ACOD
AUTOMOTIVE CLUSTER OSTDEUTSCHLAND GmbH

Handreichung

10. Januar 2018 | KC: Aus- und Weiterbildung

Kompetenzanforderungen an Data Scientisten im Shopfloor der Automobilindustrie

Handlungsempfehlungen für die ostdeutsche Automobilindustrie

Ersteller:
Dr. Felix Erler
Helena Rockser

Inhaltsverzeichnis

1 Ziel der Handreichung ... 3
1.1 Fragestellung ... 3
1.2 Methodik der Auswertung ... 3

2 Kompetenzanforderungen an Data Scientisten im Shopfloor ... 3
2.1 Analyse der Weiterbildungs- und Zertifizierungskurse von Bildungsträgern ... 4
2.2 Anforderungen aus Sicht der Data Scientisten und der Automobilindustrie ... 7
2.3 Bedeutung von Technologien ... 9
2.4 Bedeutung von Algorithmen ... 9

3 Handlungsempfehlungen ... 10
3.1 Marktlücke bei Aus- und Weiterbildung von produktionsnahen Data Scientisten ... 10
3.2 Katalog von Inhalten zur Aus- und Weiterbildung von Data Scientisten für den Shopfloor ... 10
3.2.1 Fünf notwendige Inhalte ... 10
3.2.2 Vier mögliche zusätzliche Inhalte ... 11
3.2.3 Anforderungen an Data Scientisten aus Sicht von Zulieferern ... 12

4 Kurzfazit ... 14

Abbildungsverzeichnis

Abbildung 1: Übersicht Methodik der Analyse ... 3
Abbildung 2: Charakteristische Aufgaben eines Data Scientisten ... 4
Abbildung 3: Bekanntheitsgrad des Berufs Data Scientist in der Automobilindustrie ... 4
Abbildung 4: 13 Weiterbildungs- und Zertifizierungskurse zu „Data Science“ ... 5
Abbildung 5: Beispielhafte Mindestvoraussetzungen der Bildungsangebote ... 5
Abbildung 6: Relevanz von Kompetenzen aus Sicht von Data Scientisten und Automobilzulieferern absolut ... 6
Abbildung 7: Relevanz von Kompetenzen aus Sicht von Data Scientisten und Automobilzulieferern prozentual ... 7
Abbildung 8: Einschätzung von Data Scientisten zu Studienabschluss vs. Weiterbildungs- und Zertifizierungskurse ... 7
Abbildung 9: Bedeutung eines Data Scientisten im Unternehmen in der Automobilindustrie ... 8
Abbildung 10: Bekanntheitsgrad fünf ausgewählter Technologien ... 9
Abbildung 11: Herkunft von Algorithmen in den Unternehmen der Automobilindustrie ... 10

Tabellenverzeichnis

Tabelle 1: 24 Arbeitsaufgaben & Fähigkeiten von Data Scientisten ... 8
Tabelle 2: Anforderungen an Data Scientisten aus Sicht von Zulieferern ... 12
Tabelle 3: Anforderungen an Data Scientisten - Anwenden, Kennen und Wissen ... 13

1 Ziel der Handreichung

1.1 Fragestellung

Weitreichende Änderungen sind durch die „Digitalisierung im Prozess“ zu erwarten, die auf das Leistungsprofil der gesamten Wertschöpfungskette Auswirkungen haben. Um das erforderliche Leistungsprofil vollumfänglich abbilden zu können, sind zusätzliche IT-Kompetenzen nötig, insbesondere im Umgang mit Produktionsdaten. Vor diesem Hintergrund stellt sich die zentrale Frage:

Wie sehen die Anforderungen von Unternehmen an das Berufsbild Data Scientist in der Automobilproduktion aus?

Welche Qualifikationsinhalte benötigen Shopfloor-Mitarbeiter, um als Data Scientist aktiv zu werden?

1.2 Methodik der Auswertung

Dazu wurden im ersten Schritt **(1) Qualifizierungsangebote zum Data Scientist** von 13 Bildungsträgern in Deutschland mittels Internetrecherche erhoben und Zulassungsvoraussetzungen, Dauer sowie Studieninhalte untersucht. Aus den angebotenen Studieninhalten entstand eine **(2) Shortlist mit zwölf Fachkompetenzen**. Anschließend sollte die Shortlist mit der Praxis abgeglichen werden, um daraus resultierende zukünftige Anforderungsprofile von Fachkräften zu beschreiben. Daher wurde im dritten Schritt ein **(3) leitfadengestütztes Interview** für einerseits bereits arbeitende Data Scientisten – 13 Interviewpartner - sowie andererseits bei 16 Unternehmen der Automobilzulieferindustrie in Sachsen, Sachsen-Anhalt und Thüringen durchgeführt. Damit konnten die Arbeitsaufgaben der Data Scientisten abgeglichen und Anforderungen der Unternehmen gegenübergestellt bzw. Lücken aufgedeckt werden. Anschließend wurden die gewonnenen Forschungsergebnisse mittels **(4) Kompetenzkatalogen** aggregiert dargestellt.

