

Big Data

Reichlich Potenzial

[31.01.2014] In der Verwaltung bieten sich zahlreiche Anwendungsfelder für den Einsatz von Big-Data-Methoden an, von der Stadtplanung bis hin zur Energieversorgung. Kommunale IT-Dienstleister können Big Data zudem zur Optimierung des Rechenzentrumsbetriebs nutzen.

Mit dem Begriff Big Data wird gemeinhin eine neue Qualität in der Sammlung und systematischen Analyse großer Bestände sowohl strukturierter als auch unstrukturierter Daten bezeichnet. Eine erste wichtige Triebkraft von Big Data ist, dass das Volumen und die Vielfalt der erzeugten Daten enorm zugenommen haben. Neben der wachsenden Zahl an datengenerierenden Technik-Artefakten, wie etwa mobile Endgeräte oder Sensoreinheiten, spielt dabei deren Vernetzung eine wichtige Rolle, da auf diese Weise sehr effektiv und mit geringen Kosten neue Bezüge der Daten zueinander hergestellt werden können. Zweite zentrale Triebkraft für Big Data sind Fortschritte in der Architektur der datenverarbeitenden Einheiten sowie bei der Konzeption von Algorithmen und Datenstrukturen, die zu Geschwindigkeiten bei der Analyse großer Datenbestände führen, welche vor wenigen Jahren noch undenkbar waren. Jüngstes Beispiel sind die öffentlich gewordenen Möglichkeiten der Geheimdienste bei der Analyse des weltweiten Internet-Datenverkehrs. Gerade daran wird deutlich, dass mit Big Data enorme Potenziale, aber auch erhebliche Risiken für demokratische Gesellschaften verbunden sind. Die mit der öffentlichen IT betrauten Organisationen und Unternehmen müssen aufgrund ihres Auftrags besonders sensibel für Fragen des Datenschutzes und der Informationssicherheit sein. Diese Dimension ist damit ein wesentlicher und integraler Aspekt eines jeden Big-Data-Projekts in der öffentlichen Verwaltung.

Datenmengen als Fortschritt

Der Public Sector mit seinen vielfältigen Aufgaben ist ein potenzialreiches Anwendungsfeld für Big-Data-Methoden und -Techniken. Dabei kann grundsätzlich zwischen regelmäßigem Standard-Monitoring bestimmter Daten und spezifischen, anlassbezogenen Analysen mit Projektcharakter (Ad-hoc-Analysen) unterschieden werden. Ein typisches Beispiel für eine standardisierte Auswertung aus der Privatwirtschaft ist die Analyse von Zahlungstransaktionen zur Aufdeckung von Kreditkartenbetrug (Fraud Detection). Entsprechende Verfahren der Mustererkennung (Pattern Recognition) könnten auch auf Daten der öffentlichen Verwaltung übertragen werden, etwa um die Aufdeckung von Steuer- und Sozialbetrug systematisch zu unterstützen. Der entscheidende Fortschritt von Big Data gegenüber den unter dem Begriff Business Intelligence zusammengefassten, vorhersagenden Analysen besteht in dem Anspruch, handlungsleitende Empfehlungen zu geben (Prescriptive Analysis). Derartige Analyseergebnisse und Informationen können in vielen Aufgabenbereichen der öffentlichen Verwaltung eingesetzt werden, etwa in der Verkehrsbeeinflussung, der Stadtplanung, im Gesundheitswesen, bei Energieversorgern oder im Jugend- und Sozialbereich. Ein häufig zitiertes Beispiel für eine projektbezogene Datenanalyse im kommunalen Bereich ist aus der Stadt New York bekannt. Um Problemen in der Kanalisation auf den Grund zu gehen, wurden vielfältige Datenquellen auf Basis von Geodaten miteinander in Beziehung gesetzt. Es zeigte sich, dass sich die Wahrscheinlichkeit von Verstopfungen der Kanalisation in der Nähe von Restaurants ohne adäquate Entsorgungsverträge erhöht – eine Information, auf deren Basis sinnvolle Maßnahmen zur Lösung des Problems ergriffen werden können.

