

ERFOLGREICH BIG DATA PROJEKTE REALISIEREN

WHITEPAPER

Lösen Sie sich von überholten Traditionen. Die Verfolgung Ihrer Unternehmensziele muss nicht so enden, wie die Analysten es prophezeien -als eines von zahllosen gescheiterten IT-Projekten. Sie können all Ihre Daten nutzen, um Ziele schneller und kostengünstiger zu erreichen. Entscheidend ist, dass Sie ergebnisorientiert arbeiten, indem Sie sich bei Ihrem nächsten Projekt mit der richtigen Technologie wappnen.



Contents

Kurzfassung	1
Es ist an der Zeit, Datenintegrationsprojekte neu zu überdenken	2
Technologische Hürden	
Denkblockaden	
Die Hindernisse	
Risikominderung durch Überdenken bisheriger Technologie-Entscheidungen	4
Erfolgsfaktor Mindset	
Erfolgsfaktor Technologie	
Erfolg für Ihr Unternehmen	5
Von Erfahrungen profitieren	6
Implementierung eines landesweiten Datenintegrationsprojekts mit beispiellosem Umfang	
Entwicklung eines skalierbaren Data Hubs für die Echtzeit-Bereitstellung integrierter Informationen	
Schnelle Implementierung eines zentralen, sicheren globalen Trade Stores in sechs Monaten	
Neuer Maßstab für die dynamische, bedarfsgerechte Bereitstellung von Inhalten	
Unternehmen verändern sich in rasantem Tempo – die IT sollte mithalten können	12
Weitere Informationen	

“ Aus Interviews und Schätzungen von Experten geht hervor, dass Data Scientists bis zu 80 Prozent ihrer Zeit mit Aufgaben wie Datensammlung und Aufbereitung von Daten verbringen.

– Steve Lohr, New York Times

KURZFASSUNG

Volumen, Heterogenität und Komplexität von Daten nehmen angesichts der Menge an Informationen, die tagtäglich erstellt und genutzt werden, schneller als je zuvor zu. Von Datenknappheit kann keine Rede sein. Tatsächlich sollen bis zum Jahr 2020 für jeden Menschen auf der Welt rund 1,7 MB neue Daten pro Sekunde erzeugt werden.¹ Betrachtet man all die Herausforderungen, die durch diese anhaltende *Informationsfluten entstehen*, führt das Aufkommen von Cloud Computing, mobilen Plattformen und dem *Internet of Things (IoT)* dazu, dass Unternehmen sich immer mehr mit einem effizienten Umgang mit der Masse an unstrukturierten Daten und ständig wechselnden Anforderungen beschäftigen. Es überrascht nicht, dass 85 Prozent der Unternehmen vorhandene Daten nicht richtig nutzen können und daraus keinen entscheidenden Wettbewerbsvorteil erlangen.²

Die Implementierung komplexer Unternehmensprojekte birgt Risiken. *ComputerWorld* berichtete kürzlich von einer Studie über Websites, die über 10 Millionen US-Dollar gekostet haben. Die Studie zeigte, dass von den 3.555 Projekten aus dem öffentlichen und privatwirtschaftlichem Bereich nur 6,4 Prozent erfolgreich waren. Bei über der Hälfte dieser Projekte gab es Probleme wie z. B. die Budget-Überschreitung, nicht eingehaltene Termine oder unzufriedene Benutzer.³

Laut Gartner wird bis 2018 bei der Implementierung neuer Großprojekte rund die Hälfte des Budgets in die Datenintegration fließen.⁴

Wenn IT-Projekte auf veralteter Technologie und mit veralteten Denkansätzen moderne Systeme schaffen sollen, ist ein Scheitern schon fast vorprogrammiert. Tatsächlich prognostiziert Gartner, dass bis 2017 60 Prozent der Big-Data-Projekte schon durch die Pilot- und Testphase fallen und aufgegeben werden.⁵

Wie werden Sie Ihr nächstes Datenintegrationsprojekt abwickeln, um sicherzustellen, dass Sie damit Ihre Geschäftsziele auch wirklich erreichen? Wie können Sie vermeiden, dass Ihr nächstes Datenprojekt scheitert? Die kurze Antwort lautet: Stellen Sie ein Implementierungsteam mit der *richtigen Einstellung* und der *richtigen Technologie* für den Erfolg zusammen.

Die gute Nachricht ist: Mit einer flexiblen Einstellung und Technologie, die sich ständig an die Anforderungen und Gegebenheiten anpassen, kann Ihr Unternehmen erfolgreich Datenintegrationsprojekte implementieren, damit Sie Ihre Geschäftsziele erreichen.

Dieses Whitepaper soll das Risiko eines Scheiterns verringern. Sie finden hierin praxisorientierte Tipps, um komplexe Datenintegrationsinitiativen erfolgreich abzuschließen. Der Generalinspektor des US-Gesundheitsministeriums untersuchte die Entwicklung der Website *Healthcare.gov* in einer ausführlichen Fallstudie, die im Februar 2016 veröffentlicht wurde. Sie werden bemerken, dass dieses Whitepaper sich häufig auf diesen öffentlich verfügbaren Bericht bezieht, um besseren Einblick in die komplexen Herausforderungen und deren Bewältigung bei der Implementierung von *Healthcare.gov* zu geben.

1 IDC: Digital Universe Study: Big Data, Bigger Digital Shadows and Biggest Growth in the Far East. Dezember 2012. <http://www.whizpr.be/upload/medialab/21/company/Media_Presentation_2012_DigiUniverseFINAL1.pdf>

2 Gartner: Information Innovation: Innovation Key Initiative Overview. April 2012. <<https://www.gartner.com/doc/1991317/information-innovation-innovation-key-initiative>>

3 Computerworld: Healthcare.gov website 'didn't have a chance in hell'. Oktober 2013. <http://www.computerworld.com.au/article/529631/healthcare_gov_website_didn_t_chance_hell_/>

4 Gartner: Predicts 2013 Application Integration, 2013. <<https://www.gartner.com/doc/2238015/predicts--application-integration>>

5 Gartner: Advancing Business with Advanced Analytics. <<https://www.gartner.com/user/registration/prospect?resId=3090420&srclId=1-6470978268>>

ES IST AN DER ZEIT, DATENINTEGRATIONSPROJEKTE NEU ZU ÜBERDENKEN

Unternehmen jeder Art arbeiten intensiv an der Konsolidierung und Auswertung von Datenbeständen für IT-Strategien, um Innovation und Wachstum zu ermöglichen. Allerdings ändern sich Daten oftmals und wachsen ständig, wodurch die Datenintegration zu Herausforderung wird. In dem Moment, indem Sie glauben, endlich die Anforderungen Ihres aktuellen Projekts im Griff zu haben, kommen noch mehr Daten hinzu - und damit neue Komplexitäten. Obwohl Datenarchitekten unablässig an neuen Wegen für die Verwaltung und Nutzung von Unternehmensdaten arbeiten, profitieren die meisten Unternehmen kaum von den Daten.

TECHNOLOGISCHE HÜRDEN

Bestehende Altsysteme, die für einen speziellen Zweck entwickelt wurden, verhindern oftmals Einblicke und haben zu großen isolierten Daten-Silos geführt. In diesen Daten-Silos befinden sich mehrere Kopien der Unternehmensdaten, die zu Herausforderungen bei der Data Governance und Richtigkeit der Daten führen. Das wiederum hat zur Folge, dass ein Unternehmen die Anforderungen seiner Kunden nicht auf kosteneffiziente Weise erfüllen kann.

