



Data Science im Kontext der digitalen Transformation

Nicht nur Data Scientists gesucht

Ein Beitrag
von Tatjana Wiebusch
und Michael Zimmer

Im Zuge der digitalen Transformation hat sich der Wert von Daten und der daraus abzuleitenden Informationen für Unternehmen drastisch erhöht. Lag in der Vergangenheit der Fokus auf einer überschaubaren Anzahl betriebswirtschaftlicher Kennzahlen, so steigen durch die fortschreitende Digitalisierung und Industrie 4.0 die Anzahl der Datenquellen und somit auch die Datenmenge. Die Aufbereitung und Analyse dieser Daten kann für Unternehmen einen entscheidenden Wettbewerbsvorteil darstellen. Gerade der analytische Bereich, der in der Vergangenheit von Aktuaren oder Statistikern besetzt wurde, gewinnt unter neuen Schlagworten wie beispielsweise Data Science oder Head of Algorithms an Bedeutung. Wie kann Data Science in diesem Umfeld von Nutzen sein und welche Voraussetzungen werden hierfür benötigt? Im folgenden Artikel wird die Bedeutung von Data Science in der digitalen Transformation beschrieben, sich daraus ergebende Berufsfelder und grundlegende analytische Rollen definiert sowie kritische Erfolgsfaktoren bei der Rekrutierung und Weiterbildung für den Aufbau von Data-Science-Kompetenzen im Unternehmen diskutiert.

Die Erfolgsfaktoren einer digitalen Transformation im Sinne der Veränderung der Geschäftswelt durch neue Technologien können unter vier wesentlichen Aspekten zusammengefasst werden: Fähigkeit zur Identifikation neuer Geschäftsmodelle, digitale Führungsqualität, Verfügbarkeit und Auswertbarkeit von Daten sowie effizientes Talentmanagement [RoB 17; Lin 17; Hal 16]. So sind unter anderem weitreichende Kenntnisse über verfügbare Datenquellen, Auswertungsmöglichkeiten mit Hilfe von Big-Data-Technologien und die Anwendung von Data Science erforderlich, um Visionen und daraus abgeleitete neue Geschäftsmodelle zu entwickeln. Realisierbar sind diese Geschäftsmodelle meist aber nur, wenn die Verfügbarkeit und Auswertbarkeit der Daten gewährleistet ist.

Aus diesem Grund arbeiten beispielsweise viele Automobilhersteller an der Vernetzung ihrer unterschiedlichen „Datentöpfe“, um auf dieser Basis als Mobilitätsanbieter neue Geschäftsmodelle erschließen zu können [Brü 15]. Hier spielt die industrialisierte und wiederholbare Auswertung der Daten mit Hilfe von Data Science eine zentrale Rolle. Auch das Talentmanagement ist in diesem Umfeld gemeinsam mit der Führungsqualität im Zuge der digitalen Transformation (digitale Führungsqualität) für die Gewinnung neuer und die Förderung bestehender Mitarbeiter von großer Relevanz. Doch was verbirgt sich hinter dem Schlagwort Data Science?

Schlagwort: Data Science

Der Begriff Data Science geht in seiner heutigen Form auf einen Vortrag von C. F. Jeff Wu aus dem Jahre 1997 zurück, in dem die Arbeit eines Statistikers unter dem Begriff Data Science subsumiert wurde. Schwerpunkt war hier die Erfassung, Bereitstellung und Analyse von Daten. Innerhalb der letzten Jahre hat sich Data Science als Begriff in der Wirtschaft etabliert und das Berufsbild eines Data Scientist geprägt. Ziel dieses Berufsfeldes ist es, vorhandene Datenquellen zu kombinieren, zu analysieren und daraus unternehmerisch nutzbares Wissen zu generieren [Pre 13].

Die Aufgabe eines Data Scientist besteht darin, relevante Informationen aus großen Datenmengen zu extrahieren und Entscheidungen vorzubereiten bzw. automatisiert anzustoßen. Dabei werden innovative Analyse-Tools eingesetzt, um Hypothesen abzuleiten, statistisch zu überprüfen und als Entscheidungsgrundlage für das Management aufzubereiten.

