

# Methodenbericht

## Future Skills für die Region Nordschwarzwald



**Autoren:**

**Lennart Bolwin, Jan Engler, Dr. Henry Goecke, Dr. Vanessa Hünнемeyer, Dr. Armin Mertens**

Juni 2024

# Inhalt

1	Einführendes zum Studiendesign.....	3
2	Analyse der Online-Stellenanzeigen .....	4
2.1	Datengrundlage und Datenbasis.....	4
2.1.1	Online-Stellenanzeigen .....	4
2.1.2	Abgrenzung der Branchen.....	5
2.2	Datenauswertung .....	7
2.2.1	Identifikation und Extraktion der nachgefragten Kompetenzen.....	7
2.2.2	Extraktion der Kompetenzen .....	8
2.2.3	Evaluation der Kompetenzextraktion.....	11
2.2.4	Clustering der Kompetenzen .....	11
2.2.5	Forecasting der Kompetenzbedarfe .....	12
2.2.6	Future-Skills-Cluster .....	14
3	Leitfadengestützte Interviews.....	17
3.1	Ziel und Zweck im Forschungsdesign .....	17
3.2	Auswahl der Gesprächspartner und -partnerinnen .....	17
3.3	Leitfaden und Gesprächsinhalte.....	18
4	Gruppendiskussionen im Workshop-Format.....	19
4.1	Ziel und Zweck im Forschungsdesign .....	19
4.2	Workshop 1 am 22. März 2024 in Stuttgart.....	20
4.3	Workshop 2 am 25. April 2024 in Stuttgart.....	20
4.4	Workshop 3 am 08. Mai 2024 in digitaler Form.....	21
5	Standardisierte Befragung.....	22
5.1	Ziel und Zweck im Forschungsdesign .....	22
5.2	Stichprobe.....	22
5.3	Fragebogenkonzeption.....	22
	Anhang .....	24
A.1	Ergänzende Abbildungen .....	24
A.2	Quellen- und Datenverzeichnis .....	29
A.3	Abbildungsverzeichnis.....	31
A.4	Tabellenverzeichnis.....	31

# 1 Einführendes zum Studiendesign

Die Automobil- und Zulieferindustrie ist eine der volkswirtschaftlich wichtigsten Branchen in Deutschland. Auch in der Region Nordschwarzwald sind fast 22.500 Erwerbstätige in der Automobilwirtschaft tätig (Kempermann et al. 2021). Um den automobilen Strukturwandel, der getrieben wird von Digitalisierung und Dekarbonisierung, in betroffenen Unternehmen zu unterstützen, erarbeitet das automobile Transformationsnetzwerk Nordschwarzwald Handreichungen und Orientierungshilfen für Unternehmensentscheider sowie Beschäftigte.



Dieser Methodenbericht zeigt die Vorgehensweise zur Identifizierung der 39 Future-Skills-Cluster auf, die als relevant für die Automobil- und Zulieferindustrie beschrieben werden.

Die Future-Skills-Cluster basieren auf einer Analyse von Online-Stellenanzeigen und sind somit empirisch fundiert. Mittels Machine-Learning-Verfahren werden die mehr als 12.000 Einzelkompetenzen geclustert und unter Berücksichtigung von inhaltlichen Inputs aus Workshops und leitfadengestützten Interviews in ein Klassifikationssystem – bestehend aus übergeordneten Kompetenzkategorien und Future-Skills-Cluster – überführt. Unter Berücksichtigung von Einschätzungen aus der Praxis wird der zukünftige Bedarf in statistischen Modellrechnungen bis zum Jahr 2030 geschätzt und jene Skills-Cluster als Future-Skills-Cluster identifiziert, die wachsen.

Die Analyse von Online-Stellenanzeigen ist eine relativ junge, allerdings mittlerweile etablierte Methode zur Identifizierung von Kompetenzbedarfen (siehe Kapitel 2.1). Gleichwohl gibt es bei der Einordnung und Bewertung der Ergebnisse auch Einschränkungen zu beachten:

- Andere Kanäle der Personalgewinnung, etwa über Recruiting-Dienstleister, werden nicht erfasst. Somit stellen Online-Stellenanzeigen nur eine Teilmenge dar.
- Um im Zeitalter des Fachkräftemangels ausreichend Bewerbungen zu generieren, ist es möglich, dass Stellenanzeigen bevorzugt breit formuliert sind und sehr spezialisiertes Fachwissen dagegen nicht explizit genannt wird.
- Es kann keine Aussage getroffen werden, inwiefern die tatsächliche Stellenbesetzung den zuvor formulierten Anforderungen und Wünschen der Arbeitgeber entspricht.

Unter Berücksichtigung dieser Einschränkungen stellt die Analyse der Online-Stellenanzeigen den Ist-Zustand der Kompetenzbedarfe in den Unternehmen der Untersuchungsregion dar. Um darauf aufbauend die Zukunftsrelevanz einzuschätzen, schließen sich weitere quantitative und qualitative Verfahren an. Es ist zu beachten, dass diese methodische Vorgehensweise eine Makro-Perspektive einnimmt. D. h. die tatsächliche Zukunftsrelevanz der hier vorgeschlagenen Future-Skills-Cluster in den individuellen Unternehmen muss immer im konkreten Unternehmenskontext individuell evaluiert werden.

## 2 Analyse der Online-Stellenanzeigen

### 2.1 Datengrundlage und Datenbasis

#### 2.1.1 Online-Stellenanzeigen

Als zentrale Datengrundlage für die Analyse der Zukunftskompetenzen wurden Online-Stellenanzeigen verwendet. Online-Stellenanzeigen haben sich in den letzten Jahren vermehrt als informative und objektive Quelle für die Analyse von Kompetenzbedarfen bewährt (siehe Acemoglu et al. 2022; Borgnovi et al. 2023; Buchmann et al. 2022; Büchel und Engler 2024). Denn die Volltexte der Ausschreibungen enthalten neben einer Beschreibung des ausschreibenden Unternehmens in der Regel konkrete Formulierungen mit den gewünschten Kompetenzen, über die der zukünftige Bewerber oder die zukünftige Bewerberin verfügen sollte. Weitere Vorteile von Online-Stellenanzeigen sind ihre zeitliche Verfügbarkeit bis hin zum aktuellen Rand und nicht zuletzt auch ihre große Menge, mit der sich der tatsächliche Kompetenzbedarf auf dem Arbeitsmarkt gut approximieren lässt (Khaouja et al. 2021).

Für die Identifikation der Zukunftskompetenzen wurden Daten des Datenanbieters Textkernel verwendet. Textkernel sammelt und verarbeitet Online-Stellenanzeigen aus fast 60.000 unterschiedlichen Quellen (siehe Mertens et al. 2023). Da die meisten Stellenanzeigen aber mehrfach geschaltet werden, etwa auf verschiedenen Plattformen oder wiederholt zu verschiedenen Zeitpunkten, werden die Daten um Duplikate bereinigt (dedupliziert). Zusätzlich extrahiert Textkernel mittels Machine-Learning-Verfahren weitere Informationen aus den Ausschreibungen wie den Namen und den Wirtschaftszweig (WZ-Klasse) des ausschreibenden Unternehmens, die Homepage-URL, die Klassifikation der Berufe (KldB) und Informationen zum Arbeitsort.

Um eine ausreichende Anzahl an Daten für die Identifikation von Zukunftskompetenzen zu gewährleisten, wurde ein Untersuchungszeitraum von Januar 2018 bis Dezember 2023 festgelegt, da für diesen Zeitraum die Textkernel-Daten in hoher Qualität vorliegen. Weiterhin ermöglicht ein längerer Untersuchungszeitraum eine bessere Identifikation von Trends sowie die Fortschreibung dieser Trends hin zu zukünftigen Bedarfen. Darüber hinaus wurden die Daten geografisch gefiltert, um passgenaue Analysen für die Zielregion zu ermöglichen. Dafür wurden zunächst nur solche Stellenanzeigen betrachtet, die der Region Nordschwarzwald bzw. dem Land Baden-Württemberg zugeordnet werden konnten. Diese Zuordnung erfolgte über den Arbeitsort, der in den Stellenanzeigen angegeben ist. Über den Untersuchungszeitraum von Januar 2018 bis Dezember 2023 wurden so mehr als 10 Millionen (10.361.230) Stellenanzeigen in Baden-Württemberg und 402.657 deduplizierte Stellenausschreibungen für die Region Nordschwarzwald gesammelt.

Online-Stellenanzeigen sind in der Regel in drei Abschnitte gegliedert. Zunächst stellt sich das Unternehmen selbst vor. Dieser Abschnitt enthält etwa Informationen über die Branche des ausschreibenden Unternehmens oder darüber, welche weiteren Branchen beliefert werden. Im nächsten Abschnitt wird die ausgeschriebene Stelle und das Aufgabenfeld beschrieben. Zuletzt werden die Anforderungen an den Kandidaten formuliert. In diesen beiden Abschnitten befinden sich die Tätigkeiten, Wissensgebiete und Kompetenzen, die für den Job benötigt werden.



## 2.1.2 Abgrenzung der Branchen

### *Automobil- und Zulieferindustrie*

Die aus Baden-Württemberg vorliegenden Stellenanzeigendaten stammen aus Ausschreibungen aller Branchen. Für die Analyse der Zukunftskompetenzen in der Automobil- und Zulieferindustrie müssen daher zunächst diejenigen Anzeigen herausgefiltert werden, die einen direkten Bezug zu dieser haben. Hierfür bietet sich für den ersten Schritt die Klassifikation der Wirtschaftszweige (WZ; siehe Statistisches Bundesamt 2008) des Statistischen Bundesamtes an. Die WZ-Klasse beschreibt die Zugehörigkeit eines Unternehmens zu einem Wirtschaftszweig bzw. einer Branche. Für die vorliegende Analyse sind primär alle Klassen interessant, die der Automobil- und Zulieferindustrie zugerechnet werden können. Dies umfasst die WZ-Klassen 29 (Herstellung von Kraftwagen und Kraftwagenteilen), 30 (Sonstiger Fahrzeugbau) und 45 (Handel mit Kraftfahrzeugen; Instandhaltung und Reparatur von Kraftfahrzeugen). Weitere WZ-Klassen, wie z. B. 77.1 (Vermietung von Kraftwagen), wurden geprüft und auf Grund fehlender Nähe zur Industrie (Industriefokus) nicht mitbetrachtet.

Tabelle 1 zeigt die Anzahl der Stellenanzeigen im Untersuchungszeitraum (2018 bis 2023) für die oben genannten WZ-Klassen. Für Baden-Württemberg wurden knapp 300.000 Ausschreibungen und für den Nordschwarzwald knapp über 6.200 Ausschreibungen identifiziert.

**Tabelle 1: Anzahl der Stellenausschreibungen in der Automobil- und Zulieferindustrie nach WZ-Klassen**

Anzahl der Stellenausschreibungen in Baden-Württemberg und in der Region Nordschwarzwald im Untersuchungszeitraum (2018-2023)

WZ-Klasse	Anzahl Baden-Württemberg	Anzahl Region Nordschwarzwald
29	216.437	2.290
30	17.054	621
45	64.746	3.304
Insgesamt	298.237	6.215

Quelle: Institut der deutschen Wirtschaft auf Basis von Textkernel-Daten 2024

Die Automobil- und Zulieferindustrie kann in ihrer Gesamtheit aber nicht durch die oben genannten WZ-Klassen abgebildet werden. Denn viele Unternehmen, die beispielsweise der WZ-Klasse 25 (Herstellung von Metallerzeugnissen) zugeordnet sind, produzieren (mittlerweile) fast ausschließlich Produkte für die Automobilindustrie. Die Ausschreibungen dieser Unternehmen würden bei einer ausschließlichen Betrachtung der WZ-Klassen wie in Tabelle 1 in der Analyse fehlen. Aus diesem Grund wurde die Auswahl der relevanten Online-Stellenanzeigen um eine hauseigene Datenbank ergänzt (siehe Kempermann et al. 2021). Die Datenbank enthält Unternehmen außerhalb der WZ-Klassen 29, 30 und 45, die jedoch wichtiger Teil des Wertschöpfungsnetzwerkes der Automobil- und Zulieferindustrie sind. Große Unternehmen wurden in einem aufwändigen, händischen Prozess identifiziert und KMUs automatisiert über eine Analyse ihrer Homepage gefunden. Da in der Datenbank ebenfalls die URL der Unternehmenswebsite und Geoinformationen zu den Unternehmen

erfasst wurden, können die in Baden-Württemberg ansässigen Unternehmen in den Stellenausschreibungen erkannt werden. Weil die in der Datenbank enthaltenen Unternehmen neben dem Fokus auf die Automobilindustrie auch weitere Tätigkeitsbereiche haben können, wurden die Stellenausschreibungen hinsichtlich ihres Bezugs zur Automobil- und Zulieferindustrie überprüft. Hierzu wurden die Anzeigen mit einem regelbasierten Verfahren nach Schlagwörtern durchsucht, die häufig im Zusammenhang mit der Automobil- und Zulieferindustrie verwendet werden. Durch die Datenbank wurden für Baden-Württemberg zusätzlich 142.533 Stellenausschreibungen und für die Region Nordschwarzwald 5.272 relevante Stellenausschreibungen identifiziert. Die meisten auf diesem Weg hinzugefügten Ausschreibungen stammen aus den WZ-Klassen 17, 25, 26, 28 und 46.