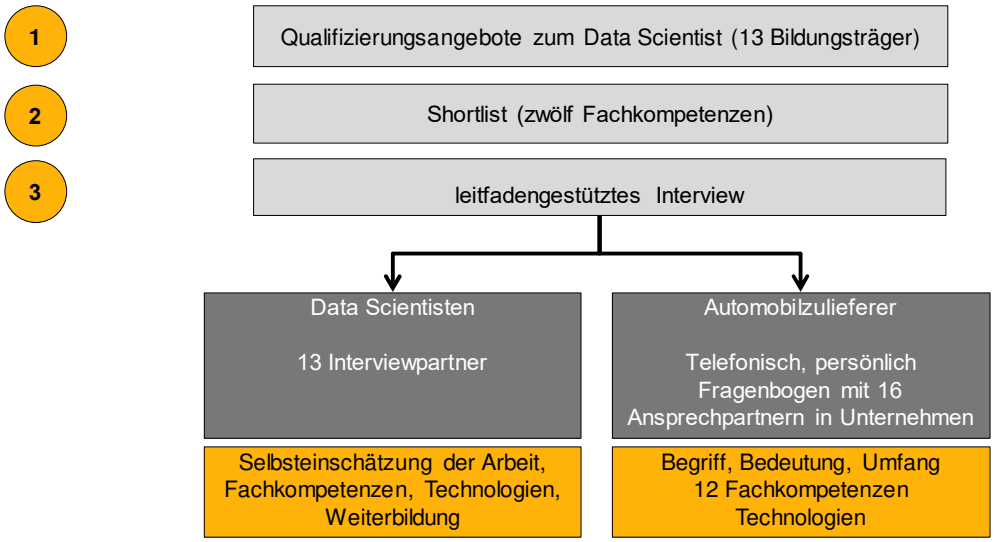

Abbildung 1: Übersicht Methodik der Analyse

2 Kompetenzanforderungen an Data Scientisten im Shopfloor

Nachdem vor allem in den USA der Data Scientist aus der Finanzwirtschaft bekannt ist, setzt sich der Beruf des Data Scientist zunehmend in Europa durch – wobei der Beruf aus Sicht der

Automobilindustrie eher auf industriellen Mehrwert abzielt. Deshalb werden Personen, die sich damit befassen, entweder als Data Scientist, Data Analyst, Data Engineer und Data Statistician bezeichnet. Jedoch weisen diese Berufsbezeichnungen Unterschiede und Gemeinsamkeiten auf, wie bspw. im Jobportal Stepstone zu sehen ist. Daher werden die Unterschiede und Gemeinsamkeiten der Berufe in der Abbildung 2 kurz zusammengefasst. Um eine eindeutige Abgrenzung der Arbeitsaufgaben eines Data Scientisten vorzunehmen, werden im Folgenden charakteristische Aufgaben dargestellt.

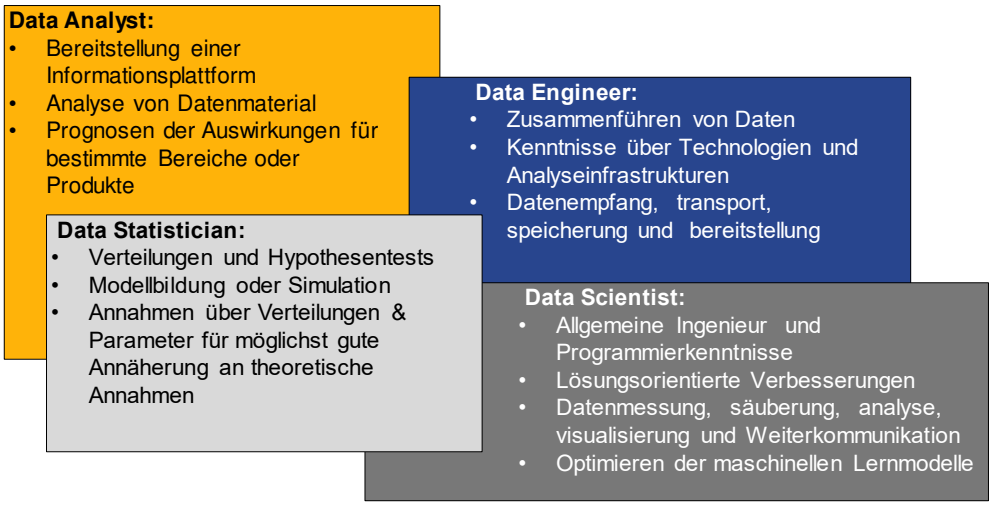

Abbildung 2: Charakteristische Aufgaben eines Data Scientisten

Insgesamt 63% der befragten Unternehmen gaben an, den Beruf des Data Scientist zu kennen und konnten Anforderungen in diesem Berufsprofil benennen. Die anderen 37% gaben an, mit diesem Beruf noch nicht in Verbindung gekommen zu sein (siehe Abbildung 3).

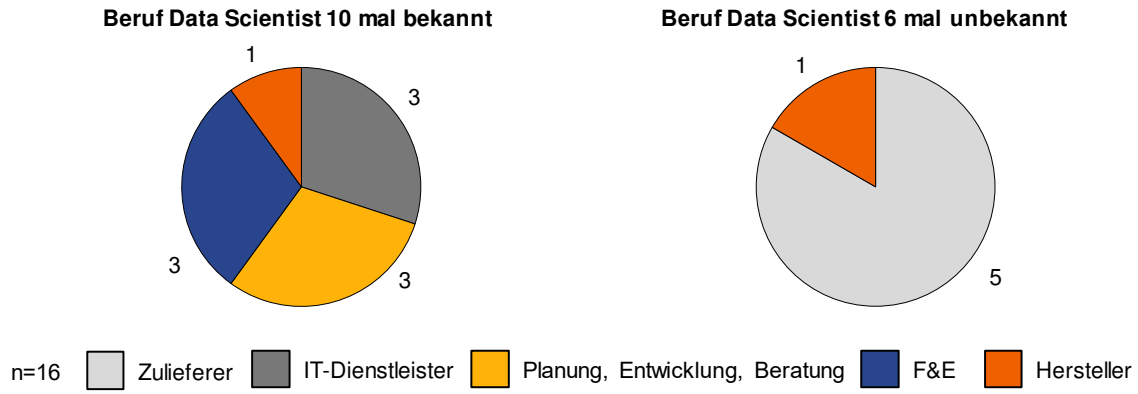

Abbildung 3: Bekanntheitsgrad des Berufs Data Scientist in der Automobilindustrie

2.1 Analyse der Weiterbildungs- und Zertifizierungskurse von Bildungsträgern

Zum Zeitpunkt der Untersuchung gab es 13 Bildungsträger in Deutschland, welche ihre Module mit dazugehörigen Vorlesungen und Lerninhalten im Bereich Data Science angeboten haben. In Ostdeutschland sind vier dieser Anbieter aktiv (siehe Abbildung 4).