Rollen und Aufgabebereiche

Bei der Betrachtung des in Abbildung 1 dargestellten Analytics-Prozesses [Roo 12] von der Datenaufbereitung über die Analyse der Daten bis hin zur Erkenntnisgenerierung und Erfolgsmessung der daraus abgeleiteten Entscheidungen lassen sich mehrere erforderliche Rollen identifizieren. Voraussetzung ist unter anderem eine geeignete Datenar-



TATJANA WIEBUSCH ist Direktorin bei Deloitte im Bereich Analytics & Information Management und hat zahlreiche Analytics-Lösungen bei ihren Kunden implementiert. Ihr Schwerpunkt liegt auf Data Warehousing, dem Entwurf von Analytics-Architekturen und der Program Delivery. Als COO des Digital-Delivery-Bereichs einer weltweit agierenden IT-Beratung war sie in der DACH-Region für ca. 400 gut ausgebildete Data Engineers, Data Architects, Solution-Designer, Business-Analysten und Data Scientists verantwortlich. In dieser Rolle konnte sie weitreichende

Erfahrungen im Bereich des Talentmanagements für Data & Analytics sammeln.

E-Mail: Twiebusch@deloitte.de

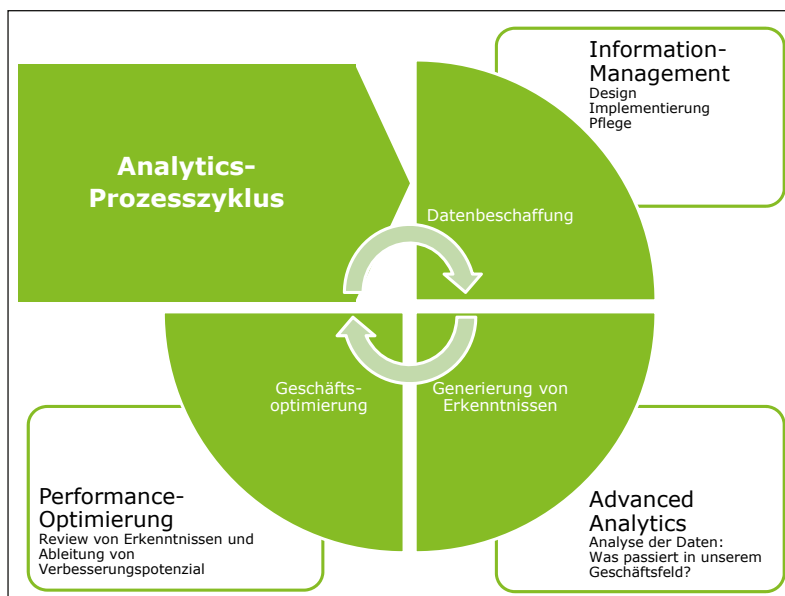
DR. MICHAEL ZIMMER ist Senior Manager bei Deloitte im Bereich Analytics & Information Management und beschäftigt sich mit dem Aufbau von (Big-)Data- und Analytics-Architekturen. Seine Schwerpunkte liegen auf Strategie und Governance, einer ganzheitlichen Delivery sowie agilen Analytics-Architekturen. Als SAS Capability Lead (DACH) einer weltweit agierenden IT-Beratung konnte er Führungserfahrung mit Data Architects, Data Engineers und Data Scientists sammeln.

E-Mail: mizimmer@deloitte.de



chitektur (Architekt) mit adäquater Datenqualität (DQ-Verantwortlicher) oder Kennzahlen (Data Owner). Diese Rollen werden innerhalb des Analytics-Prozesszyklus im Schritt „Datenbeschaffung“ vertretet. Auf sie wird in diesem Artikel nicht detailliert eingegangen. Weitere Informationen dazu finden sich bei [Glu 16].

Abb. 1: Analytics-Prozesszyklus



Die Kernaufgabe von Data Science betrifft den analytischen Teil des gesamten Prozesses. Im Prozessschritt „Generierung von Erkenntnissen“ werden Datenanalysen durchgeführt. Die hier gewonnenen Erkenntnisse dienen zur Entscheidungsunterstützung oder zur (teil-)automatisierten Entscheidung in Unternehmen.