Zuletzt wurde vom Transformationsnetzwerk Nordschwarzwald noch eine Liste mit den wichtigsten Kfz-Autohäusern und Automobilzulieferern aus der Region Nordschwarzwald bereitgestellt. Die Ausschreibungen dieser Unternehmen wurden ebenfalls zur Datenbasis hinzugefügt, sofern diese nicht bereits enthalten waren. Dadurch konnten weitere 4.741 Stellenausschreibungen für Baden-Württemberg und 3.358 für die Region Nordschwarzwald identifiziert werden.

## **Tabelle 2: Anzahl der Stellenausschreibungen in der Automobil- und Zulieferindustrie nach WZ-Klassen**

Anzahl der Stellenausschreibungen in Baden-Württemberg und in der Region Nordschwarzwald im Untersuchungszeitraum (2018-2023)

	<b>Anzahl Baden-Württemberg</b>	<b>Anzahl Region Nordschwarzwald</b>
WZ	298.237	6.215
+ eigene Datenbank	142.533	5.272
+ Unternehmensliste	4.741	3.358
= Insgesamt	445.511	14.845

Quelle: Institut der deutschen Wirtschaft auf Basis von Textkernel-Daten 2024

Die finale Datenbasis enthält alle Ausschreibungen der Unternehmen aus der Automobil- und Zuliefererindustrie in Baden-Württemberg. Tabelle 2 zeigt die Ergebnisse der oben beschriebenen Arbeitsschritte. Insgesamt wurden über 445.000 Ausschreibungen in Baden-Württemberg identifiziert. Knapp 15.000 dieser Ausschreibungen können der Region Nordschwarzwald zugeordnet werden.

### **Weitere relevante Branchen**

Für die Extraktion der Kompetenzen (siehe Kapitel 2.2) wurden zusätzlich noch die WZ-Klassen 24 (Metallerzeugung und -bearbeitung), 25 (Herstellung von Metallerzeugnissen), 26 (Herstellung von Datenverarbeitungsgeräten, elektronischen und optischen Erzeugnissen), 27 (Herstellung von elektrischen Ausrüstungen) und 28 (Maschinenbau) betrachtet, da auch hier Tätigkeiten ausgeführt werden können, die ebenso für die Automobil- und Zulieferindustrie von Bedeutung sind. Die Anzahl der relevanten Stellenausschreibungen ist in Tabelle 3 dargestellt. Diese Stellenausschreibungen wurden lediglich dafür genutzt, das Kompetenzwörterbuch anzureichern. Für die sonstigen durchgeführten Analysen spielen die zusätzlichen Branchen aber keine weitere Rolle.

### Tabelle 3: Anzahl der Stellenausschreibungen in der Metall-, Maschinenbau- und Elektroindustrie

Anzahl der Stellenausschreibungen in Baden-Württemberg und in der Region Nordschwarzwald im Untersuchungszeitraum (2018-2023)

WZ-Klasse	Anzahl Baden-Württemberg	Anzahl Region Nordschwarzwald
24	59.813	5.088
25	128.174	7.406
26	80.330	2.303
27	100.632	4.788
28	237.889	8.810

Quelle: Institut der deutschen Wirtschaft auf Basis von Textkernel-Daten 2024

## 2.2 Datenauswertung

### 2.2.1 Identifikation und Extraktion der nachgefragten Kompetenzen

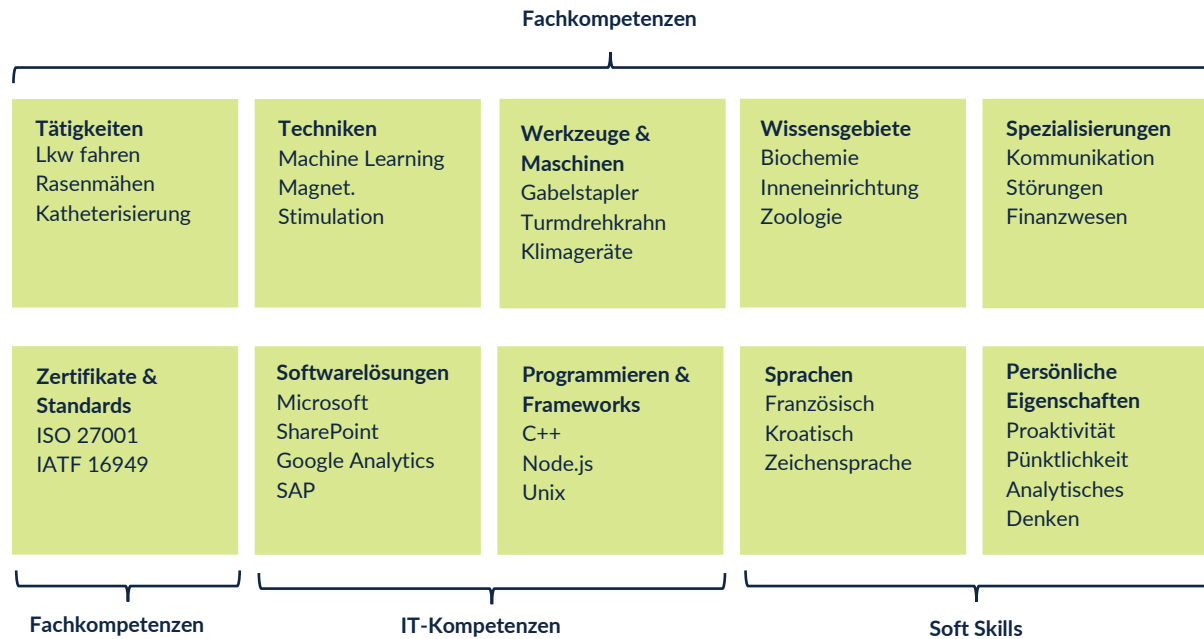
#### *Taxonomie*

Um die Nachfrage nach Kompetenzen über den gesamten Untersuchungszeitraum hinweg bis zum aktuellen Rand abbilden zu können, müssen die Kompetenzen zunächst in den Online-Stellenanzeigen identifiziert und aus diesen extrahiert werden. Für die strukturierte Extraktion wurde eine Taxonomie verwendet, die von Textkernel entwickelt wurde. Die Kompetenztaxonomie unterscheidet auf der höchsten Ebene zwischen Fachkompetenzen, IT-Kompetenzen und überfachlichen Fähigkeiten (Soft Skills) (siehe Abbildung 1)

Die Gruppe der Fachkompetenzen umfasst verschiedene Tätigkeiten, Techniken, Werkzeuge und Maschinen, Wissensgebiete, Spezialisierungen und Zertifikate. IT-Kompetenzen setzen sich aus Softwarelösungen sowie Programmiersprachen und Frameworks zusammen und die Soft Skills beinhalten persönliche Eigenschaften und Sprachen.

## Abbildung 1: Kompetenztaxonomie

Taxonomie der Kompetenzen mit den Übergruppen „Fachkompetenzen“, „IT-Kompetenzen“ und „Soft Skills“



Quelle: Institut der deutschen Wirtschaft 2024; Textkernel 2024

### 2.2.2 Extraktion der Kompetenzen

Für die Identifikation der Future-Skills-Cluster mussten zunächst alle nachgefragten Kompetenzen aus den relevanten Online-Stellenanzeigen der Automobil- und Zulieferindustrie extrahiert werden. Textkernel stellt bereits extrahierte und standardisierte Kompetenzen zur Verfügung, die mittels verschiedener Machine-Learning-Verfahren aus den Online-Stellenanzeigen gezogen wurden (Textkernel 2024). Im Datensatz von Textkernel finden sich mehr als 8.600 einzigartige Kompetenzen. Da die Qualität dieser Auswahl bei fach- und berufsspezifischen Kompetenzen stark variiert (siehe Mertens et al. 2023), wurden die Textkernel-Kompetenzen um Kompetenzen ergänzt, die in einem Wörterbuch-basierten Verfahren ermittelt wurden. Das Wörterbuch wird aus fünf Quellen befüllt.

#### (1) BERUFENET

BERUFENET ist ein Informationsportal der Bundesagentur für Arbeit (Bundesagentur für Arbeit 2024). Es enthält ausführliche Informationen zu allen anerkannten Ausbildungsberufen. Diese umfassen unter anderem Beschreibungen zu Tätigkeiten und benötigten Kompetenzen zur Ausführung dieser Berufe.

Durch Webscraping wurden alle Kompetenzen aus dem BERUFENET-Portal extrahiert. Hierfür wurden zunächst alle „Kernkompetenzen“, „weitere Kompetenzen“ und die



„Kompetenzgruppen“ aus der Tätigkeitsbeschreibung jeder KldB extrahiert.<sup>1</sup> Daraus wurde ein Datensatz mit allen 8.216 eindeutigen, in BERUFENET vorkommenden Kompetenzen erstellt. Da die Extraktion von Kompetenzen mittels regelbasierter Verfahren aber verschiedene Probleme mit sich bringt (z. B. die Häufung falsch positiver Extraktion durch Homonyme) und dadurch ein erheblicher Mehraufwand bei der Verarbeitung der Skills entstanden wäre, wurde die Auswahl der Kompetenzen auf relevante Jobs der Automobil- und Zuliefererindustrie und der anliegenden Branchen beschränkt. Hierfür wurden die 30 in den Stellenausschreibungen am häufigsten vorkommenden Berufe (KldB-5-Steller) für die Automobil-, Elektro-, Maschinen- und Metallbranche erfasst. Im Anschluss wurden nur solche Kompetenzen in das regelbasierte Wörterbuch aufgenommen, die diesen KldBs zugeordnet werden konnten.

Auf diese Weise konnten 2.500 eindeutige Kompetenzen erfasst werden. Im Anschluss wurden diese Skills so verarbeitet, dass sie in dem Format vorlagen, wie sie in Stellenausschreibungen tatsächlich zu finden sind. Z. B. wurde bei BERUFENET-Kompetenzen wie „Programmiersprache APT, EXAPT, EXAPTplus“ das Beschreibungswort „Programmiersprache“ entfernt und die Aufzählung aufgesplittet. Dadurch entstanden aus einer Kompetenz die drei Kompetenzsuchwörter „APT“, „EXAPT“ und „EXAPTplus“. Insgesamt wurden so 3.188 unterschiedliche Kompetenzen in das Wörterbuch mitaufgenommen.

## *(2) Green- und IT-Skills*

Nach einer ersten Analyse der identifizierten Kompetenzen wurde festgestellt, dass einige sehr aktuelle Kompetenzen aus den Feldern KI- und Green-Skills mit dem bestehenden Wörterbuch nur unzureichend erfasst werden. Um diese potenziellen Future Skills nicht zu unterschätzen, wurden die beiden Kompetenzgruppen durch bestehende Wörterbücher ergänzt.

Zunächst wurden alle relevanten Green Skills aus Binnewitt und Schnepf (2021) dem Wörterbuch hinzugefügt. Die Studie enthält ein Wörterbuch mit 1.800 grünen Kompetenzen, die aufwändig händisch validiert und gewichtet wurden. Die Gewichtung erfolgte durch ein Ranking der Kompetenzen hinsichtlich ihres Bezugs zu Nachhaltigkeit und erfolgte in einem ordinalen Ranking mit den Ausprägungen „gering“, „mittel“ und „hoch“. In das Wörterbuch der vorliegenden Analyse wurden nur solche Kompetenzen aufgenommen, die mindestens die Ausprägung „mittel“ hatten. Dadurch konnte das Wörterbuch um 1.059 weitere grüne Kompetenzen ergänzt werden.