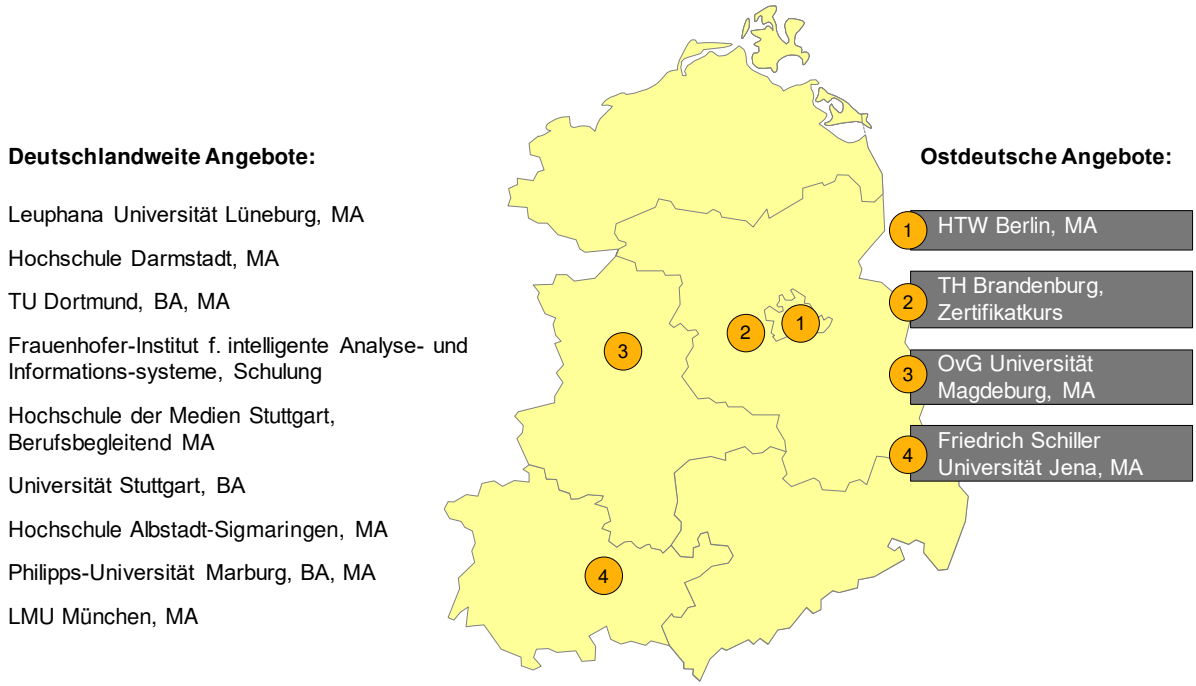

Abbildung 4: 13 Weiterbildungs- und Zertifizierungskurse zu „Data Science“

Abbildung 5 zeigt eine Aufteilung der Weiterbildungsangebote nach der Art der Angebote, ihren Inhalten, der Dauer und den Mindestvoraussetzungen. Damit soll verdeutlicht werden, dass sich die Angebote der Bildungsträger zum größten Teil an hochqualifizierte Personen mit akademischer Vorbildung richten. Fundierte theoretische Fachkenntnisse in Mathematik, Informatik und Statistik mit geeigneter berufliche Erfahrung bilden eine Basis für die Qualifizierung zum Data Scientist. Dies macht umso mehr deutlich, dass eine Vereinfachung für die Shopfloor-Berufe unabdingbar ist. Auch Englischkenntnisse gehören häufig zu den Mindestvoraussetzungen, denn bei zehn von 13 Bildungsträgern findet die Weiterbildung auf Englisch statt. Dem gegenüber geben jedoch 77% der tätigen Data Scientisten an, dass Englisch für ihren Beruf zwar wünschenswert ist jedoch als nicht notwendig angesehen wird.

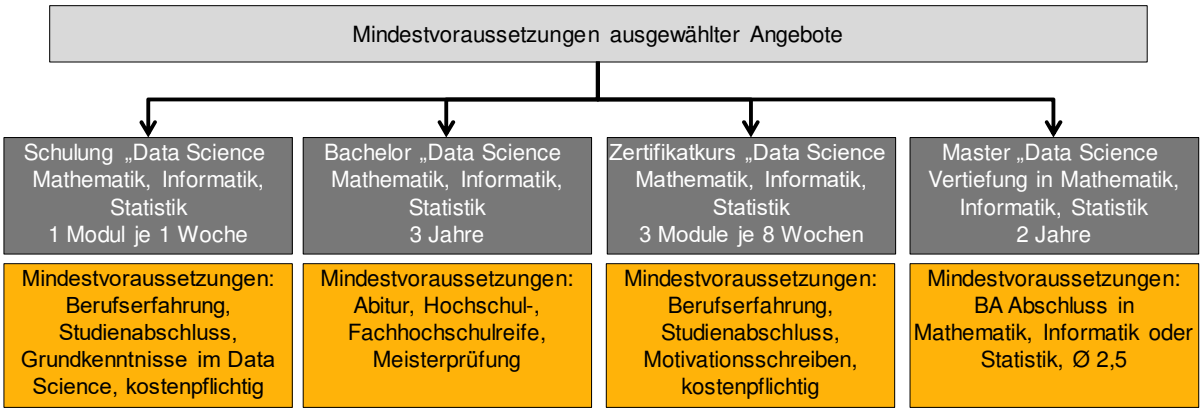

Abbildung 5: Beispielhafte Mindestvoraussetzungen der Bildungsangebote

Nun wurde das fachliche Angebot der Bildungsträger analysiert. Im Ergebnis wurde eine Shortlist mit zwölf Kompetenzen erstellt, welche von allen Bildungsträgern angeboten werden. Es handelt sich um folgende zwölf Kompetenzen:

(1) Datenintegration, Datenhaltung und Datenqualität
(2) Datenanalyse und Wissensmanagement
(3) Datenvisualisierung und Informationsbereitstellung

(4) Codierung
(5) Fortgeschrittene Programmiertechniken
(6) Methoden des Data Mining
(7) Algorithmen
(8) Software und Programmierung
(9) IT-Sicherheit, Ethik und Gesetze
(10) Logische Programmierung
(11) Funktionale Programmierung
(12) Geometrisches Modellieren

Das Ergebnis macht umso mehr deutlich, dass sich das derzeitige Angebot ausschließlich an hochqualifiziertes Personal richtet. Doch wie schätzen einerseits bereits arbeitende Data Scientisten und andererseits Automobilzulieferer in Sachsen, Sachsen-Anhalt und Thüringen das fachliche Angebot der Bildungsträger ein?

