

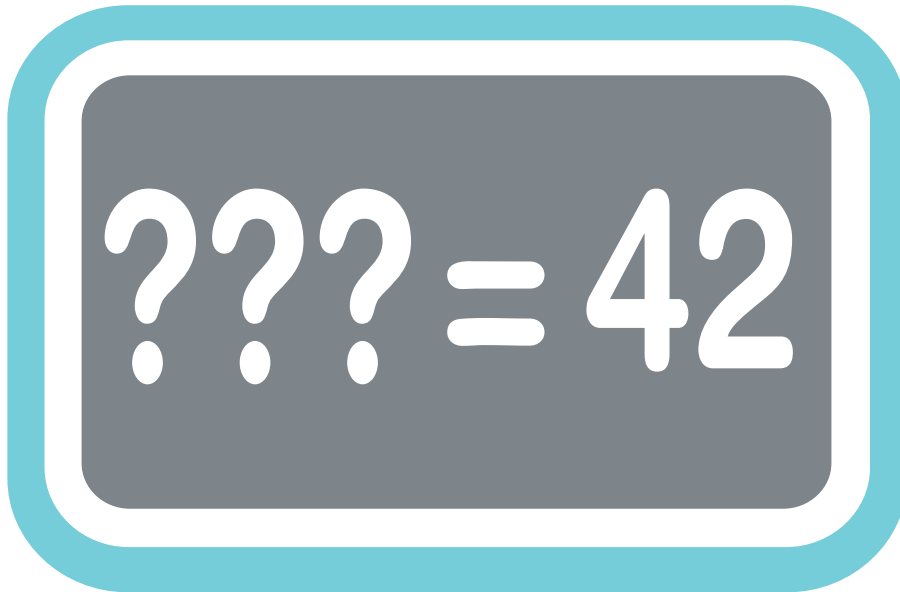
..... MACH EXPERTISE

BIG DATA

in der öffentlichen Verwaltung

.....





Big Data in der öffentlichen Verwaltung:

VIELFACH FEHLT NOCH DAS PROBLEM ZUR LÖSUNG

Big Data ist einer der meist diskutierten IT-Trends der letzten Jahre. Jedes namhafte Research-Unternehmen bespielt das Thema, die antizipierten Vorteile und Erwartungen durch Nutzung von Big Data steigen ins Unermessliche, kein Superlativ wird ausgelassen:

„Big Data überrollt die Welt“ titelt die Computerwoche im April 2014 – was bei durchschnittlichen jährlichen Wachstumsraten von 52 % auf erwartete 1,6 Mrd. Euro Marktvolumen¹ im Jahr 2016 in Deutschland sicher nicht übertrieben ist. „Daten sind das Öl des 21. Jahrhunderts“, schreibt BARC.

Der BITKOM sagt: „Neben Boden, Kapital und Arbeit sind Daten der vierte Produktionsfaktor.“²

Datability war das Leitthema der CeBIT 2014, der weltweit größten Messe für Informationstechnik. Das Kunstwort Datability (eine Zusammensetzung aus data und sustainability) beschreibt die zentralen Herausforderungen im Umgang mit Big Data, nämlich: „Die Fähigkeit große Datenmengen in hoher

¹Quelle: http://www.bitkom.org/de/presse/8477_78804.aspx

²BITKOM: Leitfaden: Management von Big-Data-Projekten, 2013

Geschwindigkeit verantwortungsbewusst und nachhaltig zu nutzen.“ Neu ist das Thema übrigens nicht. Bereits 2008 titelte das Magazin Nature erstmals mit dem Begriff „Big Data“ und machte ihn so einer breiteren Öffentlichkeit bekannt.³