Aufgrund ihrer Komplexität sind Datenintegrationsprojekte in Unternehmen riskant, zeitaufwändig und mit einer kostspieligen Implementierung verbunden. Unternehmen haben versucht, diese Komplexität mit verschiedenen technologiebasierten Strategien zu verringern, die sich allesamt als suboptimal herausgestellt haben. Dazu gehört z. B. das Enterprise Data Warehouse (EDW), das Daten von verschiedenen Geschäftsbereichen zusammenführt, um nachgeschaltete Analysen zur Entscheidungsfindung zu unterstützen. EDWs bleiben oft hinter den Erwartungen zurück, weil sie langfristig die Verlagerung der Daten von den Altsystemen in das EDW erfordern. Dazu kommt, dass EDWs meistens auf inflexiblen relationalen Datenstrukturen basieren, die mit wechselnden Ursprungsstrukturen und nachgeschalteten Anforderungen nur schwer zurecht kommen.

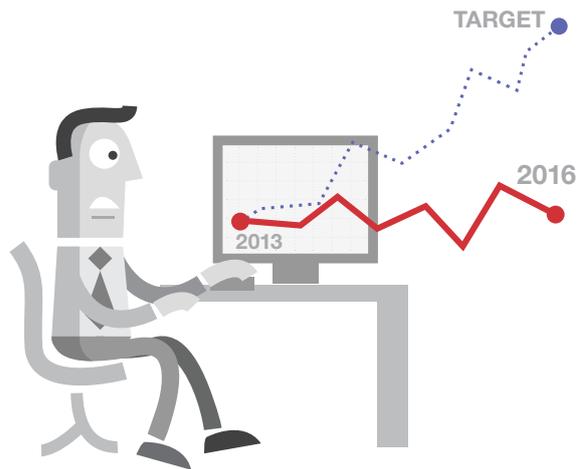


Abbildung 1: Komplexe Datenintegrationsprojekte sind riskant.

Dann gibt es die Enterprise Application Integration (EAI) und die Service-oriented Architecture (SOA) zur Unterstützung von Geschäftsprozessen. Diese Strukturen sind zwar für die unternehmensweite Integration von Systemen und Anwendungen ausgelegt. Leider können aber viele Anwendungen in solchen Systemen nicht miteinander kommunizieren, was dazu führt, dass die gleichen Datenbestände mehrfach an vielen verschiedenen Stellen gespeichert werden. Das führt letzten Endes nicht selten zu unübersichtlichen Point-to-Point-Verbindungen, die das Management erschweren.

Dazu kommt, dass zur Überwindung dieser Hürden einige Unternehmen mit Data Lakes (wie Hadoop) für große Datenspeicher mit strukturierten und unstrukturierten Daten experimentieren. Die Gefahr bei diesen „Daten-Seen“ besteht darin, dass sie schnell zum „Sumpf“ werden können, wenn Unternehmen sehr große Datenmengen hochladen, ohne dass das per Governance richtig geregelt wird. Das Ergebnis ist eine unbrauchbare Datenmasse, die keinerlei Wert besitzt, geschweige denn Einblicke ermöglicht.⁶

DENKBLOCKADEN

Heute ist es leider oftmals an der Tagesordnung, dass viele IT-Projekte weder innerhalb des vorgesehenen Zeitrahmens noch Budgets bleiben. Hohe Kosten und gescheiterte Projekte sind heute „normal“. Laut McKinsey & Company sieht die Realität bei der Hälfte der IT-Projekte mit einem Budget von über 15 Millionen US-Dollar so aus:

⁶ icrunchdatanews.com: How to Keep Your Data Lake from Becoming a Data Swamp. April 2016.
<<https://icrunchdatanews.com/how-keep-your-data-lake-from-becoming-data-swamp/>>

“ Die Hälfte der IT-Projekte mit Budgets über 15 Millionen US-Dollar überschreiten das Budget um 45 %, liegen zu 7 % hinter dem Zeitplan und bieten 56 % weniger Funktionalität als erhofft.”

– McKinsey & Company

- Das Budget wird zu 45 Prozent überschritten.
- Der Zeitplan wird zu 7 Prozent überschritten.
- Die IT-Projekte bringen 56 Prozent weniger Funktionalität als erwartet.

Was noch schlimmer ist: Rund 17 Prozent der IT-Projekte verschlingen Summen, die die Existenz des Unternehmens gefährden können.⁷

Warum werden diese Fakten nicht behoben, sondern einfach hingegenommen? Eine Theorie ist, dass alle Unternehmensebenen häufig Veränderungen ablehnen, wenn damit neue Technologien und Prozesse verbunden sind. Tatsächlich bestätigt Andrea Simon, Ph.D., Cultural Anthropologist und CEO von Simon Associates Management Consultants, in einem Forbes-Interview, dass sowohl das Management als auch die Mitarbeiter Probleme damit haben zu akzeptieren, dass „die Grenzen des Wachstums hausgemacht sind“ und durch Gewohnheiten und Praktiken entstehen, die einst zum Geschäftserfolg führten. Das mag auf den ersten Blick widersprüchlich klingen, aber die Crux ist, dass „die Wahrnehmung des Vergangenen verhindert, die Gegenwart klar zu erkennen.“⁸

Diese Tendenz zum Zurückblicken kann blind gegenüber neueren Lösungen machen. Im Falle der Integration moderner Daten tendieren Enterprise Architects dazu, sich auf die Altechnologie zu verlassen, die für sie in der Vergangenheit funktioniert hat. Dies gilt insbesondere für relationale Datenbanken, die als einzige Möglichkeit zum Erstellen neuer Systeme betrachtet werden.

DIE HINDERNISSE

Nehmen wir zur Veranschaulichung den US-Bundesstaat Oregon, den die gescheiterte

Entwicklung eines Marktplatzes für Krankenversicherungen 250 Millionen US-Dollar gekostet hat – das Projekt *Cover Oregon*, das auf einer relationalen Datenbanktechnologie basierte. Vor der Ankündigung, dass die Marktplatz-Website vom Web genommen wird und Bürger wieder auf den vollfunktionstüchtigen landesweiten Krankenversicherungsmarktplatz *Healthcare.gov* umgeleitet werden, wurde auf *Obamacarefacts.org* erklärt: „Die Umstellung wurde von einem Beratergremium nach gründlicher Prüfung empfohlen, da die Fehlerbehebung der Website bis zu 78 Millionen US-Dollar kosten könnte, während die Umstellung auf *Healthcare.gov* nur rund 7 Millionen US-Dollar kosten durfte.“⁹ Fortune erklärt dazu: „Der Staat sieht die Schuld für das Scheitern der Krankenversicherungs-Website *Cover Oregon* in der inkompetenten Arbeit des Software-Anbieters.“ Weiter wird behauptet, dass die erfolgreiche Einführung dieses Krankenversicherungsmarktplatzes durch den Staat von vornherein wegen technischer Probleme zum Scheitern verurteilt gewesen sei. Deshalb wirft der Staat dem Technologie-Anbieter vor, die Überzeugung der staatlichen Entscheidungsträger, *Hunderte Millionen Dollar in Produkte und Dienstleistungen zu investieren, die dann nicht die versprochene Leistung bringen, käme einem Betrug gleich*. Der Technologie-Anbieter wiederum behauptete, dass der Staat selbst am Scheitern des Datenprojekts Schuld gewesen sei.¹⁰

Das Ergebnis dieser Rechtsstreitigkeiten ist noch offen. Doch unabhängig vom Ausgang ist, dass Hunderte Millionen Steuergelder bei einem wichtigen Datenprojekt in den Sand gesetzt wurden. Damit drängen sich einige Fragen zur Gesamtplanung des Projekts auf: Warum wurde bei

⁷ McKinsey & Company: Delivering large-scale IT projects on time, on budget and on value. Oktober 2012.