Die nachfolgende Abbildung 6 zeigt, wie relevant die zwölf Kompetenzen der Shortlist eingeschätzt werden. Tätige Data Scientisten und Experten aus der Automobilzuliefererlandschaft schätzen die einzelnen vorgegebenen Kompetenzen mit einer Abweichung von < 10% gleich wichtig ein. Daraus lässt sich vermuten, dass die Arbeitsaufgaben der Data Scientisten und die Anforderungen der Unternehmen der genannten Kompetenzen weitgehend deckungsgleich scheinen. Auffällig ist, dass fünf Kompetenzen als besonderes wichtig eingeschätzt werden: **(1) Datenanalyse & Wissensmanagement, (2) Datenintegration, -haltung & -qualität, (3) Datenvisualisierung & Informations-bereitstellung, (4) Methoden des Data Mining** und **(5) Algorithmen**.

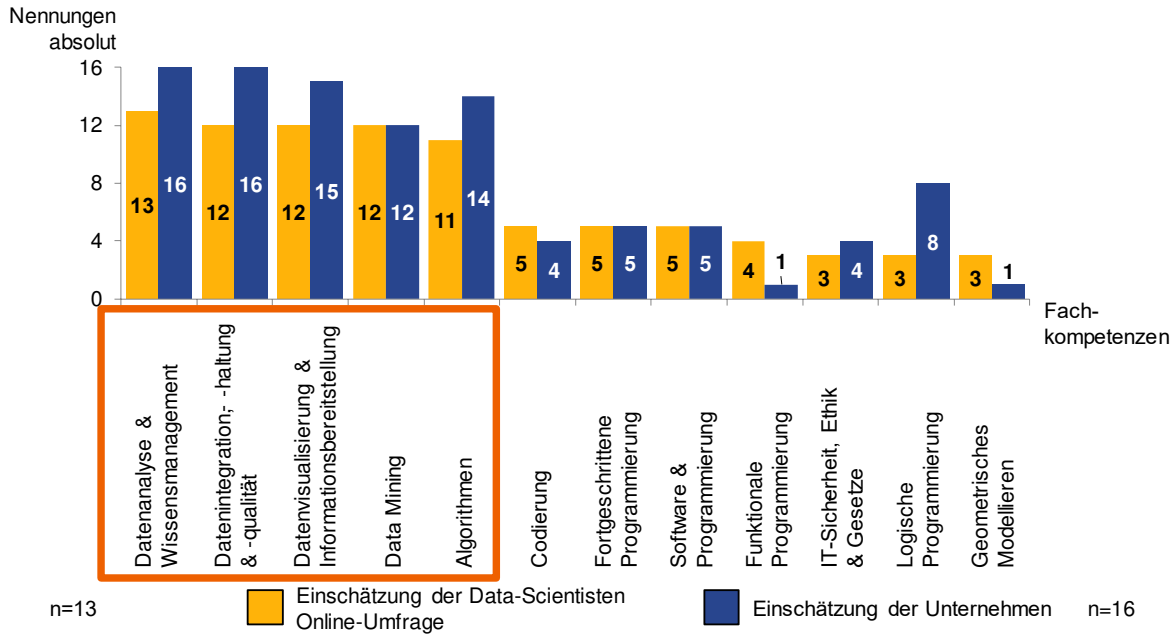

Abbildung 6: Relevanz von Kompetenzen aus Sicht von Data Scientisten und Automobilzulieferern absolut

Die Ergebnisse lassen vermuten, dass neben den fünf besonders wichtigen Kompetenzen weitere sieben Kompetenzen nur für bestimmte Bereiche der Automobilindustrie relevant sind, bspw. F&E, IT-Dienstleistungen und/oder Produktionsplanung (siehe Abbildung 7).

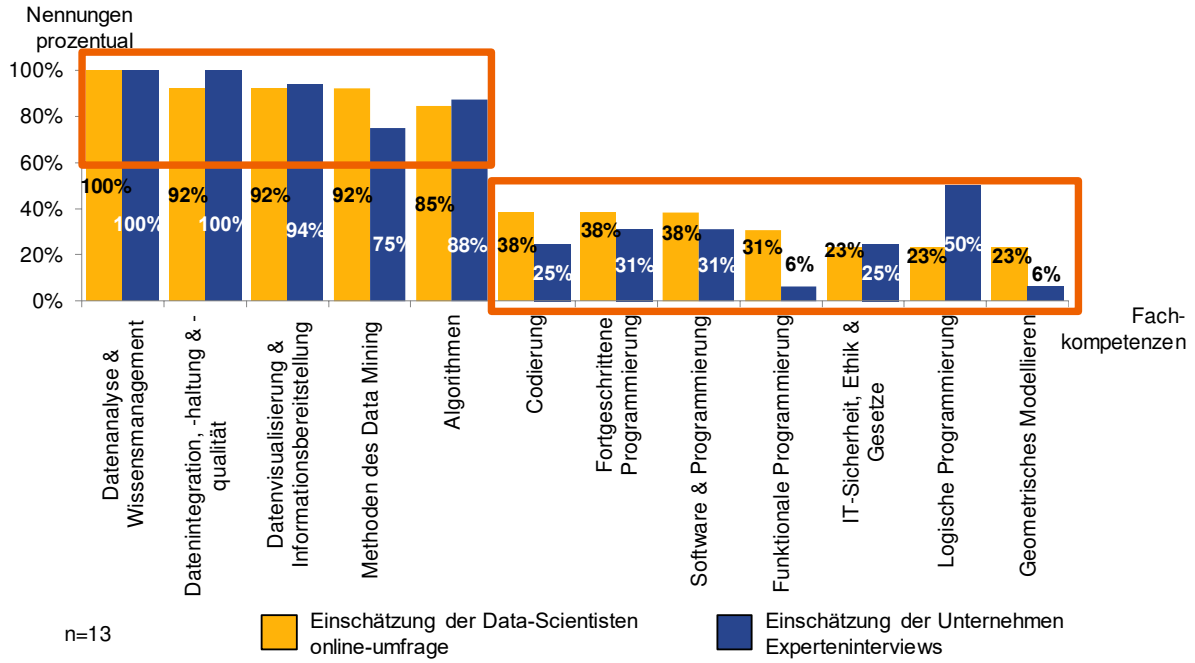

Abbildung 7: Relevanz von Kompetenzen aus Sicht von Data Scientisten und Automobilzulieferern prozentual

Die Frage danach, ob für Facharbeiter ein Studienabschluss oder Zertifikationskurs nötig ist, um den Beruf Data Scientist zu erlernen, haben die Befragten Data Scientisten wie folgt beantwortet. Weiterbildungs- und Zertifizierungskurse sind für 69% der Befragten ausreichend. Dagegen ist aus der Sicht von 31% der Befragten ein Studienabschluss erforderlich.