Wie bei jedem gehypten Trendthema gibt es viele Trittbrettfahrer, die altbewährte (klassische) BI-Lösungen im neuen Gewand verkaufen. Und diejenigen, die für jede größere BI-Anforderung auf Big Data verweisen (und dann doch ihre klassische BI-Lösung implementieren, da diese völlig ausreicht). Dies führt zu einer Verunsicherung am Markt, speziell in der öffentlichen Verwaltung, wie folgende Anfrage eines Interessenten zeigt:

„Unser ERP-Datenbestand ist jetzt 50 GB groß! Wir brauchen eine Big Data Lösung! Stellen Sie uns mal Ihre vor.“ Diese Anfrage zeigt den Aufklärungsbedarf bezüglich Big Data. Braucht ein durchschnittlicher Kunde in der öffentlichen Verwaltung aktuell überhaupt eine Big Data Lösung?

Zur Beantwortung dieser Frage gilt es, sich zunächst dem Thema Big Data zu nähern. Soviel schon jetzt: Nicht jeder große Datenbestand ist ein Fall für Big Data!

Was genau bezeichnet Big Data überhaupt? Die Definition des BITKOM lautet: „Big Data bezeichnet den Einsatz großer Datenmengen aus vielfältigen Quellen mit einer hohen Verarbeitungsgeschwindigkeit zur Erzeugung wirtschaftlichen Nutzens.“

Häufig ist im Umfeld von Big Data von den drei V die Rede. Diese stehen für Volume (Datenmenge), Variety (Datenvielfalt) und Velocity (Geschwindigkeit), die sich auch in der BITKOM-Definition wiederfinden. Treffen diese zusammen, spricht man von Big Data:

Datenmenge: Große Datenmengen im Umfeld von Big Data bezeichnen eine völlig andere Dimension als in anderen Bereichen. Terabyte, wenn nicht sogar Petabyte (1 Petabyte = 1.000 Terabyte) sind die Mengeneinheiten, in denen gerechnet wird. Das zugrundeliegende Problem (die Verdoppelung der Datenmengen alle zwei Jahre) ist schon lange bekannt, war aber aufgrund parallel sinkender Speicherkosten bisher kaum ein Thema. In den nächsten Jahren wird das Wachstum der Datenmenge das Wachstum an Speicherkapazität übertreffen. Daher müssen neue Lösungen zum Umfang mit und zur Nutzung von Daten gefunden werden.

Datenvielfalt: Die Vielzahl an Datenquellen und Datenformaten nimmt ständig zu. Diese lassen sich grob in unstrukturierte (z. B. Präsentationen, Texte, Tweets, Blogs, Facebook-Inhalte und Videos bzw. Bilder), semistrukturierte (z. B. Sensordaten) und strukturierte Daten



Nicht jeder große Datenbestand ist ein Fall für Big Data!

Die drei V im Kontext von Big Data stehen für Volume (Datenmenge), Variety (Datenvielfalt) und Velocity (Geschwindigkeit).

³Nature Magazine (nature.com), September 2008

(z. B. Daten aus ERP-Systemen) unterteilen. Nach einer Veröffentlichung von IDC⁴ machen unstrukturierte Daten mehr als 90 % der gesamten Daten aus. Ein Potenzial, das erst durch Big Data genutzt werden kann. Dies verwundert nicht, da allein bei Facebook täglich über 500 TB an neuen Daten in verschiedensten Formen (Bilder, Videos, Kommentare, Likes) entstehen.

Geschwindigkeit: Die Verarbeitung und Auswertung der Daten soll schnell nach ihrer Entstehung passieren, wenn nicht sogar in Echtzeit. Online-Shops wollen unmittelbar – und nicht erst in einer Woche – wissen, welche Artikel gerade besonders gefragt sind. Dies stellt besondere Anforderungen an die Datenaufbereitung und -bereitstellung. Ebenso muss die Applikation selber sehr performant funktionieren.