<http://www.mckinsey.com/business-functions/business-technology/our-insights/delivering-large-scale-it-projects-on-time-on-budget-and-on-value>

⁸ Forbes: Why We're So Afraid of Change – and Why That Holds Businesses Back. April 2013.

<http://www.forbes.com/sites/womensmedia/2013/04/08/why-were-so-afraid-of-change-and-why-that-holds-businesses-back/#65d63d865cd0>

⁹ *ObamaCareFacts.com*: Cover Oregon Marketplace Changes. Aktualisierung vom 25. April 2015.

<http://obamacarefacts.com/insurance-exchange/oregon-health-insurance-exchange/>

¹⁰ Fortune.com: Oregon sues Oracle over failed Obamacare website. August 2014. <http://fortune.com/2014/08/22/oregon-oracle-spat-200-million-lawsuit/>

“ Wir alle kennen das. Wenn man macht, was man immer macht, bekommt man das, was man immer bekommt.”

– Tony Robbins

einem gesetzlichen Programm, das die Integration großer, unterschiedlichster Datenmengen vorsieht, automatisch davon ausgegangen, dass die relationale Datenbanktechnologie hierfür die beste Wahl sei? Und was den Technologie-Anbieter betrifft: Erhielt er den Projektzuschlag wegen seiner Fähigkeit und Bereitschaft, sich an die Ziele und Vision des Staates anzupassen? Oder beruhte die Entscheidung des Staates für diesen Anbieter einfach auf der Vertrautheit mit einer Technologie, die in der Vergangenheit gut funktioniert hat? Fragen wie diese zwingen Organisationen, genauer die Beweggründe unseres Handelns zu prüfen.

Jedes Unternehmen will seine Daten nutzen, um von Innovationen und Wettbewerbsvorteilen im Markt zu profitieren. Ihre Chancen auf Erfolg steigen und fallen mit Ihrer Fähigkeit, Risiken zu minimieren. Und dafür müssen Sie sicherstellen, dass die Technologie die Unternehmensziele vorantreiben und die Weichen auf Erfolg stellen.

RISIKOMINDERUNG DURCH ÜBERDENKEN BISHERIGER TECHNOLOGIE-ENTSCHEIDUNGEN

ERFOLGSFAKTOR MINDSET

Haben Sie sorgfältig geprüft, dass Sie die richtigen Experten ausgewählt haben, damit Ihr Programm ein Erfolg wird? Der erste Schritt ist die Bewertung der Einstellung innerhalb Ihres Unternehmens sowie der möglicher Partner, um sicherzustellen,

dass Sie sich für die Umsetzung Ihrer Vision für die Richtigen entscheiden. Die richtigen Personen werden sich problemlos an den Innovationsgeist Ihres Unternehmens anpassen und diesen mit Ihren heutigen und künftigen Zielen in Einklang bringen – und zugleich Risiken gering halten. Um die Wahrscheinlichkeit einer gelungenen Implementierung erfolgskritischer Datenprojekte zu erhöhen, ist es entscheidend, dass von Anfang an Mitarbeitern und Partnern das richtige Mindset kommuniziert wird: eine flexible, moderne Einstellung. Dies ist genauso wichtig wie die Wahl der richtigen Technologie für das Projekt.

ERFOLGSFAKTOR TECHNOLOGIE

Verlassen Sie sich auf Technologie von gestern, um Probleme von heute zu lösen? Flexibilität und Agilität gehören zu den Hauptkomponenten einer modernen Systementwicklung, weil sie das System und Ihre Investitionen zukunftssicher machen. Niemand weiß, was genau die Zukunft bringen mag. Aber mit einem flexiblen, agilen System, das leicht an wechselnde Anforderungen angepasst werden kann, sind Sie optimal darauf vorbereitet.

RDBMS

Data-Management-Systeme (RDBMS) gibt es seit über 30 Jahren. Dennoch kann der jahrzehntelange Einsatz in der Praxis nicht den grundsätzlichen Mangel an Agilität wettmachen, der bei der relationalen Technologie systemimmanent ist. Die Entwicklung jedes neuen, agilen Systems wird ein flexibles System für das Data Management erfordern.

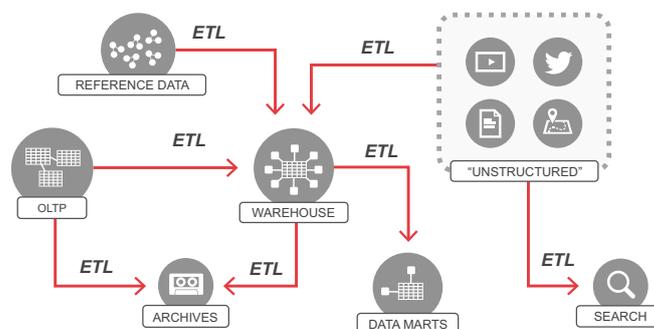


Abbildung 2: ETL ist komplex und zeitaufwändig.

Die relationale Technologie beschränkt die Möglichkeiten von Unternehmen, die eine Datenplattform für mehrere Datenquellen benötigen. Das liegt daran, dass die relationale Technologie zu einer Zeit entwickelt wurde, als es noch papierbasierte Prozesse (z. B. bei der Gehaltsabrechnung oder im ERP-Bereich) gab, die automatisiert werden mussten. Die Prozessautomatisierung ist nicht mehr das Ziel. Heutzutage müssen Unternehmen in der Lage sein, mehrere Datenquellen zu kombinieren, um cloudbasierte mobile Anwendungen zu erstellen, die bei Bedarf schnell geändert werden können. Die starren Strukturen, auf denen die RDBMS-Technologie basiert, können sich nicht an das schnelle Tempo und ständig wechselnde Anforderungen heutiger Unternehmen anpassen. Feste Datenmodelle in Kombination mit ständig wechselnden nachgeschalteten Systemen erzeugen eine sofortige Inkompatibilität zwischen den Anforderungen des Unternehmens und der Architektur des Projekts.