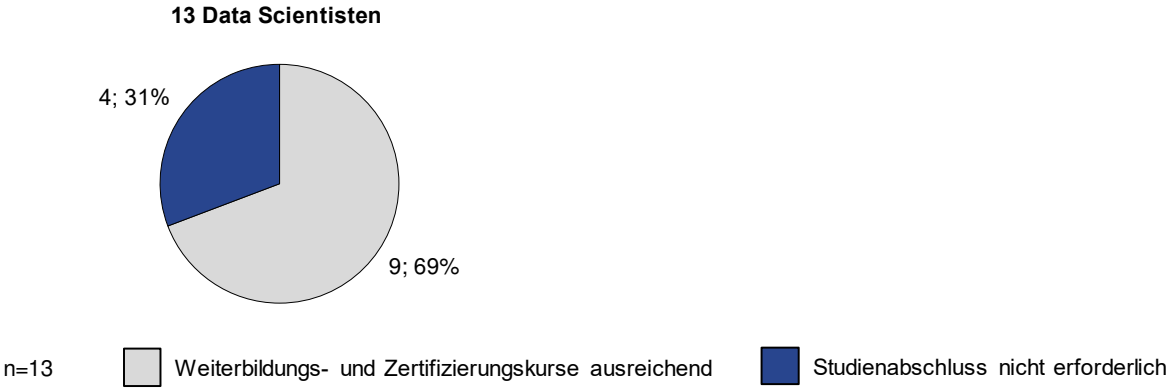

Abbildung 8: Einschätzung von Data Scientisten zu Studienabschluss vs. Weiterbildungs- und Zertifizierungskurse

Darüber hinaus fallen Kurs- bzw. Studiengebühren in unterschiedlicher Höhe an. Die Hochschule Albstadt-Sigmaringen erhebt 18.580 € für sechs Semester. Die TH Brandenburg verlangt 1.999 € pro Modul. Im Fraunhofer-Institut für intelligente Analyse- und Informationssysteme liegt der Preis für eine Basisschulung bei 4.950 €.

2.2 Anforderungen aus Sicht der Data Scientisten und der Automobilindustrie

Anschließend wurden die Anforderungen an einen Data Scientist aus Sicht der tätigen Data Scientisten und der Automobilindustrie betrachtet. Gerade die Automobilzulieferindustrie schätzt die Bedeutung eines Data Scientisten im Unternehmen als bedeutend ein. 31% der Befragten schätzen diese als sehr hoch ein, gefolgt von 44% als hoch (siehe Abbildung 9). Dies verdeutlicht umso mehr, dass der Digitalisierungsgrad in den Unternehmen stetig steigt, diese auf Daten angewiesen sind und gleichzeitig auch Mehrwerte in den Daten gesehen

werden. Eine hohe Bedeutung der Data Scientisten geht natürlich auch mit einer entsprechenden Qualifizierung einher.

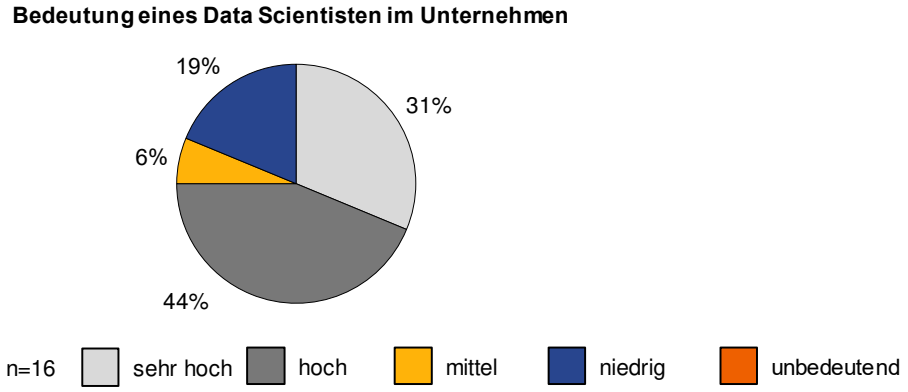

Abbildung 9: Bedeutung eines Data Scientisten im Unternehmen in der Automobilindustrie

Welche Kenntnisse, Softskills, Programmiersprachen sollten Data Scientisten beherrschen und inwieweit ist Prozessverständnis nötig? Um das Angebot der Bildungsträger (shortlist) mit dem echten Bedarf in der Automobilindustrie zu vergleichen, wurde die tägliche Arbeit eines Data Scientisten betrachtet. Insgesamt 24 relevante Arbeitsaufgaben & Fähigkeiten wurden aus der qualitativen Frage gewonnen (siehe Tabelle 1).