Des Weiteren wurde das Wörterbuch um Begriffe rund um Künstliche Intelligenz (KI) erweitert. Hierbei wurden 877 Kompetenzen von KI-Berufen (siehe Büchel und Mertens 2021) und

---

<sup>1</sup> Die KldB ist eine offizielle Klassifikation aller Berufe in Deutschland (Härpfer und Neuhauser 2021). Für die Analyse wird der sogenannte KldB-5-Steller verwendet, der 1.300 Berufsgattungen umfasst.

weitere 147 dezidierte Kompetenzen zu generativer KI (siehe Büchel und Engler 2024) hinzugefügt.

Zuletzt wurden einzelne Kompetenzen im Wörterbuch ergänzt, die in den verschiedenen Expertenworkshops und -interviews häufig genannt wurden und bisher nicht im Wörterbuch enthalten waren. Dabei handelt es sich vor allem um Soft Skills.

Tabelle 4 zeigt die Anzahl eindeutiger Kompetenzen, die mit dem erstellten Wörterbuch in den Online-Stellenanzeigen gefunden wurde. Von 877 KI-Kompetenzen wurden demnach 753 in den untersuchten Stellenanzeigen genannt. Die Spalte „Anzahl zusätzlicher Kompetenzen“ zeigt die Anzahl der Kompetenzen, die zusätzlich zu den bereits genannten Quellen gefunden wurden.

#### **Tabelle 4: Anzahl der Kompetenzen in den Stellenanzeigen**

Anzahl eindeutiger Kompetenzen nach Quelle und Anzahl zusätzlicher Kompetenzen, die bisher nicht im Wörterbuch enthalten waren

Quelle	Anzahl eindeutiger Kompetenzen	Anzahl zusätzlicher Kompetenzen
Textkernel	8.684	8.684
BERUFENET	2.978	2.429
KI-Kompetenzen (Büchel und Mertens 2022)	753	642
Generative KI-Kompetenzen (Büchel und Engler 2024)	26	24
Green Skills (Binnewitt und Schnepf 2021)	640	579
Expertenworkshops und -interviews	81	73
Insgesamt	13.162	12.441

Quelle: Institut der deutschen Wirtschaft 2024

Alle 12.441 im Wörterbuch enthaltenen Kompetenzen wurden anschließend in computerlinguistische Regeln überführt und aus den Stellenanzeigen extrahiert. Beispielhaft wurde die übergreifende Kompetenz „analytisches Denken“ in die Regel `{'LEMMA': analytisch}, {'LEMMA': Denken}}` überführt. Ein Lemma beschreibt die Wörterbuchform eines Wortes. Die Suche nach dieser Form hat den Vorteil, dass das Wort unabhängig von den Endungen und der Konjugation gefunden wird. So werden die Wortketten „analytisches Denken“, „analytischen Denkens“, „analytischem Denken“ etc. alle von der oben definierten Regel erkannt und extrahiert. Weitere computerlinguistische Regeln können auch Negationen (ein Wort darf nicht im Kontext mit einem bestimmten weiteren Wort vorkommen) oder bestimmte notwendige grammatikalische Konstellationen festlegen (ein Substantiv muss im Zusammenhang mit einem bestimmten Verb genannt werden). Zusätzlich werden reguläre Ausdrücke verwendet, um etwa mögliche Sonderzeichen zwischen Begriffen zu erfassen. Insgesamt wurden durch das Verfahren ca. 14 Millionen Kompetenzen aus den Stellenausschreibungen im Untersuchungszeitraum extrahiert.

### 2.2.3 Evaluation der Kompetenzextraktion

Im nächsten Schritt wurden die extrahierten Kompetenzen mit verschiedenen Metriken an einem Testdatensatz evaluiert. Hierfür wurden 50 zufällig ausgewählte Stellenanzeigen annotiert, d. h. es wurden alle genannten Kompetenzen manuell markiert. Diese Markierungen wurden im Anschluss mit den Kompetenzen abgeglichen, die durch das Modell automatisiert extrahiert wurden. In den Testdaten wurden durch das regelbasierte Verfahren 737 Kompetenzen extrahiert. Tabelle 5 zeigt die Evaluationsmetriken Precision, Recall und F1, die nach der regelbasierten Extraktion aller Kompetenzen in den Testdaten berechnet wurden. Die Werte können zwischen 0 und 1 liegen.

**Tabelle 5: Evaluationsmetriken**

Precision-, Recall- und F1-Werte des regelbasierten Verfahrens (Testdaten: N = 50)

	Werte
Precision	0,928
Recall	0,882
F1	0,905

Quelle: Institut der deutschen Wirtschaft 2024

Die Precision (Spezifität) der Extraktion ist mit einem Wert von 0,928 sehr hoch. Der Wert wird berechnet, indem alle korrekt durch das Modell extrahierten Kompetenzen durch alle vorhergesagten Kompetenzen dividiert werden. Der Recall (Sensitivität) des Modells ist mit 0,882 ebenfalls hoch, aber etwas niedriger als die Precision. Der Recall gibt den Anteil aller korrekt erfassten Kompetenzen an der Gesamtheit aller theoretisch zu extrahierenden Kompetenzen an. Demnach werden einige Kompetenzen durch das regelbasierte Verfahren nicht erfasst. Zuletzt beschreibt der F1-Wert das harmonische Mittel der beiden Metriken Precision und Recall. Mit 0,905 ist auch dieser Wert sehr hoch und somit die Qualität der Kompetenzextraktion insgesamt sehr hoch.<sup>2</sup>

### 2.2.4 Clustering der Kompetenzen

Anschließend wurden die extrahierten Kompetenzen automatisiert in Cluster gruppiert. Für das Clustering wurden computerlinguistische Methoden und Clustering-Verfahren kombiniert. Da es sich bei den Kompetenzen nicht um Zahlen, sondern um Wörter handelt, müssen diese

<sup>2</sup> Für eine detaillierte Beschreibung der berechneten Evaluationsmetriken siehe beispielhaft Büchel und Mertens (2021, S. 19-21).

zunächst in ein computerlesbares Format umgewandelt werden. Dafür wurden sogenannte Worteinbettungen (word2vec, Bojanowski et al. 2017) für jede Kompetenz erstellt. Unter Worteinbettungen versteht man die numerische Darstellung eines Wortes als Vektor, der die Bedeutung dieses Wortes repräsentiert. Semantisch ähnlichen Wörtern werden ähnliche Vektoren zugewiesen. Dadurch lassen sich Berechnungen mit Wörtern durchführen.

Da die deutschsprachigen Worteinbettungsmodelle schon seit einigen Jahren nicht mehr aktualisiert wurden, wurde das Modell zunächst mit eigenen Daten auf einen neuen Stand gebracht. Hierfür wurde das Training des Modells mit aktuellen Stellenausschreibungen weitergeführt. Dadurch konnte das Modell die Bedeutung von Wörtern, Modellen und Konzepten lernen (wie z. B. *generative KI*), die erst in den letzten Jahren populär geworden sind.

Durch die Überführung in Worteinbettungen lassen sich Kompetenzen gruppieren, die einander semantisch sehr ähnlich sind bzw. in ähnlichen Kontexten verwendet werden. Dafür wurden die Vektoren anhand ihrer Ähnlichkeit mit dem Clustering-Verfahren *k*-Means-Algorithmus gruppiert. Die Ähnlichkeit wurde mit der Cosinus-Ähnlichkeit gemessen. Der optimale Wert für *k* wurde mit der minimalen Summe der quadrierten Abstände jeder Beobachtung von ihrem Mittelpunkt bestimmt. Dabei hat sich eine Zahl von 300 Clustern als optimal herausgestellt. Die Ergebnisse der Clusterung wurden anschließend in Experteninterviews und -workshops validiert (siehe Kapitel 3 und 4). Basierend auf dem inhaltlichen Input zur praktischen Nähe der Kompetenzen innerhalb der Cluster wurde die Detailtiefe einzelner Skills-Cluster erweitert, indem diese in verschiedene Cluster aufgebrochen wurden oder mehrere Skills-Cluster zusammengelegt wurden. Die 300 datengetriebenen Cluster wurden unter Berücksichtigung ihrer inhaltlichen Nähe zu 40 Clustern kondensiert.

## 2.2.5 Forecasting der Kompetenzbedarfe

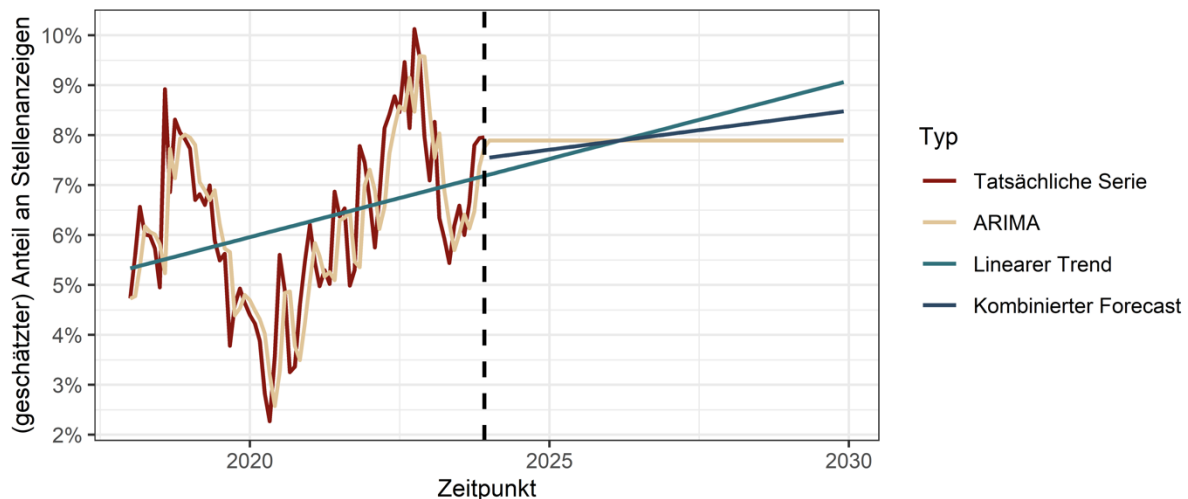
Zur Identifikation von zukünftig besonders relevanten Kompetenzen (Future Skills) wurde für jede identifizierte Kompetenz ein statistisches Zeitreihenmodell trainiert. Die Datengrundlage für jedes Zeitreihenmodell stellt der relative Anteil einer Kompetenz an allen Stellenanzeigen eines Monats im Zeitraum von Januar 2018 bis Dezember 2023 dar. Die Betrachtung des relativen Anteils wurde der Betrachtung der absoluten Anzahl vorgezogen, um für den insgesamt wachsenden Bestand an und den Einbruch von Stellenanzeigen während der Corona-Pandemie zu kontrollieren. Das statistische Zeitreihenmodell lernt systematische Informationen wie Trend und Saisonalität der betrachteten Kompetenz im Trainingszeitraum (2018-2023) und extrapoliert diese im Vorhersagezeitraum (2024-2030). Im Rahmen von modellbasierten Vorhersagen wird häufig auf zusätzliche exogene erklärende Variablen verzichtet, da auch diese im Vorhersagezeitraum zunächst vorhergesagt werden müssten, um anschließend die eigentliche Vorhersage der betrachteten Kompetenz durchzuführen. Diese doppelte Vorhersage würde die mit der Schätzung verbundenen Unsicherheit potenzieren. Folglich wurde auch im Rahmen der Vorhersage der Kompetenzbedarfe auf exogene Variablen (z. B. gesamtwirtschaftlicher Digitalisierungsfortschritt) verzichtet und stattdessen auf ein autoregressives Modell mit gleitenden Durchschnitten (ARIMA) sowie ein lineares Zeittrendmodell zurückgegriffen. Das ARIMA(*p*,*i*,*q*)-Modell hat folgende funktionale Form:

$$y_t = \underbrace{\varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p}}_{AR(p)} + \underbrace{\epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}}_{MA(q)}$$

