge reduziert wird.¹⁷⁰

Lowercase conversion ersetzt alle Großbuchstaben eines *tokens* durch die Kleinbuchstaben. Dies reduziert die Datenmenge, da Satzanfänge nicht mehr als eigene *types* geführt werden. Es kann jedoch auch die Zweideutigkeit erhöhen, indem Eigennamen nicht mehr von anderen Wörtern unterschieden werden können. Ein Beispiel hierfür ist Apple, was in diesem Zusammenhang entweder als die Firma verstanden werden kann oder als die Frucht.¹⁷¹ Die Zweideutigkeit bestimmter Wörter stellt für Text Mining Methoden ein generelles Problem dar und ist nicht nur eine Folge der *lowercase conversion*.

Lemmatizing oder auch *Stemming* reduziert semantisch verwandte Wörter auf deren Grundform.¹⁷² Dies umfasst, dass verschiedene Zeitformen eines Verbes auf die Infinitivform und Substantive auf Singular- und Nominativform gebracht werden.¹⁷³ Diese Methode wird oft auch als *inflectional stemming* bezeichnet, welches sich vom

¹⁶⁷ Vgl. Weiss et al. (2010), S. 16f.

¹⁶⁸ Vgl. ebd., S. 17.

¹⁶⁹ Vgl. Kobayashi et al. (2018b), S. 740.

¹⁷⁰ Vgl. Weiss et al. (2010), S. 22.

¹⁷¹ Vgl. Camacho-Collados und Pilehvar (2017), S. 2.

¹⁷² Vgl. Kobayashi et al. (2018b), S. 740.

¹⁷³ Vgl. Weiss et al. (2010), S. 19.

stemming to a root unterscheidet, da bei letzterem alles auf den Wortstamm reduziert wird und somit auch Vor- und Nachsilben des Wortes entfernt werden, um die Datenmenge noch weiter zu reduzieren.¹⁷⁴

Als ein Musterbeispiel für die Textdatenreinigung wird der Vorgang hier exemplarisch für einen Satz aus den Stellenanzeigen durchgeführt.

Tabelle 4: Exemplarisches Beispiel der Textreinigung¹⁷⁵

Ausgangssatz	Supporting installation and project work on equipment.
nach tokenizing	'supporting', 'installation', 'and', 'project', 'work', 'on', 'equipment'
nach stop word removal	'supporting', 'installation', 'project', 'work', 'equipment'
nach lemmatizing	'support', 'installation', 'project', 'work', 'equipment'

Multiword grouping soll einzelne aufeinanderfolgende *tokens* zu einem einzelnen *token* zusammenfassen, um eine zusammengehörige Bedeutung richtig darzustellen.¹⁷⁶ Beispiele hierfür sind „Projekt Management“ oder auch „Vereinigte Staaten“. Werden diese zusammenhängenden Wörter nicht als solche erkannt ist eine korrekte Interpretation des Textes nicht mehr möglich und diese Information würde verloren gehen und zu falschen Schlüssen führen. Dafür gibt es zwei Varianten der Umsetzung, entweder die Erkennung von Eigennamen mit Hilfe von Name Entity Recognition (NER) oder andererseits eine statistische Auswertung. NER arbeitet mit hinterlegten Eigennamen, welche aus trainierten Modellen stammen, wie Vereinigten Staaten von Amerika. Die statistische Auswertung betrachtet, wie oft zwei oder mehr Wörter im gesamten Korpus nebeneinander vorkommen. Kommen diese Wörter oft genug in gleicher Reihenfolge vor wird diese als Wortgruppe erkannt.

Textdatentransformation hat als Ziel, den Text in eine mathematische Struktur zu überführen. Zumeist kommt hier das *vector space model* (VSM) zum Einsatz, welches die Dokumente und die einzelnen Terme in einem Vektorraum darstellt.¹⁷⁷ Ein sehr häufig verwendetes VSM ist das „*document-by-term matrix*“, bei welchem die *tokens* in den Spalten dargestellt werden und Dokumente in den Zeilen. Der zugehörige Matrizeneintrag steht dabei für die Häufigkeit des Wortes (*term frequency – tf*) in dem jeweiligen Dokument. Jedes Dokument wird somit als Vektor dargestellt und bildet die einfachste Form einer Gewichtung ab.¹⁷⁸ Eine zusätzliche Gewichtung der einzelnen Matrizeneinträge anhand deren Aussagekraft des Textes ist sinnvoll, denn kommt ein Wort beinahe in allen Dokumenten vor, hat dieses nur einen

¹⁷⁴ Vgl. ebd., S. 19f.

¹⁷⁵ Eigene Darstellung.

¹⁷⁶ Vgl. Camacho-Collados und Pilehvar (2017), S. 2.

¹⁷⁷ Vgl. Raghavan und Wong (1986), S. 280.

¹⁷⁸ Vgl. Kobayashi et al. (2018b), S. 741.

geringen Mehrwert für den Inhalt.^{179,180} Eine Möglichkeit dies mit zu berücksichtigen, ist die inverse Dokumentenhäufigkeit (*inverse document frequency*).¹⁸¹

Formel 1: Inverse Dokumentenhäufigkeit

$$IDF(i) = \log\left(\frac{N}{df(i)}\right)$$

N...Gesamtanzahl der Dokumente

df(i)...Anzahl der Dokumente welche den *token* beinhalten

Das Produkt aus *term frequency* und *invers document frequency* stellt die am häufigsten verwendete Gewichtung dar und erzielt gute Ergebnisse auch im Vergleich zu aufwendigeren Algorithmen.¹⁸² Das VSM ignoriert den Satzbau, was Einfluss auf den gewinnbaren Informationsgehalt hat, jedoch besticht es durch eine einfache Anwendbarkeit und Effektivität.¹⁸³ Gerade hierbei ist die Anwendung des *Multiword groupings* sinnvoll, um zusammengehörige Wörter nicht zu trennen und somit noch mehr Information zu verlieren.

2.3.3.3 Mining Methoden – Statistische Auswertung

Die Textdatentransformation ist eng mit der Anwendung der Mining Methoden verbunden und hat einen großen Einfluss auf die gewonnenen Ergebnisse. Die *document-by term matrix* dient für die meisten Mining Methoden als Ausgangsbasis. Üblicherweise werden auch unterschiedliche Kombinationen aus Textdatentransformation und Mining Methode ausprobiert und jene mit den besten Ergebnissen letztendlich implementiert.¹⁸⁴

Die meisten Text Mining Methoden lassen sich in eine dieser fünf Kategorien zuordnen:¹⁸⁵

- *Dimensionality reduction*
- *Distance and similarity computing*
- *Clustering*
- *Topic modelling*
- *Classification*

Dimensionality reduction hat zum Ziel, die Gesamtdatenmenge nach der Textdatenreinigung und -transformation noch weiter zu reduzieren, um die benötigten Ressourcen, wie Berechnungszeit und Speicherplatz in der Mustererkennung weiter

¹⁷⁹ Vgl. Kobayashi et al. (2018a), S. 770.

¹⁸⁰ Vgl. Lan et al. (2009), S. 722.

¹⁸¹ Vgl. Algarni und Tairan (2014 - 2014), S. 336.

¹⁸² Vgl. Lan et al. (2009), S. 734.

¹⁸³ Vgl. Kobayashi et al. (2018a), S. 771.

¹⁸⁴ Vgl. Kobayashi et al. (2018b), S. 742.

¹⁸⁵ Vgl. Solka (2008), 97-104.

verringern zu können.¹⁸⁶ Der Name ist aus der Dimensionsbezeichnung von Matrizen abgeleitet. Die Dimensionsreduzierung kann entweder durch das Entfernen von irrelevanten Dimensionen oder durch das Zusammenfassen zu neuen übergeordneten *tokens* erfolgen. Eine Möglichkeit namens *thresholding* eliminiert dabei alle Elemente, welche entweder eine hohe inverse Dokumentenfrequenz oder eine niedrige Dokumentenfrequenz haben, da diese keinen großen Einfluss auf die nachfolgenden Schritte haben. Gleichzeitig werden hierbei auch gegebenenfalls seltene *token* reduziert, welche aufgrund von Schreibfehlern vorhanden sind.¹⁸⁷ Eine zweite Möglichkeit ist das Zusammenfassen von *tokens* zu übergeordneten *tokens*, um somit synonyme Verwendungen zusammenfassen zu können. Dies geschieht mit Hilfe von *feature transformation methods*, wie *principal component analysis* (PCA), *latent semantic analysis* (LSA) oder *nonnegative matrix factorization*.¹⁸⁸ Dennoch gilt es zu berücksichtigen, dass schwache Indizien durchaus auch Aufschluss über anstehende Veränderungen geben können, welche noch nicht in der breiten Masse wahrgenommen werden.¹⁸⁹ Dies macht wiederum eine Überprüfung der reduzierten *tokens* sinnvoll, um diese schwachen Indizien nicht einfach aus der Datenmenge zu entfernen.

Distance and similarity computing hat zum Ziel, die Distanz bzw. Ähnlichkeit zwischen zwei Dokumenten festzustellen und findet vorrangig im Dokumentenvergleich oder in Vorschlagsystemen, wie sie in Onlineshops zu finden sind, Anwendung. Es gibt eine Vielzahl an Ähnlichkeits- und Distanzmaßen. Ein Beispiel dafür ist die Kosinus Ähnlichkeit (*cosine similarity*), welche die Ähnlichkeit anhand der gleichen Worthäufigkeiten zweier Dokumente vergleicht. Sie kann Werte zwischen -1 (perfekte negative Korrelation zwischen den Vektoren), 0 (keine Korrelation zwischen den Vektoren; die Vektoren stehen orthogonal aufeinander) und 1 (perfekte positive Korrelation zwischen den Vektoren) annehmen.¹⁹⁰

Clustering hat zum Ziel, Text in unterschiedliche Gruppen einzuteilen. Dieses Vorgehen basiert auf Grundlage von Ähnlichkeiten zwischen den einzelnen Dokumenten und den Gruppen. Befinden sich zwei Dokumente in einer Gruppe, sind diese sich zueinander ähnlicher als zu jenen Dokumenten, welche sich in einer anderen Gruppe befinden.¹⁹¹ Es gibt zwei unterschiedliche Zugänge zum Clustern, entweder es passiert hierarchisch oder partitionierend.¹⁹² Hierarchische Cluster Methoden funktionieren entweder agglomerativ oder trennend. Agglomerative Methoden betrachten zu Beginn jedes Dokument als einzelnes Cluster und führen

¹⁸⁶ Vgl. Forman (2003), S. 1289f.

¹⁸⁷ Vgl. Kobayashi et al. (2018a), S. 771.

¹⁸⁸ Vgl. ebd., S. 772.

¹⁸⁹ Vgl. Schoemaker und Day (2009), S. 88.

¹⁹⁰ Vgl. Kobayashi et al. (2018b), S. 744.

¹⁹¹ Vgl. Jain et al. (1999), S. 265.

¹⁹² Vgl. Karypis et al. (2000), S. 4.

die Dokumente Schritt für Schritt zusammen, womit die Anzahl an Cluster reduziert wird. Trennende Methoden betrachten zu Beginn alle Dokumente in einem Cluster und spalten Schritt für Schritt einzelne Dokumente ab.¹⁹³ Bei partitionierenden Clustermethoden gibt der Anwendende eine bestimmte Anzahl an Clustern vor, welche anhand einer Zielfunktion optimiert werden.¹⁹⁴ Die Ergebnisse des Clusters sind stark von den Eingabeparametern abhängig.¹⁹⁵ Es gibt sowohl interne als auch externe Validierungsindexe, wie den Dunn Index (intern) oder das F-Maß (extern).¹⁹⁶

Topic Modeling hat zum Ziel, Dokumente anhand von Wahrscheinlichkeiten einzelnen Themengebieten zuzuordnen.^{197,198} Jedes Themengebiet wird von verschiedenen Wörtern charakterisiert und jedes Dokument enthält verschiedene Themengebiete.¹⁹⁹ Das wahrscheinlichste Themengebiet wird somit über die im Dokument enthaltenen Wörter bestimmt. Die häufigsten verwendeten Methoden sind *latent dirichlet allocation* (LDA) und das *correlated topic model* (CTM).²⁰⁰ LDA unterscheidet sich von CTM dadurch, dass LDA Themen immer als unkorreliert annimmt, wohingegen bei CTM Themen korrelieren können. CTM liefert oft bessere Ergebnisse, kann aber bei einer Erweiterung der Dokumentenmenge nicht so einfach die Anzahl der Themengebiete erhöhen wie LDA.²⁰¹ LDA ist ein *unsupervised* Algorithmus und benötigt aus diesem Grund keine Trainingsdaten.²⁰²

Classification hat zum Ziel, Dokumente vorgegebenen Klassen oder Kategorien zuzuordnen bzw. die Klasse eines Dokumentes zu bestimmen.²⁰³ In der Regel basieren diese Modelle auf dem logarithmischen Regressionsmodell.²⁰⁴ Ein typisches Klassifizierungsmodell wird mit Hilfe von kategorisierten Übungsdaten trainiert und soll daraufhin ein neues Dokument der ähnlichsten Kategorie zuordnen.²⁰⁵ Eine der bekanntesten Anwendungen dieser Methode ist die Klassifizierung von E-Mails (*spam or ham*).

2.3.3.4 Interpretation und Evaluation

Die gewonnenen Ergebnisse müssen in der empirischen Forschung immer auf ihren Wahrheitsgehalt geprüft werden. Mögliche Fehler in den Mustern können aus zwei Kategorien stammen. Einerseits, ob die gefundenen Muster tatsächlich vorhanden sind oder diese nur aufgrund der großen Datenmenge gefunden wurden, dies kann

¹⁹³ Vgl. ebd., S. 3.

¹⁹⁴ Vgl. Kobayashi et al. (2018b), S. 745.

¹⁹⁵ Vgl. Rendón et al. (2011), S. 27.

¹⁹⁶ Vgl. ebd., S. 28f.

¹⁹⁷ Vgl. Hofmann (2001), S. 178.

¹⁹⁸ Vgl. Blei et al. (2003), S. 1001.

¹⁹⁹ Vgl. Kobayashi et al. (2018b), S. 745.

²⁰⁰ Vgl. ebd., S. 745.

²⁰¹ Vgl. Blei und Lafferty (2007), S. 31.

²⁰² Vgl. Banerjee und Basu (2007), S. 432.

²⁰³ Vgl. Li und Jain (1998), S. 537.

²⁰⁴ Vgl. Kobayashi et al. (2018b), S. 746.

²⁰⁵ Vgl. Li und Jain (1998), S. 537.

zum Beispiel mit der Bonferroni-Korrektur überprüft werden. Andererseits, ob die Daten und Ergebnisse auch valide, verlässlich und glaubwürdig sind. Dabei gilt es, die Ergebnisse und nicht die Methode zu überprüfen.²⁰⁶

Die Ergebnisse müssen auf jeden Fall einer Datentriangulation unterzogen werden. Die Überprüfung der Ergebnisse kann mit Hilfe von Experteninterviews oder Umfragen evaluiert werden. Erst wenn sich die mit Hilfe von Text Mining gewonnenen Erkenntnisse mit jenen aus Experteninterviews oder Umfragen decken, können diese als hinreichend zulässig angesehen werden und für die Entscheidungsfindungen verwendet und als Wissen angesehen werden.

2.3.3.5 Anwendung

Aus validierten Ergebnissen kann Wissen abgeleitet werden und für die Beeinflussung von Aktionen, wie in der Wissenspyramide (siehe Abbildung 8), herangezogen werden. Diese Aktionen können beispielhaft eine Integration oder Anpassung von gegebenen Prozessen, Theorien oder Rahmenbedingungen sein.²⁰⁷ Die Anwendung ist facettenreich und stark abhängig von der untersuchten Problemstellung und den daraus ableitbaren Aktionen.

2.3.4 Synonymerkennung

Ein wichtiger Punkt bei der Analyse von unstrukturierten Textdaten ist die Erkennung von synonymen Verwendungen von Begriffen und Wortgruppen. Das Problem besteht darin, dass bei einer Analyse, welche keine Synonyme berücksichtigt, die statistische Auswertung verfälscht ist, da die einzelnen Wortgruppen / Begriffe nicht als zusammengehörig erkannt werden.²⁰⁸ Eine besondere Schwierigkeit stellen Synonyme dar, welche nur auf kontextueller Basis erkannt werden können.²⁰⁹

Für diese Problemstellung gibt es drei grundsätzliche Zugänge, die Synonymerkennung mit Hilfe von Ontologien, vektorbasierte Wortähnlichkeiten und kontextuelle Synonymerkennung.²¹⁰ Für spezifische Anwendungen gilt es abzuwiegen auf welcher Ebene die Synonymerkennung passieren soll.

Ontologische Synonymerkennungen arbeiten mit hinterlegten Synonymen und zumeist auf Wortbasis. Diese Ontologien z.B. können nicht vollumfänglich sein und Wörter nicht erkennen, welche nicht hinterlegt sind. Gründe dafür können die Aktualität der Ontologien, die Unvollständigkeit der Synonymlisten oder

²⁰⁶ Vgl. Kobayashi et al. (2018b), S. 747.

²⁰⁷ Vgl. Kobayashi et al. (2018b), S. 747.

²⁰⁸ Vgl. Schumacher und Dredze (2019), S. 538.

²⁰⁹ Vgl. ebd., S. 539.

²¹⁰ Vgl. ebd., S. 539.

unterschiedliche Bedeutung eines Wortes in Bezug auf dessen Kontextes (Homonym) sein.^{211,212}

Vektorbasierte Wortähnlichkeiten basieren darauf, dass jedem Wort in einem Textkorpus einen Vektor zugeordnet wird. Wenn zwei Wortvektoren bis auf einen vorzugebenden maximalen Unterschied ähnlich sind, werden diese Wörter als Synonyme angenommen.²¹³ Mit dieser Methode können, falls gewünscht, auch nur Vorschläge für potenzielle Synonyme hervorgebracht werden. Dieser Ansatz basiert ebenfalls auf Wortebene.

Kontextuelle Synonymerkennungen, wie BERT^{214,215}, bestehen aus Neuronalen Netzwerken und sind dafür ausgelegt bidirektionale Repräsentationen von umklassifiziertem Text, meist auf Satzbasis, zu generieren.^{216,217} Diese Repräsentationen werden als Vektor dargestellt.²¹⁸ Diese Vektoren können wiederum als Ähnlichkeitsmaß herangezogen werden, um wie bei der vektorbasierten Wortähnlichkeit, Ähnlichkeiten auf Satzebene erkennen zu können. Mit Satzebene ist in diesem Kontext gemeint, dass der gesamte Satz als Basis für die Erstellung der Vektoren dient. Diese Vektoren sind wortabhängig, wobei im Gegensatz zur vektorbasierten Wortähnlichkeit der Vektor eines Wortes auch vom Kontext des Satzes beeinflusst wird und somit erhalten auch gleiche Wörter in verschiedenen Sätzen im Korpus unterschiedliche Vektoren.²¹⁹ Diese Modelle müssen im Allgemeinen für die jeweilige Anwendung angepasst werden. Das Anpassen funktioniert über klassifizierte Trainingsdaten.²²⁰ Diese Modelle funktionieren auf Satzebene und sind deshalb für die Synonymerkennung auf Wortgruppenebene nicht dienlich.

Bei den im ESCO Standard vorkommenden Kompetenzen handelt es sich in der Regel um Wortgruppen, deshalb haben alle drei hier vorgestellten Methoden ihre Hürden. Kontextuelle Synonymerkennung ist auf Grund des Aufbaus auf Satzbasis und zusätzlich der nicht zur Verfügung stehenden Trainingsdaten, nur mit hohem Arbeitsaufwand, manuelle Klassifizierung von Daten zu Trainingsdaten, möglich. Sowohl die ontologiebasierte als auch die vektorbasierte Synonymerkennung arbeiten auf Wortebene und bergen dadurch die Gefahr, die Bedeutung der ganzen Wortgruppe zu beeinflussen und keine validen Ergebnisse zu liefern.

²¹¹ Vgl. Yuan und Duan (2019 - 2019), S. 848.

²¹² Vgl. Schumacher und Dredze (2019), S. 539.

²¹³ Vgl. Yuan und Duan (2019 - 2019), S. 849.

²¹⁴ URL: <https://github.com/google-research/bert>, Stand: 28.08.2021.

²¹⁵ URL: [https://en.wikipedia.org/w/index.php?title=BERT_\(language_model\)&oldid=1039012807](https://en.wikipedia.org/w/index.php?title=BERT_(language_model)&oldid=1039012807), Stand: 28.08.2021.

²¹⁶ Vgl. Devlin et al. (2018), S. 1.

²¹⁷ URL: <https://devopedia.org/bert-language-model>, Stand: 28.08.2021.

²¹⁸ Vgl. Devlin et al. (2018), S. 4.

²¹⁹ Vgl. ebd., S. 4.

²²⁰ URL: <https://devopedia.org/bert-language-model>, Stand: 28.08.2021.

3 Stand der Forschung

In diesem Kapitel wird der derzeitige Stand der Forschung im Zusammenhang mit der Problemstellung dieser Arbeit zusammengefasst und die relevantesten Arbeiten in diesem Bereich werden vorgestellt. Die Literatur wurde mit Hilfe einer systematischen Literaturrecherche gefunden. Dafür wurden auf den Suchplattformen Google Scholar, Science Direct, Springer Link, Taylor and Francis, Emerald Insight, Scopus mit verschiedenen Kombinationen der Suchbegriffe, die gefundene Literatur eingegrenzt. Die Suchbegriffe waren *Text Mining*, *Job Advertisement* (Stellenausschreibung), *Competence / Competency* (Kompetenz), *KSAO*, *Employee* (Beschäftigte(r)) und *Maintenance* (Instandhaltung), welche durch unterschiedliche Kombinationen aus logischen und- bzw. oder-Verbindungen verknüpft wurden. Die verschiedenen Kompetenzbegriffe (*Competence*, *Competency*, *KSAO*) wurden immer als logische oder Beziehung angenommen. Alle gefundenen Ergebnisse wurden aufgenommen und anhand der Relevanz der Zusammenfassungen eingegrenzt. Das Vorgehen und die genauen Sucheingaben sind in Tabelle 5 und im Anhang in Tabelle 28 im Großformat ersichtlich.

Tabelle 5: Systematische Literaturrecherche ²²¹ - Großformat im Anhang (Tabelle 28)

Search Term_ID	Language	Date of search	Used Search-Term (as entered in database)	Set search filters	Number of results after first	Search strategy	examined papers	Number of papers moved into	Name of folder v. SLR_XX (Search Term_ID)	Number of papers cannot be downloaded
SLR_01	EN	27.11.2020	"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO")	Search in: Any field Time: 2010 - Present exclude citations and patents	11	Screening until result 11	5	5	SLR_01	0
SLR_02			"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		12	Screening until result 12	5	5	SLR_02	0
SLR_03			"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		630	Screening until result 60	4	4	SLR_03	0
SLR_04			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		831	Screening until result 70	4	4	SLR_04	0
SLR_05			"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		31	Screening until result 31	10	10	SLR_05	0
SLR_06			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		38	Screening until result 4	20	15	SLR_06	5
SLR_07	EN	03.1.2022	"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO")	Search in: Any field Time: 2010 - Present	2	Screening until result 2	1	1	SLR_07	0
SLR_08			"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		4	Screening until result 4	2	2	SLR_08	0
SLR_09			"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		48	Screening until result 48	2	2	SLR_09	0
SLR_10			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		100	Screening until result 50	3	3	SLR_10	0
SLR_11			"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		5	Screening until result 5	2	2	SLR_11	0
SLR_12			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		8	Screening until result 8	2	2	SLR_12	0
SLR_13	EN	04.1.2022	"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO")	Search in: Any field Time: 2010 - Present exclude preview-only content	1	Screening until result 1	0	0	SLR_13	0
SLR_14			"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	0	0	SLR_14	0
SLR_15			"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		18	Screening until result 18	0	0	SLR_15	0
SLR_16			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		34	Screening until result 34	0	0	SLR_16	0
SLR_17			"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	0	0	SLR_17	0
SLR_18			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		2	Screening until result 2	0	0	SLR_18	0
SLR_19	EN	05.1.2022	"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")	Search in: Any field Time: 2010 - Present	5	Screening until result 5	0	0	SLR_19	0
SLR_20			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		18	Screening until result 18	0	0	SLR_20	0
SLR_21			"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	1	0	SLR_21	1
SLR_22			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	1	0	SLR_22	1
SLR_23	EN	06.1.2022	"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO")	Search in: Any field Time: 2010 - Present	9	Screening until result 9	3	1	SLR_23	2
SLR_24			"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		10	Screening until result 10	3	1	SLR_24	2
SLR_25			"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		101	Screening until result 50	1	0	SLR_25	1
SLR_26			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		100	Screening until result 50	5	1	SLR_26	4
SLR_27			"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		13	Screening until result 13	3	1	SLR_27	2
SLR_28			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		27	Screening until result 27	3	1	SLR_28	2
SLR_29	EN	07.1.2022	"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")	Search in: Any field Time: 2010 - Present	1	Screening until result 1	1	0	SLR_29	1
SLR_30			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	0	0	SLR_30	0

Von insgesamt 2.064 Suchergebnissen wurden 81 Artikel als aussagekräftig erfasst, wobei von 21 nur die Zusammenfassung verfügbar war. Bei 17 Artikeln handelte es sich um Artikel, welche bei mehreren Suchdurchläufen gefunden wurden. Weitere 23 wurden anhand der Zusammenfassung ausgeschlossen und aus den restlichen 20 werden die relevantesten Fünf in den folgenden beiden Abschnitten vorgestellt. Außerdem werden weitere in Bezug auf diese Arbeit interessante Artikel, welche

²²¹ Eigene Darstellung.

anhand weitergehender Recherche gefunden wurden, jedoch nicht mit Hilfe der systematischen Literaturrecherche, vorgestellt.

3.1 Existierende Lösungsansätze für die Problemstellung

Mit den verwendeten Suchbegriffen konnte kein Lösungsansatz für die genaue Problemstellung, nachgefragte Kompetenzen im Instandhaltungsbereich, gefunden werden. Es gibt jedoch einige Studien, welche mit Hilfe von Text Mining Methoden die nachgefragten Kompetenzen in anderen Branchen analysiert haben.

Pejic-Bach et al. (2020) verwendeten Stellenausschreibungen, welche im Zusammenhang mit „Industry 4.0“ oder „Smart Factory“ auf der Plattform LinkedIn gefunden wurden. Dabei wurden 2.566 Stellenausschreibungen gefunden, wobei nur 1.460 zur Auswertung herangezogen wurden, da diese in englischer Sprache verfasst waren. Bei diesen Stellenausschreibungen handelte es sich um semistrukturierte Daten. Es wurde eine deskriptive Analyse und eine Text Mining Analyse durchgeführt. Die deskriptive Analyse untersuchte die Zusammenhänge zwischen Einstellungsart (Vollzeit / Teilzeit), Dienstort (Land) und Einstellungsniveau (Berufseinstieg / Führungspositionen). Die Text Mining Analyse fokussierte sich auf die Häufigkeit der verwendeten Phrasen in den Stellenausschreibungen. Dazu wurde die Methode des Produkts aus der Wortfrequenz und der inversen Dokumentfrequenz angewandt. Um interessante Erkenntnisse aus selten vorkommenden Phrasen zu extrahieren und somit auch neu entstehende Trends früher erkennen zu können, wurden Phrasen, welche in einem Bereich von 10-30 Nennungen lagen, durchgesehen und die aus der Sicht der Forschenden relevantesten extra genannt. Ein Beispiel dafür wäre, die Fertigkeit komplexe Ideen gut kommunizieren zu können. Des Weiteren wurde mit Hilfe des Jaccard-Index die Nähe der Phrasen analysiert und die Phrasen geclustert. Es wurden acht unterschiedliche Cluster gefunden und diese von den Forschenden einzelnen Jobprofilen zugeteilt. Die häufigsten Phrasen und Jobcluster wurden auch anhand des Einstellungsniveaus analysiert. Einschränkungen dieser Arbeit waren zeitlicher und umfänglicher Natur. Es wurden nur Stellenanzeigen von LinkedIn als Datenbasis herangezogen. Des Weiteren können Stellenanzeigen selbst einen Bias haben, da sie zumeist von Beschäftigten der Personalabteilungen erstellt werden, welche zumeist keine spezifische Spezialisierung in genau dieser Stellenkategorie besitzen. Einschränkungen auf Grund der verwendeten Parameter der Text Mining Methoden z.B. wurden nur Phrasen für das Clustern berücksichtigt, welche 50-mal oder öfter in den Stellenanzeigen vorgekommen sind.²²²

²²² Pejic-Bach et al. (2020).