OPEN SOURCE

Zu Zeiten kostspieliger IT-Projekte, die nie realisiert werden, können Open-Source-Technologien kostenbewussten Entscheidungsträgern äußerst attraktiv erscheinen. Anders als die relationale Technologie bietet der Markt für agile NoSQL-Datenbanken (nicht nur SQL) ein flexibleres Data Management und mehr Wirtschaftlichkeit für sehr große Datenvolumen. Der Markt ist groß, genauso wie die Fülle der Anbieter. Hier besteht das Risiko darin, dass diese neuen Open-Source-Technologien nicht entwickelt wurden, um erfolgskritische Anwendungen mit Enterprise-Funktionen (wie höchste Sicherheitszertifizierungen, ACID, hohe Verfügbarkeit oder Disaster Recovery) zu unterstützen, die aber für Abläufe und den Schutz Ihres Unternehmens unverzichtbar sind. Professionelle Schutzfunktionen sind ein Muss. Konsumenten mag es egal sein, wenn

mal ein Tweet oder eine Statusmeldung bei Facebook verloren geht. Doch bei einer falschen Buchung auf dem Bankkonto dürften die Alarmglocken schrillen. Bevor Sie sich also für einen Open-Source-Anbieter im NoSQL-Markt entscheiden, sollten Sie wissen, wie Sie diese notwendigen Enterprise-Funktionen einbinden möchten. Die Verknüpfung mehrerer Open-Source-Technologien führt zu höheren Entwicklungskosten und dürfte wahrscheinlich in einer fragilen, unzuverlässigen Infrastruktur resultieren, die mit Ihren wechselnden Anforderungen langfristig nicht mithalten kann. An diesem Punkt offenbart Open Source ihre Tücken. Was Sie vorher bei der Lizenzierung sparen, ist in Wirklichkeit nicht mehr als eine Kostenverlagerung, weil anschließend neue Software entwickelt und laufend gepflegt werden muss.

ERFOLG FÜR IHR UNTERNEHMEN

Gegenüber starren RDBMS- und riskanten Open-Source-NoSQL-Lösungen kann die richtige Datenbanktechnologie Daten-Silos auflösen, die die Lücke zwischen Analyse und operativem Geschäft im Unternehmen schließen und dort Erkenntnisse liefern, wo sie im Kundenkontakt gebraucht werden. Diese Technologie kann zudem unterschiedlichste Arten von Daten verarbeiten, kommt mit unterschiedlichsten Auslastungen klar und passt sich an die Commodity-Infrastruktur und Cloud an, um die Anforderungen heutiger Unternehmen zu erfüllen – ohne dass Geschäftsabläufe gestört werden. Das fehlende Teil in diesem Puzzle ist die MarkLogic® Enterprise NoSQL-Datenbank, die all dies übernimmt und die Datenintegrität und Zuverlässigkeit sicherstellt. Dadurch können Sie einen flexiblen Ansatz bei Datenprojekten verfolgen und mit weniger Aufwand technische Meilensteine erreichen.

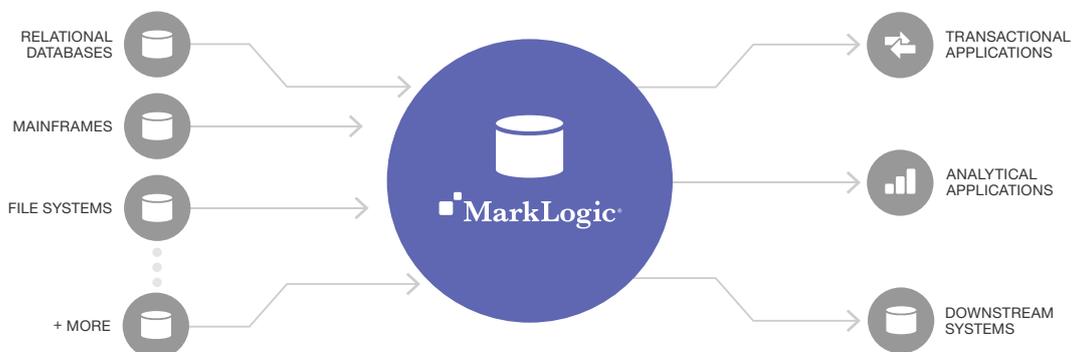


Abbildung 3: Eine professionelle NoSQL-Lösung harmonisiert Daten unternehmensweit.

Als zentrale Plattform für die unternehmensweite Harmonisierung von Daten ist MarkLogic so entwickelt, dass sich die Lösung flexibel – ohne Unterbrechung der Geschäftsabläufe – in Ihre vorhandene Infrastruktur einpasst. Diese agile Enterprise-NoSQL-Datenbank ermöglicht eine schrittweise Implementierung, ohne dass Sie erst jede einzelne Datenquelle identifizieren und vorher ein gemeinsames Datenmodell entwickeln müssen. Da MarkLogic speziell zur Verbesserung und Vereinfachung Ihrer Datenintegrationsmaßnahmen entwickelt wurde, ist diese Lösung *datenorientiert* und nicht funktionspezifisch. Das bedeutet, dass Sie Ihren Datenbestand zentral für viele Zwecke nutzen können und ein Kopieren überflüssig ist.

Entwickelt speziell für Unternehmen, vereint MarkLogic die besten RDBMS- und NoSQL-Funktionen, um die Anforderungen Ihres Unternehmens an höchste Sicherheitszertifizierungen, ACID-Transaktionen, Hochverfügbarkeit, Disaster Recovery, Agilität, Skalierbarkeit und Performance zu erfüllen.

Als beste Datenbankplattform für die Wertschöpfung aus *allen* Daten – unabhängig von der Art und Quelle der Daten oder dem Daten-Silo – löst MarkLogic Ihre Datenherausforderungen des 21. Jahrhunderts mit modernen Funktionen.

VON ERFAHRUNGEN PROFITIEREN

Dank der Technologie und Fachwissen von MarkLogic haben viele Unternehmen weltweit einige der komplexesten Datenintegrationsherausforderungen erfolgreich gemeistert. Im Folgenden wollen wir näher auf die Herausforderungen und Ergebnisse ausgewählter Datenprojekte eingehen, bei denen viel auf dem Spiel stand und die durch die Implementierung dieser neuen Datenbanktechnologie gelöst werden konnten.

IMPLEMENTIERUNG EINES LANDESWEITEN DATENINTEGRATIONSPROJEKTS MIT BEISPIELLOSEM UMFANG

Im Gegensatz zum Scheitern der Technologie beim Krankenversicherungsmarktplatz *Cover Oregon* hatte Healthcare.gov zwar einen äußerst schwierigen Start, konnte aber nach dem ersten Online-Jahr

die Gesundheitsversorgung von 7,1 Millionen US-Amerikanern¹¹ innerhalb von 14 Wochen sicherstellen und brauchte bis zum Go Live nur 10 Wochen. Healthcare.gov bietet uns einzigartige Einblicke darin, wie die richtige Technologie-Entscheidung gleich zu Anfang zu einem gewaltigen Vorteil werden kann, wenn bei einem sehr großen Projekt etwas richtig schief läuft.

