

BIG DATA – FLUCH ODER SEGEN FÜR MASTER DATA MANAGEMENT (MDM)?

ABHÄNGIGKEITEN UND HERAUSFORDERUNGEN FÜR DAS DATA HANDLING

Eines der Top IT-Themen im Jahr 2013 war der Begriff „Big Data“ und sicher wird dieses Thema auch 2014 wieder die IT-Zeitschriften füllen. Was sich hinter dem Begriff großer Daten verbirgt und welche Implikationen dieser für andere datenrelevante IT-Disziplinen mit sich bringt wird jedoch kontrovers diskutiert. Als IT-Dienstleister sollte man für seine Kunden einen realistischen Blick auf Reifegrad und Mehrwert werfen. Dabei wird man unweigerlich über Abhängigkeiten und Zusammenhänge zu einem weiteren Thema aus dem Leistungsportfolio sprechen müssen: Master Data Management (MDM).

Unsere Kunden sehen sich vermehrt mit einem wachsenden Datenvolumen konfrontiert. Nach einer Studie von BARC¹ erwarteten 49 % der befragten Unternehmen in Europa 2013 ein Wachstum von über 25 Prozent. Die IDC² prognostiziert bis 2020 ein Datenwachstum auf 40.000 Exabytes, das entspricht von heute an einer Verdopplung der Datenmengen alle zwei Jahre.

Ein großer Treiber für die Zunahme der Datenmengen ist die Intensivierung der Anforderungen der Märkte auf Anbieter- und Kundenseite an Aktualität und Granularität von Informationen. Früher wurden etwa Publikationszyklen von Produktkatalogen auf jährlicher Basis durchlaufen, wohingegen heute ad hoc neue Kataloge oder Auszüge in diversen Formaten erstellt werden müssen. Bei Herstellern von Kraftfahrzeugen, Motoren und Getrieben explodieren die Produktvarianten aufgrund der größeren Notwendigkeit, den Kunden noch individuellere Produkte anbieten zu können, um sich von der Konkurrenz abzuheben. Im Bereich der Forschung werden Messverfahren immer besser, sodass bei Versuchsreihen die Datenmenge sofort höchste Ansprüche an die IT in Bezug auf schnelle und informative Analysen stellt. Neben diesen strukturierten Daten kommen vermehrt unstrukturierte aus B2B- und B2C-Applikationen hinzu, die es zu verwalten gilt. Unternehmen wollen insbesondere aus Social Media-Plattformen anonymisierte Informatio-

nen beispielsweise zu Kaufverhalten und Mobilität ihrer Kunden gewinnen, um hieraus Rückschlüsse für Entwicklung, Marketing und Support zu ziehen.

Der Begriff Big Data wird demnach treffend mit dem „5 V“ Modell charakterisiert. **V**olume steht dabei für die Menge an Daten, **V**ariety für deren Strukturvielfalt, **V**elocity für die Anforderungen an hohe Geschwindigkeit bei Analysen, **V**eracity bezieht sich auf die Vertrauenswürdigkeit und **V**alue auf den eigentlichen Wert der Information.

Der mengenmäßige Anstieg an Stammdaten im Kontext der Dimension „Volume“ ist der Entwicklung zu schulden, dass eine große Nachfrage nach immer aktuelleren und detaillierteren Informationen nachgekommen werden soll, aber dieser ist um Größenordnungen kleiner als bei Bewegungsdaten. Große Abhängigkeiten zwischen Big Data und MDM lassen sich also vielmehr anhand der anderen „V's“ (Variety, Velocity, Veracity, Value) erkennen. Bedingt durch die Datenflut und



¹BARC „BIG DATA Survey Europe“, 2012 · ²IDC „The digital universe in 2020“, 2012



insbesondere durch die steigende Menge der unstrukturierten Daten wird es immer schwieriger Informationen transparent und konsistent zu verwalten. Eine Aussage über Informationsgehalt und Qualität dieser unstrukturierten Daten wie auch deren Analyse, um zuverlässige Unternehmensentscheidungen generieren zu können, bleibt weiter eine große Herausforderung. Das klingt stark nach bekannten Fragestellungen aus dem Stammdatenmanagement. Mangelt es hier an Erfahrung, kann ein Big Data Projekt nicht immer die gewünschten Ergebnisse erzielen. Im Schnitt erhöhen CIO's nach schlechten Erfahrungen die Anstrengungen bei Data Quality und MDM um bis zu 50 %³.