Uhm et al. (2017) untersuchten mit Hilfe von Text Mining Methoden die nachgefragten Kompetenzen von Stellen im Bereich der Gebäudeinformationsmodellierung (BIM = engl. *building information model*). Sie analysierten 242 Stellenausschreibungen, welche in etwa gleichmäßig aus den USA, Vereinigtes Königreich und China stammten. Sie nutzten die O*NET Taxonomie für die Einteilung der Kompetenzen. Die 242 Stellenausschreibungen enthielten 35 verschiedene formale Jobtitel, welche sie mit der Konvergenz von *convergence of iterated correlations* (CONCOR) Methode zu 8 unterschiedlichen Jobkategorien zusammengefasst haben. Die 242 Stellenausschreibungen enthielten ca. 6.000 Kompetenznennungen, wovon 136 unterschiedliche Kompetenzen vorkamen. Diese wurden mit Hilfe des Pareto Prinzips auf die 43 relevantesten Kompetenzen eingegrenzt, den 8 unterschiedlichen Jobkategorien zugeteilt und statistisch ausgewertet. Mit Hilfe der O*NET Taxonomie und dem Pareto Prinzip wurden die 43 Kompetenzen in drei Kategorien geclustert. Die Kategorien wurden nach essenziellen, typischen und jobspezifischen Kompetenzen eingeteilt. Diese Studie unterstützt drei Hauptbereiche: (1) es wird ein wissenschaftlicher Grundstein für die Klassifizierung von Kompetenzen im BIM Bereich gelegt, welcher nicht auf Expertenmeinungen beruht, (2) das Personalwesen wird dabei unterstützt, Weiterbildungs- und Einstellungspolitiken abzuleiten und (3) im Bildungswesen können nationale Standards entwickelt werden, welche an diese Studien angelehnt sind, sodass Bildungseinrichtungen spezifische Weiterbildungsmöglichkeiten in Bezug auf BIM Jobs zuschneiden können. ²²³

Wu et al. (2017) untersuchten die Unterschiede in Stellenausschreibungen zwischen *Business Intelligence* und *Big Data* Jobs in China. Sie verwendeten dazu 4.380 online Stellenausschreibungen der Plattform Zhaopin.com, welche entweder „*Business Intelligence*“ oder „*Big Data*“ enthielten. Von diesen 4.380 wurden noch Duplikate und lückenhafte Stellenausschreibungen ausgeschlossen und so wurden 2.942 Stellenausschreibungen mithilfe *Structural Topic Modeling* (STM = Strukturelle Themen Modellierung) analysiert. Die Autorenschaft verwendeten die Perplexität als Optimierungsgröße und kamen damit auf 10 unterschiedliche Themenbereiche, welche durch den Algorithmus einzelnen Jobschwerpunkten, wie Projektmanagement, Data Mining oder Kommunikationsstärke, zuordenbar sind. *Business Intelligence* Jobs hatten zumeist einen Fokus auf Datenbanken und die Manipulation von Data warehouses. *Big Data* Jobs waren wiederum auf verteilte Programmierfähigkeit, den Aufbau von Programmarchitektur, und tiefgehende Kenntnisse von der Programmiersprache Python und Linux ausgerichtet. Eine Limitation der Studie war die einseitige Verteilung der Stellenausschreibungen, denn es wurden im Verhältnis zirka 6-mal so viele Stellenausschreibungen aus dem *Big*

²²³ Uhm et al. (2017).

Data Bereich analysiert im Vergleich zu jenen aus dem Bereich der *Business Intelligence*.²²⁴

Wu et al. (2020) untersuchten anhand von 2.239 Stellenausschreibungen die nachgefragten Kompetenzen von Jobs im Bereich der *Data Science* in China. Die Kompetenztaxonomie wurde für die Studie selbst entwickelt und ist an die Schlüsselworte von Veröffentlichungen, Hintergründe der einzelnen Institute, Forschungsschwerpunkte und die Kernkurse der iSchool²²⁵ angelehnt, welche in einer vorangegangenen Arbeit mit Hilfe von qualitativen und quantitativen Inhaltsanalysen gefunden wurden. Die Jobs wurden bezüglich deren Kompetenzen, Erfahrung, Gehalt und formalen Stellenbezeichnung statistisch bewertet. Die Autorenschaft unterteilten die Kompetenzen in folgende Hauptkategorien: Datenfertigkeiten, Basiskompetenzen, Komplexität und nicht kognitive Fertigkeiten. In absteigender Reihenfolge waren die Datenfertigkeiten, Basiskompetenzen, Komplexität und nicht kognitiven Fertigkeiten. Sie analysierten aus den Hauptkategorien die Häufigkeit der zugehörigen granularen Kompetenzen und leiten daraus einen Leitfaden für die Anpassung von Lehrveranstaltungen ab. Des Weiteren analysierten sie die nachgefragten Werkzeuge in Bezug auf *Data Science* und leiteten ein menschenzentriertes *Data Science* Kompetenzmodell ab, welches jedoch als Basis wiederum das Modell von Barison and Santos (2011) verwendet.²²⁶

Chung und Chen (2019) analysierten taiwanesishe Stellenausschreibungen im Bereich des Personalwesens in Bezug auf die nachgefragten KSAOs. Dafür sammelten sie 2.869 Stellenausschreibungen, wobei nach dem Ausschließen von Duplikaten und nicht auf das Personalwesen bezogenen Stellenausschreibungen, nur 842 für die Analyse herangezogen wurden. Sie erstellten mit Hilfe des Text Mining Programms „SAS Text Miner“ Kategorien und Cluster aus den Stellenausschreibungen mit Hilfe eines LSA Modells. Dabei stellten sie fest, dass in diesen Stellenausschreibungen vor allem nach Fertigkeiten, Ausbildungsniveau, Berufserfahrung und spezifischen Fähigkeiten nachgefragt wird. Des Weiteren sind bereichsübergreifende Kompetenzen und Polylingualismus besonders gefragt.²²⁷

Die restlichen Artikel in diesem Abschnitt wurden nicht mit Hilfe der systematischen Literaturrecherche gefunden, jedoch sind sie für die gegebene Problemstellung von Relevanz.

Die Arbeit von Kobayashi et al. (2018b) „*Text Mining in Organizational Research*“ zeigt die einzelnen adressierbaren Fragenstellungen mit Hilfe von Text Mining Methoden auf. Die Forschenden stellten die einzelnen Schritte vom *Preprocessing*

²²⁴ Wu et al. (2017).

²²⁵ URL: <https://ischools.org/>, Stand: 13.12.2020.

²²⁶ Wu et al. (2020).

²²⁷ Chung und Chen (2019).

bis zum *Postprocessing* vor, gingen auf die zugehörigen Methoden ein und zeigten praktische Empfehlungen auf. Als praktisches Beispiel wenden sie die Text Mining Methoden auf online Jobausschreibungen an und analysierten diese anhand von den nachgefragten Aktivitäten und Fertigkeiten. Für die Unterscheidung, ob es sich um eine Aktivität oder Fertigkeit handelt, wurden Listen für die jeweilige Kategorie erstellt. Anhand dieser Listen kategorisiert ein *random forest* Algorithmus die einzelnen Textpassagen. Basis für diese Listen waren vorangegangene Stellenanalysen und eine statistische Analyse der Begriffe. Im Bereich der Krankenpflege clusterten sie die einzelnen Aktivitäten und Fertigkeiten zu Gruppen und validierten diese Gruppen mithilfe von Bereichsexperten, womit festgestellt wurde, dass die Ergebnisse aus den online Jobausschreibungen sich hinreichend gut mit jenen der Bereichsexperten deckten. Im zweiten Schritt analysierten sie alle Jobausschreibungen, ohne Einschränkung auf die Krankenpflege, mit Hilfe von Topic Mining unter Verwendung der *latent dirichlet allocation* (LDA) Methode und stellten fest, dass die meisten Gruppen von Aktivitäten und Fertigkeiten gut zu einzelnen Jobs (Managementbezug, Programmierbezug, ...) passen, jedoch waren einige der 140 Themen rein aus generischen Aktivitäten und Fertigkeiten zusammengesetzt. Die einzelnen Themen wurden darauf folglich wieder geclustert und zu übergeordneten Kategorien zugeordnet, wie beispielweise Jobs im Gesundheitsbereich.²²⁸

Tabelle 6: Vor- und Nachteile von Kompetenzidentifikationsmethoden²²⁹

Typ	Vorteile	Nachteile
Kompetenzanzahl	einfach umzusetzen	vermehrt falsche positive Kompetenzen
Topic Modeling	einfach umzusetzen	Themengebiete müssen von Experten validiert werden
Skill Embedding	Genauigkeit	Trainingsdaten benötigt; vermehrt falsch negative Kompetenzen
Machine Learning Methods	Genauigkeit	viele Trainingsdaten benötigt; aufwendige Feinabstimmung

Eine gute Zusammenfassung der angewendeten Methoden im Bereich der Kompetenzidentifikation aus online Stellenausschreibungen stellt die systematische Literaturrecherche von Khaouja et al. (2021) dar. Als Suchbegriffe dienten *job ad*, *job market* und *skill*. Alle Artikel müssen im Zeitraum (2010-2021) publiziert worden sein, die Artikel mussten in Englisch verfasst sein, die online Verfügbarkeit gegeben und *peer-reviewed* sein, im Zusammenhang mit der Kompetenzidentifikation und dem Stellenmarkt stehen. Ausschlussgründe waren Kompetenzidentifikation

²²⁸ Kobayashi et al. (2018b).

²²⁹ Eigene Darstellung angelehnt an Khaouja et al. (2021).

ausschließlich aus Lebensläufen, aus Büchern oder grauer Literatur oder Konferenzzusammenfassungen. Sie analysierten 108 verschiedene Studien und kategorisierten diese in vier methodische Ansätze:

3.2 Ähnliche Lösungsansätze für die Problemstellung

Lee et al. (2014) untersuchten und verglichen aus drei Perspektiven die wichtigsten Kompetenzen im IT-Bereich in Nordwest Florida. Zum einen die persönlichen Erwartungen von Studierenden im IT-Bereich, welche Kompetenzen im späteren beruflichen Umfeld von Ihnen gefordert werden, zum anderen die Realität, welche jungen Berufstätige erleben und als Drittes die nachgefragten Kompetenzen aus Stellenausschreibungen. Alle Perspektiven betrachten technische Positionen in der ersten Phase der Berufstätigkeit und schließen Führungspositionen aus. Ein weiteres Ziel wäre gewesen, die Unterschiede zwischen städtischen und ländlichen Gebieten zu evaluieren, wobei dies aufgrund der geringen Zahl an Stellenausschreibungen im ländlichen Raum nicht möglich war. Die verwendete Methodik baut auf folgenden drei Säulen auf: (1) statistischen Auswertung von 225 Stellenausschreibungen, (2) Interviews mit inskribierten Studierenden und (3) Interviews mit jungen Berufstätigen. Die Taxonomie der Kompetenzen erfolgte aufbauend auf dem *Competency Model for IT Program Management*²³⁰ des *Chief Human Capital Officers Council (CHCO)* und besteht aus 25 generischen Kompetenzen, einer physischen Kompetenz und 33 technischen Kompetenzen. Die statistische Auswertung zeigte auf, dass, ohne mehrfach Zählung bei mehrmaliger Nennung in einer Ausschreibung, 1333 generische Kompetenzen, 63 physische Kompetenzen und 1134 technische Kompetenzen nachgefragt wurden. Die Autorenschaft stellte fest, dass es eine große Diskrepanz zwischen den Erwartungen der Studierenden, den Anforderungen der Firmen und der Wahrnehmung der jungen Beschäftigten gab. Die größte Abweichung dieser drei Perspektiven war in Bezug auf den Kundenservice und die kommunikativen Fähigkeiten.²³¹

3.3 Stand der Praxis

Dieser Teil der Arbeit soll Methoden und Umsetzungen im Kontext von Unternehmen und Organisationen aufzeigen. Auf Grund des Eigeninteresses der meisten Unternehmen und Organisationen, nicht zu viele Informationen ihrer derzeitigen Forschung offen zu legen, ist der Zugang zu angestrebten Zielen, Methoden und Umsetzungen nur eingeschränkt möglich. Im Zuge der Recherche zu dieser Arbeit konnten dennoch zwei Beispiele gefunden werden.

²³⁰ Berry (2011).

²³¹ Lee et al. (2014).

Die Arbeiten von Zhao et al. (2015) wollten für das Karriereportal Careerbuilder mit Hilfe von Text Mining Methoden eine automatische Taxonomie aus Lebensläufen und Stellenanzeigen der Plattform erstellen.²³² Diese Taxonomie und das zugehörige Vorgehen ist nicht für die Allgemeinheit zugänglich, jedoch könnte ihre Umsetzung zu Kompetenzerkennung und -abgleich auch im Zusammenhang mit dieser Arbeit relevant sein.

Die Arbeit von Djumalieva und Sleeman (2018) ist öffentlich zugänglich und erstellt eine hierarchische Taxonomie basierend auf drei Ebenen aus Jobausschreibungen der Firma Burning Glas Technologies. Als Hauptgruppen dienen berufliche Felder, wie „health and social care“ und „business administration“, die zweite Gruppe richtet sich nach einzelnen Themenbereichen und die unterste Ebene basiert auf einzelnen Fertigkeiten.²³³

Das online Stellenportal LinkedIn stellt seit 2018 den verwendenden Personen mehrere Vergleichsfaktoren zur Verfügung, mit welchen die Übereinstimmung der KSAOs ihres Lebenslaufes mit den KSAOs aus einer spezifischen Stellenanzeige abgeglichen werden. Über diesen Vergleich wird es den Personen erleichtert zu eruieren, wie gut sie zu einer spezifischen Stelle passen. Gleichzeitig wird es den Beschäftigten der Personalabteilungen erleichtert festzustellen, wie gut eine sich bewerbende Person zu der verfügbaren Stelle passt. Premium Mitgliedschaften erhalten darüber hinaus eine Information, wie gut sie im Vergleich zu allen Bewerbern, welche sich über das Stellenportal beworben haben, abschneiden.²³⁴

Mit Hilfe von Taxonomien versuchen Initiativen wie das amerikanische „*Occupational Information Network*“ (O*NET)²³⁵, das europäische „*European Skills, Competences, Qualifications and Occupations*“ (ESCO)²³⁶ oder auch das österreichische Berufsinformationssystem (BIS)²³⁷ die verwendeten Begriffe bestimmter KSAOs zu vereinheitlichen. Weiters zu einzelnen Jobs zuzuordnen und damit spezifische Kompetenzkataloge abzuleiten. Diese Kompetenzkataloge clustern unterschiedliche KSAOs zu einem Job und sollen dabei helfen, dass ein einheitliches Bild eines Jobs entsteht.

²³² Vgl. Zhao et al. (2015), S. 4012.

²³³ Vgl. Djumalieva und Sleeman (2018), S. 10.

²³⁴ URL: <https://blog.linkedin.com/2018/may/14/introducing-how-you-match-on-linkedin-jobs>, Stand: 28.02.2022.

²³⁵ Vgl. National Research Council (2010).

²³⁶ Vgl. Directorate-General for Employment, Social Affairs and Inclusion (2017).

²³⁷ Vgl. Arbeitsmarktservice.

4 Design und Umsetzung

In diesem Kapitel wird das abstrakte Konzept dieser Arbeit anhand des Text Mining Prozesses abgeleitet (Abschnitt 4.1) und Einflusspunkte definiert und erläutert. Das spezifizierte Konzept ist in Abschnitt 4.1.6 dargestellt. In Abschnitt 4.2 werden Maßnahmen, welche zur Verbesserung der Ergebnisse anhand der ersten Evaluation durchgeführt wurden, vorgestellt. Bei den einzelnen umgesetzten Arbeitsschritten werden die relevanten Programmpakete kurz vorgestellt. Die einzelnen Auswertungen sind in Kapitel 5 gezeigt.

In Abbildung 10 wird der gesamte Text Mining Prozess dargestellt und die für dieses Kapitel relevanten Abschnitte aufgezeigt. Eine detaillierte Darstellung des Prozesses ist in Kapitel 2.3, Abbildung 9, ersichtlich.

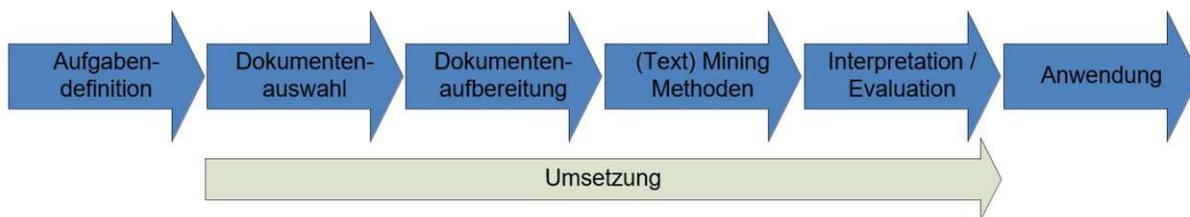


Abbildung 10: Umsetzung des Text Mining Prozess ^{238,239}

Für die Umsetzung einer der Text Mining Methoden, Kompetenzanzahl, wird ein Konzept in abstrakter und spezifischer Form erarbeitet, welches an den oben dargestellten Text Mining Prozess angelehnt ist. Das abstrakte Konzept wird in Abbildung 16 und im Anhang unter Abbildung 31 (Großformat) gezeigt. Es dient als Leitfaden für alle Unterpunkte in diesem Kapitel. Des Weiteren werden die Einflüsse der einzelnen Punkte des abstrakten Konzeptes beschrieben. Die Punkte wurden in Anbetracht der möglichen Umsetzung konkretisiert und zu einem spezifischen Konzept überführt. Das spezifische Konzept wird in Abbildung 19, in Abschnitt 4.1.6, und im Anhang unter Abbildung 32 (Großformat) gezeigt.

Für die Umsetzung der zweiten Text Mining Methode, Topic Modeling, wird auf ein Konzept verzichtet, da es sich dabei um eine standardisierte Methode des Text Mining handelt und in dieser Arbeit nur bereits verfügbare Modelle verwendet werden.

4.1 Vom Text Mining Prozess zum abstrakten Konzept

In diesem Abschnitt wird die Ableitung des abstrakten Konzeptes aus dem Text Mining Prozess vorgestellt. In den darauffolgenden Unterpunkten (4.1.1 - 4.1.5) werden die jeweiligen Prozessschritte aus Sicht des Text Mining Prozesses in Bezug

²³⁸ Eigene Darstellung angelehnt an Hippner und Rentzmann (2006), S. 288.

²³⁹ Eigene Darstellung angelehnt an Kobayashi et al. (2018b), S. 739.

auf das abstrakte Konzept präzisiert. Im ersten Teil sollen Ausschnitte des abstrakten Konzeptes bei der jeweiligen Zuordnung zum Standardprozess helfen.

Die Aufgabe dieser Arbeit ist es, ein Konzept für die Observierung von Veränderung des Kompetenzenmixes in der Instandhaltung zu realisieren und folglich einen Proof-of-Concept durchzuführen, siehe Abbildung 11.

Die Dokumentenauswahl, Abbildung 12, wird auf online verfügbare Stellenanzeigen begrenzt und mit diversen Filtern auf den Instandhaltungsbereich und spezifische Regionen eingegrenzt. Diese Stellenanzeigen wurden mit Hilfe von Webscraping für die darauffolgende Auswertung verfügbar gemacht. Die Stellenanzeigen werden vor der statischen Auswertung zwischengespeichert.

Die Dokumentenaufbereitung nimmt eine Datenreinigung und Datentransformation vor. Dies findet vor der Auswertung im Konzept statt und ist nicht explizit im Konzept dargestellt. Ein exemplarisches Beispiel für dieses Vorgehen ist in Tabelle 8 ersichtlich.

Die (Text) Mining Methoden, Abbildung 13, sind der Kernteil der Auswertung und beinhalten den Kompetenzabgleich, zwischen dem ESCO Standard und den Stellenanzeigen, sowie die Synonymerkennung. Die Ergebnisse werden in tabellarischer Form abgespeichert.



Abbildung 11: Aufgabendefinition²⁴⁰



Abbildung 12: Dokumentenauswahl²⁴¹

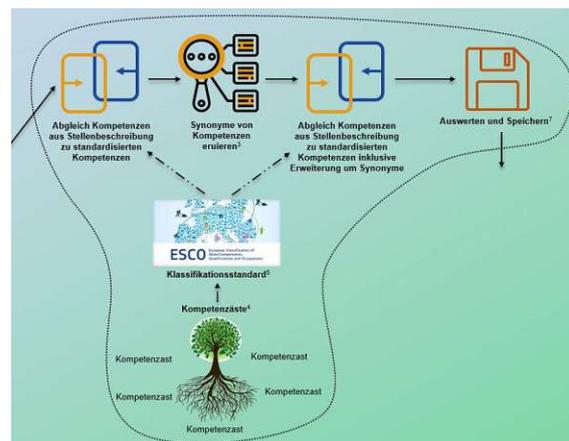


Abbildung 13: (Text) Mining Methoden²⁴²

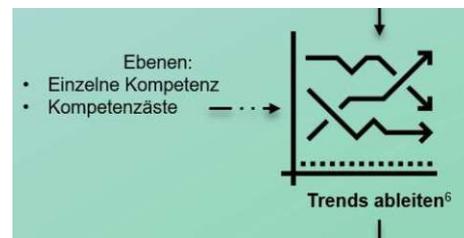


Abbildung 14: Interpretation / Evaluation²⁴³

²⁴⁰ Eigene Darstellung.

²⁴¹ Ebd.

²⁴² Ebd.

²⁴³ Ebd.

Die Evaluierung, Abbildung 14, stellt durch einen Vergleich zwischen den Kompetenzen, welche auf automatisierte Weise gefunden wurden, einer manuellen Auswertung gegenüber und gibt somit Aufschluss über die Qualität der Ergebnisse. Die Evaluierung geschieht im Übergang zwischen der Auswertung und dem Ableiten von Trends.

Die Interpretation, Abbildung 14, leitet Trends aus mehreren zeitlich abgegrenzten Auswertungen ab. Diese werden auf ihre Plausibilität geprüft.

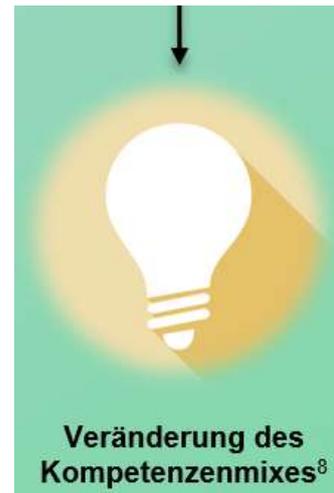


Abbildung 15: Anwendung²⁴⁴

Die Anwendung, Abbildung 15, beschäftigt sich mit der Verwendung der gewonnenen Erkenntnisse. Mögliche Anwendungen können das Ableiten von Weiterbildungsmaßnahmen für die Beschäftigten sein, die Adaptierung von Curricula von Bildungsinstitutionen sein oder verschiedenen Organisationen Einblick in Marktveränderungen von gesuchten Mitarbeitern geben.

In Abbildung 16 ist das abstrakte Konzept dargestellt. Eine großformatige Darstellung befindet sich im Anhang, Abbildung 31. Die einzelnen abstrakten Einflussfaktoren dieses Konzepts werden in den folgenden Unterpunkten (4.1.1 - 4.1.5) präzisiert und in Abschnitt 4.1.6 wird das spezifizierte Konzept dargestellt.

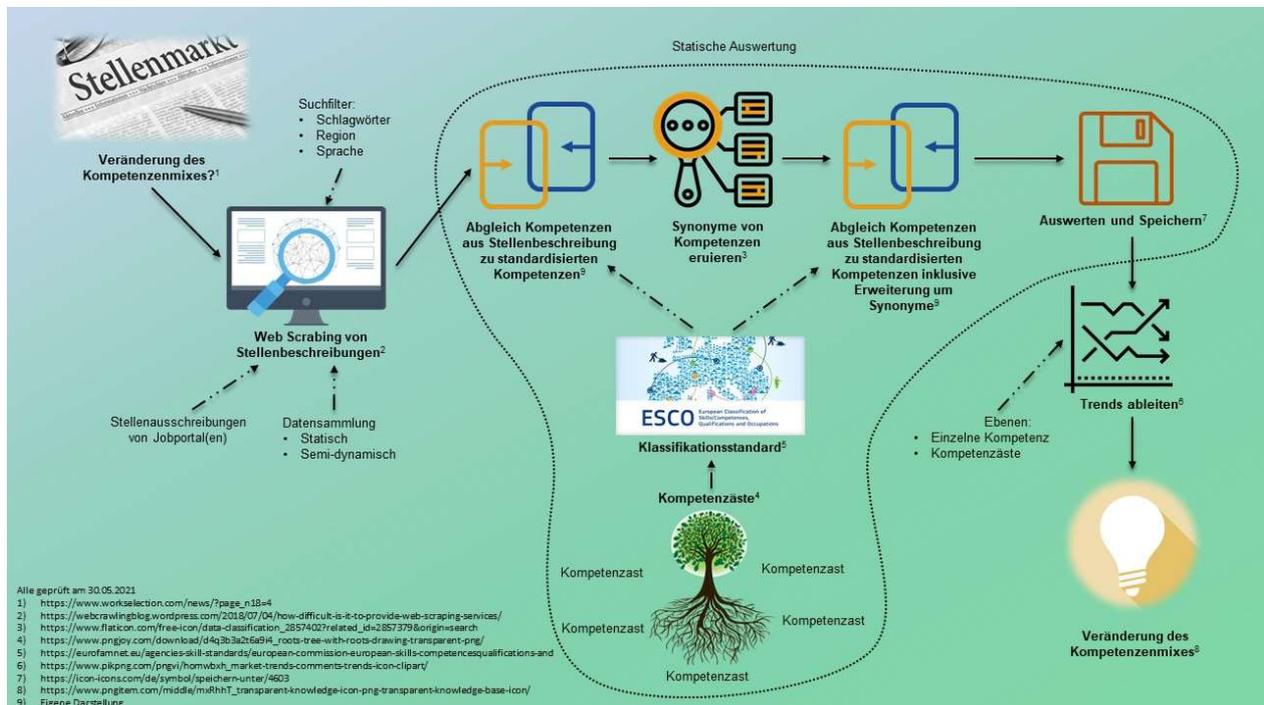


Abbildung 16: Abstraktes Konzept der Umsetzung²⁴⁵

²⁴⁴ Ebd.

²⁴⁵ Eigene Darstellung.

Für genauere Informationen zu den einzelnen verwendeten Python Paketen wird in den jeweiligen Abschnitten auf deren Website bzw. auf deren GitHub-Link verwiesen.

4.1.1 Dokumentenauswahl

Die Spezifizierungen für die Dokumentenauswahl sind in den nachfolgenden Absätzen erläutert und in Tabelle 7 zusammengefasst.