Gerade in kundenzentrierten Analytics-Projekten gewinnt das Thema Datenschutz durch die Datenschutz-Grundverordnung (DSGVO) an Bedeutung. Hier handelt es sich jedoch nicht um eine Kernaufgabe von Data Science, sondern eher um ein Problem der Rechtsgrundlage sowie der Datenhygiene, das von Datenarchitekten und Data Owners in Verbindung mit den Datenschutzbeauftragten diskutiert werden muss.

Um den vollständigen Analyseprozess End-to-End abzudecken und eine Geschäftsoptimierung zu ermöglichen, sind außerdem zusätzliche Rollen zur Visualisierung und Verbreitung der Ergebnisse notwendig, die aber in diesem Artikel nicht thematisiert werden (vgl. hierzu auch [VDS17]). Im Schritt „Geschäftsoptimierung“ werden Analytics-Ergebnisse und Maßnahmen hinsichtlich ihrer Effektivität bewertet und Optimierungsmaßnahmen abgeleitet. Dabei können auch neue Geschäftsmodelle evaluiert und gegebenenfalls wiederum die Grundlage für weitere neue Geschäftsmodelle gelegt werden.

Ordnert man der Disziplin Data Science als Kernaufgaben die Kombination, Aufbereitung und Analyse von Daten zu, lassen sich daraus verschiedene Rollen ableiten (vgl. Tabelle 1) [Agh17].

Diese Rollen sind natürlich auch miteinander kombinierbar. Gerade in der Praxis entwickeln sich Data Engineers mit mathematischem Hintergrund über die Zeit häufig zu Data Scientists. Dies ist zu empfehlen, da auch im Bereich der spezifischen Datenaufbereitung (Data Preparation) für Data Sci-

ence Anforderungen an eine professionelle ETL-Aufbereitung zu erfüllen sind.

Im täglichen Austausch mit Data Scientists wird jedoch sichtbar, dass die Aufbereitung der Daten nicht als wichtige Grundlage für die Qualität und Industrialisierbarkeit der Modelle angesehen wird, sondern vielmehr als notwendiges Übel. Oftmals werden bei der Data Preparation von Data Scientists allgemeine Programmierrichtlinien wie Parametrisierung oder Modularisierung nicht eingehalten und Ausprägungen oder Parameter hart kodiert. Idealerweise ist ein Data Scientist also nicht nur für die Analyse von Daten, sondern auch für die Industrialisierung im Sinne wartbarer und automatisch auslieferbarer Modelle und Datenaufbereitungsprozesse verantwortlich. Nur so wird eine Wiederholbarkeit sowie Quantifizierbarkeit der Ergebnisse und letztlich eine nachhaltige Weiterentwicklung ermöglicht.

Entsprechend herausfordernd ist es für ein Unternehmen im Sinne des Talentmanagements und der digitalen Führungsqualität, für eine ganzheitliche Data Science motivierte Mitarbeiter mit der entsprechenden Expertise zu finden und langfristig zu binden.

In der Theorie haben viele Unternehmen erkannt, welches Potenzial die Aufbereitung und Nutzung der ihnen zur Verfügung stehenden Daten birgt. Jedoch fehlt häufig noch die richtige Strategie, um Data-Science-Kompetenzen im Unternehmen aufzubauen. Welches Umfeld muss geschaffen werden? Wie können neue Analytics-Prozesse in die bestehenden Prozesse integriert werden? Kann das Unternehmen eine Arbeitsumgebung schaffen, die Innovation und Kreativität fördert und in der ein Data-Science-Team seine vollständige Kompetenz entfalten kann? Im Folgenden werden diese Herausforderungen näher betrachtet und Lösungsansätze aufgezeigt.