wobei  $y_t$  die  $i$ -fach differenzierte Zeitreihe des relativen Anteils einer Kompetenz zum Zeitpunkt  $t$  darstellt. Die Differenzierung ist notwendig, um Stationarität und damit die unverzerrte Modellierung der Zeitreihe sicherzustellen. Der autoregressive Teil des Modells (AR) basiert auf der Überlegung, dass vergangene Werte der Zeitreihe bei der Schätzung der zukünftigen Werte hilfreich sein können. Ist für eine Kompetenz historisch zum Beispiel ein aufsteigender Trend zu beobachten, wird diese Information mittels des autoregressiven Teils in die Zukunft extrapoliert. Der gleitende Durchschnittsteil des Modells (MA) basiert auf der Überlegung, dass der aktuelle Wert einer Zeitreihe durch vergangene Prognosefehler beeinflusst werden kann. Die konkrete Ausprägung der autoregressiven Ordnung  $p$ , des gleitenden Durchschnitts  $q$  sowie der Differenzierung  $i$  sind als Modellparameter zu verstehen, die sich mittels Informationskriterien bestimmen lassen. Zusätzlich zum oben beschriebenen Modell wurde ein lineares Trendmodell geschätzt, das den relativen Anteil einer Kompetenz auf Basis des Zeitpunktes schätzt und damit den linearen Trend im Zeitverlauf abbildet. Als finaler Forecast wurde der gleichgewichtete Mittelwert aus linearem Trendmodell und ARIMA-Modell verwendet. Theoretisch bestünde die Möglichkeit, einen gewichteten Mittelwert zu verwenden, in dem die Gewichte des Trendmodell und des ARIMA-Modells durch eine zusätzliche Schätzung bestimmt werden. In der empirischen Anwendung schneiden kombinierte Modelle mit geschätzten Gewichten allerdings häufig schlechter ab als Modelle, die den einfachen Mittelwert zwischen verschiedenen Modellen verwenden. Dieser Befund wird unter dem Begriff des „Forecast Combination Puzzle“ geführt und ist in der Forschung zu Vorhersagemodellen bereits seit Jahrzehnten bekannt (siehe z. B. Bates und Granger 1969). Die folgende Abbildung visualisiert die gewählte Forecasting-Methodik anhand der Kompetenz „Analytisches Denken“.

## Abbildung 2: Beispielhafte Forecasting-Methodik

Kompetenz „Analytisches Denken“



Quelle: IW Consult 2024

Die rote Linie stellt die tatsächliche Serie dar. Man erkennt, dass der Anteil der Kompetenz Analytisches Denken zu Beginn der Zeitreihe knapp unterhalb von 5 % und zum Ende der Zeitreihe knapp unterhalb von 8 % lag. Zwischen diesen Randpunkten verhält sich die Zeitreihe dagegen wenig systematisch. Das ARIMA-Modell (gold) ist in der Lage den tatsächlichen Verlauf der Serie im Trainingszeitraum bis zum Ende 2023 gut nachzubilden,



das lineare Trendmodell (grün) ist deutlich weniger flexibel und bildet nur den insgesamt positiven Trend der Zeitreihe ab. Beide Modelle liefern unterschiedliche Prognosen für den Zeitraum von 2024 bis 2030: Während das ARIMA-Modell aufgrund der geringen Systematik im Trainingszeitraum annähernd ein Nullwachstum prognostiziert, schreibt das Trendmodell den erlernten Trend im Vorhersagefenster weiter fort. Der kombinierte Forecast ergibt sich als Mittelwert beider Ansätze und prognostiziert im Jahr 2030 einen Anteil von gut 8 % an allen Stellenanzeigen. Ausgehend von dem Anteil von knapp 7 % zum Ende des Trainingszeitraums entspricht dies einem Gesamtwachstum von gut 20 %.

Zur Sicherstellung einer hinreichend großen Datenbasis wurde die oben beschriebene Vorhersage auf Kompetenzen beschränkt, die im Trainingszeitraum in mindestens 50 Stellenanzeigen identifiziert wurden. Nach dieser Filterung lagen damit geschätzte Wachstumsraten bis 2030 für 3.311 Kompetenzen vor. Die berechneten Wachstumsraten wurden zusätzlich auf das Intervall von -100 % bis +100 % beschränkt. Ursächlich für diesen Ansatz ist die Tatsache, dass andernfalls aufgrund des langen Vorhersagezeitraums für bestimmte Kompetenzen sehr große absolute Wachstumsraten resultieren würden. Im Sinne einer konservativen Schätzung und um statistische Ausreißer zu entfernen, wurden alle Kompetenzen mit einem Wachstum unterhalb (oberhalb) von -100 % (+100 %) auf dem unteren (oberen) Limit von -100 % (+100 %) fixiert. Die untere Intervallgrenze impliziert, dass eine Kompetenz im Jahr 2030 keine Relevanz mehr hat, die obere Intervallgrenze sagt aus, dass sich die Relevanz einer Kompetenz bis 2030 mindestens verdoppeln wird. Da die verwendete Online-Stellenanzeigen-Datenbasis einer Vollerhebung aller nachgefragten Fähigkeiten in Baden-Württemberg gleicht, betreffen auch die geschätzten Wachstumsraten die Grundgesamtheit aller Unternehmen und müssen nicht mehr durch den Einsatz von Gewichtungsfaktoren repräsentativ hochgerechnet werden.

### 2.2.6 Future-Skills-Cluster

Die Vorhersage der Kompetenzbedarfe stellt nur den ersten von vier Schritten zur Identifizierung von Future-Skills-Clustern dar. Im zweiten Schritt wurden die Wachstumsraten der Kompetenzen zu Cluster-Wachstumsraten kondensiert. Anschließend wurde eine standardisierte Onlinebefragung unter Experten und in Baden-Württemberg ansässigen Unternehmen durchgeführt. In der Befragung haben über 200 Experten und Unternehmer ihre Einschätzung bzgl. der bis 2030 wichtigsten Skills-Cluster abgegeben. Im abschließenden Schritt wurden die berechneten Cluster-Wachstumsraten und die Experteneinschätzungen durch Verfahren der bayesianischen Statistik miteinander verbunden, das Resultat stellt die identifizierten Future-Skills-Cluster dar.

#### Bestimmung von Cluster-Wachstumsraten

Zunächst wurden die Wachstumsraten der Kompetenzen gemäß der in Kapitel 2.2.4 beschriebenen Clusterbildung zu 40 Clustern kondensiert. Hierbei wurde die unterschiedliche Bedeutung der verschiedenen Kompetenzen durch die Berechnung von mengengewichteten Wachstumsraten auf Clusterebene berücksichtigt. Die Vorgehensweise lässt sich gut an einem Beispiel verdeutlichen: Die Kompetenz Analytisches Denken gehört zusammen mit 18 weiteren Fähigkeiten wie etwa Lösungsorientierung, Problemlösungskompetenz und Durchhaltevermögen dem übergeordneten Cluster Problemlösungsfähigkeit an. Die Wachstumsraten dieser Kompetenzen wurden gemäß ihrem Anteil innerhalb der Stellenanzeigen im Trainingszeitraum zu einem gewichteten Mittelwert zusammengeführt. Die Kompetenz Analytisches Denken war etwa mit gut 70.000 Stellenanzeigen für knapp 26 %

aller Stellenanzeigen im Cluster Problemlösungsfähigkeiten verantwortlich. Die berechnete Wachstumsrate dieser Kompetenz von gut 20 % fließt folglich auch mit einem Gewicht von knapp 26 % in die Wachstumsrate des zugehörigen Clusters ein. Die Verwendung der beschriebenen Gewichte bietet gegenüber einem gleichgewichteten Ansatz zwei entscheidende Vorteile: Erstens wird der absoluten Bedeutung einer Kompetenz innerhalb eines Clusters Rechnung getragen. Zweitens haben es Kompetenzen mit einer gegenwärtig geringeren Bedeutung per se leichter, eine hohe prognostizierte Wachstumsrate zu erzielen, da das Potenzial von einem geringeren Ausgangsniveau bestimmt wird. Im Rahmen der gewichteten Durchschnittsbildung erhalten aktuell wenig bedeutsame Kompetenzen allerdings ein geringes Gewicht, wodurch die Wachstumsrate des Clusters nicht zu stark durch die hohen Wachstumsraten der besagten Kompetenzen beeinflusst wird. Die resultierenden geschätzten Wachstumsraten der Cluster haben aufgrund der beschriebenen Kondensation eine geringere Bandbreite und rangieren zwischen -48 % und +37 %.

### Unternehmer- und Expertenbefragung

Zur qualitativen Validierung der berechneten Cluster-Wachstumsraten wurde zudem eine standardisierte Unternehmens- und Expertenbefragung unter  $n = 217$  Unternehmen und Experten durchgeführt. Die Teilnehmer haben für alle 40 Cluster eine Einschätzung über deren Relevanz bis 2030 auf der Skala von -100 % bis +100 % gegeben. Die Ergebnisse der Befragung rangieren damit auf der gleichen Skala wie die Wachstumsrate auf Basis der Stellenanzeigen und werden im nächsten Schritt zur Plausibilisierung der berechneten Cluster-Wachstumsraten verwendet (siehe auch Kapitel 4.4 und 5).

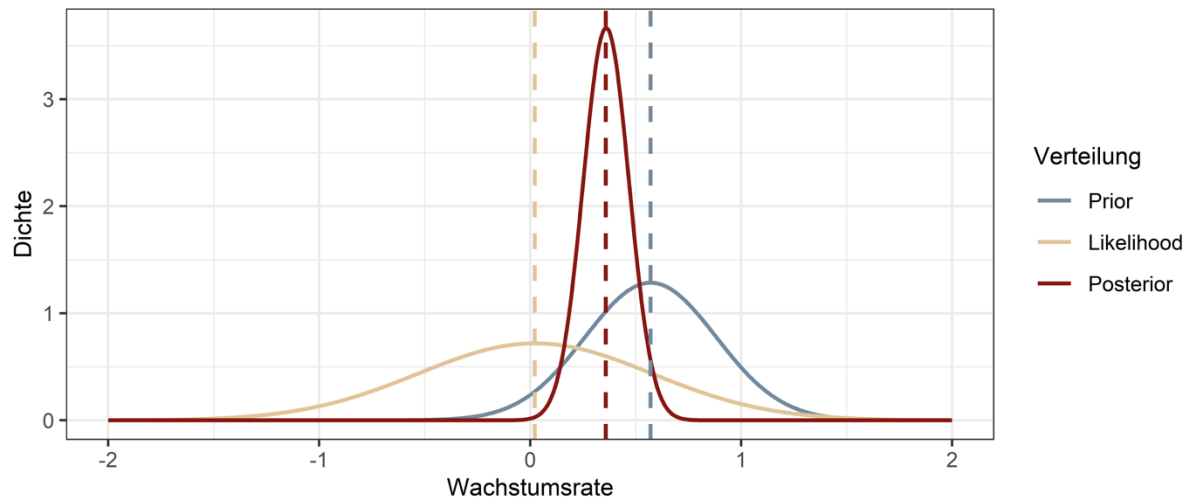
### Identifikation der Future-Skills-Cluster

Im letzten Schritt wurden beide Cluster-Wachstumsraten zu einer finalen Wachstumsrate vereint, wobei Cluster mit einer positiven finalen Wachstumsrate als sogenannte Future-Skills-Cluster bezeichnet wurden. Analog zur Verbindung der in Kapitel 2.2.5 beschriebenen Vorhersagemodelle und der Bestimmung der Cluster-Wachstumsraten kann ein gleichgewichteter oder ein gewichteter Mittelwert gebildet werden. In diesem Kontext erscheint die Berechnung eines gewichteten Mittelwerts sinnvoller, denn sowohl die Einschätzungen der Experten als auch die berechneten Wachstumsraten sind mit Unsicherheit behaftet, die sich in Form der Varianz ausdrücken lässt. Eine hohe Varianz der Experteneinschätzung zeugt von Uneinigkeit bzgl. der zukünftigen Bedeutung eines Clusters. Analog deutet eine hohe Varianz der berechneten Wachstumsraten aus den Stellenanzeigen darauf hin, dass die enthaltenen Kompetenzen eines bestimmten Clusters eine hohe Bandbreite an prognostizierten Wachstumsraten aufweisen. Aus diesem Grund wurden die Gewichte beider Bestandteile mittels Verfahren der bayesianischen Statistik bestimmt. Im Rahmen der bayesianischen Statistik wird eine vorherrschende Einschätzung mit Evidenz in Form von Daten konfrontiert und gemäß dem Satz der bedingten Wahrscheinlichkeit angepasst. Im Kontext der Future-Skills-Cluster lässt sich die Einschätzung der Experten als sogenannte A-priori-Einschätzung (auch Prior genannt) durch die berechneten Cluster-Wachstumsraten (auch Likelihood genannt) aktualisieren. Die resultierende A-posteriori-Wachstumsrate (auch Posterior genannt) ist als gewichteter Mittelwert aus Prior und Likelihood zu interpretieren, wobei das Gewicht von Prior respektive Likelihood umso größer ist, je geringer die Varianz der Experteneinschätzung respektive der Kompetenzwachstumsraten aus den Stellenanzeigen. Anders ausgedrückt: Besteht unter den Experten eine hohe Einigkeit oder in den Daten eine geringe Varianz bzgl. der zukünftigen Bedeutung eines Clusters, fließt die entsprechende

Wachstumsrate mit einem höheren Gewicht in die finale Wachstumsrate ein, als wenn eine hohe Uneinigkeit (große Varianz) besteht. Die folgende Abbildung visualisiert das bayesianische Vorgehen anhand des bereits bekannten Clusters der Problemlösungsfähigkeiten.