Die meisten Kunden der MACH AG sehen sich diesen Anforderungen bisher (noch) nicht ausgesetzt. Ihr Berichtswesen weist, verglichen mit dem der Privat-Wirtschaft, aktuell noch einen geringeren Reifegrad auf. Legt man das Evolution of Analytics Services Modell zur Strukturierung des Berichtswesens von Crisp Research zugrunde, beschäftigen sich viele Kunden noch mit der vergangenheitsorientierten Auswertung von strukturierten Daten, z. B. aus ERP-Systemen. Der Einsatz von in-memory-Technologie zur Berichtsbeschleunigung ist bei den großen Kunden im Aufbau befindlich, ebenso wie die konsequente Verknüpfung und Nutzung aller zur Verfügung stehenden strukturierten Daten in einer zentralen Data-Warehouse-Umgebung. Aber dafür reicht eine klassische BI-Lösung allemal aus!



Der strikte deutsche Datenschutz spielt eine wichtige Rolle bei der weiteren Verbreitung von Big-Data-Lösungen in der öffentlichen Verwaltung

Zudem spielt der strikte deutsche Datenschutz eine wichtige Rolle bei der weiteren Verbreitung von Big-Data-Lösungen in der öffentlichen Verwaltung. Erst wenn die offenen rechtlichen Fragen aus dem Bundesdatenschutzgesetz zum Umgang mit Big-Data-Daten geklärt sind, (speziell das Verbotprinzip⁵ und die Zweckbindung⁶) wird sich die öffentliche Verwaltung stärker für das Thema öffnen und Lösungen vorantreiben.

Und das wird passieren. Zwar nicht heute und morgen, aber auch für die öffentliche Verwaltung wird Big Data eine wichtige Rolle spielen. Denn die öffentlichen Verwaltungen sitzen auf riesigen Datenschätzen, die es zu heben gilt, z. B. im Meldewesen oder durch die Sammlung umfassender Daten durch die Statistikämter.

⁴ IDC: *Extracting Value from Chaos*, Juni 2011

⁵ Dieses besagt, dass die Erhebung, Verarbeitung und Nutzung von personenbezogenen Daten grundsätzlich nicht erlaubt ist. Sie ist nur bei klarer Rechtsgrundlage erlaubt oder wenn die betroffene Person ausdrücklich (d. h. schriftlich) ihre Zustimmung zur Nutzung gegeben hat.

⁶ Dieses Gebot soll sicherstellen, dass Daten nur für den Zweck verarbeitet werden, für den sie erhoben worden sind.

Das Thema Big Data kann den Open-Data-Trend verstärken und die technischen Voraussetzungen zu dessen komfortabler Nutzung schaffen. Allein durch die Auswertung von öffentlich zugänglichen Daten und Informationen im Rahmen von Open Data in Berlin entsteht in den kommenden Jahren ein Nutzen von über 30 Mio. Euro⁷. Wer kann dazu schon nein sagen!?

Der spannende Nebeneffekt: Manche Steuerverwaltung wird sich durchaus dafür interessieren, ob einer ihrer Bürger gerade auf Facebook über seine Zweitwohnung am Meer schreibt, ohne als Zweitwohnungssteuerpflichtiger angemeldet zu sein ...

⁷Quelle: *Digitales Gold. Nutzen und Wertschöpfung durch Open Data für Berlin, Februar 2014*



Über den Autor



Christopher Decker, Jg. 1980, ist bei der MACH AG als Leiter Produktmanagement verantwortlich für die strategische Ausrichtung des kompletten Lösungsportfolios. Der

Diplom-Kaufmann ist seit über 7 Jahren im Unternehmen und hat als Berater diverse Verwaltungsmodernisierungsvorhaben und Software-Einführungen aktiv begleitet. Decker hat bereits auf zahlreichen Messen und Fachtagungen zum Thema ERP referiert und ist Mitglied in mehreren BITKOM-Arbeitskreisen zu IT-Themen.

Impressum

MACH AG
Wielandstraße 14
23558 Lübeck
Tel. 0451/70 64 70
Fax 0451/70 64 73 00

mailbox@mach.de
www.mach.de