Tabelle 1: 24 Arbeitsaufgaben & Fähigkeiten von Data Scientisten

Bereiche	Arbeitsaufgaben & Fähigkeiten von Data Scientisten
Technologie & Daten	Problemlösung und Optimierung
	Datenexploration/-verarbeitung
Prozesse	Tiefes Verständnis der angewendeten Algorithmen aus mathematischer Sicht
	Entwicklung/Erstellung von Algorithmen zur Entscheidungsunterstützung
Organisation	Mit Daten Geschäftsprozesse/-entscheidungen unterstützen
	Grafische Präsentation der Ergebnisse, einfache, klar verständliche Sprache
	Bildung von Handlungsalternativen auf Basis von Daten und Knowhow
	Modellieren & Prognose
Fach-kenntnisse	Statistische Kenntnisse über Datenanalysis
	Skript-Programmierung
	PC-Grundkenntnisse
	Mathematische Denkweise
	Statistik, Simple Thinking, Programmierung und Methodik
	Data Mining
	Machine Learning
Softskills	Erklären und Dokumentieren
	Simulieren und Analysieren
	Kommunikation
	Präsentation
	Neugier
	Teamfähigkeit
Programmier-sprachen	R, Python, Scala, Java, SQL
	Sicherer Umgang mit Datenbanken
	Umgang mit Tools zur schnellen Datenauswertung

Besonders häufig werden folgende 15 Arbeitsaufgaben genannt: Problemlösung und Optimierung, Datenexploration/-verarbeitung, grafische Präsentation der Ergebnisse, einfache, klar verständliche Sprache, Data Mining, PC-Grundkenntnisse, statistische Kenntnisse über Datenanalysis, Skript-Programmierung, Statistik, Simple Thinking, Programmierung und Methodik, Teamfähigkeit, Kommunikation, Programmiersprachen R, Python, Scala, Java, SQL.

2.3 Bedeutung von Technologien

Zur Frage nach dem Bekanntheitsgrad von fünf vermeintlich relevanten Technologien haben sich neun tätige Data Scientisten und neun Unternehmen der Automobilindustrie geäußert (siehe Abbildung 10). Ein Vergleich zwischen Data Scientisten und Unternehmen in der Automobilindustrie zeigt: Apache SPARK ist am bekanntesten, gefolgt von Apache HIVE und SAP HANA. Jedoch bilden die Ergebnisse nur einen Überblick und lassen keine allgemeinen Schlüsse zu, denn außer den fünf vorgegebenen Technologien wurden zusätzlich folgende Technologien genannt: von fünf Teilnehmern Hadoop und Apache STORM; von vier Teilnehmer Splunk Enterprise.

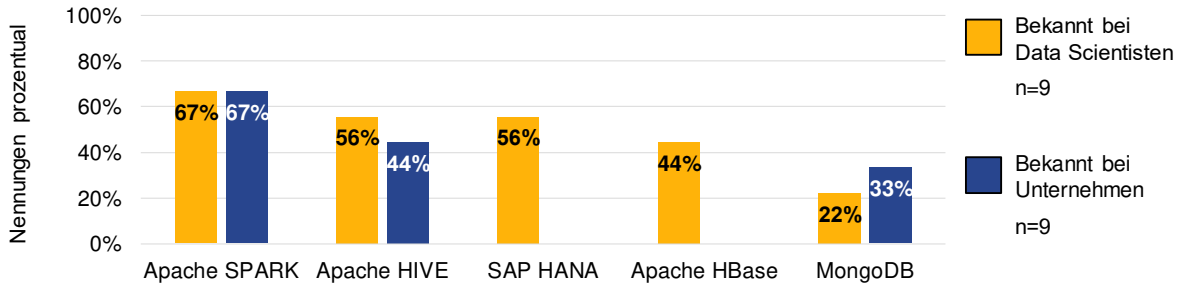

Abbildung 10: Bekanntheitsgrad fünf ausgewählter Technologien

Vereinzelt wurden u. a. folgende Technologien genannt: Cloudera, Cloudera Impala, Infobright, Tableau, Cristal Reports, power BI, Advisory Data Modeling (ADM), Rapid Miner, Omneon, ThinkWorks, Zeitreihen-Stor, QlikView, Knime, Orange, Process Mining und die REM-Software. Sie verdeutlichen, dass Technologien je nach Unternehmen unterschiedlich sind, nicht verallgemeinert werden können und daher auch nicht explizit als Qualifizierung zum Data Scientisten zu empfehlen sind.

2.4 Bedeutung von Algorithmen

Alle Teilnehmer haben angegeben, dass für die Aufbereitung, Verwaltung, Visualisierung und Weiterverwendung der Daten Algorithmen erforderlich sind. Zur Herkunft von Algorithmen in den Unternehmen wurden folgende Angaben gemacht: 44% der Unternehmen kaufen Algorithmen zu, 25% lassen diese in den eigenen Fachabteilungen programmieren und 31% variieren zwischen Zukauf und selber machen – je nach Komplexität der Aufgabe. Nichtsdestotrotz ist grundlegendes Fachwissen bei der Programmierung von Algorithmen notwendig. Die Ergebnisse sind in Abbildung 11 veranschaulicht.

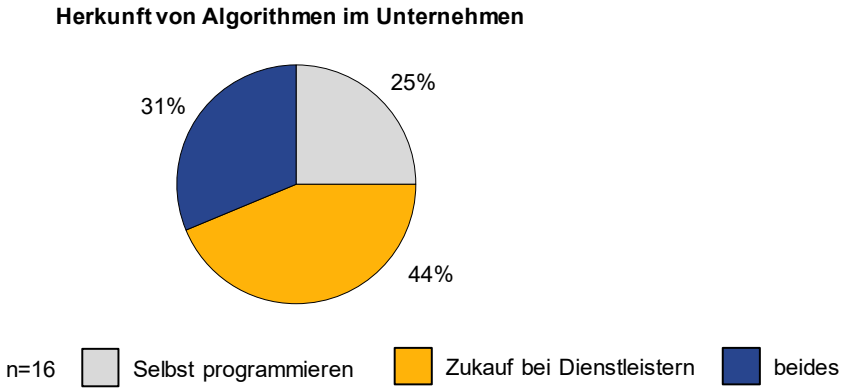

Abbildung 11: Herkunft von Algorithmen in den Unternehmen der Automobilindustrie

3 Handlungsempfehlungen

3.1 Marktlücke bei Aus- und Weiterbildung von produktionsnahen Data Scientisten

Die Untersuchung hat aufgezeigt, dass das Angebot der Bildungsträger zur Aus- und Weiterbildung von Data Scientisten nicht mit den Anforderungen der Unternehmen übereinstimmt. Das Interesse der Automobilindustrie eigene Facharbeiter zu Data Scientisten zu qualifizieren, um erfolgreich produktionsspezifische Potenziale zu heben, ist vorhanden. Aus diesem Grund sollten die Facharbeiter mit IT-Wissen ausgestattet werden und/oder umgekehrt IT-Experten Prozessverständnis entwickeln. Ein Angebot durch produktionsorientierte Bildungsdienstleister in Ostdeutschland ist wünschenswert.