Die Dokumentenauswahl wird auf die verfügbaren Stellenausschreibungen auf dem online Stellenportal indeed eingeschränkt. indeed ist ein Meta-Jobportal und verwaltet keine Stellenausschreibungen, sondern durchsucht das Internet nach Stellenanzeigen. Diese Methodik stellt einen sehr großen Pool an Stellenanzeigen zur Verfügung.

Es werden nur Stellenanzeigen in Englisch berücksichtigt, um die Vergleichbarkeit verschiedener Länder zu vereinfachen. In dieser Arbeit werden alle verfügbaren Stellenanzeigen an einem Zeitpunkt von der indeed Homepage aus Deutschland (indeed.de) und dem vereinigten Königreich (indeed.uk) herangezogen. Die Daten werden nur statisch an einem Zeitpunkt gesammelt, da die meisten Stellenausschreibungen 30 und mehr Tage online verfügbar sind und deshalb im Zeitraum dieser Arbeit es nicht möglich gewesen wäre, mehrere semi-dynamische Durchläufe ablaufen zu lassen. Semi-dynamische Datensammlung bedeutet, dass alle Stellenausschreibungen in einem vorgegebenen Zeitraum gesammelt werden. Damit wären zum Beispiel monatspezifische Auswertungen möglich.

Die Eingrenzung der Stellenanzeigen im Kontext dieser Arbeit, Instandhaltung im produzierenden Gewerbe, wird mit den Suchbegriffen „*maintenance*“ und „*manufacturing*“ erreicht. „*maintenance*“ als einziges Stichwort zu verwenden hätte zu viele Stellenanzeigen berücksichtigt, welche nicht im Kontext des produzierenden Gewerbes stünden. Eine weitere, detailliertere, örtliche Einschränkung ist auf Grund der Websitestruktur, welche länderspezifisch ist, nicht nötig. Alle Stellenanzeigen der länderspezifischen Seiten werden berücksichtigt. Die Sprache der Stellenanzeigen wird mit Hilfe eines Filters auf englische Stellenanzeigen begrenzt.

Tabelle 7: Zusammenfassung der Eingrenzungen für die Dokumentenauswahl²⁴⁶

Plattform	indeed
Regionen	Deutschland und Vereinigtes Königreich
Suchwörter	‚maintenance‘ und ‚manufacturing‘
Datensammlung	Statisch

²⁴⁶ Eigene Darstellung.

Um Duplikate der Stellenanzeigen zu vermeiden, werden die websitespezifischen IDs der Stellenanzeigen abgeglichen. Wird eine ID zwei Mal registriert, wird sie nur einmal in den Korpus aufgenommen.

Die Daten der Stellenanzeigen werden mit Hilfe von Web Scraping von der Website gewonnen. Die gesammelten Daten beinhalten in strukturierter Form den Titel, den zugehörigen Link, die ID und die Stellenbeschreibung. Für diesen Vorgang werden folgende Python Pakete verwendet:

- Selenium²⁴⁷
- BeautifulSoup²⁴⁸

Selenium ermöglicht eine automatisierte Steuerung von Websites und verfügt über eine Vielzahl von vordefinierten Funktionen, um das Navigieren auf Websites zu ermöglichen. Diese Funktionen gehen vom Öffnen eines Browsers, über das Navigieren auf einer Website bis hin zum Ausfüllen bestimmter Felder einer Website.

Beautifulsoup bietet verschiedene Möglichkeiten, den Inhalt von Websites auf dem Computer verfügbar zu machen. Ein Beispiel dafür ist das Erfassen von Text, welcher sich in bestimmten Abschnitten einer Homepage befindet. Das Navigieren erfolgt über Knotenpunkte, wie XML-Tags, der Websitedaten.

Diese Daten werden in einer EXCEL Datei abgelegt, um für die nachgelagerte Analyse der Daten immer die Datenkonsistenz zu gewährleisten.

4.1.2 Dokumentenaufbereitung

Die Dokumentenaufbereitung setzt sich, wie in Kapitel 2.3.3.2 beschrieben, aus dem Reinigen und Transformieren der Daten zusammen.

Zuerst werden die Stellenanzeigen mit Hilfe des Detector Factory²⁴⁹ Pakets auf deren Sprache überprüft, da es immer wieder Stellenanzeigen gibt, welche zwar auf der Homepage zur gewollten Sprache Englisch zugeordnet werden, dieses Kriterium jedoch nur zu einem bestimmten Anteil erfüllen. Stellenanzeigen mit einer anderen Sprache als Englisch werden in diesem Schritt ausgeschlossen und in der weiteren Auswertung nicht mehr berücksichtigt.

Die Daten werden von etwaigen Sonderzeichen bereinigt und einer *Tokenization* unterzogen. Im Zuge der *Tokenization* werden die *tokens* auch dem *Lowercasing* unterzogen, womit alle Buchstaben auf Kleinbuchstabenform gebracht werden. Als weiteren Schritt können die einzelnen *tokens* einer Rechtschreibprüfung unterzogen

²⁴⁷ URL: <https://www.selenium.dev/>, Stand: 04.09.2021.

²⁴⁸ URL: <https://beautiful-soup-4.readthedocs.io/en/latest/#>, Stand: 04.09.2021.

²⁴⁹ URL: https://github.com/Mimino666/langdetect/blob/master/langdetect/detector_factory.py, Stand: 04.09.2021.

werden, jedoch beansprucht dies viel Rechenkapazität und kann die Ergebnisse auch ungewollt verändern, deshalb wurde diese in den späteren Ergebnissen nicht berücksichtigt. Die Rechtschreibprüfung wurde mit Hilfe von Spell Checker²⁵⁰ umgesetzt.

Die *stop words* werden mit Hilfe einer Standardliste, welche im Paket Natural Language Toolkit (NLTK)²⁵¹ enthalten ist, entfernt. Des Weiteren können noch zusätzliche korpuspezifische *stop words* aus den Daten entfernt werden, jedoch wurde darauf in den Ergebnissen dieser Arbeit verzichtet.

Die einzelnen *tokens* werden mit Hilfe des in spaCy²⁵² integrierten *Lemmatizers* auf deren Stammform gebracht.

Ein explizites Beispiel der einzelnen Aufbereitungsschritte und deren jeweiliges Ergebnis ist in Tabelle 4 ersichtlich. Ein weiteres explizites Beispiel ist in Tabelle 8 gezeigt, welches auch folglich als Beispiel für die Synonymerkennung dient.

Tabelle 8: Exemplarisches Beispiel für die Textreinigung²⁵³

Ausgangssatz	and to disseminate information to Engineering Production
nach tokenizing	'and', 'to', 'disseminate', 'information', 'to', 'engineering', 'production'
nach stop word removal	disseminate', 'information', 'engineering', 'production'
nach lemmatizing	'disseminate', 'information', 'engineering', 'production'

Bei diesem Beispiel ist gut ersichtlich, dass das *Lemmatizing* nicht immer zu einer Reduktion beiträgt.

4.1.3 Text Mining Methoden

Wie in der Arbeit von Khaouja et al. (2021) gut zusammengefasst gibt es derzeit vier Hauptvorgehensweisen, wie man Kompetenzen identifizieren kann. Namentlich sind diese:²⁵⁴

- Kompetenzanzahl
- Topic Modeling
- Skill Embedding
- Machine Learning Methods

Zwei dieser Methoden werden in dieser Arbeit umgesetzt und nachgelagert in diesem Kapitel vorgestellt.

²⁵⁰ URL: <https://pyspellchecker.readthedocs.io/en/latest/>, Stand: 04.09.2021.

²⁵¹ URL: <https://www.nltk.org/>, Stand: 04.09.2021.

²⁵² URL: <https://spacy.io/>, Stand: 04.09.2021.

²⁵³ Eigene Darstellung.

²⁵⁴ Vgl. Khaouja et al. (2021), 6ff.

4.1.3.1 Kompetenzanzahl

Die Anzahl der Kompetenzen ist die am häufigsten verwendete Methode, um Kompetenzen zu identifizieren.²⁵⁵ Dies stellt das Kernelement der Umsetzung dar, welches in Abbildung 17 ersichtlich ist.

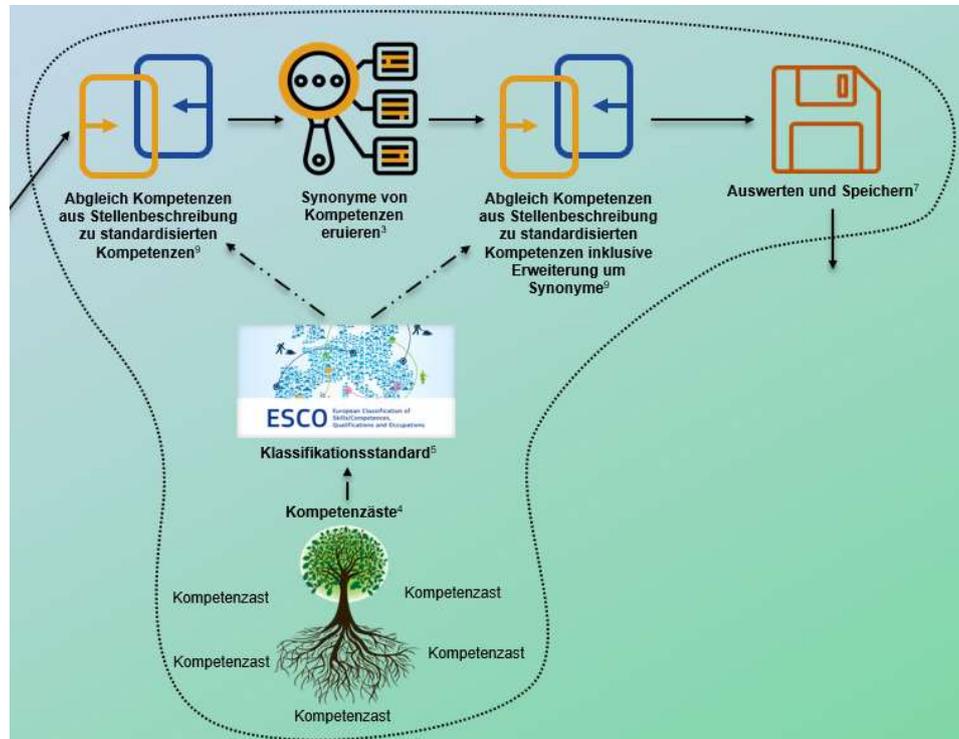


Abbildung 17: Auswertungsteil der Kompetenzanzahl²⁵⁶

Hier gibt es zwei verschiedene Zugänge, entweder werden die Kompetenzen mit einer Basis abgeglichen oder die Auswertung erfolgt manuell mit Hilfe von fachkundigen Personen, welche jede einzelne Stellenanzeige durchgehen müssen.²⁵⁷ Letzteres ist folglich äußerst arbeitsintensiv und wird nicht in dieser Arbeit angewandt. In dieser Arbeit wird aus diesem Grund der Ansatz mit einer bestimmten Basis verfolgt und dafür werden alle Kompetenzen des ESCO Standards als Basis herangezogen.

Der ESCO Standard enthält fünf Hauptäste an Kompetenzen:²⁵⁸

- *Skill* – Fertigkeiten
- *Knowledge* – Wissen
- *Digital competencies (ICT skills)* – Digitale Fertigkeiten
- *Transversal skills* – Transversale Fertigkeiten
- *Language skills* – Sprachliche Fertigkeiten

²⁵⁵ Vgl. Khaouja et al. (2021), S. 6.

²⁵⁶ Eigene Darstellung.

²⁵⁷ Vgl. Khaouja et al. (2021), S. 7.

²⁵⁸ <https://ec.europa.eu/esco> (2021).

Diese fünf Hauptgruppen werden von ESCO als Download auf ihrer Website zur Verfügung gestellt.

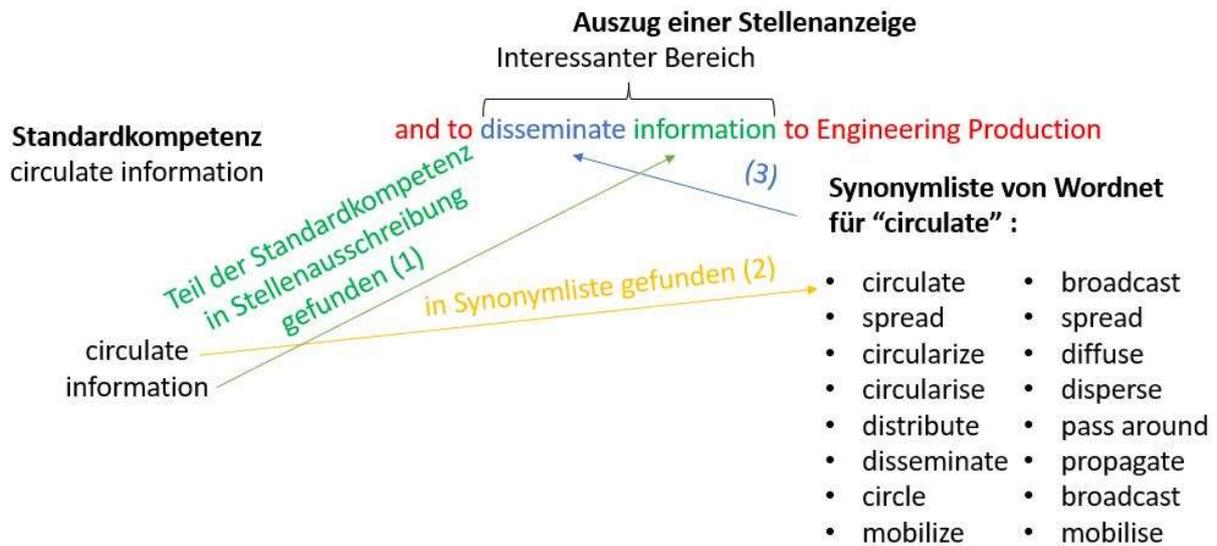


Abbildung 18: Umgesetzte Synonymerkennung²⁵⁹

Da es bei diesem Ansatz der Kompetenzidentifikation nur zu einem Treffer kommt, wenn die ganze Wortgruppe des Standards gefunden wird, wird sie in dieser Arbeit um eine ontologiebasierte Synonymerkennung (siehe auch Kapitel 2.3.4) erweitert. Sobald ein Wort der Standardwortgruppe gefunden wird und insofern die Wortgruppe länger als ein einziges Wort ist, wird jener Bereich der Stellenbeschreibung auf die Länge der Standardwortgruppe eingegrenzt. Für alle Wörter der Standardkompetenz werden die Synonyme mit Hilfe des Pakets Wordnet (Paket in NLTK verfügbar) in einer Liste gesammelt. Werden alle Wörter, bis auf das bereits gefundene, des Abschnitts der Stellenanzeige in der Synonymliste gefunden, wird es als Treffer gewertet. Ein schematisches Beispiel für die Synonymerkennung wird in Abbildung 18 gezeigt.

Jeder Treffer wird hierbei nur einmal je Stellenanzeige gewertet. Dies bedeutet, dass auch wenn eine Kompetenz in einer Stellenanzeige mehrere Male vorkommt, wird diese Kompetenz bei der Auswertung nur einmal berücksichtigt.

4.1.3.2 Topic Modeling

Wie bereits in einem Unterpunkt von Kapitel 2.3.3.3 beschrieben, gibt es mehrere gängige Methoden für das Topic Modeling. In dieser Arbeit wird ein LDA Modell von scikit-learn²⁶⁰ verwendet, welches in zwei Varianten angewandt wird. Zum einen ohne Einschränkung des Vokabulars und zum anderen mit einer Einschränkung des Vokabulars auf nur im ESCO Standard enthaltene Kompetenzen.

²⁵⁹ Eigene Darstellung.

²⁶⁰ URL: <https://scikit-learn.org/stable/>, Stand: 21.09.2021.

Auf eine Synonymerkennung wird bei diesem Ansatz verzichtet, da dies nur mit Hilfe einer Manipulation der Datenbasis möglich ist.

Die LDA Modelle werden bzgl. der optimalen Anzahl an unterschiedlichen Themengebieten bei jedem Durchlauf automatisch optimiert. Parameter dafür sind der *Perplexity*-Wert und der *log likelihood*-Wert, welche im Optimum ein Minimum (*Perplexity*) bzw. ein Maximum (*log likelihood*) erreichen.

Topic Modeling ist jedoch nicht geeignet, um zeitliche Trends ablesen zu können, da bei jedem Durchlauf neue Themengebiete erstellt werden und somit keine konsistente Zuordnung möglich ist. Es ermöglicht jedoch eine Interpretation zu spezifischen Stellen an einem bestimmten Zeitpunkt.²⁶¹

4.1.4 Interpretation und Evaluation

4.1.4.1 Evaluation

Die Evaluation der Ergebnisse erfolgt durch eine manuelle Stichprobe von fünf zufällig gewählten Stellenanzeigen. Die manuell gefundenen Kompetenzen werden in zwei Kategorien unterteilt. Zum einen Kompetenzen, welche ähnlich zum ESCO Standard verstanden werden, aber deren Detaillierungsgrad bzw. deren genaue Wortwahl nicht mit den angewandten Methoden gefunden werden. Zum anderen Kompetenzen, welche entweder zu generisch bzw. keiner Standardkompetenz entsprechen.

Die manuell gefundenen Kompetenzen werden mit den automatisiert gefundenen Kompetenzen gegenübergestellt, falsch positive identifiziert und deren Verhältnisse berechnet. Unter falsch positiven Kompetenzen versteht man Kompetenzen, welche automatisiert gefunden werden, jedoch die Bedeutung aus der Stellenanzeige sich beträchtlich von jener aus dem Standard unterscheidet.

4.1.4.2 Interpretation

Die Interpretation leitet Trends aus mehreren zeitlich abgegrenzten Auswertungen ab. Diese werden auf ihre Plausibilität überprüft.

Handelt es sich bei den festgestellten Trends um temporäre Auffälligkeiten oder sind diese dauerhaft. Dies benötigt eine Betrachtung von mehreren aufeinanderfolgenden Auswertungen. So wird klargestellt, ob ein Trend nachhaltig ist oder dieser nur zu einem bestimmten Zeitpunkt aufgetreten ist.

²⁶¹ Vgl. Pejic-Bach et al. (2020), S. 422.

4.1.5 Anwendung

Insofern die Ergebnisse valide und plausibel sind, können sie als Grundlage für Veränderungen in mehreren Bereichen sein. Die Ergebnisse könnten als Einflussfaktoren für die Einstellungspolitik von Organisationen, Anpassungen von Curricula von Bildungsinstitutionen oder auch für die Ableitung von Weiterbildungsmaßnahmen in Organisationen herangezogen werden.

4.1.6 Spezifiziertes Konzept

In Abbildung 19 sind alle offenen Punkte des abstrakten Konzeptes durch deren Spezifizierungen ersetzt. Eine großformatige Darstellung der Abbildung befindet sich im Anhang, Abbildung 32.

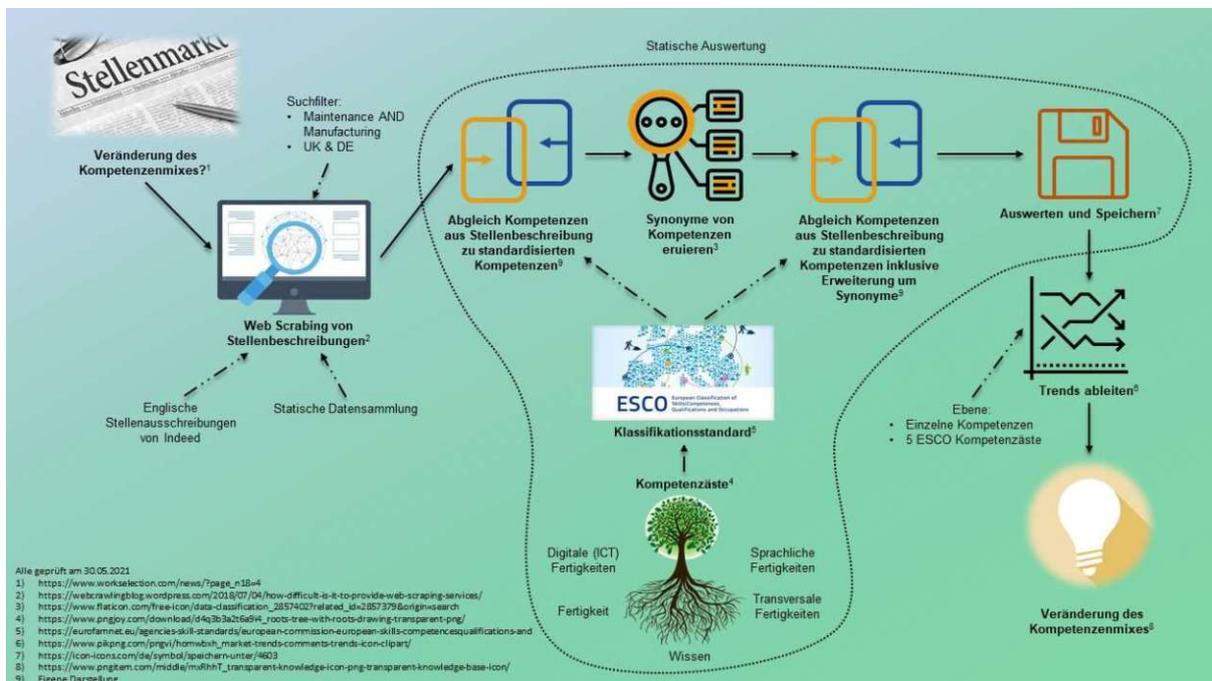


Abbildung 19: Spezifisches Konzept der Umsetzung²⁶²

4.2 Erweiterungen auf Basis der ersten Evaluierung

Die Evaluierung der Ergebnisse des Konzeptes, welches in Abbildung 19 dargestellt ist, lieferte ein Ergebnis von ca. 24%. Die rechnerische Auswertung befindet sich in Tabelle 12. Eine detaillierte Darstellung der Auswertung befindet sich im Anhang 1.

Bei dieser Evaluierung gehen drei Hauptaspekte hervor. Einerseits werden in den Stellenanzeigen ähnliche Kompetenzen zu jenen des ESCO Standard nachgefragt, jedoch nicht im nötigen Detaillierungsgrad, um die Kompetenzen mit dieser Vorgehensweise automatisiert erkennen zu können. Dies umfasst beide Richtungen, entweder waren die Kompetenzen in den Stellenanzeigen oder im Standard zu

²⁶² Eigene Darstellung.

detailliert. Andererseits werden auch generelle Kompetenzen nachgefragt, welche nicht im ESCO aufgenommen sind, bzw. mit einer manuellen Suche und Interpretation keiner im ESCO Standard enthaltenen Kompetenz zugeordnet werden können. Zumeist handelt es sich dabei um konkrete Instandhaltungskonzepte bzw. Methoden. Der dritte Aspekt ist das Erkennen von Kompetenzen, welche keinen sinnvollen Zusammenhang zu der jeweiligen Stellenanzeige erkennen lassen. Beispiel für diese Punkte sind in Tabelle 13, Tabelle 14 und Tabelle 15 angeführt.

Aus diesem Grund werden mehrere Anpassungen des Konzeptes umgesetzt. Die Stellenanzeigen werden aufgrund ihres Stellentitels eingeschränkt und um eine manuell erstellte Erweiterungsliste anhand von 15 Stellenanzeigen ausgebaut. Mit dieser Umsetzung können die Ergebnisse der Evaluierung auf ca. 75% angehoben werden.

In dieser Umsetzung mussten die Stellenanzeigen die Schlagwörter „*maintenance*“ und „*engineer*“ in deren Titel enthalten, um die Verbesserung in einem eingeschränkten Bereich zu zeigen und den Erstellungsaufwand der Erweiterungsliste möglichst gering zu halten. Es werden nur Fertigkeiten- und Wissenskompetenzen der Standardstellenanzeige „*maintenance and repair engineer*“²⁶³ des ESCO Standards herangezogen. Dies reduziert die ca. 13000 ESCO Fertigungs- und Wissenskompetenzen auf 47 Kompetenzen in diesen beiden Kategorien. Diese Maßnahme reduziert die Programmlaufzeit erheblich. Alle Kompetenzen der Kompetenzäste ICT, transversale und sprachliche Kompetenzen des ESCO Standards werden verwendet, da diese bedeutend kleinere Kompetenzgruppen darstellen und auch in vielen verschiedenen Branchen anwendbar sind.

Tabelle 9: Vergleichsbeispiele manuelle Erweiterungsliste zu ESCO Kompetenzen / Instandhaltungskonzepten²⁶⁴

Manuelle Erweiterungsliste	ESCO Kompetenz / Instandhaltungskonzept
health and safety	health and safety extra inkludiert, da nicht im Kompetenzprofil des „ <i>maintenance and repair engineer</i> “
continuous improvement	continuous improvement philosophies
grinding	operate grinding machine
TPM	Instandhaltungskonzept (total productive maintenance)
find fault	(1) inspect for faults in the vehicle's electric system (2) identify faults in utility meters
improve OEE	Instandhaltungskonzept (OEE = overall equipment effectiveness)

²⁶³ URL: <http://data.europa.eu/esco/occupation/615920c5-4f63-4eb3-8b60-afaaed3ab1ff>, Stand: 30.12.2021.

²⁶⁴ Eigene Darstellung.

Die Kompetenzen der manuellen Erweiterungsliste werden anhand von 15 Stellenanzeigen gefunden und aufgenommen, wenn diese eines von zwei Kriterien erfüllen. Entweder handelte es sich um einschlägige Instandhaltungskonzepte, wie TPM (Total productive maintenance), welche nicht im ESCO Standard enthalten sind oder um Kompetenzen, welche direkt in leicht abgewandelter Form oder auf einer anderen generischen Ebene im ESCO Standard enthalten sind. Einige Beispiele dafür sind in Tabelle 9 gezeigt. Insgesamt umfasst die Erweiterungsliste 42 Kompetenzen.

Das angepasste spezifische Konzept ist in Abbildung 20 und als Großformat im Anhang in Abbildung 33 dargestellt.

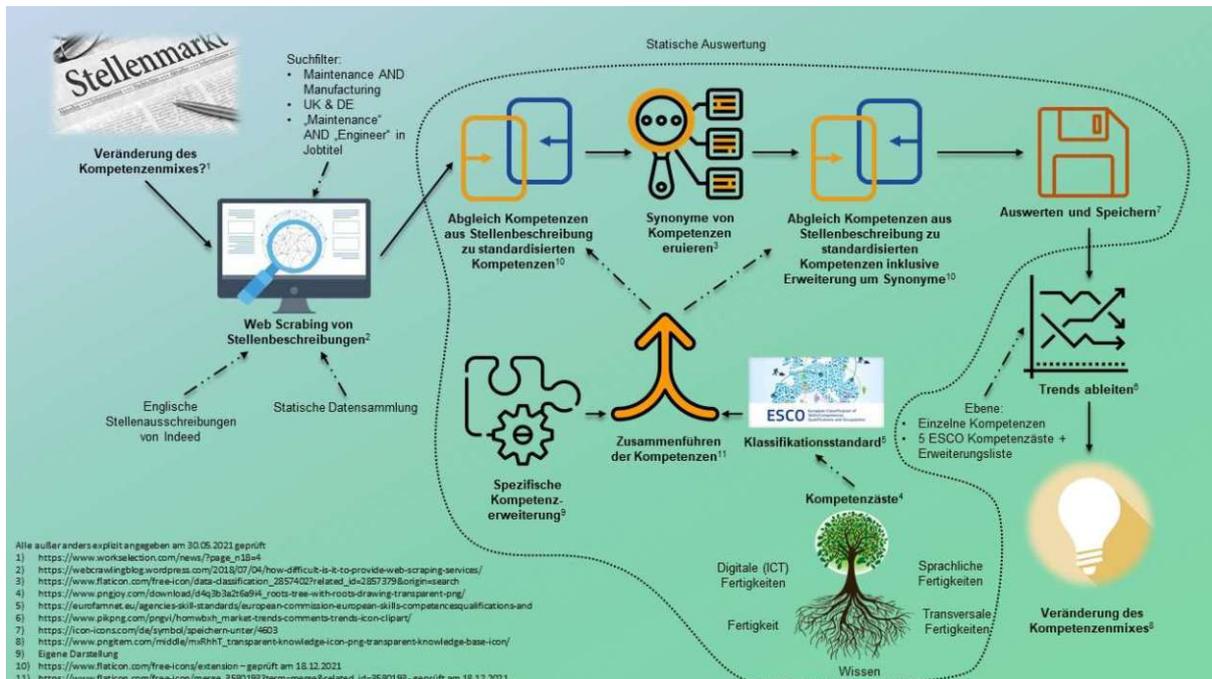


Abbildung 20: Spezifisches Konzept der Umsetzung nach der Adaptierung²⁶⁵

²⁶⁵ Eigene Darstellung.