DER KUNDE

Centers for Medicare and Medicaid Services (CMS)

DIE AUSGANGSSITUATION

Mit einem beispiellosen Umfang, neuen und sich entwickelnden Richtlinien, die erfüllt werden mussten, eskalierenden Kosten und einer vom Gesetzgeber vorgeschriebenen nahezu unerfüllbaren Deadline stand einiges auf dem Spiel. Henry Chao, CMS Deputy CIO und Director, Office of Information Services, und Technology Lead des anfänglichen Projekt-Rollouts, kann sich noch an den enormen Druck erinnern:

„Stellen Sie sich die folgende IT-Projektherausforderung vor: Erstellen eines Systems, das für Zehntausende gleichzeitig aktive Benutzer skalierbar ist, eine Uptime von 99,9 % hat und mit mindestens 20 weiteren vorhanden Großsystemen integriert werden muss, über die Sie keine Kontrolle haben. Sie müssen dieses System innerhalb eines Umfelds bereitstellen, in dem es Hunderte Interessensgruppen gibt, die alle laufend Input geben, aber keine davon über die letztliche Entscheidungsgewalt verfügt. Aufgrund dieser Bedingungen ändern sich die Anforderungen an das System laufend, während sie es entwickeln. Sie müssen das System zu einem festen Termin fertigstellen und wenn Sie scheitern, bleibt das keinem verborgen.“¹²

ZIELSETZUNG

Die Vorgabe der US-Regierung – die Einführung eines landesweiten Krankenversicherungsmarktplatzes, wie es der Affordable Care Act (CMS) vorschreibt – erforderte das Zusammenbringen von Daten aus vielen Bundesstaaten. Zahllose Bundesbehörden und Versicherungsgesellschaften mussten ein nahtloses Online-Einkaufserlebnis entwickeln, das Millionen US-Amerikanern erweiterten Zugang zu einer Gesundheitsversorgung geben würde.

¹¹ Washington Post: At least 7.1M signed up for 2015 Obamacare plans so far. Dezember 2014. <<https://www.washingtonpost.com/news/wnk/wp/2014/12/30/at-least-7-1m-signed-up-for-2015-obamacare-plans-so-far/>>

¹² InformationWeek: Healthcare.gov: Hard-Earned Lessons for CIOs. März 2016. <http://www.informationweek.com/strategic-cio/healthcaregov-hard-earned-lessons-for-cios/d/d-id/1324456?page_number=1>

“ HealthCare.gov ist nicht nur eine Website. Das wäre zu einfach ausgedrückt. Healthcare.gov ist eine komplexe Anwendung zur Überprüfung und Ermittlung der Verfügbarkeit, die in einen Data Services Hub integriert ist, der als Broker für alle Anfragen fungiert.”

– Henry Chao, *InformationWeek*

DIE HERAUSFORDERUNG

CMS arbeitete für die Entwicklung, Verwaltung und Wartung all seiner Systeme mit großen Generalunternehmern zusammen. CMS beschäftigt zudem viele hochqualifizierte Mitarbeiter und Führungskräfte, die seit Jahren zur Behörde gehören und erfolgreich sehr große Systeme implementiert haben, die jeder US-Amerikaner kennt. Bei jedem System, mit dem CMS beauftragt wurde, stand äußerst viel auf dem Spiel. Ein zentraler Teil dieser Verantwortung ist der Risikoausgleich.

Eine vernünftige GTM-Strategie für den Risikoausgleich ist der Einsatz bewährter Technologien und Prozesse, die in der Vergangenheit funktioniert haben. Die Generalunternehmer, die diese Einstellung der Behörde kannten, gaben Angebote mit traditionellen Technologien ab – darunter auch RDBMS für die Verarbeitung unterschiedlicher Daten von Marktplätzen und dem Data Services Hub (DSH). Andererseits schätzte CMS die Herausforderungen richtig ein und wusste, dass eine relationale Technologie und bekannte

Entwicklungsprozesse nicht funktionieren würden. Erstens würde die relationale Technologie nicht die Fülle an Datenformaten handhaben können, von Bundesstaaten, Versicherungsgesellschaften und anderen Bundesbehörden, die in das System fließen. Zudem war klar, dass dieses Datenbanksystem nicht schnell skaliert werden konnte, wenn am ersten Tag der Registrierung Verbraucher in großer Anzahl auf die Website kämen. Alles in allem bedeuteten diese Schwachpunkte, dass Anwendungen, die auf den vorhandenen relationalen Datenbanken basierten, nicht so funktionieren würden, wie es die CMS-Vorgaben zur Erfüllung der gesetzlichen Anforderungen des Affordable Care Act vorsahen.

CMS und Henry Chao, der stellvertretende CIO von CMS, blieb nichts anderes übrig, als diese Generalvertreter zu beauftragen und ihnen dann eine andere technische Richtung vorzugeben. Im Januar 2012 gab CMS bekannt, dass MarkLogic als Datenbank-Plattform verwendet werden soll.

Eine Studie des Office of Inspector General im US-Gesundheitsministerium unterstützte diese

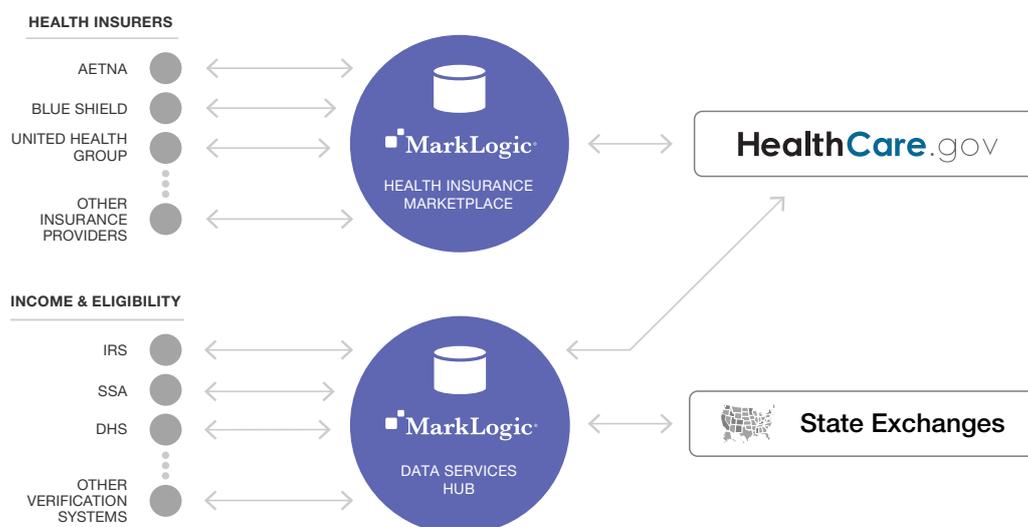


Abbildung 4: Eine Technologieplattform, die für die Registrierung von Millionen Benutzern für neue Krankenversicherungen geeignet ist

Entscheidung. Darin heißt es: „Der Hauptvorteil einer unkonventionellen NoSQL-Plattform ist deren potenziell größere Kapazität, durch die mehr Daten gleichzeitig übertragen werden können und die eine einfache Erweiterung erlaubt, um mehr Daten oder Benutzern gerecht zu werden.“¹³

Als Ersatz für die RDBMS, die die CMS-Anforderungen nicht erfüllen konnte, entschied man sich für MarkLogic. Hauptgrund war, dass MarkLogic die Flexibilität und Leistung bietet, die für die Implementierung eines komplexen Projekts dieser Größe notwendig sind. Doch der Anbieter einer neueren Technologie im Kreis etablierter Regierungslieferanten zu sein, war eine Herausforderung. So erklärte z. B. ein Generalvertreter, der keine Erfahrung mit NoSQL-Datenbanktechnologien hatte und einem Missverständnis aufsaß, MarkLogic zu einem potenziellen Risiko für die termingerechte Implementierung. Sein Argument: „Die umfassenden Fachkenntnisse von Mitarbeitern mit herkömmlichen Datenbanken lassen sich nicht vollkommen durch gleichwertige Fachkenntnisse zur NoSQL-Plattform von MarkLogic ersetzen.“ Im gleichen Zug empfahl er CMS die Ausarbeitung eines Notfallplans.¹⁴

CMS vertraute weiterhin wie geplant auf MarkLogic. Letztlich zeigten sich andere Generalvertreter offener gegenüber der neueren Datenbanktechnologie und konnten sich schnell in MarkLogic einarbeiten, um die Implementierung des Data Services Hub und das Go Live von Healthcare.gov erfolgreich abzuschließen.