Tab. 1: Aufgaben und Anforderungen an Kernrollen der Data Science

Rolle	Aufgaben	Notwendige Fähigkeiten
Data Engineer	Schaffung einer Datengrundlage für den Data Scientist durch den Entwurf des Datenmodells nach Vorgabe des Data Scientist Bereitstellung der Daten und ggf. Übergabe an den Betrieb Einhaltung von Entwicklungsrichtlinien und Vorgaben zur Sicherstellung einer qualitativ hochwertigen und anpassbaren Datenbereitstellung	Programmierkenntnisse in Big-Data-Technologien Verständnis von ETL-Tools und Datenbanken Verständnis für den funktionalen Kontext der Daten, um Harmonisierungsanforderungen fachkonform umzusetzen
Data Scientist	Modellierung von Analysemodellen Durchführung von statistischen Analysen und Ad-hoc-Auswertungen Dokumentation der Modelle und Parameter Erstellung, Training und Produktivsetzung von Modellen sowie bei Bedarf Retraining Messung der Effektivität der Modelle	Mathematik- und Statistik-Kenntnisse Programmierkenntnisse in statistischen Programmiersprachen wie beispielsweise R oder Python Kenntnisse über den fachlichen Kontext der Analysen Grundlegendes Verständnis der Daten
Data Analyst	Vorgabe von Analysebereichen und Anwendungsfällen für die Data Scientists Identifikation der relevanten Parameter Enger Austausch mit dem Data Scientist Beurteilung der Analyseergebnisse aus fachlicher Sicht Bindeglied zum Fachbereich	Tiefe Kenntnisse der fachlichen Prozesse und Zusammenhänge Überblick über analytische Auswertungsmethoden und Simulationsverfahren und deren fachlichen Nutzen sowie mögliche Anwendungsgebiete

Data Science im Unternehmen: Konkrete Erfahrungen

Anhand der Rollenbeschreibungen (vgl. Tabelle 1) lässt sich bereits erkennen, dass der Schlüssel zum wertbringenden und erfolgreichen Einsatz von Data Science in einer engen Verknüpfung von fachlichen, mathematischen und technischen Fähigkeiten liegt. Die dafür notwendigen Kompetenzen sind sehr breit gefächert und vielschichtig. Es ist daher leicht nachvollziehbar, dass die entsprechenden Fähigkeiten für eine Data-Science-Einheit nicht direkt vom Markt rekrutiert und aufgebaut werden können. Der Ausbildungspfad in diesem Umfeld ist noch relativ jung, sodass es nur eine limitierte Anzahl von ausgebildeten Fachkräften im Bewerbermarkt gibt.

Dem steht der rasante Anstieg der Nachfrage an Data-Science-Kompetenzen in den Unternehmen diametral gegenüber. Prognosen zeigen, dass auch in naher Zukunft der Bedarf stärker steigen wird als die Verfügbarkeit von ausgebildeten Ressourcen. Dadurch bleibt der Recruiting-Markt stark umkämpft und die Gehälter steigen überproportional weiter [Pre15]. Um in dieser Marktsituation entsprechende Kompetenzen aufbauen zu können, muss ein umfassendes Talentmanagement in den Unternehmen etabliert werden (vgl. Abbildung 2).

Das Talentmanagement beginnt mit dem Marketing am Recruiting-Markt und der eigenen Darstellung als potenzieller Arbeitgeber, geht mit der Gestaltung eines individuellen Anreizsystems für die Mitarbeiter weiter und baut zu großen Teilen auf einem umfassenden Aus- und Weiterbildungskonzept für die bestehenden Mitarbeiter auf.

Das Marketing für das Recruiting sollte als eigenständige Disziplin aufgebaut sein, um als potenzieller Arbeitgeber mit Karrierepotenzial positiv im Markt wahrgenommen zu werden. Ziel ist die langfristig positive Positionierung als Arbeitgeber, um eine möglichst große Anzahl an geeigneten Bewerbern anzuziehen. Geeignete Methoden sind hier beispielsweise ein bestimmtes Leistungsversprechen, offene Managementkonzepte, besondere Entwicklungsoptionen oder andere Merkmale wie eine angenehme Arbeitsatmosphäre. So hat sich in der Beratungsbranche zum Beispiel ein Studiokonzept mit geringerer Reisetätigkeit für Data-Science-Berater als Erfolgsfaktor erwiesen. Bei Studio-Ansätzen werden beispielsweise neuartige Arbeitsumgebungen mit geringeren Auslastungszielen und thematischen Freiheiten gewährt, wie sie Google in seinen Unternehmen eingeführt hat.

Eine weit verbreitete Maßnahme der Kandidatenrekrutierung ist die Zusammenarbeit mit Lehrstühlen. So gibt es in Deutschland mittlerweile mehr als 200 Angebote an deutschen Hochschulen, die das Thema Data Science behandeln. Durch die Zusammenarbeit mit Hochschulen wird direkt an der Quelle versucht, junge Talente nach der Ausbildung für einen Einstieg im eigenen Unternehmen zu motivieren. Ohne geeignete Angebote wie Praktika, CV-Checks, Soft-Skill-Trainings oder Crowd-Sourcing-Projekte und ein interessan-



tes Employer Branding haben solche Maßnahmen aber nur einen geringen Erfolg.