Weil 77% der Weiterbildungsangebote ausschließlich in englischer Sprache angeboten werden, könnte eine primär deutschsprachige Qualifizierung zum Data Scientist in Ostdeutschland für Facharbeiter etabliert werden. Zweitens sind die Zulassungsvoraussetzungen zu hoch und ein Studium nicht nötig – wie 2/3 der befragten Data Scientisten bestätigten. So ist Prozesskenntnis eine Grundvoraussetzung zur sinnvollen Analyse produktionsrelevanter Daten und damit wichtiger. Drittens könnten kompetenzorientierte Bildungsbausteine zum Data Scientisten in überschaubare Einheiten angeboten werden. Viertens sollten die Inhalte an den fünf primär genannten Schwerpunkten in der Automobilindustrie anpasst werden. Fünftens, anstatt in bestimmten Technologien zu qualifizieren, besser bestimmte zugrundeliegende, funktionale Sprachen auf Grund ihrer syntaktischen Nähe zur Mathematik zu vermitteln sowie sechstens, eine Unterstützung für Unternehmen bei der zielgerichteten Analyse bereits erfasster oder noch zu erfassender Daten in Form von Kompetenzen anbieten.

3.2 Katalog von Inhalten zur Aus- und Weiterbildung von Data Scientisten für den Shopfloor

3.2.1 Fünf notwendige Inhalte

Datenintegration, Datenhaltung und Datenqualität:

- Data Warehouse
- Technologien und Systeme (NoSQL-Datenbanken wie Apache HBase, Apache HIVE, Apache SPARK und MongoDB, SAP HANA)
- Kontinuierliches Laden und Datenströme (MessageQueuing)
- Introduction to Business Analytics: ETL-Prozesse (Extraktion, Transformation, Loading), Entwurf und Entwicklung multidimensionaler Datenmodelle (Star

Schema, Snowflake Schema, Galaxy-Schemata), Online Analytical Processing (OLAP) und dessen Operationen (roll up, drill down)

Methoden des Data Mining:

- Bagging und Boosting
- Random Forests
- Kernel Methoden, Support Vector Machines
- Generalized additive Models
- Ensemble Methoden
- Mischungsmodelle und EM-Algorithmus
- Hidden Markov models, Clustering, Naive Bayes, Decision trees, Classification

Datenanalyse und Wissensmanagement:

- Machine Learning (Knowledge Discovery)
- Data Mining (professionelles DM-Tool z.B. SAS Enterprise Miner), MARS
- Trees & Forests, Clustering, Naive Bayes, Decision trees, Classification
- Text Analytics, Informationsextraktion
- Information Retrieval & Text Mining
- Hand-On Labs mit einer State-of the Art Data Science Software
- Funktionsweise von Scale-Up und Scale-Out-Architekturen
- Nutzung von Hadoop oder In-Memory-Datenbanken
- Projektmanagement, Qualitätssicherung und Softwareevolution

Datenvisualisierung und Informationsbereitstellung:

- Informationsvisualisierung und visuelle Analytik
- Dashboards
- Direkte Simulation – Monte Carlo Methoden

Algorithmen:

- Algorithmische Methoden (Greedy-Verfahren, Dynamisches Programmieren, Divide-and-Conquer)
- Laufzeitanalysen (Worstcase, amortisiert, ausgabesensitiv, Lösung von Rekurrenten)
- Algorithmen für Mengen, Graphen, Text und geometrische Problemstellungen
- Algorithmen für Externspeicher, Datenströme, Approximation
- Verfahren in C / MPI (PVM) und in Eden (paralleles Haskell)
- Algorithmen auf großen Datenmengen (z.B. Map Reduce)

3.2.2 Vier mögliche zusätzliche Inhalte

Codierung:

- Schranken für die Güte von Codes, Hamming- und Golay-Codes, zyklische Codes, BCH- und QR-Codes, Reed-Muller und Reed- Solomon-Codes
- die Mathematik der CD, Decodierungsalgorithmen,
- Anwendungen algebraisch-geometrischer Methoden

Fortgeschrittene Programmiertechniken:

- Thread Programmierung, Design Patterns
- Objektorientierte Programmierung (GUI Programmierung, Mensch-Maschine-Interaktion)
- Modellierung (Modeling Language (UML), SDL, Petrinetze, Statecharts)
- Modellierungsoptionen im Bereich BI und Big Data (Layered Scalable Architecture, Lambda Architecture)
- Visuelle Programmierung (ProGraph, Eclipse, AndroidStudio)
- Grafikprogrammierung
- Datenhaltungsoptionen (RDBMS, MDBMS, inMemory DB, verteilte Dateisysteme)

Software und allgemeine Programmierung:

- Statistik-Software „R“
- Sprachen R, Python, SQL, MDX, Map Reduce
- SAP Business Information Warehouse on SAP HANA
- Bereich Programmierschnittstellen und Datenmodelle “Apache-Hadoop Ecosystem” (Apache HBase, Apache HIVE, Apache SPARK)
- Architekturoptionen für Big Data Architekturen (Scale Up vs. Scale Out, Commercial vs. Open Source, Cloud vs. On Premise)

Logische Programmierung:

- Logische Programmierung mit PROLOG: Horn-Klauseln, Unifikation, SLDNF-Resolution, Ausüben von Kontrolle, Inferenzmaschinen, DCG-Grammatik
- SAT-Algorithmen, Modale und Temporale Logik

3.2.3 Anforderungen an Data Scientisten aus Sicht von Zulieferern

Anforderungen an Data Scientisten aus Sicht von Unternehmen der Zuliefererindustrie haben neben den fünf notwendigen Inhalten folgende Schwerpunkte: Big Data, Informationsintegration, Unternehmensmodellierung und -analyse (siehe Tabelle 2):