Die Erfahrung von CMS wirft einige Denkanstöße hinsichtlich der Risiken auf, die mit der Angst vor Veränderungen einhergehen. In diesem Fall sah einer der CMS-Auftragnehmer in der Arbeit mit einer neueren Technologie ein Risiko. Zum Glück entschied sich Henry Chao für ein unkonventionelles Vorgehen, weil ihm das größere Risiko bewusst war, dass die Anforderungen und Terminvorgaben der Behörde nicht erfüllt werden konnten – und das die bisher eingesetzte Technologie nicht funktionieren würde. Wie sich herausstellte, eine weise Entscheidung:

- Die flexible MarkLogic-Plattform bewältigte die komplexe Integration unterschiedlicher Daten für den Marktplatz und den DSH – trotz wechselnder Datenquellen und Anforderungen.
- Selbst beim Ausfall des Netzwerk-Rechenzentrums funktionierte die MarkLogic-Datenbank reibungslos – ein sicherer Schutz für CMS mit keinerlei Datenverlusten.
- Der DSH funktioniert seit dem Go Live problemlos.
- Trotz des anfänglichen Ausfalls der Marktplatz-Anwendung für Verbraucher war CMS in der Lage, das System schnell zum Laufen zu bringen. Henry Chao erklärt: „Als alles schlecht lief, hätten wir auch mit einem Scale-out auf eine schlecht programmierte Anwendung umstellen können.“ MarkLogic „bot uns mehrere Optionen, die mit anderen Technologien nicht möglich gewesen wären.“¹⁵

ERGEBNISSE

Laut CMS haben sich 2016 über 12,7 Millionen Verbraucher für Krankenversicherungsleistungen registriert, die über Healthcare.gov angeboten werden.¹⁶ Mit MarkLogic hat CMS Folgendes erreicht:

- Erfüllung des vom Kongress vorgegebenen Termins für die Bereitstellung eines landesweiten Krankenversicherungsmarktplatzes, wie es der Affordable Care Act vorschreibt
- Einführung eines hochskalierbaren, operativen Data Hubs mit einer Datenintegrität von 100 % und der Stabilität, gleichzeitig 280.000 Benutzer zu unterstützen
- Sichere Integration unterschiedlich strukturierter und unstrukturierter Daten aus zahlreichen Systemen zu einem positiven Benutzererlebnis mit einer Verarbeitung von 6.500 Transaktionen pro Sekunde
- Besserer Zugang zu Krankenversicherungsleistungen für Millionen US-Amerikaner

¹³ U.S. Department of Health and Human Services: Healthcare.gov Case Study of CMS Management of Federal Marketplace. Februar 2016. <<http://oig.hhs.gov/oei/reports/oei-06-14-00350.asp>>

¹⁴ U.S. Department of Health and Human Services: Healthcare Gov. Case Study of CMS Management of Federal Marketplace. Februar 2016. <<http://Oig.Hhs.Gov/Oei/Reports/Oei-06-14-00350.Asp>>

¹⁵ MarkLogic: Die Technologie hinter der Rettung von HealthCare.gov. Dezember 2015. <http://www.marklogic.com/resources/the-untold-story-of-rescuing-healthcare-gov/resource_download/whitepapers/>

¹⁶ CMS.gov. Health Insurance Marketplace Open Enrollment Snapshot – Week 13. Februar 2016. <<https://www.cms.gov/newsroom/mediareleasedatabase/Fact-sheets/2016-Fact-sheets-items/2016-02-04.html>>

ENTWICKLUNG EINES SKALIERBAREN DATA HUBS FÜR DIE ECHTZEIT-BEREITSTELLUNG INTEGRIERTER INFORMATIONEN

DER KUNDE

Aetna

DIE AUSGANGSSITUATION

Aetna ist ein großer Krankenversicherer mit Leistungen für über 23 Millionen Menschen über eine Million Anbieter. Das bedeutet nicht nur, dass Aetna jährlich Milliarden Ansprüche bearbeitet, sondern auch die Personaldaten seiner 35.000 Mitarbeiter verwalten muss. Durch die Übernahme von Humana in 2015 kam eine weitere wichtige neue Datenquelle mit neuen Anforderungen hinzu.

ZIELSETZUNG

Die IT-Abteilung von Aetna stand zunehmend unter dem Druck, alle Personaldaten (wie Mitarbeiterdaten, Gehaltsabrechnungsdaten und viele andere Systeme) an zentraler Stelle zu integrieren und diese Daten für nachgeschaltete Systeme für Echtzeit-Anwendungen und Batch Analytics bereitzustellen. Das Ziel war die Entwicklung eines skalierbaren Data Hubs für den Personalbereich, um integrierte HR-Informationen in Echtzeit bereitstellen zu können.

DIE HERAUSFORDERUNG

Der Großteil von Aetnas vorhandenen Daten war in herkömmlichen relationalen Datenbankmodellen gespeichert. Die Integration dieser Daten würde normalerweise die Entwicklung komplexer Punkt-zu-Punkt-Lösungen für den Datenaustausch erfordern,

wofür hochkomplexe, vorher zu programmierende ETL-Prozesse notwendig gewesen wären. Dieser Ansatz wäre nicht nur teuer und zeitaufwändig, sondern brächte auch Beschränkungen einer relationalen Datenbank mit sich. Letztlich würde das zu einem fragilen System führen, das bei allen künftigen Änderungen umfassend überarbeitet hätte werden müssen.

ERGEBNISSE

Mit MarkLogic hat Aetna Folgendes erreicht:

- Erfolgreiche Integration von über 24 verschiedenen HR-Datenquellen, die aus rund 140 einzelnen Daten-Feeds von einem System zusammengestellt werden, das wenige Terabytes an Daten verarbeitet und einen Datendurchsatz von 50 GB pro Tag schafft
- Verwendung eines flexiblen Datenmodells, das sich für eine agile Entwicklung eignet, um die Herausforderungen komplexer ETL-Prozesse zu bewältigen.
- Einhaltung der Kosten- und Terminvorgaben für die Implementierung, ohne bei wichtigen Unternehmensanforderungen wie ACID-Transaktionen, höchsten Sicherheitszertifizierungen, Prüfbarkeit und Transparenz Kompromisse einzugehen
- Einfache Integration neuer Datenquellen und bei Bedarf Anpassung an Veränderungen mit der neuen Datenbankplattform
- Höhere Kundenzufriedenheit zu geringeren Kosten

