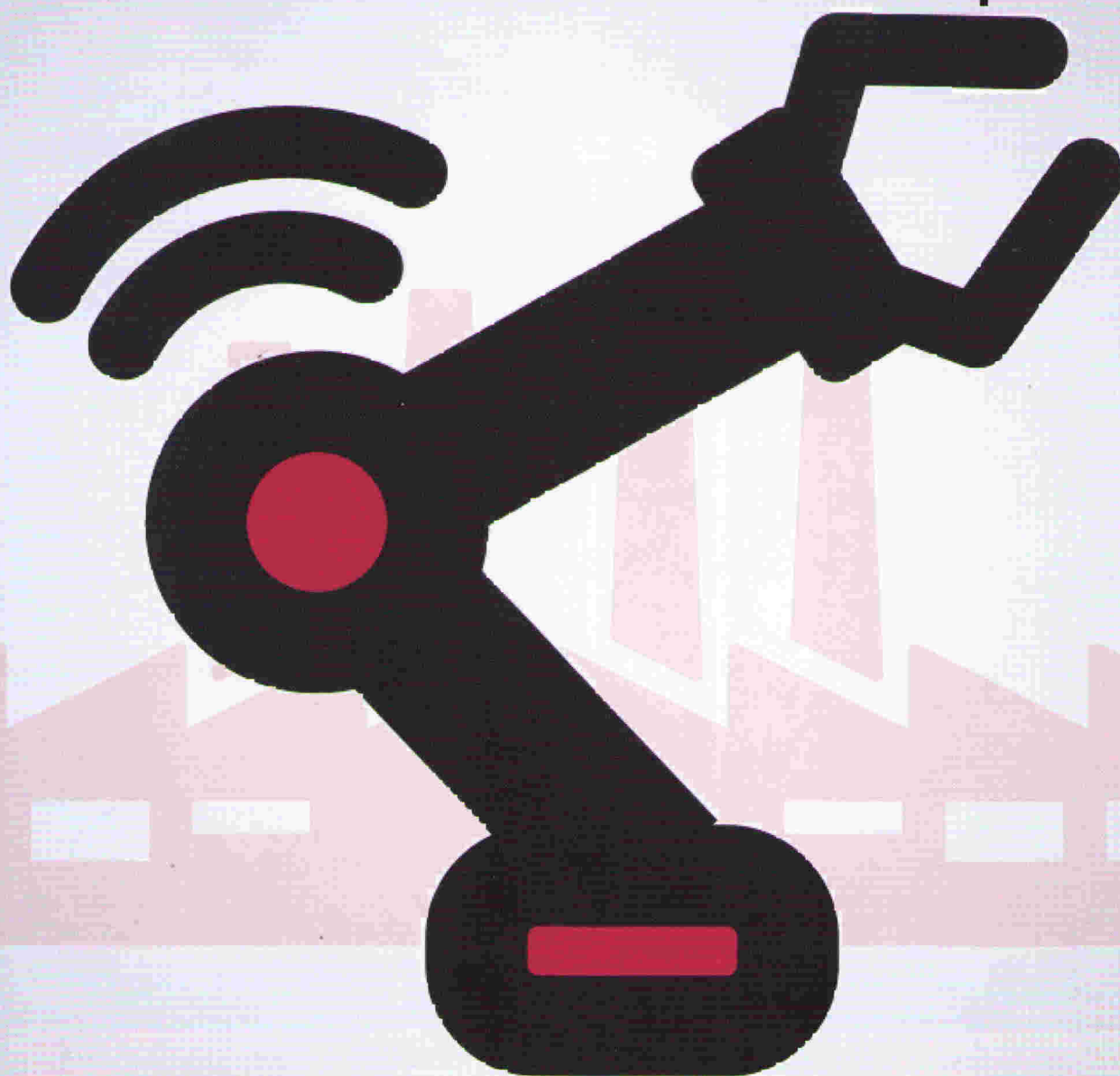


NR. 35 | MÄRZ 2016 | EINZELPREIS 14 €

FACHZEITSCHRIFT FÜR DIE ITSM-COMMUNITY

Die IT-Fabrik

SUPPORT FÜR INDUSTRIE 4.0



DAS BESTE AUS DREI WELTEN > MODERNES ITSM >
ITSM TOOLAUSWAHL > IT-SICHERHEIT 4.0 >
SERVICE MANAGEMENT BEYOND IT >
AUTOMATION TOOLAUSWAHL