Neben der reinen Erfassung, Speicherung und Nutzung von Big Data kommt es verstärkt auf ein funktionierendes Stammdatenmanagement und die Nutzung dort bekannter Methodik an. Somit hat MDM eine beachtliche Relevanz für Big Data, denn um Nutzen aus den Daten ziehen zu können, müssen die meist undurchsichtigen Datenmengen auf die wesentlichen Informationsträger wie Produkte oder Kunden qualitativ verdichtet werden. Ferner sind in den Datenmengen auch Informationen enthalten, die weitere Abhängigkeiten und Beziehungen zwischen diesen Informationsträgern

erkennen lassen, so etwa Substitutionsgüter anhand des Kaufverhaltens spezifischer Kundensegmente. Hiermit lassen sich zusätzliche „statische“ Merkmale ableiten und dienen so der mehrwertbringenderen Anreicherung von Stammdaten.

Big Data trägt folglich im Umkehrschluss auch dazu bei, dass MDM weiterhin eines der Top IT-Themen sein wird. Eine jährliche Wachstumsrate von 18% bei aktuellem Volumen von 2,17 Mrd. € weltweit spricht für sich.⁴ Gleichzeitig wird deutlich, dass sich MDM weiterentwickeln muss. So müssen etwa Data Governance Funktionen dahingehend erweitert werden, dass mit vertretbarem Aufwand Informationen aus fragwürdigen und zahlreichen Quellen regelbasiert klassifiziert und somit vertrauensvoll prozessseitig nutzbar gemacht werden. Vor allem aber müssen bestehende MDM-Lösungen selber mit größeren Datenmengen und erhöhter Performance bei Datentransfer und -abfragen zurechtkommen. Dies stellt ein Manko der meisten aktuellen Softwarelösungen auf dem Markt dar. Der Fokus ist auf das Beziehungsmanagement zwischen den Stammdaten gleicher und übergreifender Domänen zu richten, da dieses schnell mit mehreren 100 Mio. Datensätzen zu Buche schlagen kann. Entwicklungen hierzu greifen

selbst wieder auf In-Memory Technologien oder NoSQL Datenbanken, meist auf Basis des Frameworks Hadoop, zurück.

Aufgrund langjähriger Erfahrung aus realisierten Projekten im Bereich Data Handling, speziell der technologischen und fachlichen Expertise in den Bereichen Stammdatenmanagement und Datenbanktechnologien, bieten wir unseren Kunden eine ganzheitliche und professionelle Beratung. Als Dienstleister finden wir dort etablierte System- und Prozesslandschaften vor und starten deshalb nicht auf der grünen Wiese. Vielmehr suchen wir gemeinsam Synergien und schließen Lücken, wo eingesetzte Technologien an ihre Grenzen stoßen. Dabei ist es uns wichtig, zunächst unter Berücksichtigung unseres speziellen Vorgehensmodells, eine passende Big Data Strategie zu entwickeln, um sicherzustellen, dass man sich nicht im Datendschungel verliert. Voraussetzung hierzu ist eine MDM Strategie, die wesentliche Methoden und Prozesse beschreibt und nutzbar macht. Dies ermöglicht qualitativ gute Informationsträger und nachvollziehbare Datenflüsse bereitzustellen.

Der Big Data Hype wird mittelfristig dafür Sorge tragen, dass in deutschen Unternehmen der Stellenwert und Reifegrad von MDM weiter vorangetrieben wird. Eine MDM Strategie ist Basis für erfolgreiche Big Data Projekte und darauf aufbauende Analysen und Entscheidungen. Und obwohl MDM wesentlich mehr ausmacht als der reine Einsatz von Tools, werden gerade diese durch Big Data eine technologische Weiterentwicklung erfahren und so dem Umgang mit Unternehmensdaten zu Gute kommen.

KONTAKT



Holger Feigenbutz
Business Development Manager
Tel. +49 (0) 721 82800-585
holger.feigenbutz@isb-ag.de