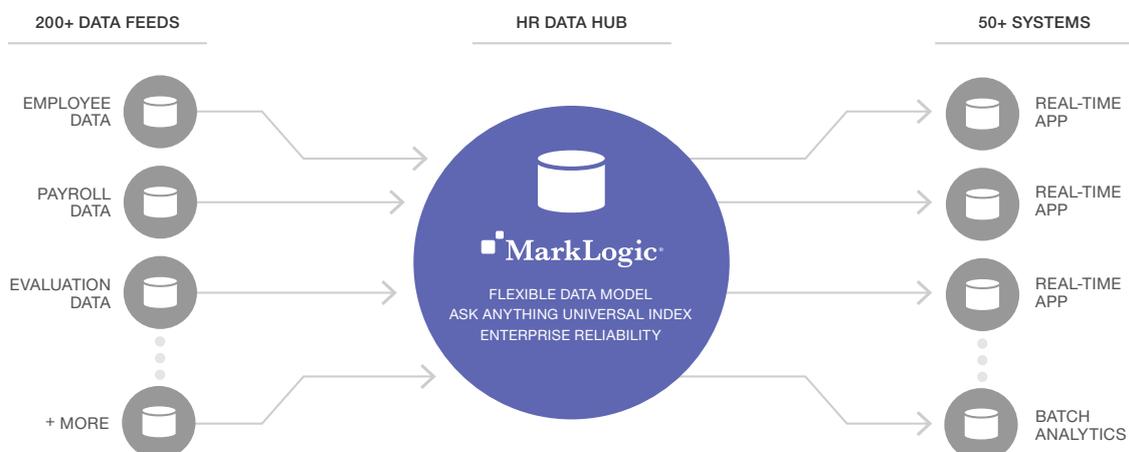


Abbildung 5: Integrierte Personaldaten mit bedarfsgerechter Bereitstellung in Echtzeit

SCHNELLE IMPLEMENTIERUNG EINES ZENTRALEN, SICHEREN GLOBALEN TRADE STORES IN SECHS MONATEN

DER KUNDE

Deutsche Bank

DIE AUSGANGSSITUATION

Mit über 100.000 Mitarbeitern in über 70 Ländern ist die Deutsche Bank ein internationaler Banking- und Finanzdienstleister mit Hauptsitz in Deutschland. Die führende Bank bietet Finanzprodukte und -dienstleistungen für Unternehmen und institutionelle Kunden sowie für Privat- und Geschäftskunden.

ZIELSETZUNG

Um eine „Single Source of Truth“ übergreifend über verschiedene Handelsarten zu erhalten, musste die Deutsche Bank ihre Daten integrieren und eine operative Handelsdatenbank (einen globalen Trade Store) für die integrierte Post-Trade-Verarbeitung entwickeln. Die nachgeschalteten Geschäftsbereiche der Bank brauchten Zugriff auf die zentralen Daten, um zahlreiche Berichts-anforderungen und auch gesetzliche Compliance-Vorgaben zu erfüllen.

DIE HERAUSFORDERUNG

Die Deutsche Bank arbeitet mit vielen handelsbezogenen Daten-Inputs. Diese werden mit einem Stovepiped-Modell mehreren Geschäftsbereichen bereitgestellt, die diese Daten auf unterschiedlichste Weise nutzen. Die Herausforderung für die Bank bestand in der Harmonisierung der Daten, dem Erstellen einer

verlässlichen, aussagekräftigen Informationsquelle (Single Version of the Truth) und der Sicherstellung der Datenintegrität. Die Deutsche Bank versuchte dies zuerst mit der RDBMS-Technologie zu erreichen. Dieses Projekt scheiterte jedoch, weil man sich nicht auf ein gemeinsames Schema einigen konnte und sich die Anforderungen öfter änderten. Folglich wurde eine flexiblere Datenintegrationstechnologie gefordert.

ERGEBNISSE

Mit MarkLogic hat die Deutsche Bank Folgendes erreicht:

- Erfolgreiche Integration von mehr als 30 Handelssystemen in einen flexiblen, zentralen globalen Trade Store – und damit Schaffung einer „Single Source of Truth“
- Kostensenkung mit einem agilen Trade Store, dessen Bereitstellungskapazität leicht bei Bedarf erweitert werden kann, während gleichzeitig umfassende ETL-Maßnahmen entfallen, die bei der vorherigen relationalen Technologie unabdingbar gewesen wären.
- Operatives Bankgeschäft mit einer zentralen, professionellen Datenbankplattform, die sichere und einheitliche Transaktionen gewährleistet
- Zeitersparnis durch schnelle Integration sämtlicher strukturierter und unstrukturierter Daten mit anschließender Inbetriebnahme der ersten Systemversion innerhalb von nur sechs Monaten

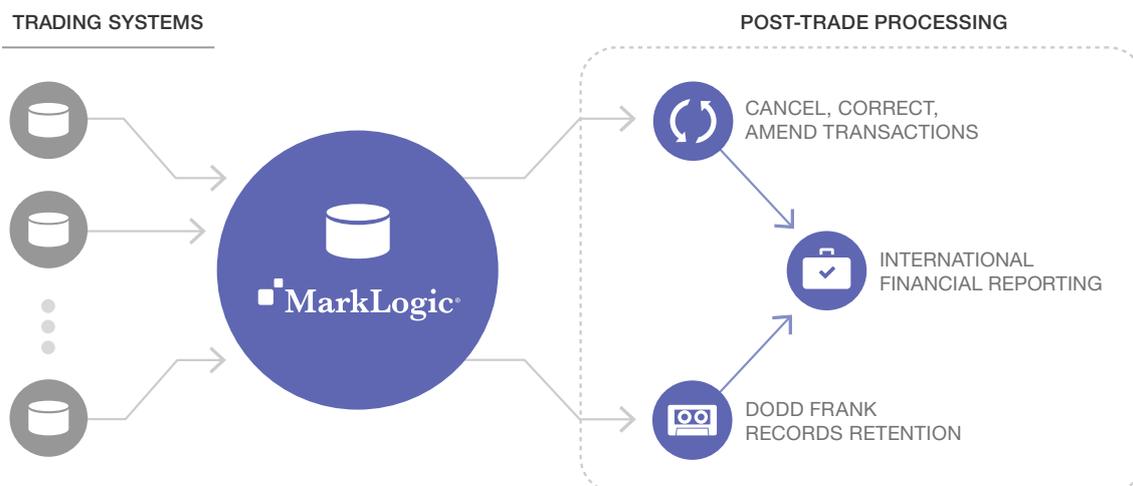


Abbildung 6: Integrierte Post-Trade-Verarbeitung

NEUER MASSSTAB FÜR DIE DYNAMISCHE, BEDARFSGERECHTE BEREITSTELLUNG VON INHALTEN

DER KUNDE

British Broadcasting Company (BBC)

DIE AUSGANGSSITUATION

Mit zehn nationalen und 40 lokalen Radiosendern, digitalen Apps und einer umfassenden Website ist die BBC eine weltweit führende öffentliche Rundfunkanstalt.

ZIELSETZUNG

Die BBC wollte für ihre Online-Berichterstattung über die Olympischen Sommerspiele 2012 eine Website mit Athleteninformationen integrieren, die jedem Besucher einen interaktiven Content mit einer einzigartigen Personalisierung in Echtzeit bieten sollte. Obwohl die BBC bislang für ihre statische Publishing-Website mit einem RDBMS gearbeitet hatte, war klar, dass diese starre Datenbankplattform keine effektive Handhabung und Bereitstellung eines kontinuierlichen Stroms unterschiedlichster ein- und ausgehender Daten bewältigen konnte. Dies war aber notwendig, damit Zuschauer der Olympiade das erweiterte und interaktive Benutzererlebnis erfahren können.