### Abbildung 3: Beispielhafte Identifikation der Future-Skills-Cluster

Cluster „Problemlösungsfähigkeit“



Quelle: IW Consult 2024

Folgende Observationen fallen auf: Die auf Basis der Stellenanzeigen berechnete Wachstumsrate (Likelihood) fällt deutlich geringer aus als die Einschätzung der Experten (Prior). Gleichzeitig weist die Einschätzung der Experten deutlich weniger Unsicherheit auf, denn die blaue Verteilung ist erkennbar stärker um den gestrichelten Mittelwert konzentriert als die goldene Verteilung. Kombiniert man die blaue und die goldene Verteilung gemäß dem Satz der bedingten Wahrscheinlichkeit von Bayes folgt die rote Verteilung (Posterior). Der Mittelwert dieser Verteilung setzt sich aufgrund des beschriebenen Unsicherheitsverhältnisses zu einem größeren Teil aus der Experteneinschätzung (61 %) als aus den Stellenanzeigen (39 %) zusammen. Insgesamt wird für das Cluster Problemlösungsfähigkeit bis 2030 eine Wachstumsrate von 36 % (0,36) prognostiziert. Die Verteilungen der restlichen Cluster finden sich im methodischen Anhang der Studie.

## 3 Leitfadengestützte Interviews

### 3.1 Ziel und Zweck im Forschungsdesign

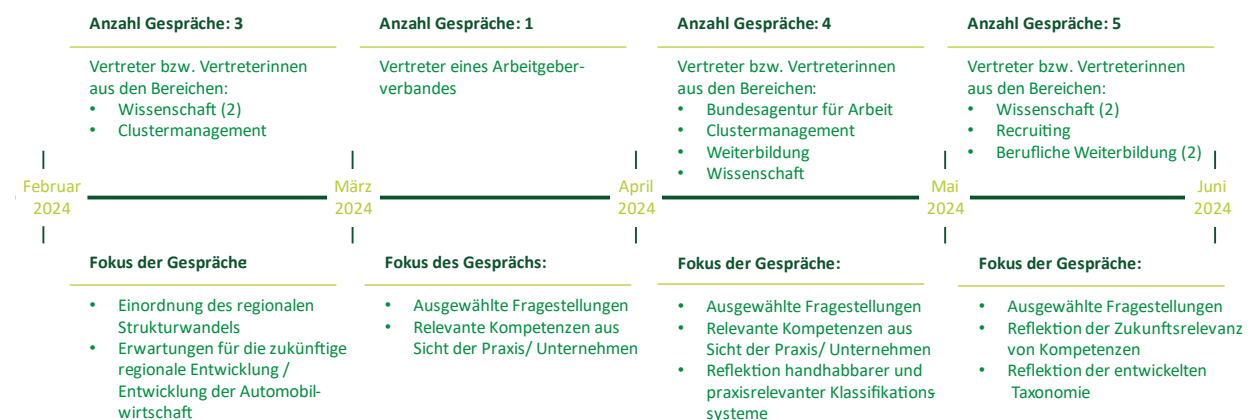
Zur Analyse der Auswirkungen der automobilen Transformation auf die benötigten Kompetenzen wird externe Fachexpertise einbezogen. Die Expertengespräche verfolgten unterschiedliche Ziele im Projektverlauf. Zu Beginn des Projektes erlaubten sie, das Studiensetting in den regionalen Kontext zu verorten und regionale Besonderheiten der Automobilwirtschaft von Anfang an mitzuberücksichtigen. Ergänzt wurde die erste Projektphase um generelle Expertise in den Bereichen Kompetenzforschung und -entwicklung. Mit Voranschreiten des Projektverlaufs sowie unter Berücksichtigung des Erkenntnisprozesses wurden die Gespräche genutzt, um Zukunftskompetenzen mit den Gesprächspartnern und -partnerinnen zu identifizieren, zu diskutieren und schließlich auch die Ergebnisse der Online-Stellenanzeigen-Analyse (vgl. Kapitel 2) zu hinterfragen.

### 3.2 Auswahl der Gesprächspartner und -partnerinnen

Insgesamt wurden 13 leitfadengestützte Interviews mit insgesamt 15 Gesprächspartnern und -partnerinnen geführt. Die Gespräche dauerten zwischen 45 und 60 Minuten und wurden digital mithilfe der Software Microsoft Teams oder face-to-face durchgeführt. Die Auswahl der Gesprächspartner und -partnerinnen erfolgte unter Berücksichtigung der Projektphase, des Erkenntnisinteresses und der mit den Gesprächen verbundenen Ziele:

#### Abbildung 4: Übersicht zu den durchgeführten leitfadengestützten Interviews

Ablauf, Anzahl und Gesprächsinhalte



Quelle: IW Consult 2024

Im Februar 2024 konkretisierte das durchführende Projektteam die regionale Problemstellung. Gespräche mit drei regionalen Akteuren bzw. Fachexperten für den automobilen Strukturwandel in Baden-Württemberg und in der Region Nordschwarzwald konnten das Verständnis für die regionalen automobilspezifischen (Wirtschafts-) Strukturen erhöhen. Insbesondere die Diskussion der von den Gesprächspartnern und -partnerinnen antizipierten Entwicklungspfade für die ansässigen Unternehmen der Automobil- und Zulieferindustrie stellte für die Einordnung der datenbasierten Identifikation von Kompetenzen relevante Anhaltspunkte dar.

- Fünf Gespräche im März und April fundierten die datenbasierte Kompetenzextraktion. Die Gespräche trugen dazu bei, die extrahierten Kompetenzen in einer praxisrelevanten und handhabbaren Taxonomie zu systematisieren. Zudem konnten ausgewählte

Fragestellungen, die für die weitere Kompetenzdarstellung und -erläuterung relevant sind, diskutiert werden und infolgedessen Argumentationswege und Erläuterungen geschärft werden.

- Die im Mai durchgeführten fünf Gespräche erfolgten zu einem fortgeschrittenen Zeitpunkt der Datenerhebung und -analyse. Sie boten daher die Gelegenheit, insbesondere die Zukunftsrelevanz ausgewählter Skills-Cluster bzw. Kompetenzen sowie die bisherige methodische Vorgehensweise zu reflektieren.

### 3.3 Leitfaden und Gesprächsinhalte

Mit den Interviews sollten das datenbasierte Vorgehen punktuell gestützt, die Datenanalyse der Online-Stellenanzeigen justiert sowie die Interpretation und Einordnung der datenbasierten Ergebnisse angereichert werden. Vor diesem Hintergrund wurden je nach Projektphase und Gesprächspartner bzw. -partnerin die Themenblöcke in den Interviews individuell gewählt. Zudem bot das Gesprächssetting Freiraum für die individuelle Schwerpunktsetzung durch die Interviewees, um ihrer jeweiligen Expertise Rechnung zu tragen.

Berücksichtigt wurden folgende Themenbereiche:

- Betroffenheit vom automobilen Strukturwandel
  - ▷ Betroffenheit der Automobilbranche
  - ▷ Betroffenheit auf Unternehmensebene bzw. auf Ebene von Unternehmensbereichen
  - ▷ Betroffenheit der Belegschaft
  - ▷ Einfluss des automobilen Strukturwandels auf Berufsbilder/Berufe in der Automobilbranche
- Einfluss der Digitalisierung auf Kompetenzbedarfe
  - ▷ Größte Bedarfe im Bereich digitaler Kompetenzen
  - ▷ Verfügbarkeit digitaler Kompetenzen auf dem Arbeitsmarkt und Fehlstellen
  - ▷ Unterschiede in den digitalen Kompetenzbedarfen nach Unternehmenstypen, Geschäfts- und Unternehmensbereichen sowie Qualifikationsniveau/Tätigkeiten der Belegschaften
- Einfluss des automobilen Strukturwandels auf Kompetenzbedarfe
  - ▷ Wichtige Kompetenzen für die Bewältigung und Gestaltung des automobilen Strukturwandels
  - ▷ Verfügbarkeit dieser Kompetenzen auf dem Arbeitsmarkt und Fehlstellen
  - ▷ Unterschiede nach Unternehmenstypen, Geschäfts- und Unternehmensbereichen sowie Qualifikationsniveau/Tätigkeiten der Belegschaften
- Kompetenzforschung
  - ▷ Restriktionen des methodischen Vorgehens und der Datengrundlage
  - ▷ Konzeption und Verständnis des Begriffs „Zukunftskompetenz“
  - ▷ Verhältnis von Fachkompetenzen und überfachlichen Fähigkeiten



## 4 Gruppendifkussionen im Workshop-Format

### 4.1 Ziel und Zweck im Forschungsdesign

Die Autoren der Studie gehen davon aus, dass der Wandel in der Automobilindustrie die Kompetenzanforderungen in der Branche sowie bei Zulieferern, Dienstleistern und in der Infrastruktur beeinflusst. Um die Auswirkungen des automobilen Strukturwandels auf die Belegschaften in Unternehmen der Automobil- und Zulieferindustrie angemessen zu berücksichtigen sowie die Zukunftsrelevanz von Kompetenzen zu erörtern, ergänzen Gruppendifkussionen im Workshop-Format das empirische Erhebungsdesign.

Gruppendifkussionen im Workshop-Format bereichern die datengetriebene Analyse. Mit Blick auf die Arbeitsphase der Teilnehmenden lassen sich empirische Arbeitsschritte validieren, wenn Teilnehmende zu ähnlichen Ergebnissen kommen. Darüber hinaus profitiert die datengetriebene Analyse von den unterschiedlichen Erfahrungen und Fachkenntnissen der Teilnehmenden. Dies erhöht das Verständnis für die Frage- und Problemstellung. Gruppendifkussionen ermöglichen es, eine breite Palette unterschiedlicher Perspektiven zu berücksichtigen. Durch den direkten Vergleich unterschiedlicher Standpunkte können Widersprüche aufgedeckt werden. Zudem erlauben der offene Austausch und die direkte Kommunikation, einen Konsens zu finden und so die Relevanz und Akzeptanz der Ergebnisse insgesamt zu stärken.

Vor dem Hintergrund, dass die Workshops zu unterschiedlichen Zeitpunkten und in unterschiedlichen Settings mit Blick auf die Teilnehmenden, Methoden und Inhalte durchgeführt wurden, sind mit ihnen unterschiedliche Ziele verbunden. Nachfolgend werden die drei durchgeführten Workshops vorgestellt sowie ihr Impact auf die datengetriebene Analyse von Online-Stellenanzeigen, die Entwicklung einer Kompetenztaxonomie sowie für die Beurteilung der Zukunftsrelevanz der identifizierten Future-Skills-Cluster vorgestellt.