5 Auswertung und Resultate

In diesem Kapitel werden alle Auswertungen und Resultate in Anbetracht der direkten Ergebnisse, in Bezug auf die Forschungsfrage und in Bezug auf den möglichen weiteren Nutzen dieser Arbeit vorgestellt.

In Abschnitt 5.1 werden die Ergebnisse der einzelnen Auswertungen inklusive jener mit manuellen Erweiterungslisten beschrieben. In Abschnitt 5.2 werden die Ergebnisse in Bezug mit der Forschungsfrage gebracht und in Abschnitt 0 werden mögliche Anwendungen der Ergebnisse aufgezeigt.

5.1 Auswertung der angewendeten Methoden

In diesem Abschnitt werden alle Ergebnisse der angewandten Methoden gezeigt. Die Methoden werden auf alle Daten anhand mehrerer Eingrenzungen, wie in Abschnitt 4.1.1 vorgestellt, gesammelt. Die Daten für beide Modelle wurden, wie in Abschnitt 4.1.2 vorgestellt, bereinigt und transformiert.

In Abschnitt 5.1.1 werden die Ergebnisse anhand der Kompetenzanzahl und in Abschnitt 5.1.2 die Ergebnisse anhand des Topic Modells aufgezeigt. In Abschnitt 5.1.1 sind die weiteren Evaluierungen zu den einzelnen Ausbaustufen enthalten.

5.1.1 Auswertung Kompetenzanzahl

Anhand der Basis des ESCO Standards werden diese Kompetenzen in den einzelnen Stellenanzeigen gefunden. Als Ergebnis erhält man die Anzahl wie oft eine Standardkompetenz in den Stellenanzeigen vorkommt und in welchen Stellenanzeigen diese Kompetenz vorkommt.

Man kann daraus, in Bezug auf die Hauptäste des ESCO Standards, die häufigsten nachgefragten Kompetenzen auswerten. Des Weiteren können die einzelnen Kompetenzen anhand der ESCO Äste summiert werden, um die Wichtigkeit der einzelnen Äste darzustellen. Mit zeitlichem Abstand zweier Durchläufe lassen sich so Veränderungen erkennen. Zum einen auf granularer Ebene und zum anderen auf abstrakter Ebene (Kompetenzäste). Die einzelnen Kompetenzäste enthalten unterschiedliche viele Kompetenzen, siehe Tabelle 10.

Tabelle 10: ESCO Kompetenzäste und jeweilige Anzahl an Standardkompetenzen (Version 1.0.8)²⁶⁶

Ast	Skills	Knowledge	ICT	Transversal	Language
Anzahl verschiedener Kompetenzen	10199	2833	21	107	345

Über die Zuordnung der Kompetenzen zu den einzelnen Dokumenten lässt sich auswerten, wie viele Kompetenzen in den einzelnen Stellenanzeigen nachgefragt werden.

Diese Auswertungen an zwei Zeitpunkten (12.7.2021 & 30.8.2021) werden folglich vorgestellt. Weitere Ergebnisse, wie der Kompetenzkatalog und erkennbare Trends, werden in Kapitel 5.2 vorgestellt.

5.1.1.1 Auswertung vom 12.7.2021

Es wurden gesamt 1228 Stellenanzeigen von beiden indeed Websites gesammelt. Davon stammten 653 von indeed UK und 575 von indeed DE. Die automatische Spracherkennung schloss zwei Stellenanzeigen aus, da diese nicht zum überwiegenden Teil in englischer Sprache sind. Es wurden 1226 Stellenanzeigen der Auswertung unterzogen.

Die Top 10 der nachgefragten Kompetenzen gegliedert nach den ESCO Hauptästen sind in Tabelle 11 ersichtlich.

Tabelle 11: Top 10 Kompetenzen je Kompetenzast -12.07.2021²⁶⁷

	Skill		Knowledge		ICT		Transversal		Language	
Summe	2087		5024		9		244		789	
Top 1	give grant	153	communication	457	computer programming	5	work independently	65	english	406
Top 2	troubleshoot	107	communication science	247	solve technical problem	4	attend detail	39	german	301
Top 3	promote	98	textile material	189			work team	26	french	22
Top 4	operate control system	77	pneumatic	175			work shape space	25	dutch	13
Top 5	breed stock	57	hydraulic	146			work efficiently	15	spanish	10
Top 6	develop design plan	53	energy	129			identify opportunity	14	write english	6
Top 7	management skill	39	c	120			make decision	14	swedish	5
Top 8	select subject matter	39	electrical engineering	109			meet commitment	6	japanese	4
Top 9	control production	36	mechanical engineering	102			computer programming	5	italian	4
Top 10	work shift	33	electronic	101			manage time	5	danish	3

Es wurden gesamt 8153 ESCO Kompetenzen in den 1226 Stellenanzeigen gefunden. Die summierten Kompetenzen sind in Abbildung 21 als Balkendiagramm,

²⁶⁶ Eigene Darstellung.

²⁶⁷ Ebd.

gegliedert nach den Kompetenzästen, dargestellt. Dadurch wird gut ersichtlich, dass in diesen Stellenanzeigen primär nach Wissen nachgefragt wird.

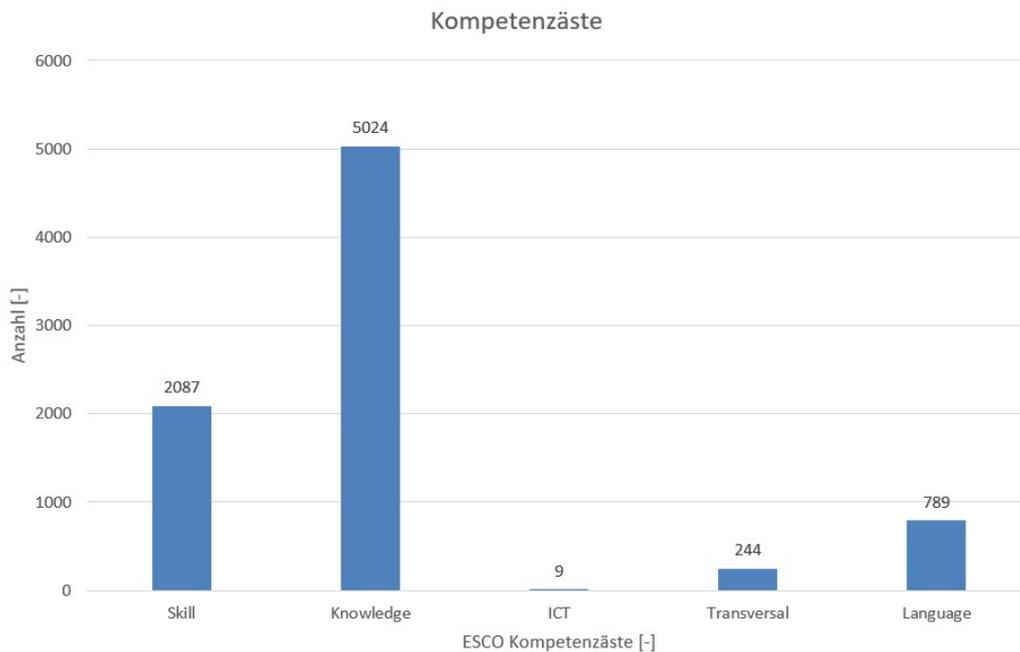


Abbildung 21: Kompetenzäste – 12.07.2021²⁶⁸

Die Kompetenzverteilung wird in Abbildung 22 gezeigt. Im Schnitt werden 6,65 Kompetenzen je Stellenanzeige gefunden, jedoch wird in 133 Stellenanzeigen keine einzige zuordenbare Kompetenz gefunden. Die möglichen Problematiken und Gründe werden in der nachfolgenden Evaluierung, Abschnitt 5.1.1.2, aufgezeigt.

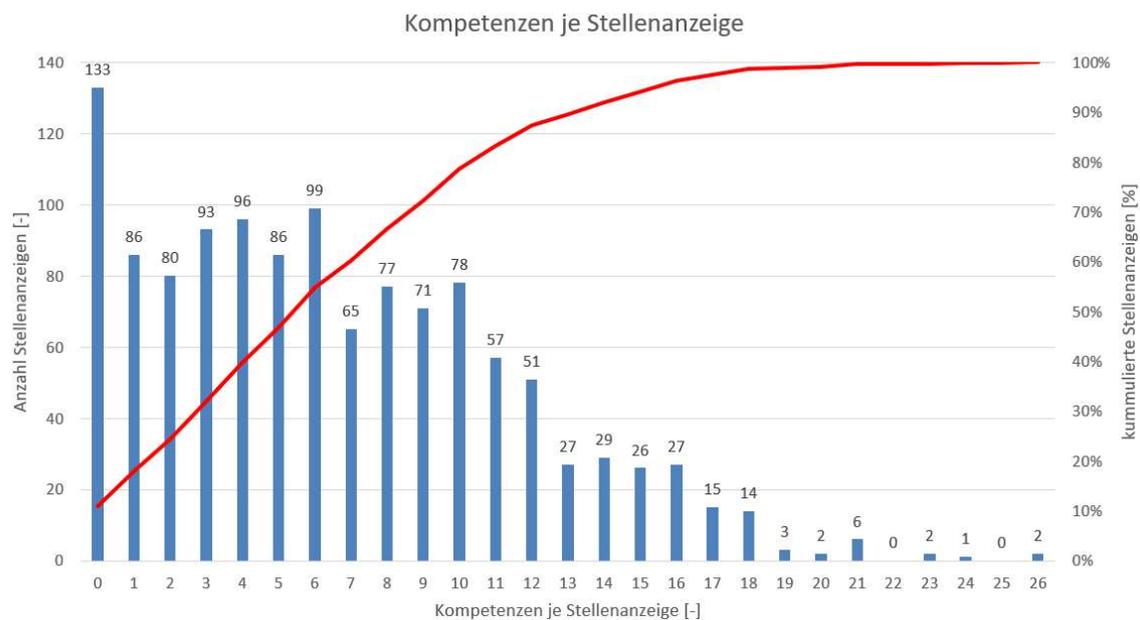


Abbildung 22: Kompetenzverteilung – 12.07.2021²⁶⁹

²⁶⁸ Eigene Darstellung.

²⁶⁹ Ebd.

5.1.1.2 Evaluierung der Auswertung vom 12.07.2021:

Die Evaluierung wurde wie in Kapitel 4.1.4 beschrieben durchgeführt. Dafür wurden fünf zufällig gewählte Stellenanzeigen aus dem Korpus manuell ausgewertet. Bei der manuellen Auswertung wurde unterschieden, ob es eine ähnliche Kompetenz im ESCO Standard gibt oder nicht.

Wie bereits beschrieben war in 133 Stellenanzeigen keine automatisiert gefundene Kompetenz enthalten, deshalb ist auch eine der fünf Stellenanzeigen ohne eine automatisch gefundene Kompetenz. In Tabelle 12 sind die Ergebnisse zusammengefasst. Die Details der Evaluierung bei welchen die Kompetenzen in den einzelnen Stellenanzeigen markiert sind, befinden sich im Anhang 1.

Tabelle 12: Evaluation²⁷⁰

Stellenanzeige / Kategorie	26	419	629	886	902
Kompetenzen ähnlich zum ESCO Standard	3	1	0	11	16
Kompetenzen ohne Bezug zum ESCO Standard	3	9	2	7	8
∑ manuell gefundener Kompetenzen	6	10	3	18	24
Automatisiert gefundene Kompetenzen	0	1	1	6	10
Falsch positive automatisiert gefundene Kompetenzen	0	0	1	0	0
Automatisiert gefundene Kompetenzen	0%	10%	33%	33%	42%
∑ manuell gefundener Kompetenzen					
Falsch positive automatisiert gefundene Kompetenzen	-	0%	100%	0%	0%
Automatisiert gefundene Kompetenzen					

Im arithmetischen Mittel wurden rund 24% der enthaltenen Kompetenzen in den Stellenanzeigen mit diesem Vorgehen gefunden.

In einer der fünf Stellenanzeigen wurde eine falsche positive Kompetenz gefunden. „breed stock“ befindet sich unter den Top 10 der gefundenen Kompetenzen, aber es kann angenommen werden, dass diese Kompetenz im Zusammenhang mit der Instandhaltung nicht häufig nachgefragt wird, sondern in einem landwirtschaftlichen Kontext verwendet wird bzw. im ESCO Standard als Fertigkeit im Kontext der Zucht von Insektenstöcken steht.

²⁷⁰ Eigene Darstellung.

Gründe für Probleme beim Erkennen von Kompetenzen sind:

- Verwendung von unterschiedlichen Wörtern für eine dem Standard ähnliche Kompetenz, jedoch nicht ähnlich genug, um diese mit Hilfe der Synonymerkennung zu erfassen. Beispiel:

Tabelle 13: Beispiel des Ähnlichkeitsproblems aus einer Stellenanzeige²⁷¹

Stellenanzeige	ESCO
repair and maintenance of printing and bookbinding machinery	maintenance of printing machines

- Generische Kompetenzen, welche in dieser Abstraktheit nicht im ESCO Standard vorkommen. Beispiel:

Tabelle 14: Beispiel des generischen Kompetenzproblems aus einer Stellenanzeige²⁷²

Stellenanzeige	ESCO
work safely	work safely with machines

- Falsch Positive erkannte Kompetenzen, welche über die Synonymerkennung einer ESCO Kompetenz zugeordnet werden, jedoch kein sinnhafter Zusammenhang erkannt werden kann. Beispiel:

Tabelle 15: Beispiel einer falsch positiven Kompetenz aus einer Stellenanzeige²⁷³

Stellenanzeige	ESCO
upgrading equipment	climb equipment
employee stock	breed stock

5.1.1.3 Auswertung vom 30.8.2021

Es wurden gesamt 1277 Stellenanzeigen von beiden indeed Websites gesammelt. Davon stammten 646 von indeed UK und 631 von indeed DE. Die automatische Spracherkennung schloss eine Stellenanzeige aus, da diese nicht zum überwiegenden Teil in englischer Sprache war. Es wurden 1276 Stellenanzeigen der Auswertung unterzogen.

Die Top 10 der nachgefragten Kompetenzen gegliedert nach den ESCO Hauptästen sind in Tabelle 16 ersichtlich.

²⁷¹ Eigene Darstellung.

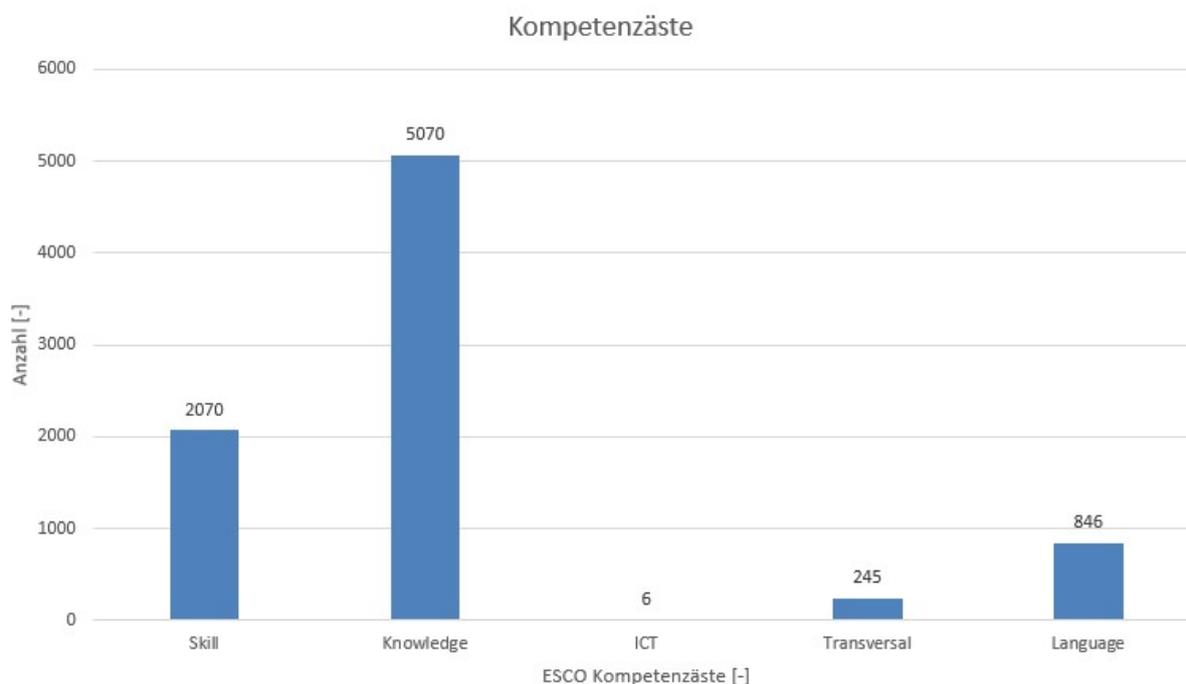
²⁷² Ebd.

²⁷³ Ebd.

Tabelle 16: Top 10 Kompetenzen je Kompetenzast -30.08.2021²⁷⁴

	Skill		Knowledge		ICT		Transversal		Language	
Summe	2070		5070		6		245		846	
Top 1	give grant	161	communication	472	solve technical problem	4	work independently	69	english	437
Top 2	troubleshoot	111	communication science	254	computer programming	2	attend detail	41	german	328
Top 3	promote	87	textile material	194			work team	27	french	19
Top 4	breed stock	64	pneumatic	167			work efficiently	19	spanish	9
Top 5	operate control system	64	hydraulic	135			work shape space	14	dutch	9
Top 6	develop design plan	57	energy	129			identify opportunity	13	write german	9
Top 7	management skill	44	electronic	122			make decision	13	italian	7
Top 8	application process	34	electrical engineering	119			adapt change	10	write english	5
Top 9	provide documentation	32	model	117			meet commitment	5	japanese	5
Top 10	control production	30	mechanical engineering	116			solve technical problem	4	czech	5

Es wurden gesamt 8237 ESCO Kompetenzen in den 1276 Stellenanzeigen gefunden. Die summierten Kompetenzen sind in Abbildung 23 als Balkendiagramm, gegliedert nach den Kompetenzästen, dargestellt. Dadurch wird gut ersichtlich, dass in diesen Stellenanzeigen primär nach Wissen nachgefragt wird.

Abbildung 23: Kompetenzäste – 30.08.2021²⁷⁵

Die Kompetenzverteilung wird in Abbildung 24 gezeigt. Im Schnitt werden 6,5 Kompetenzen je Stellenanzeige gefunden, jedoch wird in 151 Stellenanzeigen keine

²⁷⁴ Eigene Darstellung.

²⁷⁵ Ebd.

einzigste zuordenbare Kompetenz gefunden. Die möglichen Problematiken und Gründe sind in der Evaluation der Ergebnisse vom 12.7.2021, Tabelle 12, ersichtlich.

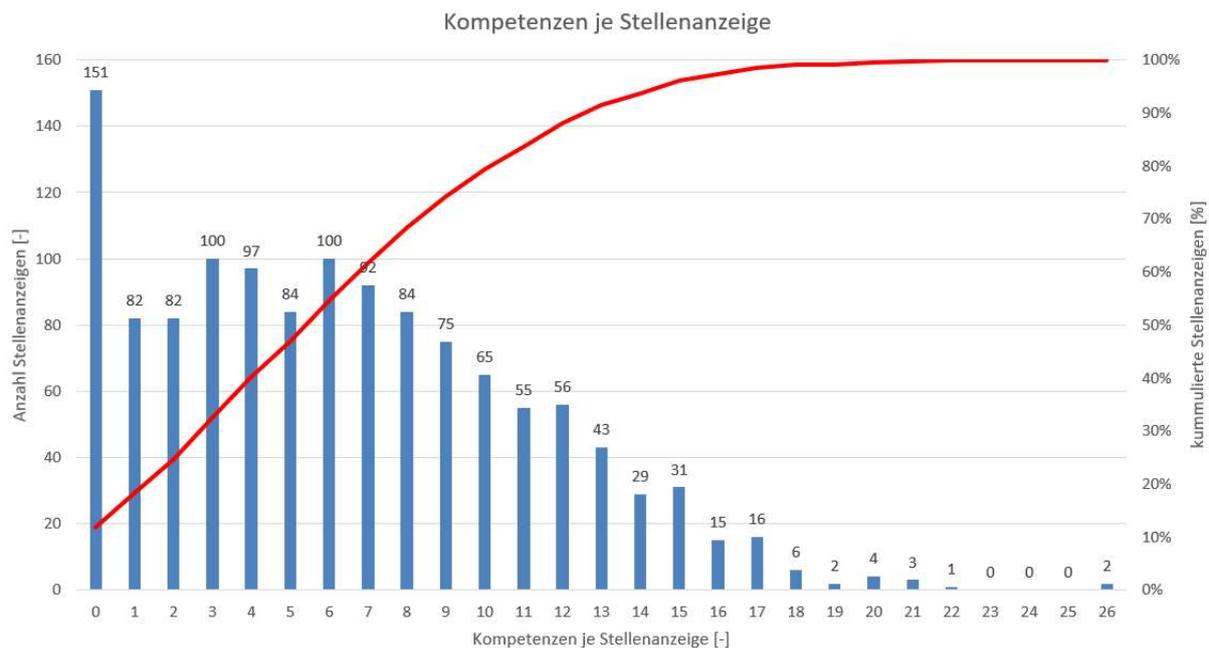


Abbildung 24: Kompetenzverteilung – 30.08.2021²⁷⁶

Auf eine Evaluierung der Ergebnisse wurde zu diesem Zeitpunkt verzichtet.

5.1.1.4 Auswertung nach Adaptierungen auf Basis der ersten Evaluation

Die Änderungen der Umsetzung sind Abschnitt 4.2 vorgestellt. In diesem Abschnitt werden die Auswertung und Evaluation der adaptierten Version vorgestellt.

Es wurden gesamt 1228 Stellenanzeigen von beiden indeed Websites gesammelt. Durch die Einschränkung, wie in Abschnitt 4.2 beschrieben, dass im Stellentitel die Schlagwörter „*maintenance*“ und „*engineer*“ enthalten sein mussten, wurden bei der Auswertung 371 Stellenanzeigen herangezogen. Die automatische Spracherkennung schloss keine Stellenanzeigen aus.

Die Top 10 der nachgefragten Kompetenzen gegliedert nach den ESCO Hauptästen sind in Tabelle 17 ersichtlich. Die Top 10 der nachgefragten Kompetenzen, welche aufgrund der Erweiterungsliste erkannt wurden, sind in Tabelle 18 ersichtlich.

²⁷⁶ Eigene Darstellung.

Tabelle 17: Top 10 Kompetenzen je Kompetenzast nach Adaptierung ESCO Kompetenzen - 12.07.2021²⁷⁷

	Skill		Knowledge		ICT		Transversal		Language	
Summe	149		871		5		155		22	
Top 1	troubleshoot	78	electricity	269	solve technical problem	3	work team	42	english	11
Top 2	maintain equipment	15	mechanic	242	computer programming	1	attend detail	31	german	6
Top 3	inspect machinery	13	communication	101	identify need technological response	1	work efficiently	15	french	3
Top 4	conduct routine machinery check	12	pneumatic	94			work independently	12	danish	1
Top 5	use computerised maintenance management system	6	hydraulic	76			make decision	12	write german	1
Top 6	apply technical communication skill	4	automation technology	30			identify opportunity	6		
Top 7	conduct quality control analysis	3	electronic	18			manage time	6		
Top 8	maintain machinery	3	maintenance repair	13			make use personal robot practical support	4		
Top 9	provide technical documentation	3	engineering process	7			solve technical problem	3		
Top 10	advise efficiency improvement	2	mechatronic	7			develop strategy solve problem	3		

Tabelle 18: Top 10 nach Adaptierung Erweiterungsliste - 12.07.2021²⁷⁸

	Added	
Summe	1069	
Top 1	fault find	170
Top 2	continuous improvement	130
Top 3	preventative maintenance	94
Top 4	health safety	91
Top 5	ppm	78
Top 6	reactive maintenance	73
Top 7	install equipment	56
Top 8	plan maintenance	48
Top 9	report	40
Top 10	maintenance activity	37

Es wurden gesamt 2271 ESCO Kompetenzen in den 371 Stellenanzeigen gefunden. Die summierten Kompetenzen sind in Abbildung 25 als Balkendiagramm, gegliedert nach den Kompetenzästen, dargestellt. Dadurch wird gut ersichtlich, dass in diesen Stellenanzeigen primär nach Wissen und den zusätzlich aufgenommenen Kompetenzen nachgefragt wird.

²⁷⁷ Eigene Darstellung.

²⁷⁸ Ebd.

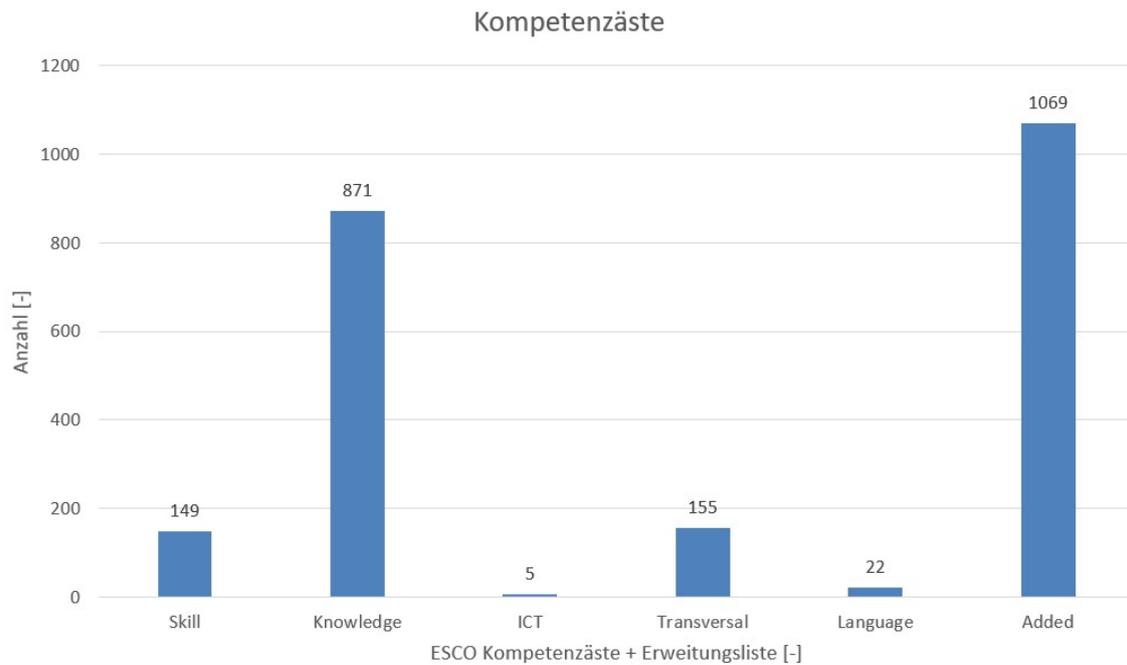


Abbildung 25: Kompetenzzäste Erweiterung – 12.07.2021²⁷⁹

Die Kompetenzverteilung wird in Abbildung 26 gezeigt. Im Schnitt werden 6,1 Kompetenzen je Stellenanzeige gefunden, jedoch wird in 24 Stellenanzeigen keine einzige zuordenbare Kompetenz gefunden.

