

DEN MEISTEN UNTERNEHMEN FEHLT EINE DATENSTRATEGIE



Bild: Bowie 15-123rf

eine behördliche Meldung zu erfolgen hat. Nicht nur für mittelständische Unternehmen eine echte Herausforderung.

Die Fokussierung auf Big Data geht am Ziel vorbei

Die Beschäftigung mit der EU-DSGVO bringt in vielen Unternehmen ein grundsätzliches Problem ans Licht: Es fehlt eine Strategie für den Umgang mit Daten. Hinter Big Data steckte bisher vor allem eine oft planlose Datenerfassung; an Datenintegration, Zentralisierung, Analyse und den notwendigen Tools mangelte es zumeist. Denn das Thema Big Data war eher technologiegetrieben. In den Projekten ging es weniger um geschäftsunterstützende Strategien, sondern eher um technische Details und Algorithmen zur Verarbeitung großer Datenmengen. So haben zum Beispiel auf MapReduce basierende Technologien wie Hadoop – im Einsatz, um datenintensive Rechenprozesse auf Clustern mit verteilt arbeitender Software durchführen zu können – Datensilos entstehen lassen.

Unternehmen, die nicht genau wissen, wo ihre Daten gespeichert und welche Daten für welchen Zweck erfasst werden, können diese auch nicht nutzen. Der eigentliche Mehrwert von Big Data kann sich demnach erst durch Data Analytics eröffnen. Dabei dürfen Unternehmen mehr erwarten, als den monatlichen Management-Report mit Geschäftszahlen im Rückblick: Die gezielt geschäftsorientiert durchgeführte Datenanalyse optimiert Unternehmensprozesse und bietet die Basis für kundengenaue Geschäftsentscheidungen.

Data Analytics auf Big-Data-Projekte aufzusetzen ist – nicht erst seit die DSGVO gilt – der von der falschen Seite gedachte Weg. Ein Unternehmen wird dann sinnvolle, datengetriebene Entscheidungen treffen können, wenn es weiß, welche Metriken aus Business-Sicht wichtig sind. Typisch ist die Zusammenstellung von Umsätzen, Forecasts und Produkti-

Datenbasierte Analysen können enorme Optimierungs- und Einsparpotenziale in Unternehmen freisetzen. Nicht umsonst sind Big Data und Data Analytics im Rahmen von Digitaler Transformation und Industrie 4.0 viel diskutierte Themen, erst recht im Licht der DSGVO. Die Praxis sieht in den meisten Unternehmen jedoch anders aus.

Autor: Mathias Golombek **Redaktion: Axel Pomper**

► Das Inkrafttreten der EU-DSGVO Mitte des Jahres hat viele Unternehmen Zeit und Nerven gekostet. Denn die Vorgaben machen vieles, was bisher beim Umgang mit Daten üblich war, unmöglich. So dürfen Kundendaten beispielsweise ohne entsprechende Berechtigung kaum mehr gespeichert, geschweige denn anderweitig genutzt werden. Es muss klar geregelt sein, wer Zugriff auf welche Daten hat und wo sich diese befinden. Zudem sind Prozesse vonnöten, die klar regeln, wann ein Datenschutzverstoß vorliegt und wann

DATA ANALYTICS BRAUCHT PROZESSE UND AUTOMATISIERUNG.

onsstanddaten in regelmäßigen Abständen, monatlich etwa, bisweilen wöchentlich. Obwohl die notwendigen Daten alle digital vorliegen, dauert deren Auswertung Stunden, manchmal Tage, es müssen mehrere Systeme befragt und die Ergebnisse mühselig kumuliert werden. Abweichende Abfragen oder gar das Herstellen neuer Zusammenhänge sind kaum denkbar.

Data Analytics auf Basis der Geschäftsanforderungen gedacht, kann hingegen die Flexibilität und die Performance deutlich steigern. Strategische Überlegungen beginnen mit einer Auflistung der Daten und Auswertungen, die aktuell gemacht werden. Welche Daten werden dafür benötigt und wo sind sie gespeichert? Demgegenüber steht der Wunschzettel: Management-Reports flexibel nach Bedarf, mit Echtzeit-Daten, Forecast-Einberechnungen, Fachabteilungs-Input und so wandelbar, dass jederzeit neue Datenrelationen hergestellt werden können.