### Abbildung 5: Übersicht zu den durchgeführten Workshops

Ablauf, Teilnehmende und Inhalte

Workshop 1	Workshop 2	Workshop 3
<p>Vertreter bzw. Vertreterinnen aus den Bereichen:</p> <ul style="list-style-type: none"> <li>• Arbeitgebervertretung</li> <li>• Arbeitnehmervertretung</li> <li>• Wissenschaft</li> <li>• Unternehmerschaft</li> </ul>	<p>Vertreter bzw. Vertreterinnen aus den Bereichen:</p> <ul style="list-style-type: none"> <li>• Arbeitgebervertretung</li> <li>• Arbeitnehmervertretung</li> <li>• Clustermanagement</li> <li>• Unternehmerschaft</li> </ul>	<p>Vertreter bzw. Vertreterinnen aus den Bereichen:</p> <ul style="list-style-type: none"> <li>• Wissenschaft</li> <li>• Arbeitgebervertretung</li> </ul>
März 2024	April 2024	Mai 2024
		Juni 2024
<b>Fokus des Workshops:</b>	<b>Fokus des Workshops:</b>	<b>Fokus des Workshops:</b>
Diskussion und Brainstorming notwendiger Kompetenzen für die Gestaltung des automobilen Strukturwandels	Diskussion der Kompetenz-Taxonomie mit besonderem Fokus auf Fachkompetenzen	Diskussion der Zukunftsrelevanz der identifizierten Skills-Cluster

Quelle: IW Consult 2024

## 4.2 Workshop 1 am 22. März 2024 in Stuttgart

- Zweck

- ▷ Einbindung regionaler Akteure in die datenbasierte Analyse
- ▷ Akzeptanz- und Vertrauensbildung in die Ergebnisse

- Ziele

- ▷ Identifikation relevanter Kompetenzen für die Gestaltung des automobilen Strukturwandels
- ▷ Entwicklung erster Ansätze einer Clusterung von Kompetenzen
- ▷ Diskussion der Zukunftsrelevanz der identifizierten Kompetenzen für den automobilen Strukturwandel

- Teilnehmende

- ▷ Anzahl: 11
- ▷ Vertreter und Vertreterinnen aus den Bereichen Arbeitgeber- und Arbeitnehmervertretung, Automobilunternehmen, Branchenverband, Wissenschaft (5)
- ▷ Vertreter und Vertreterinnen des Transformationsnetzwerks (3)
- ▷ Vertreter und Vertreterinnen des Auftragnehmers (3)

- Methoden und Ablauf

- ▷ Vormittag des 22. März 2024 in Stuttgart
- ▷ Input-Vortrag sowie verschiedene Formen des Brainstormings und der Diskussion

- Ergebnisse/ Impact auf die Analyse der Online-Stellenanzeigen

- ▷ Identifikation von Kompetenzen insbesondere im Bereich der digitalen und überfachlichen Kompetenzen
- ▷ Ergänzung des Wörterbuchs als Grundlage für die semantische Analyse um weitere 73 Begriffe
- ▷ Berücksichtigung der Erkenntnisse aus den Diskussionen bei der Interpretation der Ergebnisse

## 4.3 Workshop 2 am 25. April 2024 in Stuttgart

- Zweck

- ▷ Einbindung regionaler Unternehmensvertreter und -vertreterinnen in die datenbasierte Analyse und Aufbereitung der Ergebnisse
- ▷ Akzeptanz- und Vertrauensbildung in die Ergebnisse

- Ziele

- ▷ Prüfung der Zwischenergebnisse, insbesondere die Kompetenztaxonomie, auf ihre Praxisrelevanz
- ▷ Erhalt von Impulsen für die Finalisierung der Kompetenztaxonomie, insbesondere mit Fokus auf Fachkompetenzen
- ▷ Einschätzungen zur Zukunftsrelevanz ausgewählter Kompetenzen und Skills-Cluster, insbesondere mit Fokus auf Fachkompetenzen

- **Teilnehmende**

- ▷ Anzahl: 22
- ▷ Vertreter und Vertreterinnen regionaler Unternehmen der Automobil- und Zulieferindustrie, der Arbeitnehmervertretungen sowie des Clustermanagements (18)
- ▷ Vertreter und Vertreterinnen des Transformationsnetzwerks (2)
- ▷ Vertreter und Vertreterinnen des Auftragnehmers (2)

- **Methoden und Ablauf**

- ▷ Nachmittag des 24. April 2024 in Stuttgart
- ▷ Input-Vortrag sowie Gruppenformate zur Bearbeitung ausgewählter Schwerpunktthemen
- ▷ Moderierte Diskussion im Plenum

- **Ergebnisse/ Impact auf die Analyse der Online-Stellenanzeigen**

- ▷ Anpassung der Skills-Cluster durch das Auflösen einzelner Cluster und Differenzierung einzelner Cluster
- ▷ Anpassung der Zuordnung der Kompetenzen zu Skills-Cluster
- ▷ Impulse für die schriftliche Ergebnisdarstellung sowie weiterer Auswertungslogiken

#### 4.4 Workshop 3 am 08. Mai 2024 in digitaler Form

- **Zweck**

- ▷ Externe Validierung der Kompetenztaxonomie
- ▷ Akzeptanzbildung in die Ergebnisse in der wissenschaftlichen Community

- **Ziele**

- ▷ Erhalt von Impulsen für die Finalisierung der Kompetenztaxonomie, insbesondere mit Blick auf die schriftliche Berichtlegung
- ▷ Einschätzungen zur Zukunftsrelevanz ausgewählter Kompetenzen und Skills-Cluster

- **Teilnehmende**

- ▷ Anzahl: 9
- ▷ Vertreter und Vertreterinnen aus der Wissenschaft (5)
- ▷ Vertreterin eines Arbeitgeberverbandes (1)
- ▷ Vertreterinnen des Transformationsnetzwerks (2)
- ▷ Vertreter und Vertreterinnen des Auftragnehmers (2)

- **Methoden und Ablauf**

- ▷ Mittag des 08. Mai 2024, digital
- ▷ Input-Vortrag, Mentimeter-Abfrage, moderierte Diskussion im Plenum

- **Ergebnisse/ Impact auf die Analyse der Online-Stellenanzeigen**

- ▷ Berücksichtigung der Einschätzungen zur Zukunftsrelevanz von Skills-Clustern im Forecast-Modell
- ▷ Impulse für die schriftliche Ergebnisdarstellung sowie weiterer Auswertungslogiken

## 5 Standardisierte Befragung

### 5.1 Ziel und Zweck im Forschungsdesign

Um den automobilen Wandel zu analysieren, liefert die Online-Stellenanzeigen-Analyse eine empirisch fundierte Annäherung an die heutigen Kompetenzbedarfe der Unternehmen. Eine Vorausschau auf zukünftig relevante Kompetenzen ist jedoch mit Unsicherheiten verbunden. Mit dem Ziel, diese Unsicherheiten zu reduzieren, reichert die Unternehmensbefragung das Forecasting der Kompetenzbedarfe an. Die datenbasierte Extrapolation (siehe Kapitel 2.2.4 und 2.2.5) wird so um Einschätzungen aus der Praxis justiert.

### 5.2 Stichprobe

Die 217 Rückmeldungen speisen sich zum Großteil aus Industrieunternehmen aus Baden-Württemberg sowie zusätzlich ausgewählter Experten, die am dritten Workshop teilgenommen haben (siehe Kapitel 4.4).

Unternehmensvertreter und -vertreterinnen wurden über folgende Kanäle angesprochen:

- Mittels der Unternehmensdatenbank beDirect wurden Unternehmen in Baden-Württemberg aus den Bereichen Metallindustrie, Elektroindustrie, Herstellung von Kraftfahrzeugen identifiziert und über einen E-Mail-Verteiler zur Teilnahme eingeladen.
- Über die AgenturQ wurde die Befragung unter den Mitgliedsunternehmen der IG Metall Baden-Württemberg sowie Südwestmetall gestreut.

Zu 106 Unternehmen liegen Strukturdaten vor. Knapp die Hälfte (49,2 Prozent) können der Metallindustrie zugerechnet werden, 29,5 Prozent dem Maschinenbau. Als Unternehmen der Elektroindustrie bezeichnen sich 16,4 Prozent und 4,9 Prozent sind in der Herstellung von Kraftfahrzeugen tätig. 61 Unternehmen machten Angaben zu ihrer Unternehmensgröße: 34,0 Prozent haben 249 oder weniger Beschäftigte, 26,4 Prozent weniger als 1.000 Beschäftigte und 39,6 Prozent der teilnehmenden Unternehmen 1.000 und mehr Beschäftigte.

### 5.3 Fragebogenkonzeption

Die standardisierte Online-Befragung wurde von der IW Consult konzeptioniert, durchgeführt und auf eigenen Servern gehostet. Der Befragungszeitraum erstreckte sich auf den Mai 2024.

Die Befragung forderte die Teilnehmenden dazu auf, die Zukunftsrelevanz der Skills-Cluster einzuschätzen. Dafür wurde ihnen ein Wertebereich von -10 bis +10 angeboten. Die Zuweisung des Wertes -10 bedeutet, dass im Jahr 2030 Kompetenzen im jeweiligen Skills-Cluster irrelevant sein werden. Der Wert 0 indiziert, dass die Relevanz unverändert zur heutigen Relevanz bleiben wird. Ein Wert in Höhe von +10 spiegelt die Einschätzung wider, dass die Relevanz des zu bewertenden Skills-Clusters doppelt so hoch sein wird wie heute.

Beurteilt wurden folgende Skills-Cluster:

Kompetenzen im Bereich Technologie und Digitalisierung:

- Cloud- und IT-Infrastruktur
- Data Analytics
- Data Management
- Grundlegende IT-Fähigkeiten
- IT-Systemsicherheit

- Künstliche Intelligenz
- Programmierung
- Sensorik und IOT
- Softwarearchitektur
- Softwareentwicklung
- Robotik

#### Industrielle Kompetenzen:

- Alternativer Automobylantrieb
- Konventioneller Automobylantrieb
- Autonomes Fahren
- Electrical Engineering
- Emissionsfreie Produktion
- Fahrzeugbau und Montage
- Forschung und Entwicklung
- Industrial Engineering
- Qualitätssicherung und Dokumentation
- Technisches Grundverständnis
- Wartung, Reparatur und Instandhaltung

#### Überfachliche Kompetenzen:

- Eigeninitiative
- Flexibilität
- Innovatives Denken
- Kollaboration
- Kundenorientierung
- Organisationsfähigkeit
- Problemlösungsfähigkeit
- Resilienz
- Sprachkenntnisse
- Zielorientierung

#### Kompetenzen zur Sicherstellung zentraler Geschäftsprozesse:

- Beschaffung und Lieferkettenmanagement
- Betriebliches Ressourcenmanagement
- Logistik, Distribution und Materialflussmanagement
- Marketing und Unternehmenskommunikation
- Personalmanagement
- Projektmanagement, Unternehmensführung und Leadership
- Vertrieb und Kundenbeziehungsmanagement

Die Ergebnisse der Befragung gingen, wie in Kapitel 2.2.6 erläutert, in die Vorhersage der Kompetenzbedarfe ein.