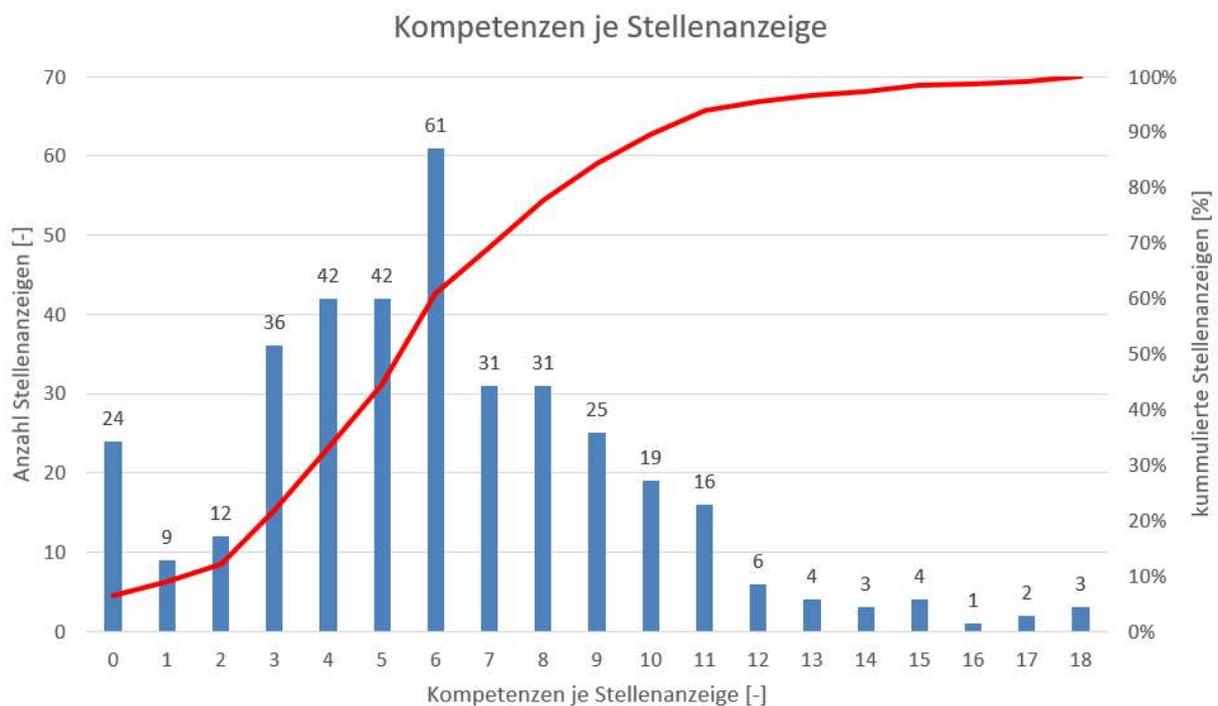


Abbildung 26: Kompetenzverteilung Erweiterung – 12.07.2021²⁸⁰

²⁷⁹ Eigene Darstellung.

²⁸⁰ Ebd.

5.1.1.5 Evaluierung nach Adaptierung vom 12.07.2021:

Die Evaluierung wurde wie in Kapitel 4.1.4 beschrieben durchgeführt. Dafür wurden fünf zufällig gewählte Stellenanzeigen aus dem Korpus manuell ausgewertet. Bei der manuellen Auswertung wurde unterschieden, ob es eine ähnliche Kompetenz im Standard gibt oder nicht.

Wie bereits beschrieben, war in 24 Stellenanzeigen keine automatisiert gefundene Kompetenz enthalten, in dieser Auswertung ist jedoch keine dieser berücksichtigt. In Tabelle 19 sind die Ergebnisse zusammengefasst. Die Details der Evaluierung bei welchen die Kompetenzen in den einzelnen Stellenanzeigen markiert sind, befindet sich im Anhang 2.

Tabelle 19: Evaluation – nach Adaptierung²⁸¹

Stellenanzeige / Kategorie	99	138	157	273	298
Kompetenzen ähnlich zum ESCO Standard	1	6	2	6	2
Kompetenzen aus Erweiterungsliste	4	3	2	12	4
Automatisiert gefundene Kompetenzen (Kompetenzen ähnlich zum ESCO Standard + Kompetenzen aus Erweiterungsliste)	5	9	4	18	6
Zusätzlich manuell identifizierte Kompetenzen	3	1	2	6	1
Manuell identifizierte Kompetenzen	8	10	6	24	7
Falsch positive automatisiert gefundene Kompetenzen	0	0	0	0	0
Automatisiert gefundene Kompetenzen	63%	90%	67%	75%	86%
Kompetenzen ähnlich zum ESCO Standard					
Falsch positive automatisiert gefundene Kompetenzen	-	0%	0%	0%	0%
Automatisiert gefundene Kompetenzen					

Im arithmetischen Mittel wurden rund 76% der enthaltenen Kompetenzen in den Stellenanzeigen automatisiert gefunden und gleichzeitig wurden keine falsch positiven Kompetenzen gefunden.

5.1.2 Topic Modeling

Für Darstellung der Ergebnisse für diese Methode wird nur ein Zeitpunkt (12.07.2021) gewählt. Gesamt wurden wieder 1228 Stellenanzeigen als Daten herangezogen, wobei zwei wegen der Sprache ausgeschlossen wurden und somit nur 1226 der Auswertung zugeführt wurden. Bei LDA Modellen ist es möglich das verwendete Vokabular einzuschränken oder das alle im Korpus enthaltenen Wörter als Vokabular herangezogen werden. Beide dieser Ansätze wurden umgesetzt und folglich werden die Ergebnisse vorgestellt.

²⁸¹ Eigene Darstellung.

Gleichbleibende Parameter sind die maximale Dokumentenfrequenz (80%), minimale Dokumentenvorkommen (mindestens in 3 Stellenanzeigen) und die Optimierung der Themenanzahl anhand von *perplexity* und *log. likelihood*.

5.1.2.1 Einschränkung auf ESCO Vokabular:

Wie in Kapitel 2.3.3.2 vorgestellt, ist es sinnvoll die Daten zuerst anhand der Worthäufigkeit, inversen Dokumentenhäufigkeit und deren Produkt (term frequency * inverse document frequency) zu betrachten. Die Summe der TFIDF für einzelne Wörter über den ganzen Korpus hinweg und die relative Betrachtung geben Aufschluss über die Wichtigkeit des Wortes in den einzelnen Dokumenten. Je höher der Anteil von $TFIDF/TF$ ist, desto wichtiger ist das Wort in den Dokumenten. Damit wird ersichtlich, dass die Kompetenzen „electrical engineering“, „mechanical engineering“ und „pneumatic“ eine höhere Wichtigkeit in den jeweiligen Dokumenten haben als die anderen Kompetenzen. In Tabelle 20 sind die Ergebnisse der Top 10 Standardkompetenzen, sortiert nach der Worthäufigkeit, zusammengefasst.

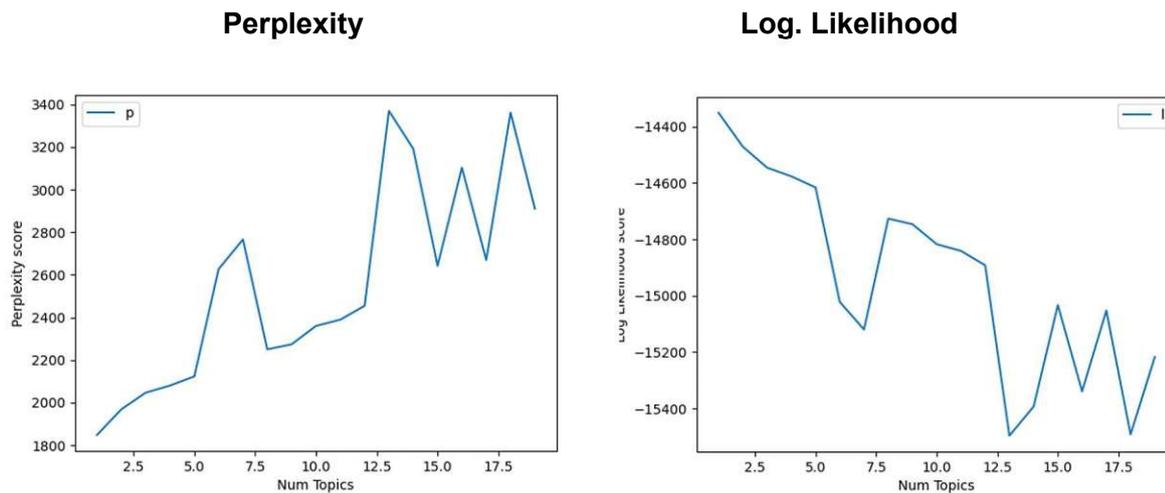
Tabelle 20: Top 10 - ESCO Vokabular²⁸²

term	summed_tfidf	summed_tfidf/tf	tf
communication	139,278	0,259	537
english	95,061	0,191	497
german	82,514	0,207	399
security	64,429	0,201	321
energy	50,355	0,24	210
electrical engineering	65,943	0,437	151
mechanical engineering	62,293	0,415	150
project management	41,069	0,314	131
pneumatic	47,451	0,413	115
account	30,83	0,283	109

Die Worthäufigkeit ist hier im Vergleich zu der Auswertung der Kompetenzanzahl höher, da hier die mehrfache Verwendung des Wortes in einer Stellenausschreibung auch mehrfach gezählt wird.

Das LDA Modell liefert nach der Optimierung, siehe Abbildung 27, jedoch nur ein Themengebiet. Dies steht im Widerspruch zum Ziel, welches mit einem LDA Modell erreicht werden sollte, Aufgliederung des Korpus in mehrere Themengebiete, deshalb kann mit Hilfe der Einschränkung des Vokabulars auf ESCO Standardkompetenzen keine weitere Zuordnung zu spezifischen Stellen erfolgen.

²⁸² Eigene Darstellung.

Abbildung 27: Optimierungsplots – ESCO Vokabular²⁸³

Das gefundene Themengebiet mit den Top 10 Standardkompetenzen ist der Vollständigkeit halber in Tabelle 21 angeführt. Des Weiteren würden die einzelnen Dokumente auch zu dem jeweiligen Thema zugeordnet, dies erübrigt sich jedoch, wenn nur ein Thema gefunden wird.

Tabelle 21: LDA Thema - ESCO Vokabular²⁸⁴

Kompetenz	Thema #0
Top 1	communication
Top 2	english
Top 3	german
Top 4	security
Top 5	electrical engineering
Top 6	mechanical engineering
Top 7	energy
Top 8	project management
Top 9	attention detail
Top 10	pneumatic

5.1.2.2 Ohne Einschränkung des Vokabulars:

Bei einer Auswertung ohne Einschränkung des Vokabulars müssen noch zwei weitere Parameter des LDA Modells festgelegt werden. Zum einen die maximale Anzahl an unterschiedlichen Wörtern bzw. Wortgruppen (*Features*), welches das Modell verwenden soll und zum anderen wie lange diese Wortgruppen mindesten und maximal sein dürfen. In dieser Auswertung wurden die *features* auf 2000 Stück und die Wortgruppenlänge zwischen einem Wort und bis zu vier Wörtern eingegrenzt. In Tabelle 22 sind die Ergebnisse bzgl. TFIDF und TF zusammengefasst.

²⁸³ Eigene Darstellung.

²⁸⁴ Ebd.

Tabelle 22: Top 10 sortiert nach tf- Ohne Einschränkung des Vokabulars²⁸⁵

term	summed_tfidf	summed_tfidf/tf	tf
team	94,478	0,039	2439
engineering	101,117	0,042	2396
working	89,489	0,043	2094
skills	77,234	0,042	1847
equipment	91,581	0,053	1734
engineer	104,382	0,061	1698
production	80,667	0,051	1588
electrical	95,433	0,061	1556
support	66,89	0,044	1513
systems	69,339	0,047	1468

Hier ist gut zu erkennen, dass das Wort „team“ im Schnitt beinahe zweimal in jeder Stellenanzeige vorkommt.

Auffällig ist, dass keine einzige Wortgruppe in den Top 10, sortiert nach der Worthäufigkeit, ist. Die erste Wortgruppe wäre „maintenance engineer“ auf Platz 24. Auch eine Auswertung bzgl. der Wichtigkeit in den einzelnen Dokumenten, sortiert nach summed_tfidf/tf, siehe Tabelle 23, liefert keinen besonderen Mehrwert.

Tabelle 23: Top 10 sortiert nach summed_tfidf/tf- Ohne Einschränkung des Vokabulars²⁸⁶

term	summed_tfidf	summed_tfidf/tf	tf
multiskilled	31,273	0,103	304
skilled	32,906	0,1	329
per	38,578	0,098	394
finding	30,305	0,098	308
technician	37,728	0,097	387
recruitment	34,451	0,097	355
client	31,521	0,097	326
preventative	30,996	0,097	321
hours	33,213	0,096	345
apprenticeship	28,775	0,096	301

Das LDA Modell liefert nach der Optimierung, siehe Abbildung 28, 16 Themengebiete. Der Optimierungsbereich wurde hier auf 50 Themengebiete erhöht, um sicherzustellen, dass es sich um ein globales Maximum bzw. Minimum handelt. Diese 16 Themengebiete könnten nun für weitere Zuordnung zu spezifischen Stellen interpretiert werden.

²⁸⁵ Eigene Darstellung.

²⁸⁶ Ebd.

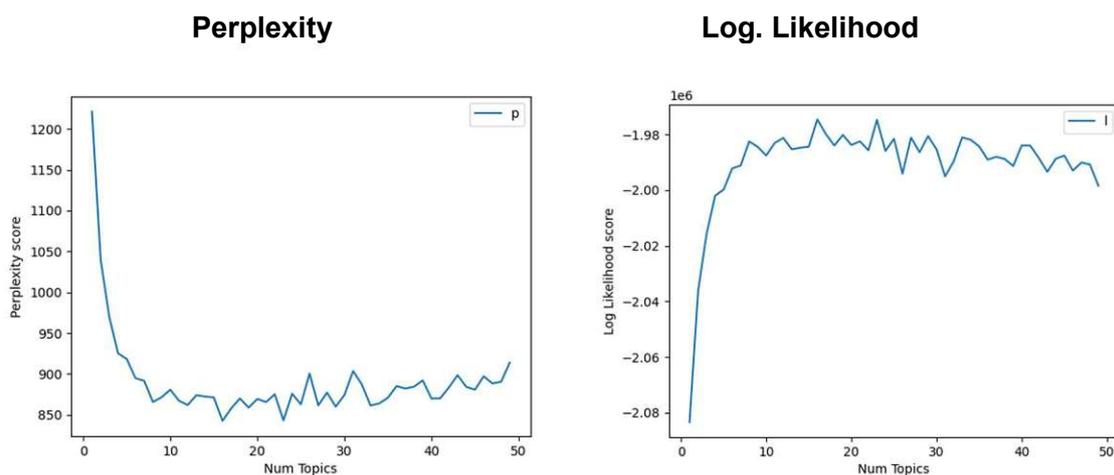


Abbildung 28: Optimierungsplots – Ohne Einschränkung des Vokabulars²⁸⁷

Die Anzahl der zugeordneten Stellenanzeigen zu den jeweiligen Themen sind in Tabelle 24 ersichtlich und die Top 10 Wörter des jeweiligen Themas werden in Tabelle 25 gezeigt.

Tabelle 24: Anzahl der zugeordneten Stellenanzeigen - Ohne Einschränkung des Vokabulars²⁸⁸

LDA	Anzahl der zugeordneten Stellenanzeigen
Thema 0	22
Thema 1	0
Thema 2	8
Thema 3	291
Thema 4	0
Thema 5	159
Thema 6	0
Thema 7	75
Thema 8	8
Thema 9	50
Thema 10	60
Thema 11	269
Thema 12	133
Thema 13	142
Thema 14	9
Thema 15	0

Die fünf Themen mit den meisten zugeordneten Stellenanzeigen (3, 5, 7, 11, 12 und 13) umfassen gesamt 1069 Stellenanzeigen. Diese Themen unterscheiden sich zwar zueinander in einem bestimmten Maß, jedoch lassen sich auch hier keine spezifischen Stellen erkennen.

²⁸⁷ Eigene Darstellung.

²⁸⁸ Ebd.

Tabelle 25: LDA Themen - Ohne Einschränkung des Vokabulars²⁸⁹

LDA	term - sortiert nach Wichtigkeit im Thema von links nach rechts (wichtigstes)
Thema 0	'growing', 'produce', 'take', 'working', 'performance', 'environment', 'infarm', 'production', 'food', 'team'
Thema 1	'going', 'improvements', 'unsolicited', 'acting', 'expected', 'first', 'resumes', 'der', 'und', 'konux'
Thema 2	'technology', 'technical', 'design', 'system', 'control', 'test', 'solutions', 'automation', 'security', 'systems'
Thema 3	'us', 'company', 'business', 'working', 'development', 'skills', 'customer', 'solutions', 'customers', 'team'
Thema 4	'operate', 'large scale', 'national', 'dynamic', 'least', 'pride', 'audit', 'manufacturing business', 'quality control', 'comprehensive'
Thema 5	'knowledge', 'team', 'requirements', 'process', 'business', 'project', 'product', 'development', 'quality', 'management'
Thema 6	'operate', 'large scale', 'national', 'dynamic', 'least', 'pride', 'audit', 'manufacturing business', 'quality control', 'comprehensive'
Thema 7	'site', 'improvement', 'working', 'production', 'plant', 'team', 'equipment', 'ensure', 'safety', 'engineering'
Thema 8	'must able', 'must', 'system', 'systems', 'space', 'aircraft', 'engineering', 'services', 'support', 'airbus'
Thema 9	'story', 'success story', 'team', 'success', 'production', 'teslas', 'tesla', 'process', 'engineering', 'equipment'
Thema 10	'join', 'company', 'pioneer', 'working', 'training', 'candidate', 'electrical', 'skills', 'maintenance engineer', 'engineer'
Thema 11	'maintenance engineer', 'mechanical', 'machinery', 'company', 'engineering', 'engineer', 'within', 'working', 'role', 'electrical'
Thema 12	'must', 'requirements', 'service', 'safety', 'technical', 'ability', 'systems', 'support', 'equipment', 'required'
Thema 13	'maintenance technician', 'engineering', 'pm', 'per', 'technician', 'working', 'electrical', 'mechanical', 'maintenance engineer', 'engineer'
Thema 14	'international ltd', 'skilled', 'correct', 'would', 'recruitment', 'successful candidate', 'check', 'international', 'food', 'candidate'
Thema 15	'calibration', 'high', 'excellent communication', 'aptitude', 'excellent communication skills', 'learn', 'production', 'spare parts', 'spare', 'parts'

Es wurde auch in Betracht gezogen die Anzahl der *features* für das LDA Modell zu variieren, jedoch lieferte weder eine Erhöhung noch ein Absenken der *feature*-Anzahl bzw. Anzahl der Themengebiete bessere Ergebnisse in Bezug auf die Zuordenbarkeit der Themen. Die Ergebnisse werden aus diesem Grund in dieser Arbeit nicht explizit dargestellt.

²⁸⁹ Eigene Darstellung.

5.2 Resultate in Bezug auf die Forschungsfragen

5.2.1 Kompetenzkatalog

Wie in Kapitel 5.1 gezeigt, lassen sich mit den angewandten Methoden keine Stellen spezifischen Kompetenzkataloge erstellen, weshalb hier der Kompetenzkataloge der Auswertungen aus Kapitel 5.1.1 als allgemeiner Kompetenzkatalog für die Instandhaltung herangezogen werden kann. Kompetenzen, welche in beiden Auswertungen in den Top 10 gelandet sind, sind in Tabelle 26 zusammengefasst und es ist davon auszugehen, dass diese einen hohen Stellenwert im Bereich der Instandhaltung haben.

Tabelle 26: Gleichbleibende Kompetenzen in den Top 10²⁹⁰

Skill	Knowledge	ICT	Transversal	Language
give grant	communication	computer programming	work independently	english
troubleshoot	communication science	solve technical problem	attend detail	german
promote	textile material		work team	french
operate control system	pneumatic		work shape space	dutch
breed stock	hydraulic		work efficiently	spanish
develop design plan	energy		identify opportunity	write english
management skill	electrical engineering		make decision	japanese
	mechanical engineering		meet commitment	italian
	electronic			

5.2.2 Trends

Anhand der beiden Durchläufe lassen sich Trends ableiten. Die Gesamtaufteilung der Kompetenzäste über diesen kurzen Betrachtungszeitraum bleibt beinahe gleich, siehe Abbildung 29. Wissen bleibt der primär nachgefragte Kompetenzast. Als Referenz dient Tabelle 10, welche die Anzahl an unterschiedlichen Kompetenzen je Kompetenzast zeigt. Es gibt im ESCO Standard 10199 Fertigkeiten (*Skill*) bezogen darauf, ca. $\frac{1}{4}$ so viele wissensbezogene Kompetenzen (2833), ca. $\frac{1}{500}$ so viele ICT Kompetenzen (21), ca. $\frac{1}{100}$ so viele transversale Kompetenzen (107) und $\frac{1}{33}$ so viele sprachenbezogene Kompetenzen (345).

²⁹⁰ Eigene Darstellung.

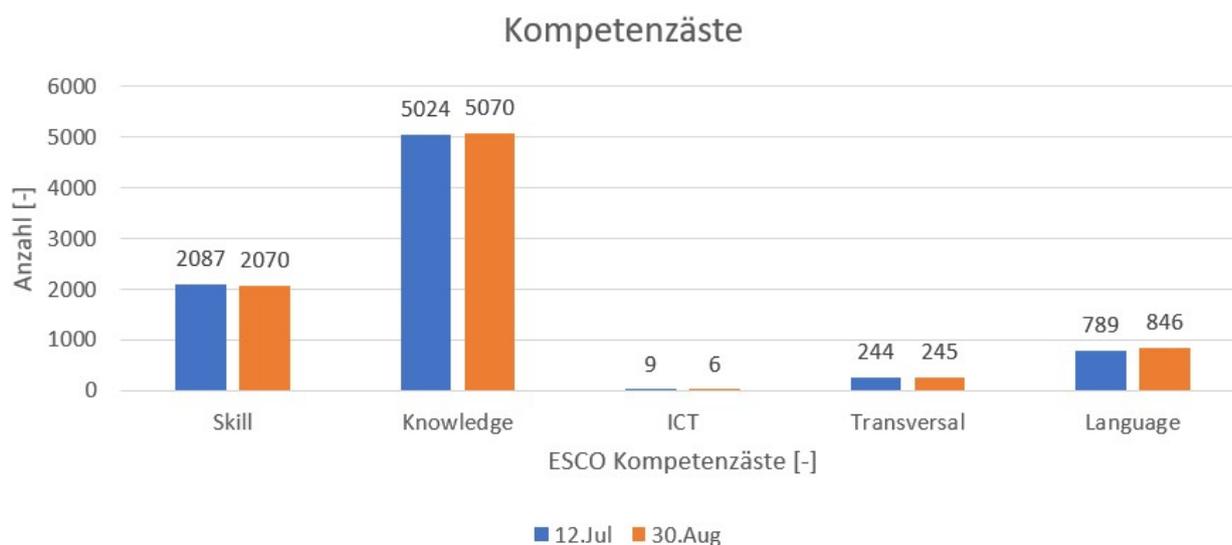


Abbildung 29: Kompetenzäste Vergleich beider Auswertungen²⁹¹

Bei den einzelnen Kompetenzen lassen sich kleinere Veränderungen in den Top 10 feststellen. Diese Veränderungen werden qualitativ in Tabelle 27 zusammengefasst.

Tabelle 27: Veränderungen bei den Top 10²⁹²

Zeitpunkt / Kompetenzast	12.07.2021	30.08.2021
Skill	select subject matter	application process
	work shift	provide documentation
Knowledge	c	model
	computer programming	adapt change
Transversal	manage time	solve technical problem
	danish	czech
Language	swedish	write german

Kompetenzen, welche in den Top 10 beider Auswertungen waren, ist in Abbildung 30 als Balkendiagramme je Kompetenzast dargestellt. Es lassen sich jedoch nur kleine quantitative Unterschiede zwischen den beiden Auswertungen ausmachen.

²⁹¹ Eigene Darstellung.

²⁹² Ebd.

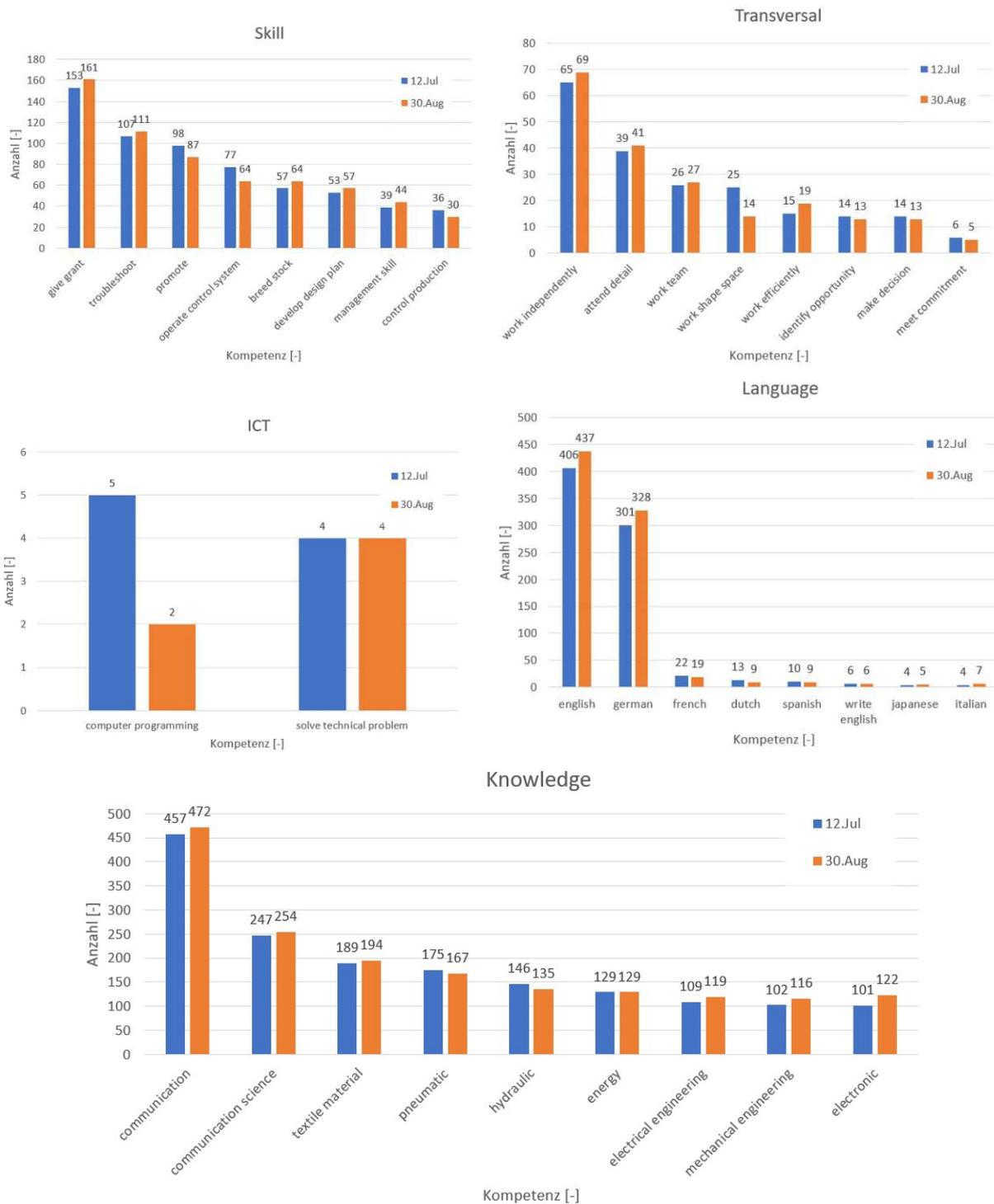


Abbildung 30: Kompetenzentwicklungen der gleichbleibenden Top 10 Skill, Transversal, ICT, Language und Knowledge²⁹³

²⁹³ Eigene Darstellung.