Gerade hier, bei den Überlegungen, welche Informationen die Geschäftsentscheidungen beeinflussen oder stützen könnten, darf und soll visionär gedacht werden: Vielleicht hängt der schwankende Umsatz von der Außentemperatur ab, obwohl zwar das Produkt unabhängig davon ist, aber die Lieferkette dadurch hakt. Vielleicht beeinflusst der sinkende Aktienkurs des Mitbewerbers den eigenen Umsatz nicht, weil sich die aktuelle Marketing-Kampagne an die falsche Zielgruppe richtet. Oder möglicherweise verursachen Lieferschwierigkeiten des Partners in naher Zukunft Probleme, denen jetzt noch entgegengesteuert werden kann.

Die Beispiele sollen verdeutlichen, dass Business-Entscheidungen, die schnell und datenbasiert getroffen werden, das Geschäft maßgeblich beeinflussen können. Das funktioniert allerdings nur, wenn der Data-Analytics-Prozess fundiert strategisch gedacht und als flexibler Prozess verstanden wird. Klare Definitionen, immer auch im Sinne der Datenschutzgrundverordnung, hinsichtlich Datenerfassung, Datenspeicherung und Datenzugriff sind notwendig und gehören in den Mittelpunkt der unternehmerischen Architektur. Nicht mehr nur Controller sollten Analysen machen können, sondern im Prinzip jeder Mitarbeiter mit seiner ganz eigenen fachlichen Sicht. Dies verlangt ein mehrstufiges Datenzugriffskonzept und einfach zu nutzende Front-End-Applikationen.

Große Datenmengen schnell analysieren

Gerade um große Datenmengen mit beliebiger Komplexität abbilden und in der entsprechenden Geschwindigkeit analysieren zu können, sind geeignete Werkzeuge notwendig. Im Zentrum einer Datenstrategie steht eine, bestenfalls zentral verwaltete Datenbank, die für Analysezwecke optimiert als Data Warehouse fungiert. Gerade in den Bereichen Big Data und Data Analytics haben sich die schnellen In-Memory-Datenbanken (IMDB) einen Namen gemacht.

Im Gegensatz zu herkömmlichen Datenbanksystemen nutzen IMDBs den Arbeitsspeicher als effizienten Beschleunigungs-Cache. Das erhöht die Zugriffsgeschwindigkeiten deutlich – eine wichtige Voraussetzung für flexibel gestaltbare und mit Echtzeitdaten arbeitende Analysen. Die höheren Kosten für den RAM bremsen die Performance der IMDBs zunächst aus. Inzwischen sind die Preise für RAM jedoch deutlich gesunken, sodass er kostenseitig kaum noch Rele-

vanz haben soll. Zumal nicht stets die gesamte Datenmenge im Hauptspeicher liegen muss, um Performance-Steigerungen zu erreichen. Das Argument der RAM-Kosten verliert ganz an Bedeutung, wenn Cloud-Infrastruktur im Spiel ist. Hier kann flexibel nach Bedarf skaliert werden, wobei Kosten nur nach tatsächlichem Verbrauch anfallen. Leistungsfähige Systeme, die heute am Markt erhältlich sind, arbeiten mit Algorithmen, die für die Analyse großer Datenmengen optimiert sind. Massive, parallele Datenverarbeitung nutzt beispielsweise Cluster-Architekturen gezielt aus.

Optimal für die Umsetzung einer Datenstrategie ist es, wenn das Datenbank- und Analysesystem zunächst nahtlos in die vorhandene Infrastruktur integriert wird. So lässt sich kurzfristig ein Mehrwert durch schnellere Analysen erzielen und mittelfristig lassen sich die Prozesse in Abstimmung mit den Fachabteilungen optimieren. Manche Systeme liefern Analysewerkzeuge und Datenschutzfunktionalitäten gleich mit – wichtige Auswahlkriterien, wenn das System ins Zentrum der Datenstrategie rücken soll. Die Integration verschiedener Datenquellen muss ebenso möglich sein wie die Unterstützung verschiedener zukunftsträchtiger Data-Science-Sprachen.

Eigentlich steckt hinter dem, was die Europäische Datenschutzgrundverordnung fordert, genau solch eine Datenstrategie. Unternehmen, die wissen, zu welchem Zweck sie welche Daten erfassen, wo diese gespeichert sind und wer darauf Zugriff hat, werden weniger Schwierigkeiten haben, die Vorgaben der DGSVO zu erfüllen. Gleichzeitig ist eine prozessual und weniger rein technisch gedachte Datenanalyse die Grundvoraussetzung, um aus den Daten echtes Potenzial für das Business zu generieren.

Mathias Golombek ist CTO bei Exasol