DIE HERAUSFORDERUNG

Um eine interaktive Website mit allen denkbaren Funktionen bereitzustellen, musste die BBC eine agile Datenbankplattform entwickeln, die skalierbar genug war, um zehntausende Transaktionen pro Sekunde zu unterstützen. Unzählige Hindernisse erschwerten dieses Vorhaben, darunter große Mengen komplexer Daten, die in unterschiedlichen Systemen gespeichert waren, und eine veraltete IT-Infrastruktur, mit der kein dynamischer Content in mehreren Formaten (z. B. Videos, redaktionelle Beiträge oder

Statistiken) dem Publikum bereitgestellt werden konnte. Außerdem fehlte eine Datenbanktechnologie, die agil und leistungsstark genug war, um Millionen Online-Besuchern jederzeit und überall Zugriff auf Informationen und Berichte über Olympiateilnehmer zu geben.

ERGEBNISSE

Mit MarkLogic hat die BBC Folgendes erreicht:

- Schnellere Inbetriebnahme durch kürzere Entwicklungszeiten und weniger Infrastrukturbeschränkungen als bei der bisherigen relationalen Technologie
- Einsatz von ACID-Transaktionen für eine stabile, unterbrechungsfreie digitale Abdeckung ohne Datenverluste, Verschlechterung der Performance oder Ausfälle
- Professionelle Abwicklung von über 25.000 Transaktionen pro Sekunde und 45 Milliarden Anfragen, um bis zu 10.000 dynamische Seiten für die Website-Besucher bereitzustellen
- Schaffung eines hochskalierbaren und personalisierten digitalen Benutzererlebnisses, mit dem eine noch nie da gewesene Menge an interaktivem Content (Bereitstellung von 2,8 Petabytes an Spitzentagen) für Millionen von Benutzern auf jedes beliebige Gerät geliefert werden konnte.

Nach diesem Erfolg mit MarkLogic bei der Implementierung der dynamischen Bereitstellungsplattform, setzt die BBC mittlerweile noch stärker auf MarkLogic, um die Performance ihres beliebten Video-on-Demand-Services iPlayer zu optimieren.¹⁷



¹⁷ Computerworld UK: BBC iPlayer sees performance uplift after relaunching on NoSQL database. Juni 2104
<<http://www.computerworlduk.com/news/data/bbc-iplayer-sees-performance-uplift-after-relaunching-on-nosql-database-3524893/>>

UNTERNEHMEN VERÄNDERN SICH IN RASANTEM TEMPO – DIE IT SOLLTE MITHALTEN KÖNNEN

Bei der Implementierung von komplexen Projekten *kann* vieles schiefgehen und *Fehler* passieren – aber Ihre Technologie muss nicht dazu gehören. Wird die Datenbank gut ausgewählt, lassen sich damit Risiken gezielt verringern und Sie leisten einen wertvollen Beitrag, damit Ihr Unternehmen das gesteckte Ziel erreicht. Entscheidend ist, dass Ihr Projekt auf Erfolg ausgerichtet ist, indem Sie die *passenden* Experten einsetzen, die die *richtige* Technologie als Entwicklungsgrundlage nutzen.

Angesichts rasanter Marktveränderungen hängt der Erfolg von Ihrer Fähigkeit ab, schneller als der Wettbewerb einen Mehrwert zu liefern. Dafür sollten Sie für erfolgskritische Projekte mit einer soliden, zugleich aber auch agilen Plattform arbeiten, die heute für Sie funktioniert und sich gemeinsam mit dem Unternehmen für künftige Anforderungen entwickelt. Vertrauen Sie bei der Entwicklung eines Systems für moderne Daten auf alte Technologien. Stattdessen empfiehlt sich ein Blick in die Zukunft und die Wahl einer Technologie, die sich an ein sich schnell und ständig veränderndes Geschäftsumfeld anpasst. Mit solch einer Technologie wird Ihr Unternehmen Datenprojekte erfolgreich meistern und damit eine leistungsstarke Grundlage für das Erreichen von Geschäftszielen schaffen.

MarkLogic ist eine operationale und transaktionale Enterprise NoSQL-Datenbank, mit der Unternehmen in jeder Branche Daten aus Silos schneller und kostengünstiger als je zuvor integrieren können. Wir tragen dazu bei, die komplexesten Datenherausforderungen zu lösen, indem wir marktführenden Unternehmen einen umfassenden Überblick über ihre Daten verschaffen.

Verabschieden Sie sich von veralteten Herangehensweisen und Technologien. Mit dem richtigen Team und der richtigen Technologie, die

eigens für heutige moderne Datenanforderungen geschaffen wurden, werden nicht nur Projekte, bei denen viel auf dem Spiel steht, zum Erfolg. Sie nutzen damit sämtliche Daten im Unternehmen, um Ihren Wettbewerbsvorteil auszubauen.

WEITERE INFORMATIONEN

Relationale Datenbanken wurden nicht für die Herausforderungen entwickelt, die mit den enormen Mengen unterschiedlichster, komplexer Daten, die wir tagtäglich verwenden und produzieren, einhergehen. Aber Sie können all Ihre Daten nutzen, um Ihre Unternehmensziele schneller und kostengünstiger als gedacht zu erreichen. Erfahren Sie, wie die neue Generation von Datenbank Ihnen den Weg zu erfolgreichen Datenprojekten ebnet:

Relational war gestern

Bei einem RDBMS können selbst geringe Veränderungen kostspielig werden. Deshalb verabschieden sich führende Unternehmen zunehmend von der relationalen Technologie und setzen auf eine flexible Datenbanktechnologie, die weniger Datenmodelle erfordert – Lesen Sie das Whitepaper:

<http://www.marklogic.com/resources/beyond-relational/>

Enterprise NoSQL für Dummies

Erfahren Sie, wie Datenbanken mit semantischer Technologie Big-Data-Herausforderungen lösen.

<http://info.marklogic.com/nosql-for-dummies.html>

Bewältigung komplexer Herausforderungen im Datenbereich

Als Technologie einer neuen Generation, die speziell für moderne Arten von Daten entwickelt wurde, ist MarkLogic die einzige NoSQL-Datenbank der Enterprise-Klasse. Sie basiert auf einem flexiblen Datenmodell und kann all Ihre Daten speichern, verwalten und durchsuchen, ohne Abstriche bei der Ausfallsicherheit oder Konsistenz zu machen, die Sie vom RDBMS kennen. Weitere Informationen erhalten Sie auf unserer Website:

<http://de.marklogic.com/what-is-marklogic/>

© 2016 MARKLOGIC CORPORATION. ALLE RECHTE VORBEHALTEN. Diese Technologie ist durch die US-Patente 7,127,469 B2, 7,171,404 B2, 7,756,858 B2 und 7,962,474 B2 geschützt. MarkLogic ist eine Marke oder eingetragene Marke der MarkLogic Corporation in den USA und/oder anderen Ländern. Alle anderen hierin erwähnten Marken sind Eigentum der jeweiligen Besitzer.

MARKLOGIC GMBH

Theatinerstr. 11, 8. Etage, München 80333 | +49 89 7 10 42 2151
Skyper Villa, Taunusanlage 1, Frankfurt 60329 | +49 69 5 05 06 0588
de.marklogic.com | germany@marklogic.com