5.3 Möglicher Nutzen in Bezug auf die Problemstellung

Durch eine zeitlich längere Beobachtung der, in Kapitel 5.2 dargelegten, Resultate und Trends lassen sich mehrere nachfrageorientierte Maßnahmen in unterschiedlichen Bereichen ableiten.

Organisationen können sich einen Überblick über die nachgefragten Kompetenzen des Stellenmarktes verschaffen und somit ihre eigenen Bedürfnisse damit abgleichen. Dieser Abgleich kann dazu genutzt werden, um ihnen dabei zu helfen noch unbekannte Trends zu erkennen und spezifische Maßnahmen für die Organisation abzuleiten. Diese Maßnahmen können wiederum einen positiven Einfluss auf unterschiedliche KPIs, wie Weiterbildungskosten oder Weiterbildungsnutzen, haben.

Die strategische Ausrichtung der Einstellungspolitik von Organisationen kann besser an den Stellenmarkt angepasst werden. Diese Anpassung kann der Organisation dabei helfen, Personen mit den zukünftig wichtigen Kompetenzen einzustellen, um für die kommenden Herausforderungen bereits das richtige Personal angestellt zu haben. Dies könnte Einfluss auf KPIs wie Entlassungsquote oder Fluktuationsrate haben.

Sekundäre und tertiäre Bildungseinrichtungen können diese nachgefragten Kompetenzen nutzen, um ihre Curricula daran anzupassen. Diese Anpassung kann zum Beispiel durch die Aufnahme eines Themengebietes zu einzelnen bestehenden Lehrveranstaltungen oder durch neue Lehrveranstaltungen erfolgen. Diese Maßnahmen wiederum könnten den Absolventen und Absolventinnen der Bildungseinrichtungen bessere Einstiegsmöglichkeiten in Organisationen bieten.

6 Diskussion und Ausblick

In diesem Kapitel werden die Ergebnisse kritisch betrachtet, Einschränkungen des Ansatzes und der Ergebnisse aufgezeigt sowie Schritte für eine mögliche Weiterentwicklung aufgezeigt.

6.1 Diskussion der Ergebnisse

Im Vergleich zu den Kompetenzen, welche nach Krason et al. (2018) in Zukunft für die Instandhaltung im Zusammenhang mit der Industrie 4.0 wichtig sein werden, gibt es nur eine Überschneidung bei den Ergebnissen dieser Arbeit. Die Kompetenzen von Krason et al. (2018) umfassen:²⁹⁴

- Generelle IT-Kompetenz, um auftretende Probleme an spezialisierte Personen kommunizieren zu können
- Umgang mit 3D Simulationen
- Arbeiten mit zusätzlichen Informationsquellen (wie Augmented Reality)
- Umgang und Analyse von großen Datenmengen
- Kommunikative und soziale Kompetenzen

Nur die kommunikative Kompetenz wird sowohl von Krason et al. als auch von dieser Arbeit hervorgehoben.

Betrachtet man nicht nur die Kompetenzen, welche Krason et al. hervorheben, sondern auch die Themengebiete, welche im Zusammenhang mit Industrie 4.0 wichtig sind, gibt es dabei keine Überschneidung. Wichtige Themengebiete für die Instandhaltung im Kontext der Industrie 4.0 nach Krason et al. (2018) sind:²⁹⁵

- Internet of services
- Big data
- Data mining
- Internet of things
- Cyber physical systems

Dabei ist besonders auffallend, dass noch nicht alle dieser Themenfelder Teil des ESCO Standards sind. Zwar sind *internet of things*, *data mining* und *(analyze) big*

²⁹⁴ Vgl. Krason et al. (2018), S.°52.

²⁹⁵ Vgl. ebd., S.°50f.

data Teil des Standards, jedoch *internet of services* und *cyber physical systems* noch nicht explizit im ESCO Standard als Kompetenzen enthalten.²⁹⁶

Grund für dieses Ergebnis ist vermutlich die unterschiedliche Betrachtungsweise der Kompetenzen. Krason et al. (2018) leiteten aus bestehender Literatur Meinungen für die Zukunft ab. Im Gegensatz dazu werden in dieser Arbeit die derzeit nachgefragten Kompetenzen betrachtet.

6.2 Einschränkungen der Ansätze und Ergebnisse

In diesem Abschnitt werden die jeweiligen Einschränkungen und bestehenden Problematiken anhand der einzelnen Teilschritte der Umsetzung näher beleuchtet und dargelegt.

6.2.1 Dokumentenauswahl

In dieser Arbeit wurden nur Stellenanzeigen aus zwei Ländern und von einem online Stellenportal als Datenbasis herangezogen.

Bei der Dokumentenauswahl ist es mit den gegebenen Einschränkungsmöglichkeiten nicht möglich mit Bestimmtheit davon auszugehen, dass nur Stellen der Instandhaltung im produzierenden Gewerbe herangezogen werden. Es muss davon ausgegangen werden, dass auch Stellen, welche die verwendeten Stichwörter enthalten, in den Korpus aufgenommen werden, jedoch diese Stichwörter nicht unbedingt im verwendeten Kontext stehen. Als Beispiel könnte eine Stellenanzeige in der Buchhaltung gefunden werden, worin der Instandhaltungsbezug (Stichwort „*maintenance*“) über eine Dateninstandhaltung bestimmt wird und der Bezug zum produzierenden Gewerbe (Stichwort „*manufacturing*“) über die Organisationsbeschreibung hergestellt wird. Aus manueller Durchsicht des Datensets ging hervor, dass sich zwar nur in geringem Maße, aber auch Stellen für Softwarearchitekt_in, Softwareingenieur_in und Hardwareingenieur_in zu finden sind. Wobei es hier zu unterstreichen gilt, dass eine eindeutige Zuordnung der einzelnen Stellen nicht immer leicht ist. Ein Beispiel hierfür wäre ein_e Softwareingenieur_in, wessen primäre Aufgabe darin besteht, ein zustandsbasiertes Überwachungssystem aufzubauen und instand zu halten.

²⁹⁶ Vgl. Chiarello et al. (2021), S. 10.

Des Weiteren kann es sinnvoll sein, spezifischere Stellen, wie in der Adaptierung gezeigt, zu betrachten, um detaillierte Ergebnisse für eine Stelle zu erhalten.

6.2.2 Dokumentenaufbereitung

Da es für fast alle der verwendeten Programmpakete mehrere Alternativen gibt, könnte eine Verwendung von verschiedenen Kombinationen dieser zu besseren Ergebnissen führen. Gerade in Anbetracht der Vielzahl an unterschiedlichen verfügbaren *Lemmatizern* und *Stemmern* können die Ergebnisse durchaus beeinflusst werden.

6.2.3 Text Mining Methoden

6.2.3.1 Kompetenzanzahl

Mit diesem Ansatz ist eine Zuordnung der Kompetenzen zu einzelnen Stellen nicht möglich. Deshalb können daraus keine stellenspezifischen Kompetenzkataloge abgeleitet werden. Eine zeitliche Ableitung von Trends über den gesamten Stellenkorpus ist jedoch sehr wohl möglich.

In Bezug auf die Ergebnisse dieses Ansatzes ist in der Evaluierung der ersten Umsetzung, in Abschnitt 5.1.1.2, zu erkennen, dass mit der umgesetzten Synonymerkennung nicht alle Kompetenzen erkannt werden. Diese Probleme beim Erkennen der Kompetenzen kommen aus der Verwendung von anderen Worten in den Stellenanzeigen als im ESCO Standard. Diese können entweder zu generisch oder detailliert in den Stellenanzeigen vorkommen und somit vom ESCO Standard abweichen. Die große Anzahl an Kompetenzen des Standards birgt als Problem, falsch positive Ergebnisse zu erhalten, also Kompetenzen, die erkannt werden, jedoch sich in der gegebenen Verwendung nicht sinnhaft mit jener des Standards decken.

Ein Beispiel für die Probleme mit der Synonymerkennung ist die oft gefundene Kompetenz von „breed stock“, welche nach ESCO Definition im Zusammenhang mit dem Zuchtbestand von Käfern und Bienen steht. Hier muss davon ausgegangen werden, dass dies nur in den seltensten Fällen bei Stellen im Instandhaltungsbereich nachgefragt wird. Da das Wort „*stock*“ jedoch sehr oft im Korpus vorkommt und über die Synonymerkennung ein Synonym für das Wort „breed“ (Synonym „employee“) erkannt wird. Das Wort „breed“ kommt kein einziges Mal direkt im Korpus vor.

Durch die Adaptierung auf spezifischere Stellen und die Erweiterung um eine spezifische Stellenliste konnten mehrere dieser oben genannten Probleme besser bewältigt werden. Es wurde damit auch gezeigt, dass über die manuell erstellte Liste die Trefferquote erhöht werden konnte und es keine falsch positiven Treffer mehr in der Evaluierung gab. In Abschnitt 5.1.1.5 sind die Details dieser Evaluierung beschrieben.

6.2.3.2 Topic Modeling

Der Ansatz des *Topic Modelings* würde zwar eine Zuordnung von einzelnen Kompetenzen zu Stellen ermöglichen, als Beispiel sei hier auf die Arbeit von Pejic-Bach et al. (2020) verwiesen, jedoch ist dies in dieser Arbeit nicht möglich. Grund dafür könnte eine zu geringe Datenbasis in Kombination mit weniger geläufigen Schlagwörtern, wie sie bei Stellenausschreibungen anderer Branchen vorkommen, sein.

Eine Ableitung von zeitlichen Veränderungen der Kompetenzen wäre mit diesem Ansatz nur eingeschränkt möglich, da die Themen bei jedem Durchlauf neu geteilt werden und die zugehörigen Stellen immer wieder neu interpretiert werden müssen.

6.2.4 Anwendung

Bisweilen sind die Ergebnisse auf keine spezifische Anwendung angepasst und die Ergebnisse müssen manuell für eine Anwendung abgeleitet werden.

6.3 Weiterentwicklungsmöglichkeiten

In diesem Abschnitt werden die jeweiligen Weiterentwicklungsmöglichkeiten anhand der einzelnen Teilschritte der Umsetzung näher beleuchtet und dargelegt.

6.3.1 Dokumentenauswahl

Eine umfassendere Datenbasis könnte für globalere Ergebnisse sinnvoll sein. Dies würde umfassen, dass sowohl die Daten von mehreren online Stellenportalen und aus mehreren Ländern herangezogen werden.

Für spezifischere Ergebnisse wäre wiederum eine weitere Einschränkung zum Beispiel auf eine spezifische Stelle sinnvoll, wie es in der adaptierten Version umgesetzt ist.

Dies ist jedoch stark von der jeweilig gewollten Anwendung abhängig.

6.3.2 Dokumentenaufbereitung

In diesem Bereich gibt es für die jeweiligen Teilschritte eine Vielzahl an verfügbaren Programmpaketen und die Ergebnisse könnten durch unterschiedliche Kombinationen dieser positiv beeinflusst werden.

6.3.3 Text Mining Methoden

6.3.3.1 Kompetenzanzahl

In Bezug auf den Ansatz der Kompetenzanzahl sollte die Synonymerkennung verbessert werden. Die Verbesserung der Synonymerkennung könnte über mehrere Ansätze erfolgen. Zum einen könnte eine zusätzliche Integration einer kontextuellen Synonymerkennung sinnvoll sein, da so auch Kompetenzen erkannt werden können, bei welchen kein einziges Wort mit dem Standard übereinstimmt.

Wie in der Adaptierung gezeigt kann die Trefferquote über zusätzliche Erweiterungen positiv beeinflusst werden. Eine gute Weiterentwicklungsmöglichkeit wäre, diesen Prozess zum Beispiel mit Hilfe von Mustererkennung zu automatisieren, um den Arbeitsaufwand der manuellen Auswertung als Basis für die umgesetzte Erweiterung zu reduzieren.

6.3.3.2 Topic Modeling

Eine Ausweitung der Datenbasis könnte positiven Einfluss auf die Ergebnisse haben, vor allem in Anbetracht des Ansatzes des Topic Modelings.

6.3.4 Anwendung

Für die jeweilig spezifische Anwendung könnte ein abgestimmtes Dashboard mit verschiedenen Visualisierungen hilfreich sein, um die Vergleichbarkeit der Ergebnisse für den Anwender zu erleichtern.

7 Anhang

Anhang 1: Competence analysis – Data from 2021.07.12

Competencies similar to ones from ESCO Standard

Potential competencies **not** similar to ones from the ESCO Standard

Automatized found competences

Competencies which were found automatized and manually with fit to ESCO were only marked as automatized found ones and only considered in calculation, see Tabelle 12.

Job post 26

Competences found automatically (0)	additional competences found manually (6)
	repair and maintenance of printing and bookbinding machinery → maintenance of printing machines
	inspection of all ancillary equipment → manage inspections of equipment
	problem solving → solving problems
	installation and project work on equipment
	work safely
	own initiative

Job description

TJ Books is fortunate to be based in beautiful Padstow The company has been around for over years and continues to develop and grow enabling it to satisfy the needs of its customers both old and new Ever increasing demand means we are extending our operating hours and therefore we need to increase our Maintenance team to ensure we can cover our night shift too The purpose of this role is to help us in our mission to make good books quickly To ensure our product remains of the high standard we expect we need to ensure that our equipment remains in the best condition possible Alongside that we are bringing new equipment into the business to help safeguard our futureThe role will operate across four nights Main Duties include Daily repair and maintenance of printing and bookbinding machinery to acceptable company standards and conforming to all legal standardsDaily repair maintenance and inspection of all ancillary equipment associated with the site and buildings Supporting installation and project work on equipment associated with the printing and book binding trade plus ancillary equipment associated with the site and buildings Work safely and under own initiative whilst maintaining high standards of health and safety at all times Competencies Behaviours Previous engineering experience in a manufacturing or similar environment Positive approach with can do attitude Good problem solving abilities Job Types Fulltime Permanent Salary per year Benefits Bike to work scheme casual dress Company pension Life insurance Onsite parking Profit sharing Sick pay Store discounts Schedule Night shift COVID considerations All customers are required to wear masks and have temperature checks Sanitising stations are at numerous locations throughout the site and social distancing is maintained wherever possible Experience maintenance years preferred

Job post 419

Competences found automatically (1)	additional competences found manually (9)
Robotics	Preplanned machine maintenance
	Repair of breakdowns
	maintain maintenance logs
	Assist team in quality issues
	continuous improvement projects
	Coordinate installation of equipment
	electrical qualification
	review and analysis production reports
	organisational skills

Job description

At Osborne we are currently recruiting for a Maintenance Electrician to join a manufacturing company based in Ebbw Vale This is a fantastic opportunity for someone who wants to join an organisation that invests in their people and culture Providing a flat structure autonomy and trust is a big part of this role You will join a manufacturing environment that have a range of modern equipment across mechanical automation and **robotics** A high growth and innovative manufacturing environment this organisation delivers to their bluechip clients across multiple sectors For You Competitive salary day working week Permanent Role Training provided Career progression Parking on site Favourable shift times Responsibilities Preplanned machine maintenance Repair of breakdowns Ensure fixing machines in time to reduce downtime and increase productivity Input and maintain maintenance logs Assist team in quality issues raised Work on designs for new product lines Work on continuous improvement projects Coordinate installation of equipment Supporting facilities such as building maintenance and additional services Requirements Essential electrical qualification Ability to review and analysis production reports and high issues for resolution Excellent organisational skills To be considered for position apply today or call Emma on Job Type Fulltime

Job post 629

Competences found automatically (1)	additional competences found manually (2)
Upgrading equipment → climb equipment (wrong positive)	work alone and in a team
	reactive and planned and preventative maintenance

Job description

JOB TITLE Maintenance Engineer Electrical Bias SHIFT PATTERN Sunday midnight am Monday am Tuesday am Wednesday am Thursday am COMMUTABLE FROM Brough South Cave Market Weighton Cottingham Beverley Hull Willerby Hessle Hedon Preston Gilberdyke and Driffield SALARY AND BENEFITS Including shift allowance days annual leave bank holidays Private pension maximum company employee Life insurance Private Health Care Overtime is paid at x hourly rate including shift rate Monday to Friday Overtime is paid at x hourly rate including shift rate for the st hours on a Saturday and at x for all other hours COMPANY INFORMATION Our client is a leading manufacturer in the Humber area they have recently had a large amount of capital investment into the site to help increase production capacity YOU MUST HAVE THE FOLLOWING HNC ONC or Apprenticeship in a relevant engineering discipline Ideally have experience working in an FMCG Manufacturing environment Ability to work alone and in a team th th or th Edition ROLE INFORMATION Our client is currently expanding its production capability by introducing a new shift Ideally my client is looking to recruit for electrical bias Engineer and mechanical bias Engineer to join the Engineering team Recently my client has invested heavily into upgrading equipment on site so successful applicants will be working with modern machinery Duties will include both reactive and planned and preventative maintenance on a variety of production equipment carrying out any project upgrade work needed to the lines To apply for the Maintenance Engineer Electrical Bias role please send your CV along with your salary details quoting ref AC by email to Alistair Curran acurranrosjobs@rosinternational.co.uk RoS International Ltd Parliament Street Hull HU 9AP Tel 01482 424242 RoS International Ltd is a well established and respected Engineering and Technical Recruitment Company Since we have been placing staff

into both permanent and contract vacancies throughout the UK We often have such a large response to our advertisements that we are unable to provide feedback to every applicant If you do not receive an email from us within the next days your application has been unsuccessful We would like to keep your details on file for future vacancies unless you state otherwise SECTOR Manufacturing

Job post 885

Competences found automatically (6)	additional competences found manually (12)
Project management	Lean Management → Lean Project Management
Electronic	Msoffice → use microsoft office
English	SAP → SAP R3
French	work to strict deadlines → meet deadlines
Give grant	Team player → work as a team
Help our customers → assist customer	Procurement of resources
	Planning of capacity and investments
	Conduct of quality analysis
	process optimization
	Continuous improvement
	Conduct layout plannings
	Flexibility responsibility motivation Self assertion resilience → generic attitudes

Job description

Apex Tool Group is an equal opportunity employer All qualified applicants will receive consideration for employment without regard to race national origin religion age color sex sexual orientation gender identity disability or protected veteran status Process Engineer Assembly Westhausen Key Responsibilities Introduction and implementation of new products in assembling as project leader Process engineering of existing products and maintenance Procurement of resources with purchasing department Planning of capacity and investments Conduct of quality analysis and if necessary suggested solution with deviations Interface position between assembling RD and quality assurance Work on process optimization Finding costsaving opportunities Leading projects Product optimization with RD and quality department Continuous improvement in process of assembling and production Conduct Layout plannings Critical Skills Qualifications Bachelor or Master degree in process engineering production engineering or manufacturing engineering least years working experience in manufacturing engineering Experience in project management necessary Experience in Lean Management Good Msoffice and SAP skills necessary Good English skills French would be a plus Behavioral Skills Ability to work to strict deadlines Team player Flexibility responsibility motivation Selfassertion resilience At Apex Tool Group www.apextoolgroup.com we build innovation Each day our associates strive to find new and exciting ways to help our customers solve their most complex challenges By harnessing our global resources unprecedented insights and spirit for service we build more than just tools for the jobsite we help build the future ATG is a global manufacturer of hand and power tools tool storage and accessories chain and electronic soldering solutions with more than billion in annual revenues Our associates around the world have built powerhouse brands like GEARWRENCH SATA Crescent Cleco Weller and APEX Our brands are trusted by the professional trades and DIY enthusiasts alike With our roots dating back to the s our tools have driven technological advancements that drive efficiency speed and safety for customers in a broad range of commercial and consumer markets As part of the ATG team you will move fast think globally learn from your colleagues and grow your career Youll enjoy competitive benefits a healthy worklife balance and have opportunities to give back to the communities we serve Our six core values Customers Come First Integrity in All We Do Continuous Improvement Innovation for Growth Passion to Succeed and Best Talent One Team drive our daily decisions Connect with us on social media to learn more Facebook Instagram LinkedIn and Twitter If your goal is to work where finding a better way never ends and your ideas become reality join us and say WeBuildATG

Job post 902

Competences found automatically (10)	additional competences found manually (14)
Laser	Create go to market strategies → implement marketing strategies
English	Stay abreast of current market trends → perform market research
Promote	product life cycle → Punctuation problem
Communication	leadership → provide leadership or leading and motivating
Electrical engineering	presentation skills → prepare presentation skills
Project management	meet stated goals → meet commitments
Electronic	Support the sales teams
Management skills	Define product strategy
Pricing strategies	Make product decisions
Optics	Calculate sales prices
	Communicate with customers
	Report activities
	demonstrate a firm understanding of the scientific principles of optics fiberoptics
	Strong commercial drive

Job description

AMADA WELD TECH a Global company is part of the Amada group and a market leader in developing manufacturing and servicing systems products and components for Laser Welding Laser Marking Resistance Welding and Hot Bar Bonding We design and manufacture precision micro joining products and systems for a wide range of industries Our products and systems are typically low volume high value designed with the latest mechanical electronic and electrical technology available and range from standalone joining products to fully automated manufacturing solutions AMADA WELD TECH GmbH located in Munich is currently looking for a Product Manager Laser Technology who will be responsible for driving the sales growth of standard laser welding and laser marking products in Europe As Product Manager you have responsibility from initial concept through development market introduction technical support product life cycle and end of life This is a newly created position within our organisation Your responsibilities Support the sales teams with technical and application related information video telephone email joint customer visits etc Define product strategy including target markets segments and ideal customer profiles for product lines Release of new products to the market Make product decisions with respect to time features resources and quality to meet the business objectives Calculate sales prices and margin based on the cost price calculations from suppliers and affiliated companies group companies and communicate to the sales team Supporting quotations for standard laser products Create and manage pricing strategies including product support and maintenance pricing Drive the laser sales actively following up on opportunities with the sales teams Establish and promote European sales actions to increase laser product sales Communicate with customers with regards to custom components on project by project basis Maintain regular communication with our sales teams and provide high level sales support Regular communication with our AMADA WELD TECH group companies regarding new applications and products Report activities and performance on a monthly basis Create go to market strategies and plans develop clear product positioning Monitor and report new Laser applications or product releases from competitors Stay abreast of current market trends Your profile Academic qualification in laser technology Ability to demonstrate a firm understanding of the scientific principles of optics fiberoptics and lasers Strong commercial drive awareness and commitment are of particular importance to this role Relevant work experience in similar high technology focused product management Knowledge of managing a product life cycle Proven leadership and management skills with the ability to influence others Excellent presentation skills both developing and delivering content at customers engineers operators and senior business manager tradeshows conferences and sales meetings Ability to analyse and track the business to meet stated goals proactively identifying when to make changes and or develop creative alternatives and or innovative product ideas to achieve goals Strong communication skills written and oral Excellent project management skills ability to identify

interdependencies overcome obstacles and able to manage complex projects on time and deliver quality results Willing to travel in Europe Preferably living in the Munich region We offer Possibility for personal and professional development in a dynamic and international environment Interesting and challenging tasks in an open motivated team If you are interested in this position please send an EMail with your full application in **English** to our HR Manager Astrid van Vliet at humanresourcesamadaweldtecheu Your profile Academic qualification in **laser** technology Ability to demonstrate a firm understanding of the scientific principles of optics fiberoptics and **lasers** **Strong commercial drive** awareness and commitment are of particular importance to this role Relevant work experience in similar high technology focused product management Knowledge of managing a **product life cycle** Proven **leadership** and **management skills** with the ability to influence others Excellent **presentation skills** both developing and delivering content at customers engineers operators and senior business manager tradeshows conferences and sales meetings Ability to analyse and track the business to **meet stated goals** proactively identifying when to make changes and or develop creative alternatives and or innovative product ideas to achieve goals Strong **communication** skills written and oral Excellent **project management** skills ability to identify interdependencies overcome obstacles and able to manage complex projects on time and deliver quality results Willing to travel in Europe Preferably living in the Munich region

Anhang 2: Competence analysis after adaptation – Data from 2021.07.12

Competencies similar to ones from ESCO Standard

Potential competencies **not** similar to ones from the ESCO Standard

Automatized found competences

Competencies which were found automatized and manually with fit to ESCO were only marked as automatized found ones and only considered in calculation, see Tabelle 19.