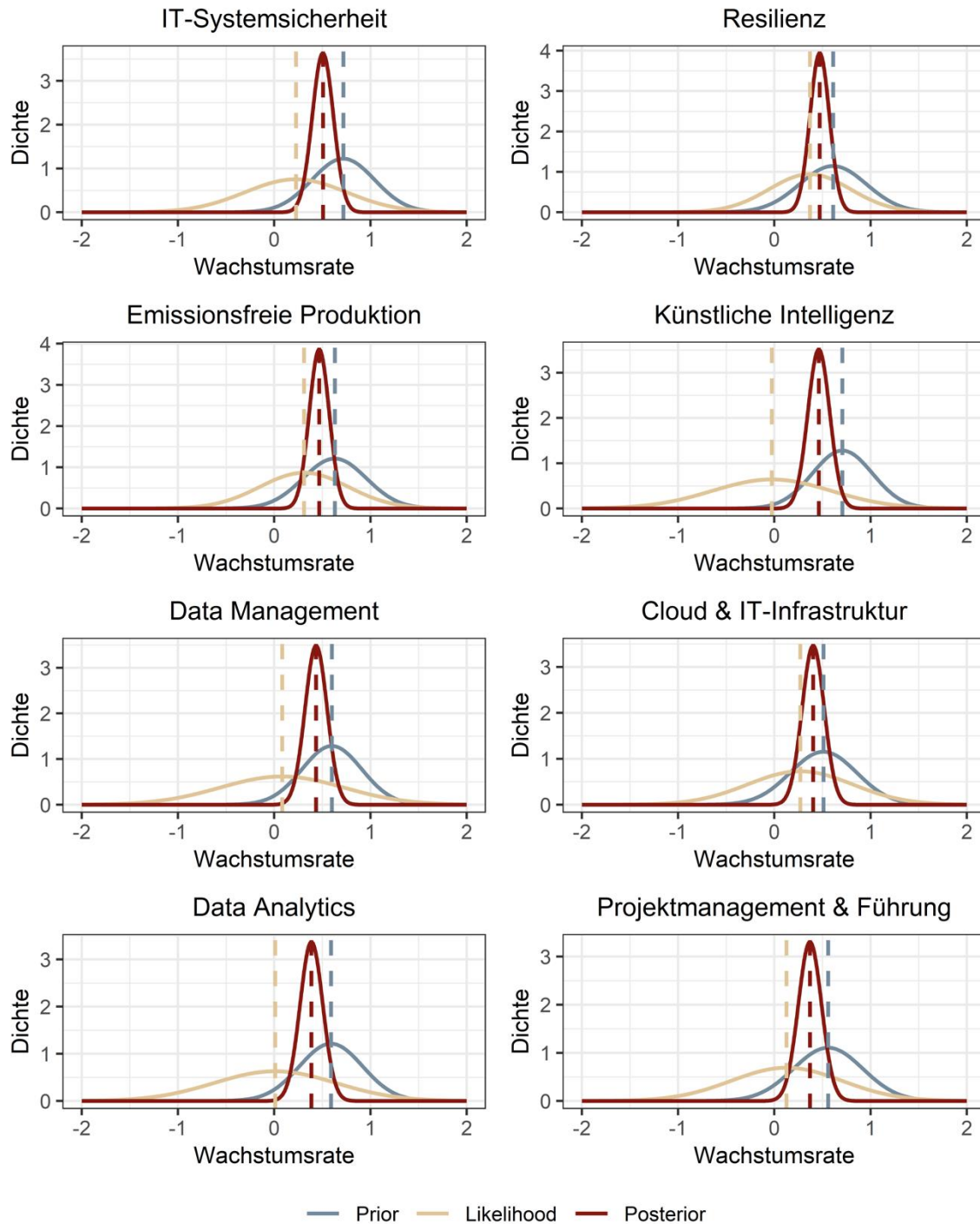
# Anhang

## A.1 Ergänzende Abbildungen

### Abbildung 6: Cluster 1-8

Absteigend sortiert nach Posterior Wachstumsrate

Prior (Expertenschätzung), Likelihood (Zeitreihenmodell), Posterior (kombinierte Schätzung)

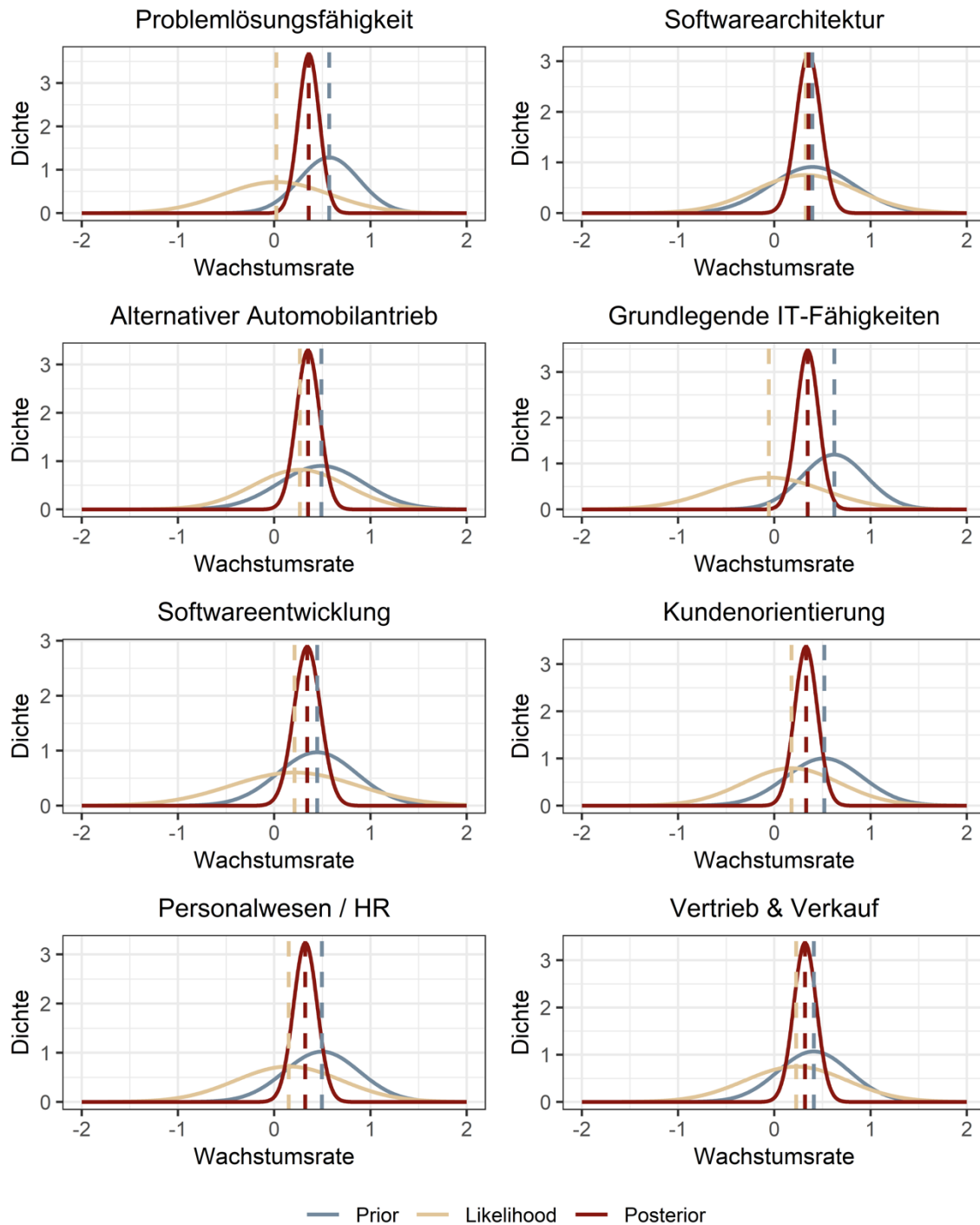


Quelle: IW Consult 2024

## Abbildung 7: Cluster 9-16

Absteigend sortiert nach Posterior Wachstumsrate

Prior (Expertenschätzung), Likelihood (Zeitreihenmodell), Posterior (kombinierte Schätzung)

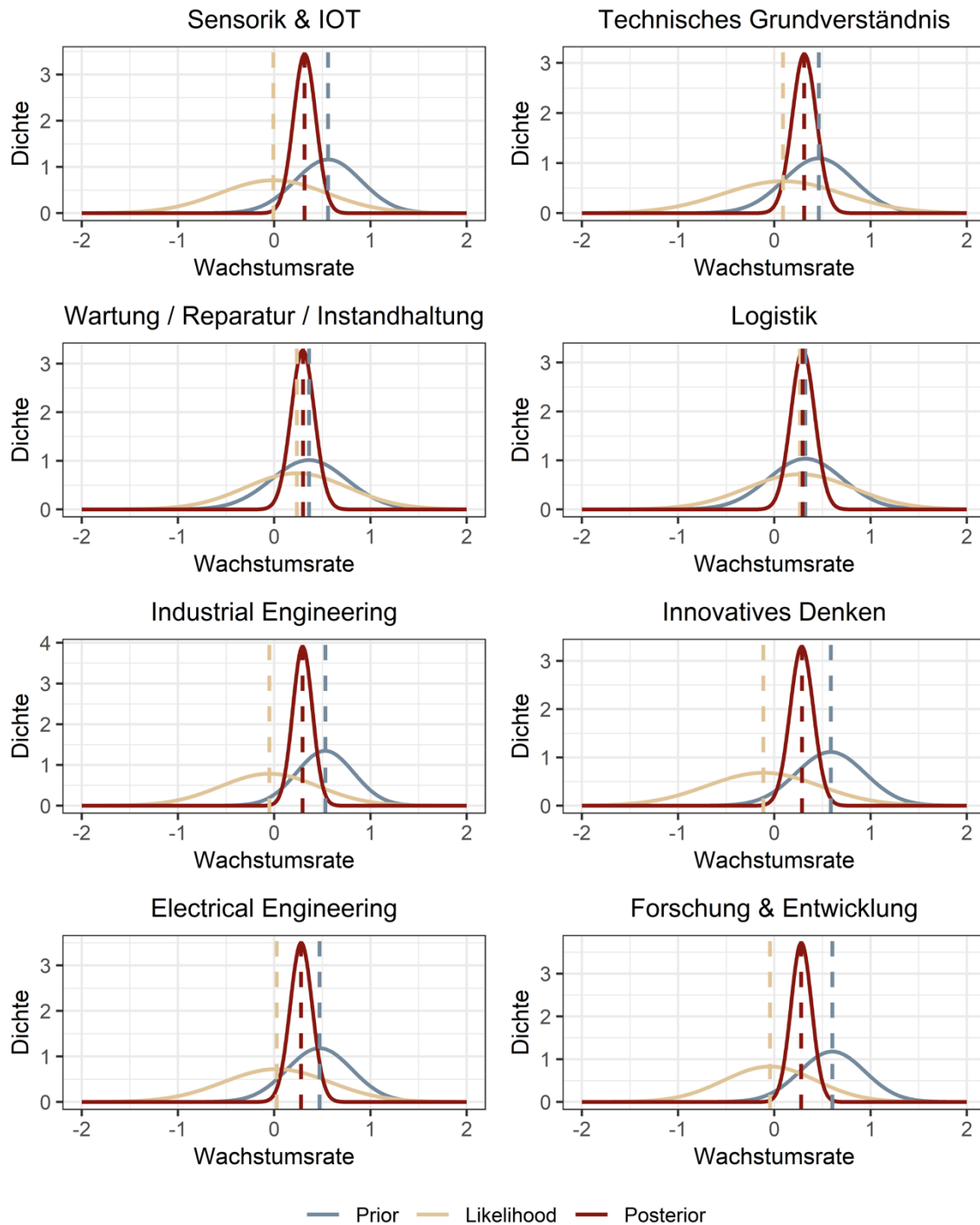


Quelle: IW Consult 2024

## Abbildung 8: Cluster 17-24

Absteigend sortiert nach Posterior Wachstumsrate

Prior (Expertenschätzung), Likelihood (Zeitreihenmodell), Posterior (kombinierte Schätzung)