Ergänzend hierzu sollten im Recruiting-Prozess auch die richtigen Kanäle genutzt werden, um auf das Unternehmen aufmerksam zu machen. Klassische Job-Anzeigen sind hier weniger erfolgversprechend, sondern eher Maßnahmen wie Direktmarketing, Direkt-Ansprachen, Recruiting-Events oder Hackathons. Diese Formate helfen, sich aus der Masse der Unternehmen herauszuheben und um die wenigen Talente zu werben.

Wird aus dem Bewerber ein Mitarbeiter, ist eine lückenlose Integration mit passenden Aufgaben wichtig. Data Scientists erwarten meist immer wieder neue Herausforderungen und sie möchten in einem Umfeld arbeiten, in dem sie sich permanent entwickeln können. Leerlauf in der Einstiegsphase und keine Weiterentwicklung führt schnell zur Frustration und somit zur Abwanderung des Mitarbeiters.

Da im Bereich Data Science die Verknüpfung mathematischer und technologischer Kenntnisse mit tiefer fachlicher Expertise so wichtig ist, kommt dem internen Fort- und Weiterbildungssystem eine entscheidende Aufgabe zu. Interne Aus- und Weiterbildungswege sind zu definieren und ein Anreizsystem ist zu schaffen, das die Entwicklung der Mitarbeiter im eigenen Unternehmen belohnt. So können beispielsweise Quereinsteiger mit tiefen Fachkenntnissen und IT-Affinität zu Data Analysts fortgebildet werden, oder Entwickler aus dem Big-Data/ETL-Umfeld verbreitern ihre Kenntnisse um Data-Science-Fähigkeiten und entwickeln sich zum Data Scientist.

Gleichzeitig muss die Fach- und Methodenkompetenz der Data Scientists gefördert werden. So kann ein Unternehmen eigene Mitarbeiter, die die notwendige Kombination von fachlichen und technischen Kenntnissen aufweisen, ausbilden. Hier können auch externe Zertifizierungen zum Data Science Professional, wie zum Beispiel vom TDWI angeboten, helfen. Durch den Aufbau eigener Kompetenzen kann der Kampf um neue Talente am Markt relativiert werden.

Das Halten dieser Mitarbeiter wird am Ende bedingt durch die geschickte Orchestrierung von individuellen Anreizsystemen, das Bieten von heraus-

Abb. 2: Überblick über die Komponenten des Talentmanagements

fordernden Tätigkeiten, stetige Weiterentwicklung sowie enge Bindung an die Unternehmensführung durch eine digitale, team-orientierte Unternehmens- und Führungskultur. Data Scientists fordern in der Regel ein kreatives, agiles Arbeitsumfeld. Vorhandene Prozesse und Strukturen im Unternehmen verhindern unter Umständen jedoch diesen Arbeitsansatz. Es ist daher empfehlenswert, ganze Unternehmensprozesse zu restrukturieren oder Studio-Ansätze in Erwägung zu ziehen. Dadurch können flexible und kreative Arbeitsmodelle unterstützt und eine positive Arbeitsatmosphäre geschaffen werden.

Ausblick

Aufgrund der aus der digitalen Transformation resultierenden Aufgaben ist die Nachfrage nach Data Scientists am Markt enorm. Immer mehr Unternehmen erkennen, welche Wettbewerbsvorteile durch die Identifikation neuer Geschäftsmodelle und die Verfügbarkeit sowie Auswertbarkeit von Daten gewonnen werden können. Die Kompetenzen in diesem Umfeld können allerdings häufig weder intern noch extern am Markt in der notwendigen Geschwindigkeit ausgebildet werden. Im Zuge eines effizienten Talentmanagements stellt sich die Fra-

ge, wie diese Lücke zwischen Bedarf und verfügbaren Ressourcen geschlossen werden kann.