Übersicht eröffnet Möglichkeiten

Für kommunale IT-Dienstleister liegt ein wichtiges Anwendungsfeld von Big-Data-Methoden in der Optimierung des Rechenzentrumsbetriebs. Um den wachsenden Anforderungen und Erwartungen der Kunden hinsichtlich Kapazität, Kontinuität, Performance, Sicherheit und Verfügbarkeit der IT-Services gerecht zu werden, sind die kommunalen IT-Dienstleister ständigem Veränderungs- und Anpassungsdruck ausgesetzt. Daher wird es immer wichtiger, frühzeitig Probleme zu erkennen und flexibel reagieren zu können. So lassen sich durch eine Echtzeitanalyse von Alerts, Events und Incidents im IT-Service-Management auftretende Ereignisse, welche die Servicequalität beeinflussen, proaktiv erkennen und entsprechend beheben. Damit eine Auswertung von Service Level Agreements (SLAs) in Form des aktuellen (Real-Time SLA) oder des zukünftigen Status (Predictive SLA) möglich ist, wird der Einsatz neuartiger Technologien notwendig. Für IT-Dienstleister müssen spezifische Anwendungsszenarien entworfen und evaluiert werden, sodass die Potenziale der Echtzeitanalysen, des proaktiven Handelns und der besseren Nutzung vorhandener Informationen klar erkennbar sind. Neben den Potenzialen müssen auch Themen wie die Amortisation von Anfangsinvestitionen, Schnittstellen-Aufwände und Nutzerakzeptanz analysiert und bewertet werden. Um die Services und die zugrundeliegende IT-Infrastruktur zu überwachen, werden in einem Rechenzentrum verschiedene Software-Lösungen für das Management und Monitoring eingesetzt, so zum Beispiel Nagios. Daneben stehen im Rechenzentrum viele Datenquellen zur Anbindung bereit, darunter Verfügbarkeitsdaten, IT-Auslastung, Daten zum Netzwerkverkehr oder Server-Temperaturen. Aufgrund der unterschiedlichen Software-Lösungen sowie Daten- und Informationsquellen fehlt in vielen Rechenzentren bislang die ganzheitliche Sicht zur Steuerung. In-Memory-Datenbanken und Big-Data-Techniken bieten sich als Lösung an und eröffnen damit vielfältige Möglichkeiten.

Lohnendes Potenzial

Um die Potenziale von Big-Data-Methoden und -Techniken im öffentlichen Sektor auszuschöpfen, sind zahlreiche Herausforderungen sowohl bei IT-Dienstleistern als auch in der Verwaltung zu bewältigen. Ein erster Aspekt betrifft den Aufbau einer entsprechenden Big-Data-Infrastruktur. Darüber hinaus sind die vielfältigen Datenquellen zu konsolidieren und zu integrieren und an geeigneter Stelle das notwendige Wissen um Big-Data-Methoden und -Techniken aufzubauen. Hierfür bedarf es spezialisierter Mitarbeiter, die im ohnehin angespannten Markt für IT-Fachkräfte angeworben werden müssen. Womöglich die schwierigste Herausforderung betrifft die Etablierung entsprechender Prozesse, die im Zusammenspiel von IT-Dienstleistern und Entscheidern in der Verwaltung zur Initiierung entsprechender Datenanalyseprojekte und zur Verwertung des generierten Wissens führen. Die Erfahrungen im Zusammenhang mit der seit Jahren verfügbaren Entscheidungsunterstützung durch Business Intelligence und Data Warehouse sind dabei wenig ermutigend, denn entsprechende Lösungen werden aus der Verwaltung bisher eher selten nachgefragt und genutzt. Die angespannte Situation der öffentlichen Haushalte und der überall spürbare Kostendruck könnten das Interesse an datengetriebener Entscheidungsunterstützung steigern. Gute Chancen dürften dabei insbesondere solche Projekte haben, die einen schnellen Rückfluss der notwendigen Investitionen in Aussicht stellen und Daten aus bereits vorhandenen Quellen analysieren.

()

Dieser Beitrag ist in der Februar-Ausgabe von Kommune21 im Schwerpunkt Big Data erschienen. Hier können Sie ein Exemplar bestellen oder die Zeitschrift abonnieren.

Stichwörter: Panorama, Big Data, Rechenzentrum, Vitako