



Bild: Developer Media/Creakom

Smart Data Developer Conference

Big, smart oder Datengrab?

Knapp 100 Software-Entwickler besuchten im Dezember 2015 die Erstveranstaltung der Smart Data Developer Conference und machten sich an zwei Tagen mit den Herausforderungen in der Verarbeitung großer Datenmengen vertraut. Den thematischen Schwerpunkt bildeten die Cloud-basierte Speicherung und Analyse umfangreicher Daten sowie die dafür bereitstehenden Tools und Frameworks. Eine wichtige Rolle spielte auf der Konferenz der praktische Einsatz der Azure-Plattform von Microsoft, die diverse Tools, Vorlagen und Dienste bereitstellt.

Die zentrale Fragestellung bei Big Data ist, wie man aus der Datenflut einen konkreten Nutzen zieht. Geschätzt wird, dass deutlich weniger als 10 Prozent der anfallenden Daten verwertbar sind. Erschwert wird die Auswertung dadurch, dass heute schon 80 Prozent der Daten vollkommen unstrukturiert anfallen – Tendenz steigend. Dies führt zu einer Renaissance auch zum Teil schon länger bekannter Verfahren aus dem Bereich der künstlichen Intelligenz (Artificial Intelligence, AI), die man mit der

heute in der Cloud verfügbaren Rechenpower aber deutlich effizienter einsetzen kann. Sowohl die für IoT-Anwendungen (Internet of Things) wichtige Stream Analytics als auch das Machine Learning bedienen sich u.a. dieser Algorithmen. Man wähnt sich hiermit nun im Zeitalter der kognitiven Systeme, die komplexe Informationen interpretieren, organisieren und Erklärungen anbieten. Die Resultate sind immer statistischer Natur, d.h. mit Wahrscheinlichkeiten versehen, nie deterministisch. Ziel ist es, mit den Ergeb-

nissen von guten zu Daten getriebenen Entscheidungen zu kommen. Insbesondere auch im industriellen Umfeld bleibt das spezifische Domänenwissen der mit der Datenauswertung betrauten Mitarbeiter essentiell, denn Mensch und Maschine kooperieren, sind erst gemeinsam stärker und besser. Damit aus big smart wird, und damit für das Unternehmen Nützliches und Gewinnbringendes entsteht, ist reichlich Verständnis der zugrundeliegenden Prozesse notwendig. Von über 100 laufend erfassten Parametern an einer Fräsmaschine z.B. sind nur wenige relevant, um auf den Verschleiß der Spindel schließen zu können. Machine Learning kann helfen, diese Merkmale herauszufinden. Dennoch muss man sich im Rahmen der Datenanalyse eingehend Gedanken machen über den Wert der Daten und die Relevanz bzw. Nähe zur Realität. Die Vorbereitung der Daten für den eigentlichen Lernprozess bleibt eine zentrale Aufgabe.

Datensicherheit im Zentrum der Bedenken

Kritisch kommentiert wurde von vielen Teilnehmern das Thema Datensicherheit. Welcher mittelständische Produktionsbetrieb möchte schon alle Produktionsdaten in der Cloud ablegen? Zwar lassen sich damit Geschäftsprozesse verbessern, man kann aber auch einen Eindruck über Stillstandszeiten und die aktuelle Auslastung des Unternehmens gewinnen. Wer möchte diese sensiblen Informationen schon in falschen Händen sehen? Dass Microsoft zusammen mit T-Systems Ende 2016 eine deutsche Cloud mit Rechenzentren in Frankfurt und Magdeburg anbieten will, dürfte viele Datenschutzkritiker besänftigen. Auf der SPS IPC Drives 2015 war jedenfalls vielfach von Anbieterseite zu hören, dass die Ankündigung schwierige Gespräche mit Cloud-kritischen Kunden erspart habe.

Intransparentes Gebührensystem

Wenig anwenderfreundlich und recht intransparent wirkt das Gebührensystem des Azure-Portals. Es ist schwierig abzuschätzen, welche Kosten für das Hosting der

Daten und für die Nutzung der diversen Dienste entstehen werden. Häufig gibt es für die Evaluierung der Azure-Funktionalität kostenlose Basisdienste, die aber leider im Menü versteckt sind. Offensichtlich unterliegen die Preise aktuell auch noch einer hohen Änderungsdynamik. Einer der Referenten sprach von geschätzt „18 Preisänderungen pro Jahr“. Microsoft lieferte hierzu live einen Nachweis, denn vom ersten auf den zweiten Konferenztage änderten sich, auch zur Überraschung der bei der Konferenz anwesenden Microsoft-Experten, diverse Preise und die mit den Diensten verbundenen Leistungsmerkmale.

Ausblick

Unter den Teilnehmern der Konferenz herrschte durchaus eine gespannte Aufbruchsstimmung, so wie man sie in den frühen Phasen der Einführung neuer (IT-)Technologien häufig erlebt. Entsprechend oft war in den Vorträgen von dawn und beginning die Rede. Man erlebe gerade einen Paradigmenwechsel in der Art und Weise, wie man Software-Projekte angehe und niemand wisse, wo man in fünf Jahren stehen werde. Dieser Umbruch generiert vollkommen neue Berufsbilder, die noch vor wenigen Jahren niemand auf der Agenda hatte. So gehören sogenannte Data Scientists aktuell zu den begehrtesten Experten in der IT-Branche. Dennoch bleibt die zentrale Frage, mit welchen Produkten und Geschäftsmodellen man aus den neuen IT-Technologien Geld machen kann? Dass allzu hochfliegende Pläne diesbezüglich von einer stabilen und leistungsfähigen Internet-Verbindung abhängen, konnten die Workshop-Teilnehmer am ersten Konferenztage live erfahren: Das am Konferenzort vorhandene WLAN erwies sich schnell als Flaschenhals sowohl für den Workshop-Leiter als auch die Teilnehmer, die versuchten Azure-Services online zu konfigurieren und dabei regelmäßig mit Fehlermeldungen über einen Verbindungsabbruch konfrontiert waren. Die Smart Data Developer Conference soll in einem jährlichen Turnus stattfinden. Die nächste Ausgabe der Veranstaltungsreihe findet bereits vom 18. bis 19. April in München statt. ■

Autor: Dipl.-Ing. Christian Demant,
Geschäftsführer,
Demant Industriesoftware GmbH
www.softwareforindustry40.com