Um die wenigen gut ausgebildeten Mitarbeiter möglichst optimal im Unternehmen einzusetzen, ist es im Rahmen digitaler Führungsqualität daher auch wichtig, Data Scientists von wiederkehrenden und langweiligen Aufgaben zu befreien, um Zeit für kreative Aufgaben zu haben. Gerade die Automatisierung und Standardisierung einzelner Schritte des Analyseprozesses haben hier großes Potenzial. Durch die Reduzierung von Standardtätigkeiten wird wiederum implizit ein attraktives, forderndes Arbeitsumfeld geschaffen.

Ergänzend entwickelt sich auch der Werkzeugkasten eines Data Scientist weiter. So fokussieren analytische Werkzeuge vermehrt auf Standardisierung, Model Deployment und Automatisierung, um den Data Scientist zu entlasten. Zukünftig stellt auch der Einsatz künstlicher Intelligenz einen vielversprechenden Ansatz zur Entlastung dar [Mo17].

Unternehmen, die in der Lage sind, Mitarbeiter mit den richtigen Kompetenzen aufzubauen und eine Data-Science-Einheit langfristig zu etablieren, haben zukünftig einen entscheidenden Wettbewerbsvorteil in einem datengetriebenen Markt für sich gesichert.

Quellen

[AGH17] Aghabozorgi, S.: Data Scientist vs. Data Engineer, what's the difference? <https://cognitiveclass.ai/blog/data-scientist-vs-data-engineer/>, abgerufen am 13.11.2017

[Brü15] Brünglinghaus, C.: Vom Automobilhersteller zum Mobilitätsanbieter. In: ATZ 10/2015

[Glu16] Gluchowski, G.: Ergebnisse der TDWI Studie Data Governance im Big-Data-Zeitalter. 2016

[Hal16] Halper, F.: Data Science and Big Data: Enterprise Paths to Success. TDWI 2016

[Lin17] Linné, H.: Kritische Erfolgsfaktoren der Digitalen Transformation. F.A.Z. Verlag 2017

[Moo17] Moore, S.: Gartner Says More Than 40 Percent of Data Science Tasks Will be Automated by 2020. <https://www.gartner.com/newsroom/id/3570917>, abgerufen am 13.11.2017

[Pre13] Press, G.: A very short history of data science. Forbes.com, 2013, <https://www.forbes.com/sites/gilpress/2013/05/28/a-very-short-history-of-data-science>, abgerufen am 13.11.2017

[Pre15] Press, G.: The supply and demand of data scientists. Forbes.com 2015, <https://www.forbes.com/sites/gilpress/2015/04/30/the-supply-and-demand-of-data-scientists-what-the-surveys-say>, abgerufen am 13.11.2017

[Ro012] Roopen, R.: Big Data - Harnessing a Game Changing Asset. <https://www.slideshare.net/sasindia/roopen-rooy>, abgerufen am 13.11.2017

[RoB17] Rowles, D. / Brown, T.: Building Digital Culture: A practical guide to successful digital transformation. 2017

[VDS17] Visual Data Science Org, <http://www.visualdatascience.org/2017/index.html>, abgerufen am 13.11.2017

Senior Data Scientist (m/w) - Analytics & Information Management

Als Senior Data Scientist übernehmen Sie eine verantwortliche Rolle in der fachlichen Bearbeitung Ihres Projektes von der Analyse über die Planung und Konzeption bis hin zur Implementierung. Sie erarbeiten Anwendungsfälle mit dem Kunden, entwerfen Analytics Modelle und können somit Ihre analytischen Kompetenzen einbringen und kontinuierlich in einem herausfordernden Umfeld ausbauen.

Die Lösung anspruchsvollster Aufgaben setzt ausgezeichnete Mitarbeiter voraus. Daher sind für uns folgende Qualifikationen wünschenswert:

- Hochschulabschluss mit den Fachrichtungen (Wirtschafts-) Informatik, Betriebswirtschaftslehre oder vergleichbaren Fächern
- Projekterfahrung in den relevanten Technologien

- Analytisches Denkvermögen, Leadership-Potential und ausgeprägte Teamorientierung
- Sehr gute Präsentations- und Kommunikationsfähigkeit

Weitere Infos zu dieser und weiteren offenen Positionen bei uns im Bereich Analytics finden Sie auf unserer Homepage <https://jobs.deloitte.de>.

Deloitte.