Tabelle 2: Anforderungen an Data Scientisten aus Sicht von Zulieferern

Gegenstand	Inhalte
Ziel	Datenerfassung für Einkauf, Umsatz & Investition
	Prozessoptimierung
	Verbesserung der Wirtschaftlichkeit
	Verschwendungen erkennen, Maßnahmen ableiten
	Wirkungen der Änderungen nachvollziehen
	Komplette Durchgängigkeit schaffen
	Rückverfolgen in alle Fertigungsstufen
	Poka Joke System
Tätigkeiten	Generalistische Vorgehensweise
	BDE, RAP, QS-Systeme verstehen
	Prozesssichere Erfassung & Auswertung von Maschinendaten, Ergebnisse
	Prozessänderungen nachvollziehen
	Managementgerechte Präsentation der Datenanalyse
Programmier-sprachen	Eher Datenbanken als Sprachen, z.B. Oracle Datenbank; R, Python, Scala, Java, SQL, C#

Gegenstand	Inhalte
Fach-kompetenzen	Datenintegration, Datenhaltung & Datenqualität
	Datenanalyse & Wissensmanagement
	Datenvisualisierung & Informationsbereitstellung
	Grundlagen der Programmierung
	Grundlagen des Data Mining
	Algorithmen
	Grundlagen der IT-Sicherheit, -Ethik & -Gesetze
	Grundlagen der Logik
Fähigkeiten & Vorkenntnisse	Elektrotechnik
	IT-Kenntnisse
	Mathematik, ins Besondere statistische Erhebungen
	Analyseverfahren
	Maschinendatenerfassung, -Steuerung, evtl. Standardprogrammierung einer Maschine
Softskills	Teamfähigkeit
	Kommunikationsfähigkeit
	Fähigkeit mit Mitarbeitern unterschiedlicher Qualifikationsniveaus zu arbeiten
	Neugierde & Wille zum Lernen

Die ABC-Ampel dient zur Orientierung bei der Verteilung der Kern- und Zusatzkompetenzen zum Beruf Data Scientist nach ihrer Wichtigkeit in den Bereichen: IT-Dienstleistung; Planung, Entwicklung, Beratung; Forschung & Entwicklung; Zulieferer; Hersteller. A – steht für besonders wichtig, im Kontext von **„Anwenden“;** B - steht für wichtig, im Kontext von „**Kennen**“; C – steht für weniger wichtig, im Kontext von „**Wissen**“.

Tabelle 3: Anforderungen an Data Scientisten - Anwenden, Kennen und Wissen

Bereiche	A - Anwenden	B - Kennen	C - Wissen
Hersteller	Datenintegration, -haltung & -qualität	Algorithmen	Logische Programmierung
	Datenanalyse & Wissensmanagement Datenvisualisierung & Informationsbereitstellung Methoden des Data Mining Codierung		Software & Programmierung
	IT-Sicherheit, Ethik & Gesetze		
	Fortgeschrittene Programmierung Funktionale Programmierung Geometrisches Modellieren		
Zulieferer	Datenintegration, -haltung & -qualität	Methoden des Data Mining	Codierung
	Datenanalyse & Wissensmanagement	Software & Programmierung	Logische Programmierung
	Datenvisualisierung & Informationsbereitstellung Algorithmen	IT-Sicherheit, Ethik & Gesetze Fortgeschrittene Programmierung	Funktionale Programmierung Geometrisches Modellieren

	Datenintegration, -haltung & -qualität	Software & Programmierung	Logische Programmierung
	Datenanalyse & Wissensmanagement	IT-Sicherheit, Ethik & Gesetze	Fortgeschrittene Programmierung
	Datenvisualisierung & Informationsbereitstellung		Funktionale Programmierung
	Methoden des Data Mining		Geometrisches Modellieren
	Algorithmen		Codierung
	Datenintegration, -haltung & -qualität	Datenvisualisierung & Informationsbereitstellung	Geometrisches Modellieren
	Datenanalyse & Wissensmanagement	Software & Programmierung	
	Methoden des Data Mining	Logische Programmierung	
	Algorithmen	Codierung	
	IT-Sicherheit, Ethik & Gesetze Fortgeschrittene Programmierung	Funktionale Programmierung	
	Datenintegration, -haltung & -qualität Datenanalyse & Wissensmanagement Datenvisualisierung & Informationsbereitstellung Methoden des Data Mining Algorithmen Logische Programmierung IT-Sicherheit, Ethik & Gesetze Fortgeschrittene Programmierung	Software & Programmierung	Codierung Funktionale Programmierung Geometrisches Modellieren

4 Kurzfazit

Es bleibt u. a. festzuhalten, dass derzeitige Data Scientisten vielfältige Bildungsabschlüsse haben können - Informatiker, Ingenieure, Elektrotechniker, Physiker, Mathematiker. Es gibt keine vordefinierte Vorgehensweise, wie man zum Data Scientist wird. Im Vordergrund stehen berufliche Erfahrungen, Grundverständnis für die Arbeit mit großen Datenmengen sowie fundiertes Wissen im Bereich Domänen und Technik.

Der Data Scientist im Shopfloor zeichnet sich gegenüber einem ausgebildeten Data Scientisten dadurch aus, dass er Prozessverständnis hat.

Die Erwartungen der Unternehmen an Data Scientisten sind hoch. Zu diesen gehören nicht nur die Analyse und Auswertung von Maschinendaten, sondern insbesondere das Prozessverständnis sowie die Prozessoptimierung. Auch das Ableiten und die Rückverfolgbarkeit der Ergebnisse sowie das Begleiten von Prozessen und Abläufen gehören dazu.

ACOD
AUTOMOTIVE CLUSTER
OSTDEUTSCHLAND

ACOD GmbH
August-Bebel-Str. 73
04275 Leipzig

Telefon: +49 (0) 341 3038 2535
E-Mail: info@acod.de

www.acod.de

Bild- und Quellennachweis:
Freeimages.com/Donald Cook (Umschlag)

Gefördert als JOBSTARTER plus-Projekt aus Mitteln des Bundesministeriums für Bildung und Forschung und des Europäischen Sozialfonds.