PREDICTIVE ANALYTICS

Wie große Datenmengen für Prognosen genutzt werden

TEXT MARKUS RAATZ

Aufzüge, die selbst den Techniker benachrichtigen, wenn eine Reparatur ansteht, Polizeistreifen, die vor einem geplanten Einbruch vor Ort sind, um ihn zu verhindern – das klingt noch etwas nach Science Fiction, ist es aber nicht. Es sind Anwendungen von Predictive Analytics – und sie sind erst der Anfang. In Zukunft werden Kombinationen aus Analyse und Modellierung immer häufiger eingesetzt, um Entwicklungen vorherzusagen und Handlungsoptionen auszuwählen.

Predictive Analytics ist im Grunde die logische Weiterentwicklung in der Datenerhebung, -analyse und -verarbeitung. In den letzten Jahren ermöglichten wachsende Rechenleistungen die Analyse riesiger Datenmengen, bekannt unter dem Begriff Big Data. Darin Muster und Zusammenhänge zu entdecken, konnte in der Vergangenheit nur ein erster Schritt sein. Im zweiten Schritt sollen solche wiederkehrenden Muster genutzt werden, um ihre künftige Entwicklung zu prognostizieren und frühzeitig Handlungsoptionen ableiten zu können. Das ist das Ziel von Predictive Analytics Modellen. Um das zu erreichen, kombiniert man Methoden aus Data Mining und Business Intelligence, Spieltheorie und Optimierung. Dabei kann zum Teil auch schon auf vorgefertigte Algorithmen zurückgegriffen werden, wie sie etwa Microsoft Azure Machine Learning bietet.

Sind die Datenmengen ausreichend groß und unterstellt man eine Fortsetzung der Entwicklung in der Vergangenheit auch für die Zukunft, so sind weite Anwendungsfelder denkbar. Praktische Umsetzungen gibt es bereits. Zu den ersten Anwendern gehörten

Banken, die die Prognoseverfahren nutzen, um künftige Entwicklungen von Wertpapieren aber auch Kreditnehmer besser einschätzen zu können. Mögliche Risiken sollten so begrenzt werden. In den USA nutzt sogar schon die Polizei Predictive Analytics teilweise zur Prävention von Verbrechen. Ein Zustand, wie im Science Fiction Film „**Minority Report**“ beschrieben, bleibt zwar eine Fiktion, aber es ist bereits möglich, die Wahrscheinlichkeit von Einbruchsdiebstählen oder Gewaltausbrüchen für bestimmte Zeiten und Straßenzüge näher einzugrenzen. Dort kann die Polizei Ihre Streifen dann gezielt verstärken.

Die Logistik, ebenfalls ein Bereich mit riesigen Datenmengen, setzt Predictive Analytics Verfahren ein, um beispielsweise unter Verwendung von System- und Echtzeitdaten die Zeit der voraussichtlichen Zustellung an private Empfänger bis auf eine Stunde eingrenzen zu können – unnötige Wartezeiten für den Empfänger zuhause werden so vermieden. Die Vorhersagen spielen auch im Vertrieb eine wichtige Rolle. Kundenrankings auf Basis dieser Daten und Methoden erweisen sich – nicht selten zur

Überraschung selbst erfahrener Sales-Verantwortlicher – als äußerst genau. Zeitschriftenverlage, für die das Abonnementgeschäft existenziell ist, nutzen ebenfalls Predictive Analytics, um frühzeitig zu erkennen, welche ihrer Kunden bestehende Abonnements voraussichtlich nicht verlängern. So können Sie Maßnahmen zur Kundenbindung genau auf diese Kunden zuschneiden. In diesen Fällen kann der Vertrieb seine Ressourcen gezielter einsetzen.

Zukunftsweisend sind außerdem Internet-of-Things-Anwendungen wie das Wartungsprogramm MAX, das die ThyssenKrupp Elevator AG in Zusammenarbeit mit Microsoft entwickelt hat. Der Aufzughersteller geht davon aus, dass bei weltweit 12 Millionen Aufzügen jedes Jahr 190 Millionen Stunden durch Wartungsarbeiten für den Betrieb verloren gehen. Daher hat sich der Hersteller entschlossen, mit der Microsoft Azure IoT Cloud ein System zur Echtzeitanalyse der Daten aller Aufzüge weltweit zu installieren. Mit Hilfe komplexer Algorithmen kann ThyssenKrupp Elevator so Restlebensdauer und notwendige Reparaturen ermitteln. Die Daten sendet die Cloud in Echtzeit an 20.000 Servicetechniker weiter. Diese können anstehende Reparaturen in Zeiten mit geringer Auslastung eintakten. Im Ergebnis sollen die Ausfallzeiten durch Wartungsarbeiten um 50 Prozent sinken.