Job post 99

Competences found automatically (5)	additional competences found manually (3)
electricity	preventative maintenance
reactive maintenance	maximum equipment availability
report	prevent repeat failure
support breakdown	
troubleshoot	

Job description

An exciting opportunity has become available for an **Electrical** Maintenance Engineer to be working for a heavy industry manufacturing business near Goole The successful Electrical Maintenance Engineer will **report** into the Maintenance Supervisor and will be responsible for preventative and **reactive maintenance** across the site Working Hours Continental shift pattern hour shifts days and nights Salary including shift allowance Overtime available Electrical Maintenance Engineer Duties and Responsibilities Include Planned corrective and preventative maintenance **Supporting breakdown** activities to ensure **maximum equipment availability** **Problem solving** activities to **prevent repeat failure** The successful candidate will need to demonstrate the following skills and experience Level or above in **Electrical** Engineering Preventative and reactive maintenance experience Worked in heavy industry manufacturing environment previously If you would like to know more about this position or any other Maintenance positions please contact Kerry Hill directly

Job post 138

Competences found automatically (9)	additional competences found manually (1)
troubleshoot	continually improve
electricity	
mechanic	
pneumatic	
solve technical problem	
manage time	
fault find	
plan maintenance	
maintenance activity	

Job description

MultiSkilled Maintenance Engineer Jobs Butler are now working with a leader in the field of plastic mouldings who have been established for almost years Situated in Brough East Riding of Yorkshire our client has gained a fantastic reputation nationwide for having the capability to produce a wide variety of complex components in an extensive range of materials Due to continued growth they are now looking to recruit a MultiSkilled Maintenance Engineer responsible for performing preventive **maintenance activities** and responding to breakdowns Often working in a clean room environment providing technical expertise to **fault find** maintain and **continually improve** automated injection moulding equipment ancillary equipment and assembly processes To be successful you will need engineering experience within a manufacturing environment with skills in fault finding and **problem solving on technical** equipment Experience with the following Phase **electrical** supply Controls and

drives Pneumatics Complex mechanical assemblies Planned maintenance software Injection Moulding machines Compressors Chillers Mould Heaters Temperature Controllers You will also need to be able to demonstrate a track record of technical competence working proactively and managing your own time effectively Ideally you will have completed a recognized Engineering apprenticeship or have a higher education qualification such as an HNC HND in a related subject What we can offer Remuneration package commensurate with qualifications and experience days holiday Bank holidays and sick pay scheme Applications will be treated in the strictest of confidence Reference ID MSME Job Type Fulltime Experience maintenance engineering years preferred Licence Certification HNC qualification preferred HND qualification preferred

Job post 157

Competences found automatically (4)	additional competences found manually (2)
troubleshoot	knowledge of drawings and schematics
electricity	keep the production plant and lines running
fault find	
maintenance activity	

Job description

Electrical maintenance engineer Location Sheffield Job Type Permanent Salary per year ANNUAL PAYMENT PENSION HOLIDAYS OVERTIME AVAILABLE GREAT PACKAGE Contact Asad Hussain Call Reference AHCF Role Electrical Maintenance Engineer Shifts Monday Friday Weekends Off Location Chesterfield South Yorkshire Salary Annual Payment Holidays Pension OT The Company We are currently working with a well established multisite manufacturing business that specialises in the building products sector This site has seen a huge investment in new machinery over the last few years and has an excellent reputation for looking after its employees which is clear to see in its excellent staff retention figures This means they don't recruit very often and this opportunity is quite rare They are now looking to add a new electrical maintenance engineer into their experienced team The Role Working a shift pattern of mornings afters and nights your main responsibility is to keep the production plant and lines running by performing planned maintenance activities and responding to emergency breakdowns There are quite a lot of machinery breakdowns on shift so this role would suit an electrical maintenance engineer who really enjoys fault finding and problem solving The plant is highly automated so you will be using PLCs to fault find Electrically you will also be working on invertors ACDC drives Motors relays and sensors There is a mix of new and old machinery to work on and you won't find this type of technology anywhere else Required Experience You will be Apprentice trained as an electrical maintenance engineer with an NVQ level or equivalent Excellent fault finding skills Experience within an automated manufacturing environment A good knowledge of drawings and schematics etc Benefits Very secure business where people like to stay long term Lots of investment in new equipment and machinery Opportunity to learn new skills and develop technically Well established and friendly maintenance team that work well together Interviews can be arranged immediately so apply by following the online instructions send a CV to asadhussain@atarecruitment.co.uk or call ATA Recruitment specialises in Manufacturing Infrastructure Civil Transportation and general Engineering recruitment on both a permanent and contract basis for more opportunities like this one visit our website By applying you accept the terms of our Privacy Notice which can be found on our website ATA Recruitment is committed to creating a diverse workforce and is an equal opportunities employer We welcome applications from all suitably qualified persons regardless of their race sex disability religion/belief sexual orientation or age ATA Recruitment specialise in recruiting the best technical and engineering talent and providing complete workforce solutions to help build and maintain infrastructure and transportation for a wide range of UK clients for more opportunities like this one visit our website By applying you accept the terms of our Privacy Notice which can be found on our website

Job post 273

Competences found automatically (18)	additional competences found manually (6)
troubleshoot	order and collect spares
electricity	use various types of testing equipment
mechanic	communicating
pneumatic	cooperate with supervision and fellow employees
work efficiently	call out duties in the event of a plant breakdown
develop strategy solve problem	carry out risk assessments
fault find	
continuous improvement	
preventative maintenance	
health safety	
install equipment	
plan maintenance	
report	
maintenance activity	
lean manufacturing	
plan work	
corrective action	
assist project work	

Job description

Electrical Maintenance Engineer Hartlebury Kidderminster DY HB We are currently recruiting for an Electrical Maintenance Engineer to be based at our Hartlebury site in Kidderminster Our Role This role will involve all aspects of trouble shooting **problem solving** and **project work** You will be required to look for and implement **continuous improvements** individually and as part of a team In addition you will also be required to plan and carry out **preventative maintenance** in order to ensure a high standard of plant maintenance is always achieved You will be required to assist in all **maintenance activities** on site to ensure factory production budgets are met or exceeded The successful candidate will be working as part of the maintenance team ensuring that all **electrical** plant equipment is **operating efficiently** with the minimum amount of downtime and in a safe order and will be required to contribute to the application of s and other **lean manufacturing** principles within the factory You will also be required to adhere to strict **health and safety** requirements at all times What Youll Do Install and or maintain all electrical electromechanical plant or machinery in line with Company requirements Operate a preventative electrical maintenance system to ensure that the plant machinery are in good working order **Fault find** and repair any electrical or electromechanical related problems on the plant or equipment Program all Siemens other PLC equipment **Use various types of testing equipment** in order to locate and remedy faults in the shortest possible time Ensure that all plant and machinery is kept in a good working and clean condition Assist with the **installation of new plant equipment** **Order and collect spares** and electrical materials for new and existing plant in keeping with the required standards Maintain and monitor the machines and equipment to enable them to achieve their budgeted outputs at the required quality levels Ensure that your workstores area is kept clean and free from obstructions and waste maintaining a good standard of general housekeeping Complete all relevant paperwork associated with the task you are performing in a timely and accurate manner **Report** any problems regarding equipment and machinery immediately to the Team Leader on shift Ensure that you are aware of the Company Health Safety and Environmental policies and procedures and that you adhere to them at all times Maintain a positive work atmosphere by acting and **communicating** in a manner that promotes cooperation with colleagues Promote Lean thinking and actively look for ways to improve your efficiency Take care of equipment used while performing the job **Cooperate with supervision and fellow employees** Actively take part in any training requested to enhance your role Safely carry out all manual handling in the correct and trained manner Work safely and follow all safety rules at all times Perform other duties as required by your line manager Accurately report on the work that you have carried out on the relevant paperwork Constantly watch for wear and tear on machinery and correct or inform the line manager and report on the relevant paperwork Be available for **call out duties in the event of a plant breakdown** Ensure you have a very good working knowledge of all plant equipment and machinery Ensure that you are adequately trained

to run the equipment after servicing or maintenance in order that the production personnel are able to start run the equipment without further modifications or adjustments Actively contribute to the creation and review of RA and SSOW Promote the quality of our products at all times Be aware of the effect that maintenance changes can have on product quality Ensure adherence to Wienerbergers quality system standards in accordance with BS EN ISO Ensure that all quality issues are dealt with immediately as per agreed procedures Ensure conformity and safe working practices in accordance with the Companys health safety and environment procedures Carry out regular housekeeping spot checks Report any accidents immediately to your line manager and take **corrective actions** Carry out **risk assessments** and liaise with HSE department where necessary Attend Health Safety meetings if required Be a health safety advocate for the Company promoting health safety at all times Ensure availability and use of Personal Protective Equipment PPE in the appropriate areas Refer any concerns near misses to the EHS Manager What Youll Bring Electrical engineering or similar apprenticeship with relevant experience is essential ONCBTEC HNCBTEC or equivalent in Electrical engineering th edition qualified Experience of working in a manufacturing environment is essential ideally heavy manufacturing clay refractory quarrying cement or FMCG Good communicator with team skills Sound knowledge of Siemens PLCs s TIA portal Excellent analytical fault finding problem solving skills Good knowledge of electric motors **pneumatics** Good sound **mechanical** skills Working within a multi skilled team Experience of working in a SafetyFirst culture Ability to work on ones own and make necessary judgments Hours of Work am pm pm am rotating hour shifts days and nights on a on off shift pattern What Youll Get Company Pension Cycle to Work Scheme Discounts cash back from various high street online retailers Health Cash Plan Life Assurance Payroll Giving SIP Share Scheme Tech Scheme Winners Club Meet our People [httpbitlyWBMeetOurPeople](http://bitly/WBMeetOurPeople) What we are Built on [httpbitlyWBMadeonBricks](http://bitly/WBMadeonBricks)

Job post 298

Competences found automatically (6)	additional competences found manually
electricity	machinery installations and commissioning
mechanic	
fault find	
reactive maintenance	
plan maintenance	
plan work	

Job description

We are currently recruiting for a experienced **electrical** maintenance engineer to join a world class manufacturing facility based near Chessington This is a Automated factory who are seeking a electrical maintenance engineer on a permanent basis The electrical maintenance engineer will be involved in all site planned and **reactive maintenance** including **project work** such as **machinery installations and commissioning** This is a fantastic opportunity for a electrical maintenance engineer to join a very well established company who are offering a long term career Skills required for the Maintenance Engineer Electrical Engineering Qualification Experience working in **manufacturing environment** Electrical experience with Inverters Drives PLCs Automation Motors **Fault Finding** Electrical experience with AC DC systems relays contactors transformers **Mechanical** hands on experience Proactive attitude and willing to learn new skills The Maintenance Engineer will benefit from Competitive benefits and pension package Training and development Chances to progress into senior engineering roles Benefits Pension Healthcare Training Holidays If you are interested in this role and feel that you have the right skills then please click apply at the bottom of this advert As a registered candidate with Pioneer Selection Ltd you automatically become eligible for our referral scheme You will receive for every candidate we place in permanent employment who has been recommended by you

Tabelle 28: Systematische Literaturrecherche - Großformat²⁹⁷

	Search Term_ID	Language	Date of search	Used Search-Term (as entered in database)	Set search filters	Number of results after first	Search strategy	examined papers	Number of papers moved into	Name of folder Y: SLR_XX (Search Term_ID)	Number of papers cannot be downloaded
Google Scholar https://scholar.google.com/	SLR_01	EN	27.11.2020	"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO")	Search in: Any field Time: 2010 - Present exclude citations and patents	11	Screening until result 11	5	5	SLR_01	0
	SLR_02			"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		12	Screening until result 12	5	5	SLR_02	0
	SLR_03			"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		630	Screening until result 60	4	4	SLR_03	0
	SLR_04			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		831	Screening until result 70	4	4	SLR_04	0
	SLR_05		"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")	31		Screening until result 31	10	10	SLR_05	0	
	SLR_06		"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")	38		Screening until result 4	20	15	SLR_06	5	
Science Direct https://www.sciencedirect.com/	SLR_07	EN	03.1.2022	"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO")	Search in: Any field Time: 2010 - Present	2	Screening until result 2	1	1	SLR_07	0
	SLR_08			"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		4	Screening until result 4	2	2	SLR_08	0
	SLR_09			"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		48	Screening until result 48	2	2	SLR_09	0
	SLR_10			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		100	Screening until result 50	3	3	SLR_10	0
	SLR_11			"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		5	Screening until result 5	2	2	SLR_11	0
	SLR_12			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		8	Screening until result 8	2	2	SLR_12	0
Springer Link https://link.springer.com/	SLR_13	EN	04.1.2022	"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO")	Search in: Any field Time: 2010 - Present exclude preview-only content	1	Screening until result 1	0	0	SLR_13	0
	SLR_14			"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	0	0	SLR_14	0
	SLR_15			"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		18	Screening until result 18	0	0	SLR_15	0
	SLR_16			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		34	Screening until result 34	0	0	SLR_16	0
	SLR_17			"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	0	0	SLR_17	0
	SLR_18			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		2	Screening until result 2	0	0	SLR_18	0
Taylor and Francis https://www.tandfonline.com/	SLR_19	EN	05.1.2022	"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")	Search in: Any field Time: 2010 - Present	5	Screening until result 5	0	0	SLR_19	0
	SLR_20			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		18	Screening until result 18	0	0	SLR_20	0
	SLR_21			"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	1	0	SLR_21	1
	SLR_22			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	1	0	SLR_22	1
Emerald Insight https://www.emerald.com/insight/	SLR_23	EN	06.1.2022	"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO")	Search in: Any field Time: 2010 - Present	9	Screening until result 9	3	1	SLR_23	2
	SLR_24			"Text Mining" AND "Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		10	Screening until result 10	3	1	SLR_24	2
	SLR_25			"Job advertisement" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		101	Screening until result 50	1	0	SLR_25	1
	SLR_26			"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		100	Screening until result 50	5	1	SLR_26	4
	SLR_27			"Text Mining" AND "Job advertisement" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")		13	Screening until result 13	3	1	SLR_27	2
	SLR_28			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		27	Screening until result 27	3	1	SLR_28	2
Scopus https://www.scopus.com/home.uri	SLR_29	EN	07.1.2022	"Text Mining" AND "Employee" AND "Maintenance" AND ("Competence" OR "KSAO" OR "Competency")	Search in: Any field Time: 2010 - Present	1	Screening until result 1	1	0	SLR_29	1
	SLR_30			"Text Mining" AND "Job advertisement" AND "Employee" AND ("Competence" OR "KSAO" OR "Competency")		1	Screening until result 1	0	0	SLR_30	0

²⁹⁷ Eigene Darstellung.

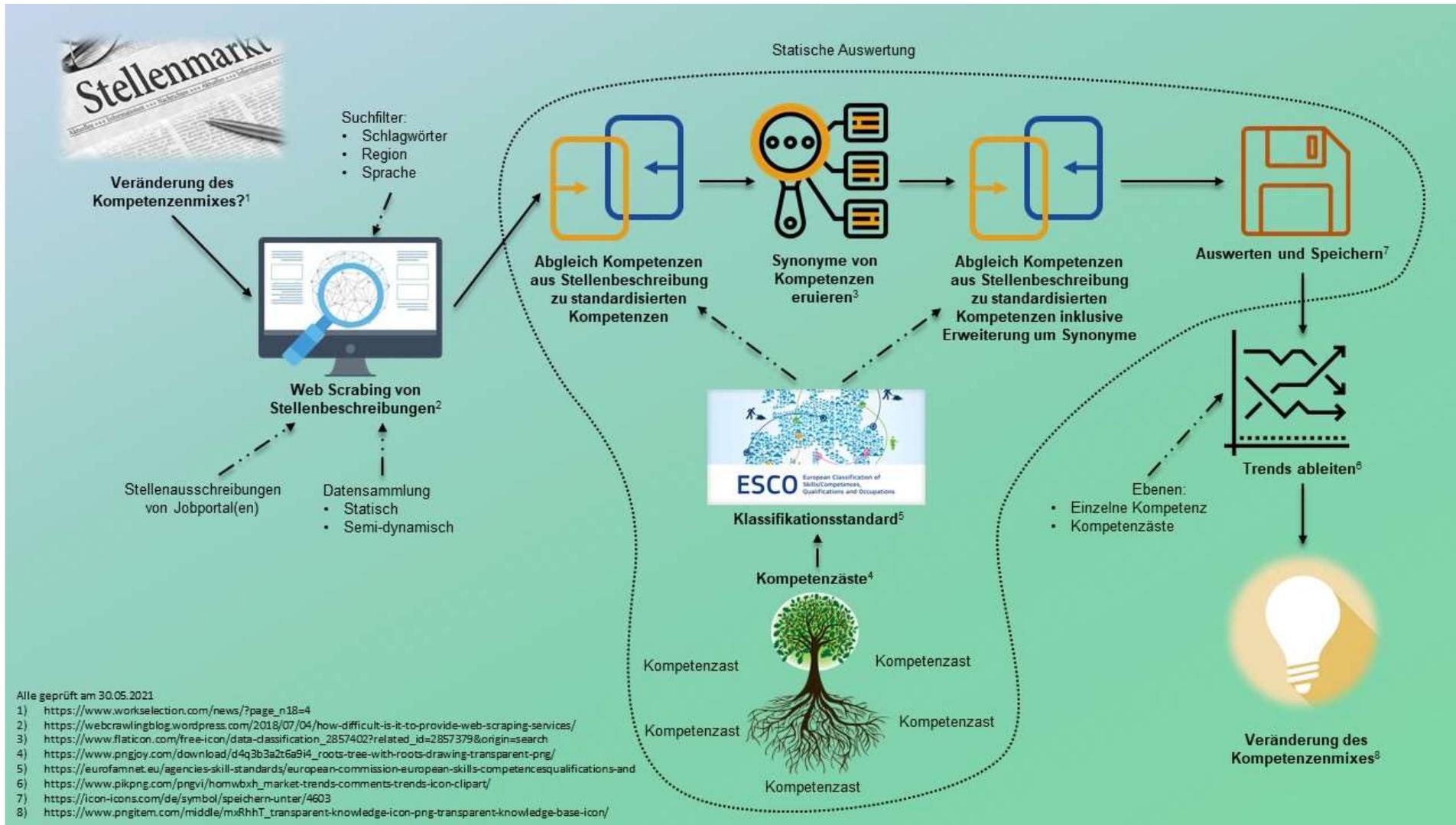


Abbildung 31: Abstraktes Konzept der Umsetzung - Großformat²⁹⁸

²⁹⁸ Eigene Darstellung.

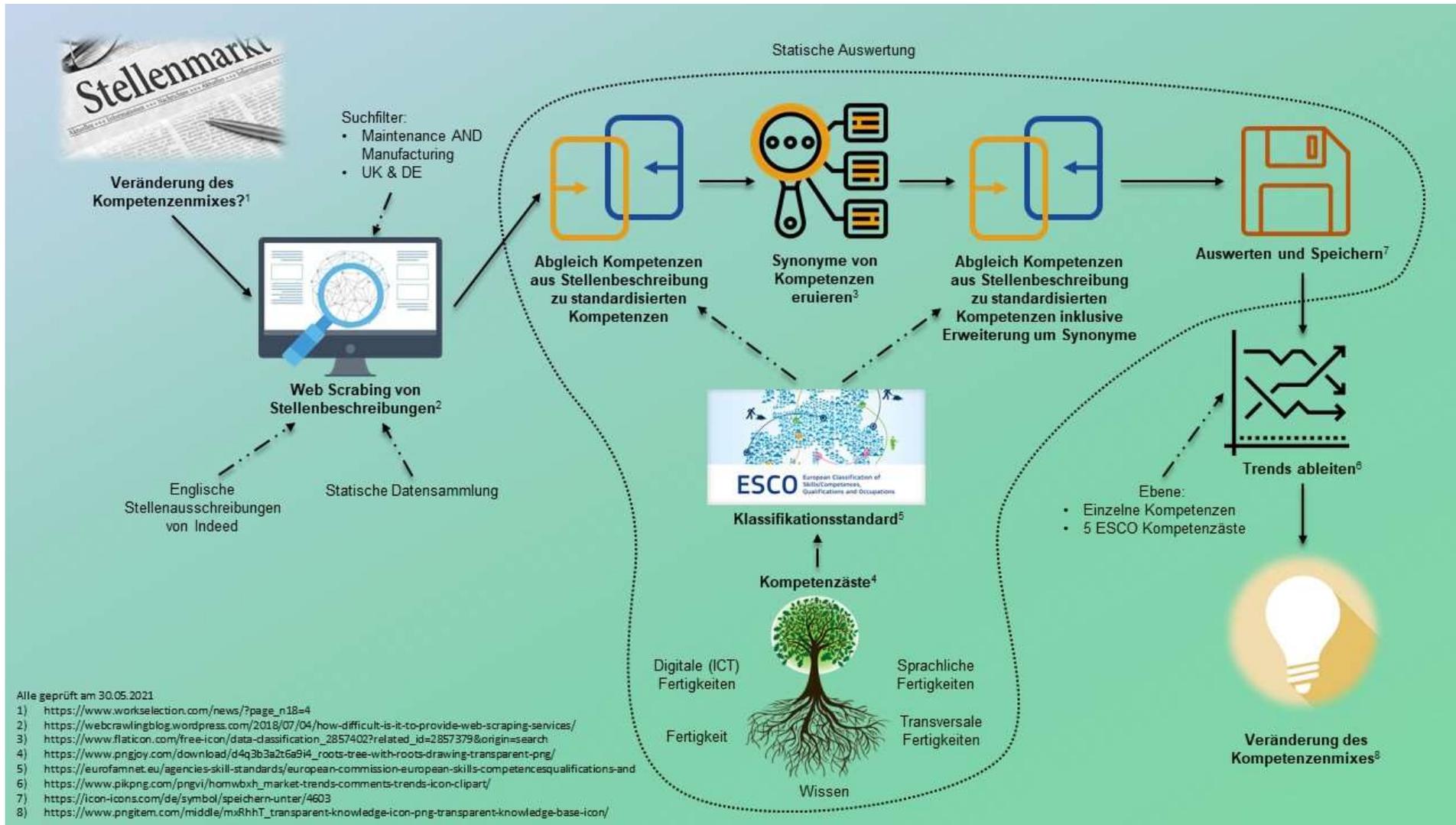


Abbildung 32: Spezifisches Konzept der Umsetzung - Großformat²⁹⁹

²⁹⁹ Eigene Darstellung.

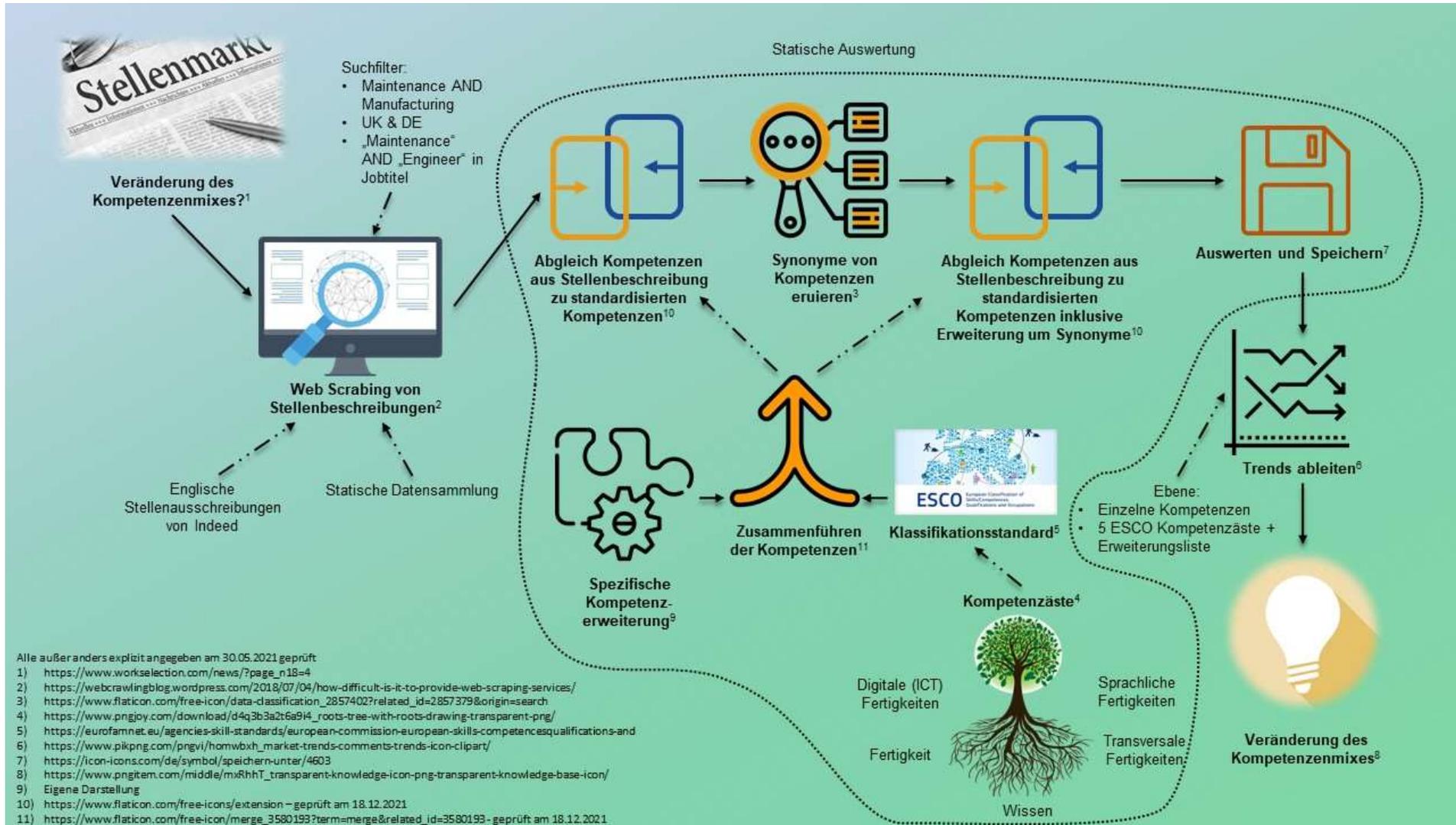


Abbildung 33: Spezifisches Konzept der Umsetzung nach Adaptierung- Großformat³⁰⁰

³⁰⁰ Eigene Darstellung.

8 Literaturverzeichnis

About O*NET at O*NET Resource Center (2021). URL:

<https://www.onetcenter.org/overview.html>, Stand: 19.11.2021.

Adams, Jack A. (1987): Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. In: Psychological bulletin 101 (1), S. 41–74.

Algarni, Abdulmohsen; Tairan, Nasser (2014 - 2014): Feature Selection and Term Weighting. In: 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT). 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT). Warsaw, Poland, 11.08.2014 - 14.08.2014. IEEE, S. 336–339.

Arbeitsmarktservice : Berufsinformationssystem. Hg. v. Arbeitsmarktservice - Dienstleistungsunternehmen des öffentlichen Rechts. URL:

<https://www.ams.at/bis/bis/>, Stand: 19.10.2009.

b.telligent (Hg.) : Evolution Industrie 4.0. Unter Mitarbeit von Sebastian Amtage.

URL: <https://www.btelligent.com/themen/industrie-40/>, Stand: 29.08.2020.

Baladi, Peter (1999): Knowledge and Competence Management: Ericsson Business Consulting. In: Business Strategy Review 10 (4), S. 20–28. DOI: 10.1111/1467-8616.00116.

Banerjee, Arindam; Basu, Sugato (2007): Topic Models over Text Streams: A Study of Batch and Online Unsupervised Learning. In: Proceedings of the 2007 SIAM International Conference on Data Mining. Philadelphia, PA. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Batini, Carlo; Cappiello, Cinzia; Francalanci, Chiara; Maurino, Andrea (2009): Methodologies for data quality assessment and improvement. In: ACM Comput. Surv. 41 (3), S. 1–52. DOI: 10.1145/1541880.1541883.

Bauernhansl, Thomas; Krüger, Jörg; Reinhart, Gunther; Schuh, Günther (2016): WGP-Standpunkt Industrie 4.0. In: Wissenschaftliche Gesellschaft für Produktionstechnik Wgp e. v., Stand: 27.07.2020.

Beautiful Soup (2021): Beautiful Soup. URL: <https://beautiful-soup-4.readthedocs.io/en/latest/#>, Stand: 04.09.2021.

- Bergenhengouwen, G. J.** (1996): Competence development - a challenge for HRM professionals: core competences of organizations as guidelines for the development of employees. In: *Jnl Euro Industrial Training* 20 (9), S. 29–35. DOI: 10.1108/03090599610150282.
- Berio, Giuseppe; Harzallah, Mounira** (2005): Knowledge management for competence management. In: *Journal of Universal Knowledge Management* 1, S. 21–28.
- Berry, John** (2011): Competency Model for IT Program Management. Hg. v. Chief Human Capital Officers Council. United States Office of Personnel Management. URL: <https://www.chcoc.gov/print/2613>, Stand: 13.12.2020.
- BERT** (2021): BERT (language model). URL: [https://en.wikipedia.org/w/index.php?title=BERT_\(language_model\)&oldid=1039012807](https://en.wikipedia.org/w/index.php?title=BERT_(language_model)&oldid=1039012807), Stand: 28.08.2021.
- Blei, David M.; Lafferty, John D.** (2007): A correlated topic model of science. In: *The Annals of Applied Statistics* 1 (1), S. 17–35.
- Blei, David M.; Ng, Andrew Y.; Jordan, Michael I.** (2003): Latent dirichlet allocation. In: *Journal of machine learning research* 3 (Jan), S. 993–1022.
- Bohlouli, Mahdi; Ansari, Fazel; Fathi, Madjid** (Hg.) (2012): Design and realization of competence profiling tool for effective selection of professionals in maintenance management. IEEE: IEEE.
- Bohnacker, Ulrich; Dehning, Lars; Franke, Jürgen; Renz, Ingrid** (2002): Textual Analysis of Customer Statements for Quality Control and Help Desk Support. In: H.-H. Bock, W. Gaul, M. Schader, Krzysztof Jajuga et al. (Hg.): *Classification, Clustering, and Data Analysis (Studies in Classification, Data Analysis, and Knowledge Organization)*. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 437–445.
- Boyatzis, Richard E.** (1982): *The competent manager. A model for effective performance*. John Wiley & Sons.
- Braconier, Henrik; Nicoletti, Giuseppe; Westmore, Ben** (2014): Policy challenges for the next 50 years.
- Camacho-Collados, Jose; Pilehvar, Mohammad Taher** (2017): On the role of text preprocessing in neural network architectures: An evaluation study on text categorization and sentiment analysis. In: *arXiv preprint arXiv:1707.01780*, S. 1–7.

