

Es lebe die Datenschleuder!

Data Science – aus Daten relevantes Wissen gewinnen – ist der neue IT-Job dieses Jahres. Damit er sinnvoll in Unternehmen eingesetzt werden kann, braucht er vor allem eines: den Zugriff auf Unternehmensdaten. Und das ist unser Job!



«Die kluge Kombination von Daten ist eine interdisziplinäre Arbeit.»

Daniel Liebhart ist Dozent für Informatik an der ZHAW (Zürcher Hochschule für Angewandte Wissenschaften) und Solution Manager der Trivadis AG.

«Daten bleiben scheu und grausam. Es ist schwierig, aus ihnen zu lernen, und noch schwieriger, das Gelernte umzusetzen», sagte der Unternehmer und Big-Data-Pionier Nicolas Bissantz bereits vor zwei Jahren im Wissensmagazin des GDI (Gottlieb Duttweiler Institut).

Inzwischen haben wir genau dafür einen neuen Berufszweig, und der nennt sich «Data Science». Renommiertere Institute wie das Fraunhofer IAIS oder die ZHAW bieten Ausbildungen zum Data Scientist an. Die Aufgabenstellung reicht von der Auswahl geeigneter Verfahren und Tools für die Analyse der Daten über die Betreuung der Analysesysteme und die Kontrolle der Verfahrens- und Datenqualität bis hin zur Aufbereitung, Präsentation und Interpretation der Daten. Die «Harvard Business Review» sagt sogar, Data Scientist werde «The Sexiest Job of the 21st Century» sein. Und der Bedarf an Spezialisten steigt stark. Gemäss Gartner schaffen Unternehmen weltweit in den nächsten zwei Jahren 4,4 Millionen neue Jobs, um Wettbewerbsvorteile durch die kluge Aufbereitung von Daten zu erreichen oder zu behaupten.

Was ein Data Scientist braucht

Die Voraussetzung für eine gute Analyse und Aufbereitung von Unternehmensdaten ist eine Gesamtansicht auf die zentralen Fragestellungen des Marktumfelds und auf die Aufgabenstellung, die sich aus der Geschäftstätigkeit ergeben. Unternehmen mit sehr vielen Kunden möchten beispielsweise gerne das Kundenverhalten sehr genau kennen und wenn möglich voraussehen, um Produkte und Logistik dar-

aufhin auszurichten. Transportunternehmen sind auf die zeitnahe Aufbereitung von Fahrzeugen und Umweltinformationen angewiesen, um ihre Lieferketten laufend anzupassen. Für die industrielle Produktion sind Warenflüsse, Lagerbedingungen, Losgrößen, Maschinenkapazitäten und andere nichtlineare Faktoren relevant, um flexibel und ressourcenschonend herstellen zu können.

Die kluge Kombination von Daten ist eine interdisziplinäre Arbeit, die Kenntnisse aus dem Bereich der Führung und Strukturierung von Unternehmen mit branchenspezifischem Know-how kombiniert und auf der Basis modernster Analyseverfahren und Technologien entsprechende Lösungen erarbeitet. Die Ausbildung ist entsprechend umfangreich und kombiniert gemäss der Expert Group Fachbereiche wie Statistik, Mathematik, Informatik, Psychologie und Medienwissenschaften. Kein Wunder, dass nur Nachdiplomlehrgänge angeboten werden. Der Beruf ist zwar interessant, erfordert jedoch breites Basiswissen und hohe Analysefähigkeit, um aus Daten Informationen zu generieren. Darüber hinaus braucht ein Data Scientist vor allem eines: den Zugriff auf unternehmensrelevante Daten. Und das ist der Job der IT – unser Job.

Was wir tun müssen

Die typische IT-Landschaft eines Unternehmens ist heute darauf ausgelegt, das betriebliche Tagesgeschäft zu unterstützen. Das bedeutet in den meisten Fällen, dass die Struktur der Informationssysteme der Struktur der Ablauforganisation des Unternehmens ent-

spricht. Die konkrete Tätigkeit bestimmt also den Aufbau der Anwendungen und die Interaktion zwischen den Systemen. Die Daten sind im besten Fall prozessbegleitend organisiert und werden entsprechend in verschiedensten Systemen auf unterschiedlichste Art und Weise gehalten. Um an alle relevanten Daten zu kommen, bauten wir bisher separate Infrastrukturen auf, die im Nachhinein einen Zusammenzug und eine Verdichtung dieser Daten zur Entscheidungsunterstützung erlaubten. Diese Trennung zwischen operativen von dispositiven Systemen macht jedoch eine zeitnahe und flexible Aufbereitung unmöglich.

Damit der Data Scientist vernünftig arbeiten kann, muss jedes Unternehmen eine vernünftige Datenlogistik umsetzen. Sie stellt die geforderte Menge benötigter Daten in der richtigen Zusammensetzung zur richtigen Zeit am richtigen Ort bereit. Damit erweitert sich die Aufgabestellung der IT von der Bereitstellung typischer zentraler Systemlandschaften hin zur Organisation, dem Betrieb und der Überwachung von Datenflüssen.

Je nach Branche variiert die genaue Aufgabestellung der Datenlogistik. Die Datenlogistik zwischen Firmen stellt beispielsweise hohe Anforderungen an die Standardisierung und Sicherheit, während die Produktion auf die sehr schnelle Kombination relevanter Produktions- und Umgebungsdaten angewiesen ist und eine Verwaltung den umfassenden Austausch zentraler Datenobjekte braucht. Allen gemeinsam ist eines: Eine gute Datenlogistik ist nur mit unternehmensweiten oder sogar unternehmensübergreifenden Mechanismen zu realisieren. Mechanismen, die Daten schnell, sicher und übergreifend austauschen helfen. Damit sind Plattformen, Infrastrukturen und Dienstleistungen gemeint, die interne und externe Systeme verbinden und Daten als Informationsobjekte unternehmensweit bereitstellen und verwalten können. Nur auf dieser Basis wird Data Science auch für Unternehmen sexy: Es lebe die Datenschleuder!