Die Bandbreite an Anwendungen, bei denen Predictive Analytics Verfahren künftig zum Einsatz kommen werden, wird weiterhin zunehmen. Es ist außerdem davon auszugehen, dass dies umgekehrt sogar die Art der Datenspeicherung beeinflussen wird. Da viele Datenzusammenhänge im Voraus noch gar nicht bekannt sein können und sich gerade durch die Analysen selbst auch neue Anwendungen ergeben, könnte sich der Ansatz durchsetzen, zunächst einmal so viele Daten wie möglich in Megaspeichern, den „Data Lakes“ für spätere Auswertungen zu sammeln. Technisch ist das mit der Lösung Microsoft Azure Data Lake bereits umsetzbar.

DIE STUNDE DER PAKETZUSTELLER – FALLBEISPIEL DPD

Der Paket- und Expressdienstleister DPD, ein langjähriger Kunde der Ceteris AG, sah sich vor der Herausforderung, das Geschäftsmodell an eine neue Entwicklung, nämlich das überdimensional hohe Wachstum im Privatkundenbereich, anzupassen. Haben bis vor einigen Jahren noch die Zustellungen an Unternehmen deutlich überwogen, so sind Privatkunden aufgrund des wachsenden Online-Handels nun

eine zunehmend größer werdende Zielgruppe. Dies bedeutet für die Logistiker, dass Kunden zu beliefern sind, die in der Regel tagsüber nicht zuhause sind und auch bei angekündigten Lieferungen nicht gut einen halben Tag auf ihr Paket warten können. Logistikanbieter, so auch DPD, schöpfen inzwischen alle Möglichkeiten aus, die sich durch geschützte Abstellmöglichkeiten oder alternative Lieferorte wie die Abgabe bei Nachbarn, an die Arbeitsstelle oder Paketshops ergeben. Betrachtet man aber den klassischen Auslieferungsfall, rückt das Zeitfenster der Auslieferung in den Fokus. In diesem Logistik-Kernbereich wollte sich DPD durch besonderen Service vom Wettbewerb absetzen: Das Zeitfenster für eine Auslieferung sollte auf eine Stunde eingegrenzt und dem Kunden vorab mitgeteilt werden.

Bei rund 1,5 Millionen Paketen pro Tag und gut 30 Millionen operativen Datensätzen war das ein ambitioniertes Ziel. Die Hürde bestand vor allem darin, dass die Daten auf unterschiedliche Systeme dezentral verteilt waren. Hinzu kamen Rechnungsdaten sowie polystrukturierte Daten wie Maschinenprotokolle, oder Videos, die ebenfalls dezentral gespeichert waren.

Die Lösung bot hier der Einsatz eines Microsoft SQL 2012 Servers Parallel Data Warehouse (PDW). Markus Raatz, Vorstand der Ceteris AG, sieht die Vorteile vor allem in der schnellen und einfachen Anwendung – auch ohne große Optimierung und den damit verbundenen Programmieraufwand ließen sich schon hohe Geschwindigkeiten erzielen. Für DPD war zudem entscheidend, dass man bereits ein Team aus 15 erfahrenen Microsoft SQL Spezialisten im Hause hatte.

Auf Basis dieser Zusammenführung und Analyse aller Daten konnte das Projekt realisiert und die Zustellung innerhalb eines einstündigen Zeitfensters umgesetzt werden.

Die geballte Rechenleistung, die hinter der Anwendung steckt, resultiert schließlich in einer einfachen Übersicht für den Endkunden: Nachdem er bereits im Vorfeld aus mehreren Optionen wählen konnte, um Ort und Tag der Zustellung seinen Wünschen anzupassen, erhält er am Zustelltag per SMS die Information, in welcher Stunde sein Paket ausgeliefert wird. Er kann bequem online verfolgen, wo sich seine Lieferung gerade befindet. Nähert sich der voraussichtliche Lieferzeitpunkt, wird das Zeitfenster zudem auf 30 Minuten eingegrenzt.