- Cedefop** (2018): Insights into skill shortages and skill mismatch. Learning from Cedefop's European skills and jobs survey (Cedefop reference series, 106). Luxembourg: Publications Office of the European Union.
- Chiarello, Filippo; Fantoni, Gualtiero; Hogarth, Terence; Giordano, Vito et al.** (2021): Towards ESCO 4.0 – Is the European classification of skills in line with Industry 4.0? A text mining approach. In: Technological Forecasting and Social Change 173, S. 121177. DOI: 10.1016/j.techfore.2021.121177.
- Chung, Chih-Hung; Chen, Lu-Jia** (2019): Text mining for human resources competencies: Taiwan example. In: EJTD ahead-of-print (ahead-of-print). DOI: 10.1108/EJTD-07-2018-0060.
- Clifford, James P.** (1994): Job Analysis. Why Do it, and How Should it Be Done? In: Public Personnel Management 23 (2), S. 321–340. DOI: 10.1177/009102609402300211.
- Cook, T. W.** (1934): Studies in cross education. III. Kinaesthetic learning of an irregular pattern. In: Journal of Experimental Psychology 17 (5), S. 749–762. DOI: 10.1037/h0075917.
- Corallo, Angelo; Lazoi, Mariangela; Margherita, Alessandro; Scalvenzi, Massimo** (2010): Optimizing competence management processes. A case study in the aerospace industry. In: Business Process Mgmt Journal 16 (2), S. 297–314. DOI: 10.1108/14637151011035615.
- David Stark** (2010): ISMI Predictive Preventive Maintenance (PPM) Data Requirements. Technology Transfer #10105120A-TR. In: International SEMATECH Manufacturing Initiative, Stand: 27.07.2020.
- Deutsches Normungsinstitut - DIN 31051, 2019-06:** Grundlagen der Instandhaltung. Berlin: Beuth Verlag GmbH. DOI: 10.31030/3048531.
- Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton; Toutanova, Kristina** (2018): BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. URL: <https://arxiv.org/pdf/1810.04805>.
- Devopedia** (2019): BERT (Language Model). In: Devopedia Foundation. URL: <https://devopedia.org/bert-language-model>, Stand: 28.08.2021.
- Dhillon, B. S.** (2002): Engineering maintenance. A modern approach. Boca Raton, FL: CRC Press. URL: <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10143275>.

- Dhillon, Balbir S.** (1986): Human reliability: with human factors. Elsevier.
- Directorate-General for Employment, Social Affairs and Inclusion** (2017): ESCO handbook. European skills, competences, qualifications and occupations. First edition. Luxembourg: Publication Office of the European Union. DOI: 10.2767/934956.
- Djumalieva, Jyldyz; Sleeman, Cath** (2018): An open and data-driven taxonomy of skills extracted from online job adverts. In: Developing Skills in a Changing World of Work: Concepts, Measurement and Data Applied in Regional and Local Labour Market Monitoring Across Europe 425.
- Dudenredaktion** (o.J.): Industrie. Duden online. URL: <https://www.duden.de/node/70642/revision/70678>, Stand: 28.08.2020.
- Eigene Darstellung** : Eigene Darstellung.
- Elkin, Graham** (1990): Competency based human resource development. Making sense of the ideas.
- Ellington, J. Kemp; Brannick, Michael T.; Levine, Edward L.** (2015): Job Specification. In: Wiley Encyclopedia of Management.
- Empl, Dg** (NaN): ESCO - ESCOpedia - European Commission. URL: https://ec.europa.eu/esco/portal/escopedia/Main_Page, Stand: 19.11.2021.
- Empl, Dg** (NaN): ESCO - Skills/competences - European Commission. URL: <https://ec.europa.eu/esco/portal/skill>, Stand: 19.11.2021.
- Empl, Dg** (2021): ESCO - Occupations - European Commission. URL: <http://data.europa.eu/esco/occupation/615920c5-4f63-4eb3-8b60-afaaed3ab1ff>, Stand: 30.12.2021.
- Employment and Training Administration** (Hg.) : Building Blocks for Competency Models. United States Department of Labor. URL: <http://www.melvinhenry.org/generic-model.pdf>, Stand: 28.09.2020.
- Eppler, Martin J.** (2006): Managing Information Quality. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/3-540-32225-6.
- Fleishman, Edwin A.; Wetrogan, Leon I.; Uhlman, Charles E.; Marshall-Mies, Joanne C.** (1995): Development of Prototype Occupational Information Network(0*NET) Content Model. Volume I: Report [and] Volume II:Appendices.Content Model (09).

- Forman, George** (2003): An extensive empirical study of feature selection metrics for text classification. In: Journal of machine learning research 3 (Mar), S. 1289–1305.
- Fuchs-Kittowski, Klaus** (2001): Organisationsinformatik und Digitale Bibliothek in der Wissenschaft. 1. Aufl. (Wissenschaftsforschung).
- Geissbauer, Reinhard; Schrauf, Stefan; Koch, Volkmar; Simon, Kuge** (2014): Industry 4.0 – Opportunities and Challenges of the Industrial Internet.
- Ghosh, Sayantani; Roy, Sudipta; Bandyopadhyay, Samir K.** (2012): A tutorial review on Text Mining Algorithms. In: International Journal of Advanced Research in Computer and Communication Engineering 1 (4), S. 7.
- GitHub** (2021): GitHub - google-research/bert: TensorFlow code and pre-trained models for BERT. URL: <https://github.com/google-research/bert>, Stand: 28.08.2021.
- GitHub** (2021): langdetect/detector_factory.py at master · Mimino666/langdetect. URL: https://github.com/Mimino666/langdetect/blob/master/langdetect/detector_factory.py, Stand: 04.09.2021.
- Grimes, Seth** (2008): Unstructured Data and the 80 Percent Rule. Hg. v. Breakthrough Analysis. URL: <http://breakthroughanalysis.com/2008/08/01/unstructured-data-and-the-80-percent-rule/>, Stand: 24.10.2020.
- Guzmán, Valeria E.; Muschard, Bernd; Gerolamo, Mateus; Kohl, Holger et al.** (2020): Characteristics and Skills of Leadership in the Context of Industry 4.0. In: Procedia Manufacturing 43, S. 543–550. DOI: 10.1016/j.promfg.2020.02.167.
- Hartl, Karin; Jacob, Olaf** (2015): Determining the Business Value of Business Intelligence with Data Mining Methods, S. 1–5.
- Heilmann, Pia; Heilmann, Jouni** (2011): Competence management in maintenance. Case – Finnish forest company. In: Management Research Review 35 (1), S. 4–13. DOI: 10.1108/01409171211190779.
- Hevner, Alan R.; March, Salvatore T.; Park, Jinsoo; Ram, Sudha** (2004): Design science in information systems research. In: MIS quarterly, S. 75–105.
- Hippner, Hajo; Rentzmann, René** (2006): Text Mining. In: Informatik Spektrum 29 (4), S. 287–290. DOI: 10.1007/s00287-006-0091-y.

- Hobsbawm, E. J. 1917-2012** (1999): Industry and empire. From 1750 to the present day. [New] edition / revised and updated with Chris Wrigley. New York : New Press : Distributed by W.W. Norton, [1999] ©1999.
- Hofmann, Thomas** (2001): Unsupervised learning by probabilistic latent semantic analysis. In: Machine learning 42 (1-2), S. 177–196.
- Houtzagers, Gijs** (1999): Empowerment, using skills and competence management. In: Part and Emp: An Int Jnl 7 (2), S. 27–32. DOI: 10.1108/14634449910271429.
- https://ec.europa.eu/esco** (2021): ESCO - Homepage. Hg. v. ESCO - European Commission. URL: <https://ec.europa.eu/esco/portal>, Stand: 28.08.2021.
- Inmon, W. H.** (1996): The data warehouse and data mining. In: Commun. ACM 39 (11), S. 49–50. DOI: 10.1145/240455.240470.
- Introducing How You Match on LinkedIn Jobs** (2022). URL: <https://blog.linkedin.com/2018/may/14/introducing-how-you-match-on-linkedin-jobs>, Stand: 28.02.2022.
- iSchools, Inc.** (2020): Homepage. URL: <https://ischools.org/>, Stand: 13.12.2020.
- Jain, Anil K.; Murty, M. Narasimha; Flynn, Patrick J.** (1999): Data clustering: a review. In: ACM Comput. Surv. 31 (3), S. 264–323.
- Kaare, Kati Körbe; Otto, Tauno** (2015): Smart Health Care Monitoring Technologies to Improve Employee Performance in Manufacturing. In: Procedia Engineering 100, S. 826–833. DOI: 10.1016/j.proeng.2015.01.437.
- Kagermann, Henninger; Wahlster, Wolfgang; Helbig, Johannes** (2013): Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Abschlussbericht des Arbeitskreises Industrie 4.0 2013 (April), Stand: 05.08.2020.
- Karypis, Michael Steinbach George; Kumar, Vipin; Steinbach, Michael** (Hg.) (2000): A comparison of document clustering techniques.
- Khaouja, Imane; Kassou, Ismail; Ghogho, Mounir** (2021): A Survey on Skill Identification from Online Job Ads. In: IEEE Access, S. 1. DOI: 10.1109/ACCESS.2021.3106120.
- Kiefer, Cornelia** (2016): Assessing the Quality of Unstructured Data: An Initial Overview, S. 1–12.
- Kobayashi, Vladimer B.; Mol, Stefan T.; Berkers, Hannah A.; Kismihók, Gábor et al.** (2018a): Text Classification for Organizational Researchers: A Tutorial. In:

Organizational research methods 21 (3), S. 766–799. DOI:
10.1177/1094428117719322.

Kobayashi, Vladimer B.; Mol, Stefan T.; Berkers, Hannah A.; Kismihók, Gábor et al. (2018b): Text Mining in Organizational Research. In: Organizational research methods 21 (3), S. 733–765. DOI: 10.1177/1094428117722619.

Krason, Paulina; Maczewska, Anna; Polak-Sopinska, Aleksandra (Hg.) (2018): Human factor in maintenance management. Springer: Springer.

Kumar, Lokesh; Bhatia, Parul Kalra (2013): Text mining: concepts, process and applications. In: Journal of Global Research in Computer Science 4 (3), S. 36–39.

Lan, Man; Tan, Chew Lim; Su, Jian; Lu, Yue (2009): Supervised and traditional term weighting methods for automatic text categorization. In: IEEE transactions on pattern analysis and machine intelligence 31 (4), S. 721–735. DOI:
10.1109/TPAMI.2008.110.

Lee, Jisue; Spears, Laura; Ambapavaru, Chandrasa Reddy; Ma, Jinxuan et al. (2014): Aligning expectation and reality about IT career preparation: Perception of job competencies by students, new professionals, and employers. In: New Professionals, and Employers (September 17, 2014), S. 1–45.

Levine, Edward L.; Ash, Ronald A.; Hall, Hardy; Sistrunk, Frank (1983): Evaluation of job analysis methods by experienced job analysts. In: Academy of management journal 26 (2), S. 339–348.

Li, Yong H.; Jain, Anil K. (1998): Classification of text documents. In: The Computer Journal 41 (8), S. 537–546.

Lorenz, M.; Rohrschneider, U. (2009): Erfolgreiche Personalauswahl: Sicher, schnell und durchdacht. Gabler Verlag. URL:
<https://books.google.at/books?id=PKyIWN4awM8C>.

Madreiter, Theresa (2020): Design and development of a prototype of a text understanding tool for maintenance 4.0 by measuring associations, readability and sentiment (TU-MARS). Unter Mitarbeit von Wilfried Sihm und Fazel Ansari Chaharsoughi.

Maisiri, Whisper; Darwish, Hasan; van Dyk, Liezl (2019): AN INVESTIGATION OF INDUSTRY 4.0 SKILLS REQUIREMENTS. In: SAJIE 30 (3), S. 90–105. DOI:
10.7166/30-3-2230.

- Matyas, Kurt** (2002): Ganzheitliche Optimierung durch individuelle Instandhaltungsstrategien. In: Ganzheitliche Optimierung. URL: <https://permalink.catalogplus.tuwien.at/AC04106306>.
- Mokyr, Joel; Strotz, Robert** (1998): The second industrial revolution, 1870-1914. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.478.7503&rep=rep1&type=pdf>.
- Morgeson, F. P.; Brannick, M. T.; Levine, E. L.** (2019): Job and Work Analysis. Methods, Research, and Applications for Human Resource Management. SAGE Publications.
- National Research Council** (2010): A Database for a Changing Economy: Review of the Occupational Information Network (O*NET). National Academies Press. URL: <https://books.google.at/books?id=pTVkAgAAQBAJ>.
- Nikfarjam, Azadeh; Emadzadeh, Ehsan; Muthaiyah, Saravanan** (2010): Text mining approaches for stock market prediction. In: 2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE). 2nd International Conference on Computer and Automation Engineering (ICCAE 2010). Singapore, 26.02.2010. IEEE, S. 256–260.
- NLTK** (2021): Natural Language Toolkit — NLTK 3.6.2 documentation. URL: <https://www.nltk.org/>, Stand: 04.09.2021.
- Nordhaug, Odd; Grønhaug, Kjell** (1994): Competences as resources in firms. In: The International Journal of Human Resource Management 5 (1), S. 89–106. DOI: 10.1080/09585199400000005.
- OECD, Publishing** (2013): OECD Skills Outlook 2013. First Results from the Survey of Adult Skills. Paris: OECD Publishing. URL: <http://gbv.ebib.com/patron/FullRecord.aspx?p=1593819>.
- Olston, Christopher; Najork, Marc** (2010): Web Crawling. In: FNT in Information Retrieval 4 (3), S. 175–246. DOI: 10.1561/1500000017.
- Parry, Scott** (1996): The quest for competencies. In: Training 33 (7), S. 48–56. URL: <http://search.proquest.com/docview/203398191/>.
- Pejic-Bach, Mirjana; Bertonce, Tine; Meško, Maja; Krstić, Živko** (2020): Text mining of industry 4.0 job advertisements. In: International Journal of Information Management 50, S. 416–431. DOI: 10.1016/j.ijinfomgt.2019.07.014.

- Pfeiffer, Sabine** (2017): The Vision of "Industrie 4.0" in the Making-a Case of Future Told, Tamed, and Traded. In: *Nanoethics* 11 (1), S. 107–121. DOI: 10.1007/s11569-016-0280-3.
- Pinzone, Marta; Fantini, Paola; Perini, Stefano; Garavaglia, Stefano et al.** (Hg.) (2017): *Jobs and skills in Industry 4.0. An exploratory research*. Springer: Springer.
- Polanyi, Michael** (1966): The Logic of Tacit Inference. In: *Philosophy* 41 (155), S. 1–18. DOI: 10.1017/S0031819100066110.
- Prahalad, Coimbatore K.; Hamel, Gary** (1997): The core competence of the corporation. In: *Strategische Unternehmensplanung/Strategische Unternehmensführung*. Springer, S. 1–15.
- Pyspellchecker** (2021): *pyspellchecker — pyspellchecker 0.6.2 documentation*. URL: <https://pyspellchecker.readthedocs.io/en/latest/>, Stand: 04.09.2021.
- Raghavan, Vijay V.; Wong, S. MichaelK** (1986): A critical analysis of vector space model for information retrieval. In: *Journal of the American Society for information Science* 37 (5), S. 279–287.
- Reinders, D.**; Thönißen, F. (2020): *Personalmanagement: für Betriebswirte & Technische Betriebswirte IHK*. Books on Demand. URL: https://books.google.at/books?id=yp_sDwAAQBAJ.
- Rendón, Eréndira; Abundez, Itzel; Arizmendi, Alejandra; Quiroz, Elvia M.** (2011): Internal versus external cluster validation indexes. In: *International Journal of computers and communications* 5 (1), S. 27–34.
- Renold, C. G.** (1928): The Nature and Present Position of Skill in Industry. In: *The Economic Journal* 38 (152), S. 593. DOI: 10.2307/2224102.
- Rogelberg, Steven G.** (Hg.) (2017): *The SAGE Encyclopedia of Industrial and Organizational Psychology*, 2nd edition. 2455 Teller Road, Thousand Oaks California 91320: SAGE Publications, Inc.
- Rojko, Andreja** (2017): Industry 4.0 Concept. Background and Overview. In: *Int. J. Interact. Mob. Technol.* 11 (5), S. 77–90. DOI: 10.3991/ijim.v11i5.7072.
- Sackett, Paul R.**; Walmsley, Philip T.; Laczó, Roxanne M. (2012): *Job and Work Analysis (Major Reference Works)*. DOI: 10.1002/9781118133880.hop212004.

- Safdar, Rehman; Waheed, Ajmal; Rafiq, Khattak Hamid** (2010): Impact of job analysis on job performance. Analysis of a hypothesized model. In: Journal of Diversity Management 5 (2), S. 17.
- Schoemaker, P. J.; Day, George S.** (2009): How to make sense of weak signals. In: Leading Organizations: Perspectives for a New Era 37, S. 80–90.
- Schumacher, Elliot; Dredze, Mark** (2019): Learning unsupervised contextual representations for medical synonym discovery. In: JAMIA open 2 (4), S. 538–546. DOI: 10.1093/jamiaopen/ooz057.
- scikit-learn** (2021): scikit-learn: machine learning in Python — scikit-learn 0.24.2 documentation. URL: <https://scikit-learn.org/stable/>, Stand: 21.09.2021.
- Secretary of the Department of Transport and Regional Development** (Hg.) : Human factors in airline maintenance: A study of incident reports. Bureau of Air Safety Investigation. URL: https://www.atsb.gov.au/publications/1997/sir199706_001/, Stand: 13.11.2020.
- Selenium** (2021): Selenium. URL: <https://www.selenium.dev/>, Stand: 04.09.2021.
- Silva, Vander Luiz; Kovalski, João Luiz; Pagani, Regina Negri** (2019): Technology Transfer and Human Capital in the Industrial 4.0 Scenario. A Theoretical Study. In: FSRJ 11 (1), S. 102–122. DOI: 10.24023/FutureJournal/2175-5825/2019.v11i1.369.
- Simpson, B. M.** (2002): The knowledge needs of innovating organisations. In: Singapore Management Review 24 (3), S. 51–60.
- Singley, M. K.; Anderson, J. R.** (1989): The Transfer of Cognitive Skill (Cognitive science series). Harvard University Press. URL: <https://books.google.at/books?id=8UrN-09ZXUsC>.
- Solka, Jeffrey L.** (2008): Text Data Mining: Theory and Methods (2). DOI: 10.1214/07-SS016.
- spaCy** (2021): spaCy · Industrial-strength Natural Language Processing in Python. URL: <https://spacy.io/>, Stand: 04.09.2021.
- Spencer, Lyle M.; Spencer, Phd Signe M.** (1993): Competence at Work models for superior performance. John Wiley & Sons.
- Steinmayr, Ricarda** (2005): Kompetenz-und eigenschaftsbasierte Anforderungsanalysen an Stichproben von Führungskräften und Mitarbeitern.

- Strunz, Matthias** (Hg.) (2012): Instandhaltung. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Suthar, B. K.; Chakravarthi, T. Latha; Pradhan, Shamyal** (2014): Impacts of Job Analysis on Organizational Performance. An Inquiry on Indian Public Sector Enterprises. In: *Procedia Economics and Finance* 11, S. 166–181. DOI: 10.1016/S2212-5671(14)00186-5.
- Taylor, Christine** (2018): Structured vs. Unstructured Data. URL: <https://www.datamation.com/big-data/structured-vs-unstructured-data.html>, Stand: 23.10.2020.
- Teodorescu, Tina** (2006): Competence versus competency: What is the difference? In: *Perf. Improv.* 45 (10), S. 27–30. DOI: 10.1002/pfi.4930451027.
- Uhm, Miyoung; Lee, Ghang; Jeon, Boyoung** (2017): An analysis of BIM jobs and competencies based on the use of terms in the industry. In: *Automation in Construction* 81, S. 67–98. DOI: 10.1016/j.autcon.2017.06.002.
- United States Small Business Administration Office of Management** (1980): Job Analysis, Job Specifications, and Job Descriptions (Business basics). U.S. Small Business Administration, Office of Management Assistance.
- van Wel, Lita; Royackers, Lambèr** (2004): Ethical issues in web data mining. In: *Ethics and Information Technology* 6 (2), S. 129–140.
- Wayne W. Eckerson** (2002): DATA QUALITY AND THE BOTTOM LINE. Achieving Business Success through a Commitment to High Quality Data, S. 1–36.
- Weinert, Franz E.** (Hg.) (1999): Concepts of competence. OFS: OFS.
- Weiss, Sholom M.; Indurkha, Nitin; Zhang, Tong** (2010): Fundamentals of Predictive Text Mining. London: Springer London. DOI: 10.1007/978-1-84996-226-1.
- Wilk, Gabriele** (2011): Stellenbeschreibungen und Anforderungsprofile (Haufe-Praxisratgeber). Freiburg: Haufe.
- Winterton, Jonathan; Delamare-Le Deist, Françoise; Stringfellow, Emma** (2006): Typology of knowledge, skills and competences. Clarification of the concept and prototype. Office for Official Publications of the European Communities Luxembourg.

- Wu, Dan; Lv, Siyu; Xu, Hao** (2020): An analysis on competency of human-centered data science employment. In: Proc Assoc Inf Sci Technol 57 (1). DOI: 10.1002/pras.2.219.
- Wu, Jun; Shi, Honglei; Yang, Jiaping** (2017): Are big data talents different from business intelligence expertise?: Evidence from text mining using job recruitment advertisements. In: Jiafu Tang, Jian Chen und Xiaoqiang Cai (Hg.): The 14th International Conference on Services Systems and Services Management (ICSSSM2017). June 16-18, 2017, Dalian, China : proceedings. 2017 International Conference on Service Systems and Service Management, 2017/06. ICSSSM; Tsinghua University; Chinese University of Hong Kong; IEEE Systems, Man, and Cybernetics Society; International Conference on Services Systems and Services Management. Piscataway, NJ: IEEE.
- Yuan, Zhengwu; Duan, Lian** (2019 - 2019): Construction Method of Sentiment Lexicon Based on Word2vec. In: 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). Chongqing, China, 24.05.2019 - 26.05.2019. IEEE, S. 848–851.
- Zhang, Yu; Chen, Mengdong; Liu, Lianzhong** : A review on text mining. In: 2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS). IEEE, S. 681–685.
- Zhao, Meng; Javed, Faizan; Jacob, Ferosh; McNa, Matt** (2015): SKILL: A System for Skill Identification and Normalization. In: Innovative Applications of Artificial Intelligence; Twenty-Seventh IAAI Conference. URL: <https://www.aaai.org/ocs/index.php/IAAI/IAAI15/paper/view/9363>.

9 Abbildungsverzeichnis

Abbildung 1: Aufbau und Struktur dieser Arbeit.....	6
Abbildung 2: Evolution der Industrie	9
Abbildung 3: Instandhaltungsstrategien und deren Merkmale	12
Abbildung 4: Zusammenhang individuelle Kompetenzen und Kernkompetenzen	14
Abbildung 5: Oberflächliche und tieferliegende Eigenschaften.....	16
Abbildung 6: Kompetenzmanagement Prozessrahmenmodell	20
Abbildung 7: Bereiche des Text Mining	31
Abbildung 8: Wissenspyramide	33
Abbildung 9: Der Text Mining Prozess	35
Abbildung 10: Umsetzung des Text Mining Prozess	52
Abbildung 11: Aufgabendefinition	53
Abbildung 12: Dokumentenauswahl	53
Abbildung 13: (Text) Mining Methoden.....	53
Abbildung 14: Interpretation / Evaluation.....	53
Abbildung 15: Anwendung.....	54
Abbildung 16: Abstraktes Konzept der Umsetzung	54
Abbildung 17: Auswertungsteil der Kompetenzanzahl.....	58
Abbildung 18: Umgesetzte Synonymerkennung.....	59
Abbildung 19: Spezifisches Konzept der Umsetzung	61
Abbildung 20: Spezifisches Konzept der Umsetzung nach der Adaptierung	63
Abbildung 21: Kompetenzäste – 12.07.2021	66
Abbildung 22: Kompetenzverteilung – 12.07.2021	66
Abbildung 23: Kompetenzäste – 30.08.2021	69
Abbildung 24: Kompetenzverteilung – 30.08.2021	70
Abbildung 25: Kompetenzäste Erweiterung – 12.07.2021	72
Abbildung 26: Kompetenzverteilung Erweiterung – 12.07.2021	72
Abbildung 27: Optimierungsplots – ESCO Vokabular.....	75
Abbildung 28: Optimierungsplots – Ohne Einschränkung des Vokabulars	77
Abbildung 29: KompetenzästeVergleich beider Auswertungen	80
Abbildung 30: Kompetenzentwicklungen der gleichbleibenden Top 10 Skill, Transversal, ICT, Language und Knowledge.....	81
Abbildung 31: Abstraktes Konzept der Umsetzung - Großformat	98
Abbildung 32: Spezifisches Konzept der Umsetzung - Großformat.....	99
Abbildung 33: Spezifisches Konzept der Umsetzung nach Adaptierung- Großformat	100

10 Tabellenverzeichnis

Tabelle 1: Vergleich der kompetenzbasierenden Taxonomien von Experten.....	29
Tabelle 2: Unterschiede zwischen Data- und Text Mining.....	32
Tabelle 3: Daten- versus Informationsqualitätsprobleme.....	34
Tabelle 4: Exemplarisches Beispiel der Textreinigung	39
Tabelle 5: Systematische Literaturrecherche - Großformat im Anhang (Tabelle 28).....	45
Tabelle 6: Vor- und Nachteile von Kompetenzidentifikationsmethoden.....	49
Tabelle 7: Zusammenfassung der Eingrenzungen für die Dokumentenauswahl	55
Tabelle 8: Exemplarisches Beispiel für die Textreinigung	57
Tabelle 9: Vergleichsbeispiele manuelle Erweiterungsliste zu ESCO Kompetenzen / Instandhaltungskonzepten.....	62
Tabelle 10: ESCO Kompetenzäste und jeweilige Anzahl an Standardkompetenzen (Version 1.0.8).....	65
Tabelle 11: Top 10 Kompetenzen je Kompetenzast -12.07.2021	65
Tabelle 12: Evaluation	67
Tabelle 13: Beispiel des Ähnlichkeitsproblems aus einer Stellenanzeige.....	68
Tabelle 14: Beispiel des generischen Kompetenzproblems aus einer Stellenanzeige	68
Tabelle 15: Beispiel einer falsch positiven Kompetenz aus einer Stellenanzeige.....	68
Tabelle 16: Top 10 Kompetenzen je Kompetenzast -30.08.2021	69
Tabelle 17: Top 10 Kompetenzen je Kompetenzast nach Adaptierung ESCO Kompetenzen - 12.07.2021	71
Tabelle 18: Top 10 nach Adaptierung Erweiterungsliste - 12.07.2021	71
Tabelle 19: Evaluation – nach Adaptierung	73
Tabelle 20: Top 10 - ESCO Vokabular	74
Tabelle 21: LDA Thema - ESCO Vokabular	75
Tabelle 22: Top 10 sortiert nach tf- Ohne Einschränkung des Vokabulars.....	76
Tabelle 23: Top 10 sortiert nach summed_tfidf/tf- Ohne Einschränkung des Vokabulars.....	76
Tabelle 24: Anzahl der zugeordneten Stellenanzeigen - Ohne Einschränkung des Vokabulars.....	77
Tabelle 25: LDA Themen - Ohne Einschränkung des Vokabulars.....	78
Tabelle 26: Gleichbleibende Kompetenzen in den Top 10	79
Tabelle 27: Veränderungen bei den Top 10	80
Tabelle 28: Systematische Literaturrecherche - Großformat	97

11 Formelverzeichnis

Formel 1: Inverse Dokumentenhäufigkeit	40
--	----

12 Abkürzungsverzeichnis

bzw.	beziehungsweise
ESCO	European classification of Skills, Competences, Qualifications and Occupations
DIN	Deutsches Institut für Normung
max.	maximal
z.B.	zum Beispiel
bzgl.	bezüglich
dbzgl.	diesbezüglich
EN	Englisch
DE	Deutsch / Deutschland
UK	United Kingdom
e.g.	for example
approx.	approximately
KSAO	Knowledge, Skills, Abilities and other Characteristics