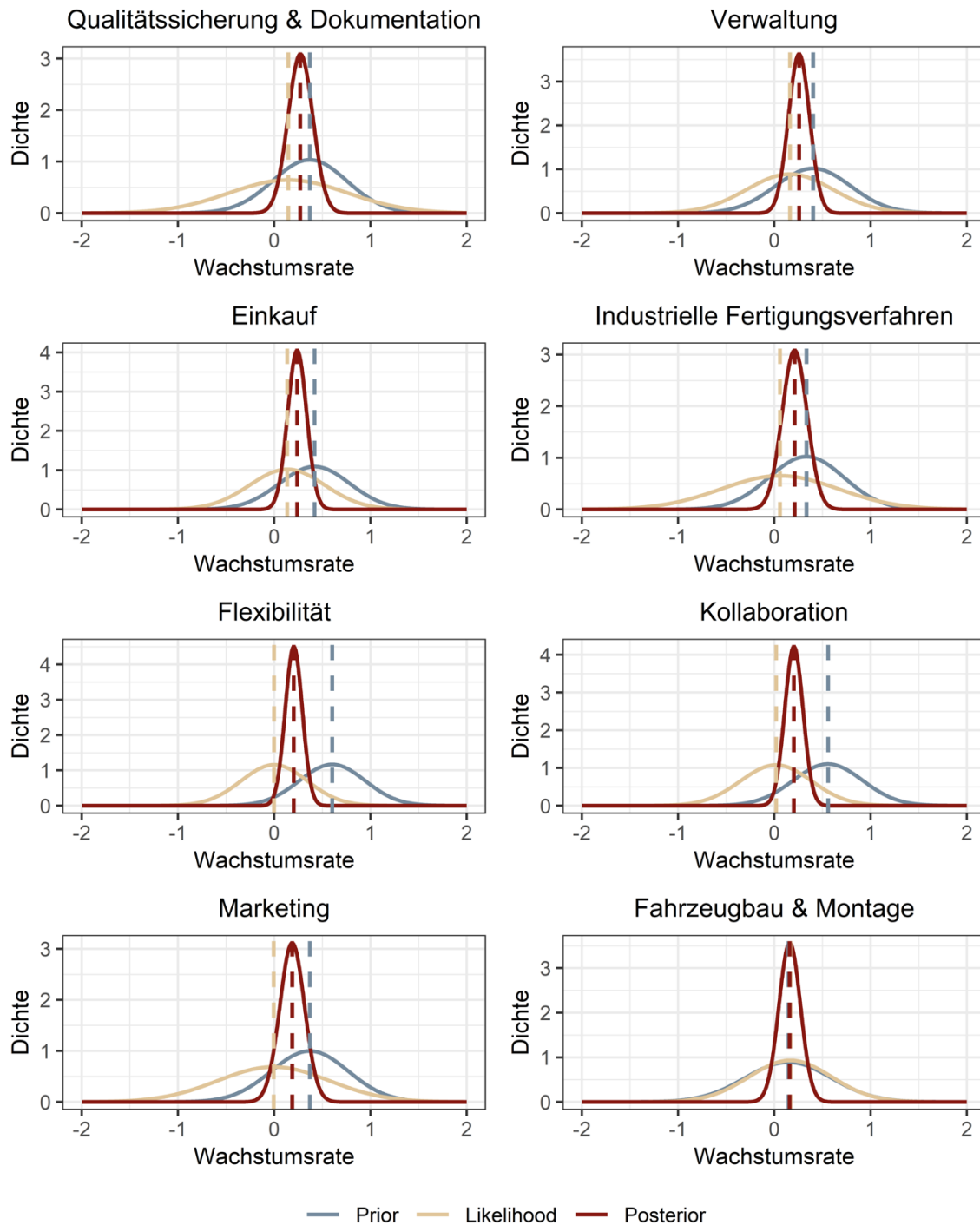


Quelle: IW Consult 2024

## Abbildung 9: Cluster 25-32

Absteigend sortiert nach Posterior Wachstumsrate

Prior (Expertenschätzung), Likelihood (Zeitreihenmodell), Posterior (kombinierte Schätzung)

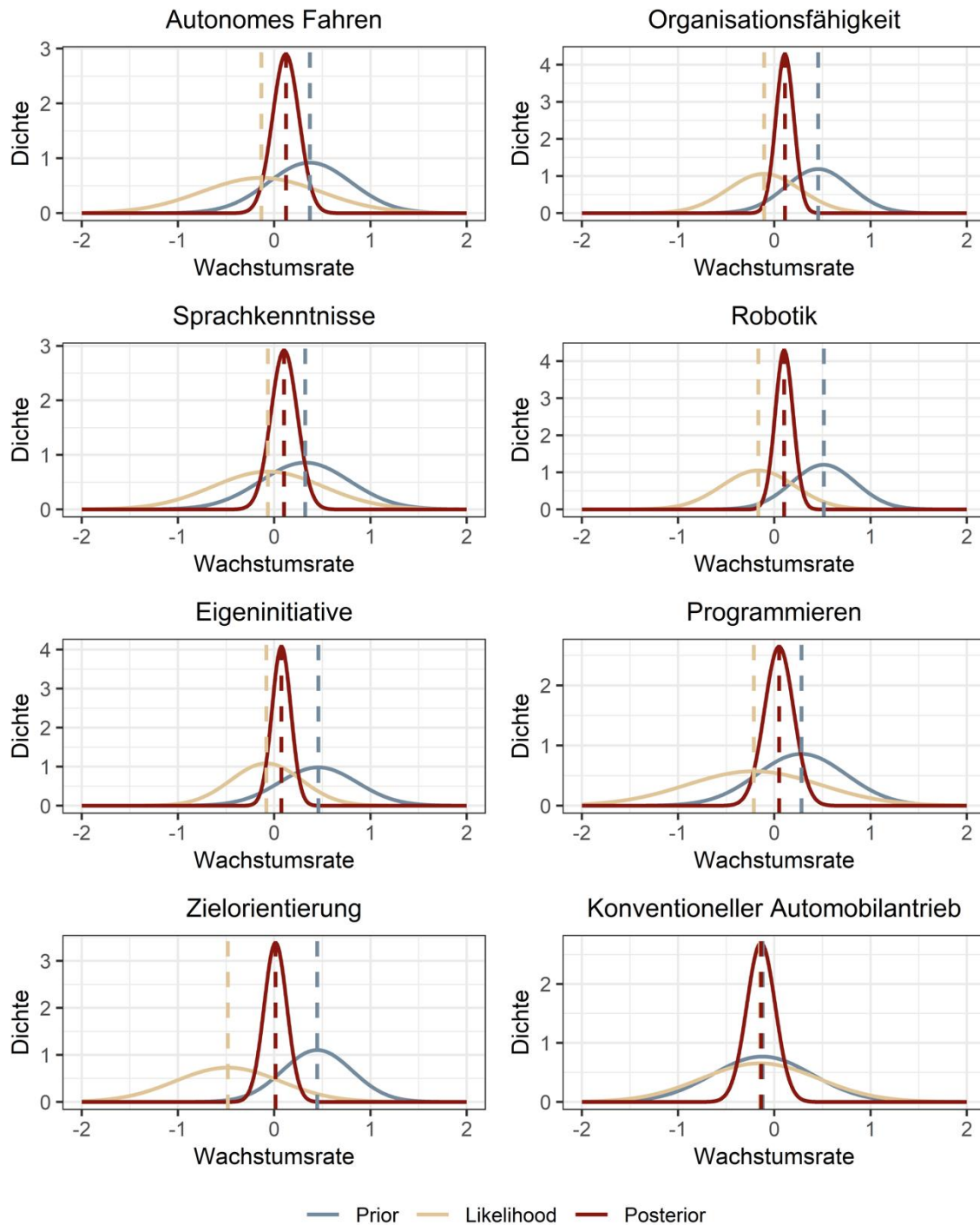


Quelle: IW Consult 2024

## Abbildung 10: Cluster 33-40

Absteigend sortiert nach Posterior Wachstumsrate.

Prior (Expertenschätzung), Likelihood (Zeitreihenmodell), Posterior (kombinierte Schätzung)



Quelle: IW Consult 2024



## A.2 Quellen- und Datenverzeichnis

Acemoglu, D.; Autor D.; Hazell J.; Resrepo P. (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies, in: Journal of Labor Economics, 40(1), S. 293-340

Bates J.; Granger C. (1969). The combination of forecasts, in: Operational Research Quarterly, 20(4), S. 451-468

Binnewitt J.; Schnepf T. (2021). Ökologische Nachhaltigkeit als Attraktivitätsdimension der Berufsausbildung– Extraktion von Aspekten der „Green Economy“ aus Online-Stellenanzeigen. Link:

[https://www.bibb.de/dokumente/pdf/AGBFN\\_Nachhaltigkeit\\_Praes\\_3.a.2\\_Binnewitt\\_Schnepf\\_Attraktivitaet.pdf](https://www.bibb.de/dokumente/pdf/AGBFN_Nachhaltigkeit_Praes_3.a.2_Binnewitt_Schnepf_Attraktivitaet.pdf)

Bojanowski et al. (2017). Enriching Word Vectors with Subword Information. Link: <https://arxiv.org/abs/1607.04606>

Borgnovoi et al. (2023). Emerging trends in AI Skill demand across 14 OECD countries, in: OECD Artificial Intelligence Papers, 2, Link: [https://www.oecd-ilibrary.org/science-and-technology/emerging-trends-in-ai-skill-demand-across-14-oecd-countries\\_7c691b9a-en](https://www.oecd-ilibrary.org/science-and-technology/emerging-trends-in-ai-skill-demand-across-14-oecd-countries_7c691b9a-en)

Buchmann et al. (2022). Swiss Job Market Monitor. A Rich Source of Demand-Side Micro Data of the Labour Market, in: European Sociological Review, 38(6), S. 1001-1014

Bundesagentur für Arbeit (2024). Link: <https://web.arbeitsagentur.de/berufenet/>

Büchel J.; Engler J. (2024). Generative KI in Deutschland: Künstliche Intelligenz in Gesellschaft und Unternehmen. Link: <https://www.iwkoeln.de/studien/jan-buechel-jan-felix-engler-kuenstliche-intelligenz-in-gesellschaft-und-unternehmen.html>

Büchel J.; Mertens A. (2022). IW-Gutachten: KI-Bedarfe in Deutschland. Link: <https://www.iwkoeln.de/studien/jan-buechel-armin-mertens-regionale-analyse-und-entwicklung-der-anforderungsprofile-in-ki-stellenanzeigen.html>

Büchel J.; Mertens A. (2021). IW-Gutachten: KI-Bedarfe der Wirtschaft am Standort Deutschland. Eine Analyse von Stellenanzeigen für KI-Berufe. Link: <https://www.de.digital/DIGITAL/Redaktion/DE/Digitalisierungsindex/Publikationen/publikation-download-ki-bedarfe-wirtschaft.pdf?blob=publicationFile&v=5>

Härpfer J.; Neuhauser P. (2021). Einführung der „Klassifikation der Berufe 2010 – überarbeitete Fassung 2020“ in die Arbeitsmarktstatistiken. Link: <https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010-Fassung2020/Publikationen/Publikationen-Nav.html;jsessionid=618900AE3B837E1CFE34E52E6DECDBFF>

Kempermann et al. (2021). IW-Gutachten: Wirtschaftliche Bedeutung regionaler Automobilnetzwerke in Deutschland. Link: <https://www.iwkoeln.de/studien/hanno-kempermann-johannes-ewald-manuel-fritsch-oliver-koppel-benita-zink-wirtschaftliche-bedeutung-regionaler-automobilnetzwerke-in-deutschland.html>

Khaouja I.; Kassou I.; Ghogo M. (2021). A Survey on Skill Identification From Online Job Ads, in: IEEE Access, 9, S. 118134-118153

Mertens et al. (2023). IW-Gutachten: Kompetenzen und Berufe in der Nahrungs- und Genussmittelindustrie. Link: <https://www.iwkoeln.de/studien/armin-mertens-alexander->

[burstedde-henry-goecke-christoph-metzler-nicole-straefling-kompetenzen-und-berufe-in-der-nahrungs-und-genussmittelindustrie.html](#)

Statistisches Bundesamt (2008). Klassifikation der Wirtschaftszweige. Link: <https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/klassifikation-wz-2008.html>

Textkernel (2024). Link: <https://www.textkernel.com/de/produkte-loesungen/skills-intelligence/>

### A.3 Abbildungsverzeichnis

Abbildung 1: Kompetenztaxonomie .....	8
Abbildung 2: Beispielhafte Forecasting-Methodik .....	13
Abbildung 3: Beispielhafte Identifikation der Future-Skills-Cluster .....	16
Abbildung 4: Übersicht zu den durchgeführten leitfadengestützten Interviews.....	17
Abbildung 5: Übersicht zu den durchgeführten Workshops .....	19
Abbildung 6: Cluster 1-8.....	24
Abbildung 7: Cluster 9-16 .....	25
Abbildung 8: Cluster 17-24 .....	26
Abbildung 9: Cluster 25-32 .....	27
Abbildung 11: Cluster 33-40 .....	28

### A.4 Tabellenverzeichnis

Tabelle 1: Anzahl der Stellenausschreibungen in der Automobil- und Zulieferindustrie nach WZ-Klassen.....	5
Tabelle 2: Anzahl der Stellenausschreibungen in der Automobil- und Zulieferindustrie nach WZ-Klassen.....	6
Tabelle 3: Anzahl der Stellenausschreibungen in der Metall-, Maschinenbau- und Elektroindustrie .....	7
Tabelle 4: Anzahl der Kompetenzen in den Stellenanzeigen .....	10
Tabelle 5: Evaluationsmetriken.....	11