



Neu in

# MarkLogic 9

WHITEPAPER • MAI 2017

Die beste Datenbank der Welt für die Datenintegration ist jetzt noch besser: Mit MarkLogic 9 präsentieren wir unser bislang anspruchsvollstes Release. MarkLogic 9 umfasst wichtige neue Funktionen zur Verbesserung von Datenintegration, Sicherheit und Management – entwickelt, um die Datenintegration aus Silos einfacher und schneller zu gestalten.



# Contents

---

<b>Einführung</b> .....	<b>2</b>
<b>Schnellere, einfachere Datenintegration</b> .....	<b>3</b>
Entity Services: Modellbasierte Datenintegration für Entitäten, mit denen Ihr Unternehmen arbeitet.....	3
MarkLogic Optic API: Verknüpfungen und Aggregationen über Dokumente hinweg.....	4
Template Driven Extraction: Spezielle Ansichten für Ihre Dokumente.....	5
ODBC-Treiber: Bessere Integration mit BI und Analytics Tools.....	5
Data Movement SDK: Massendatenbewegung mit Java.....	6
<b>Marktführende Datensicherheit</b> .....	<b>6</b>
Verschlüsselung: Erweiterte Verschlüsselung im Ruhezustand.....	6
Sicherheit auf Elementebene: Granulare Sicherheit auf Dokumenten-Unterebene.....	7
Schwärzen: Exportkontrolle verhindert das Teilen sensibler Daten.....	8
<b>Verbesserte Handhabbarkeit</b> .....	<b>8</b>
Ops Director: Verständlicher Überblick über das MarkLogic-System.....	9
<b>Noch mehr Verbesserungen</b> .....	<b>9</b>
<b>Erste Schritte mit MarkLogic 9</b> .....	<b>11</b>

## EINFÜHRUNG

MarkLogic® 9 ist das bislang anspruchsvollste Release der weltbesten Datenbank zur Integration von Daten aus Silos – im wahrsten Sinne des Wortes: MarkLogic 9 umfasst mehr neue Funktionen und Updates als jede andere Version. Die neue, verbesserte Funktionalität konzentriert sich auf drei Kernbereiche: Datenintegration, Sicherheit und Management. MarkLogic war zwar bereits die beste und sicherste NoSQL-Datenbank für die Datenintegration, doch mit den neuen robusten Admin-Tools legt MarkLogic 9 auch in diesen Punkten die Messlatte höher. Diese Verbesserungen positionieren MarkLogic als führende Datenbank einer neuen Generation, mit der Sie alle Ihre Daten aus Silos integrieren und moderne, transaktionale Anwendungen bedarfsgerecht ausführen können.

### **SCHNELLERE, EINFACHERE DATENINTEGRATION DANK DER MULTI-MODELL-FUNKTIONEN VON MARKLOGIC**

Herkömmliche relationale Ansätze gehen immer mit einem Kompromiss zwischen Flexibilität und Governance einher und begrenzen Sie bei der Integration von Daten auf ein starres, tabellarisches Datenmodell. MarkLogic 9 bringt zahlreiche Neuerungen, die Sie von diesem „Zwangskorsett“ befreien. Die Datenintegration funktioniert dadurch einfacher und iterativer, damit Sie schneller noch umfassender von Ihren Daten profitieren. Sie können alle Ihre Daten in Silos – ob relational oder nicht relational – in einer einzigen, einheitlichen Plattform zusammenführen, die eigens für Vielseitigkeit entwickelt wurde. Und mit MarkLogic 9 entscheiden Sie allein, wie auf Ihre Daten zugegriffen werden darf. Mit den neuen SQL-Funktionen können Sie Daten als Dokumente, Grafiken oder relationale Daten anzeigen. Modelle lassen sich sogar kombinieren. Dank dieser beispiellosen Flexibilität können Sie auf kostspielige, komplizierte ETL-Prozesse verzichten und profitieren von einer besseren Verwaltung der Entitäten und Beziehungen, mit denen Ihr Unternehmen arbeitet.

### **BESSERE DATA GOVERNANCE UND INTELLIGENTE SICHERHEITSFUNKTIONEN**

Was wäre der Sinn einer Integration von Daten aus Silos, wenn diese Daten nicht ordnungsgemäß gehandhabt und gesichert werden können? MarkLogic 9 wurde mit neuen Funktionen entwickelt, die Ihre integrierten Daten während des gesamten

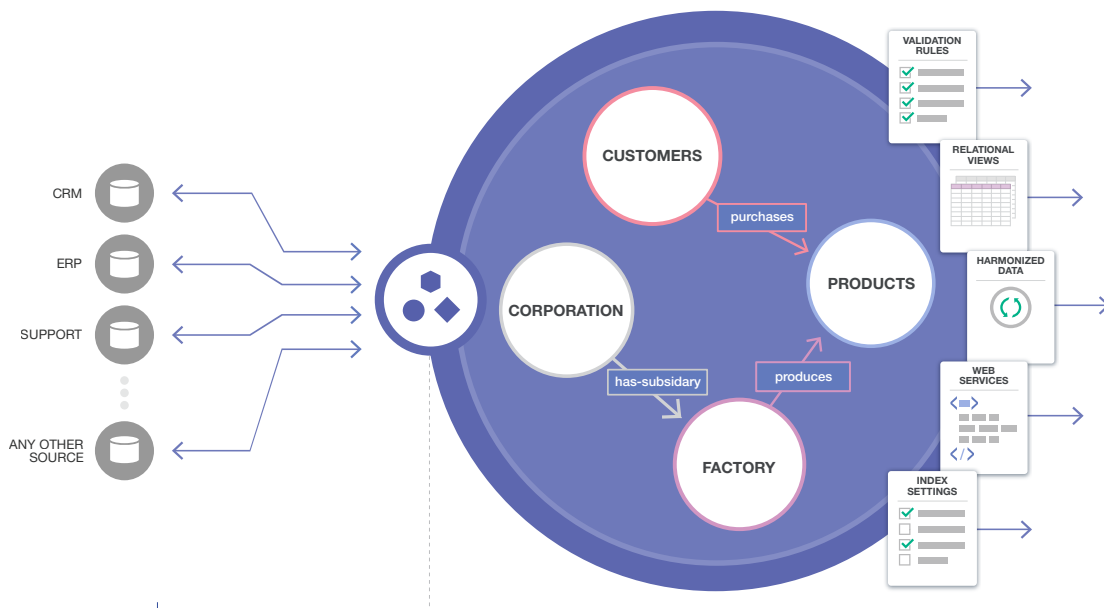
Lebenszyklus verwalten und vor zunehmenden Cyberangriffen schützen. Seit über einem Jahrzehnt dient MarkLogic als Fundament für erfolgskritische Umgebungen. Die neuen Sicherheits- und Datenschutzfunktionen festigen die Position von MarkLogic als eine der vertrauenswürdigsten NoSQL-Datenbanken für die Data Governance. MarkLogic 9 bietet jedoch weitaus mehr und ist auf dem besten Weg, zur sichersten Datenbank auf dem Markt zu werden.

### **BESSERES MANAGEMENT FÜR MEHR PRODUKTIVITÄT**

Angesichts der schnell steigenden Fülle von Anwendungen und Diensten weltweit brauchen Datenbankadministratoren (DBA) bessere Tools, um dieses rasante Wachstum zu bewältigen. MarkLogic 9 bietet neue Funktionen, die eine einheitlichere Verwaltung weltweit verteilter MarkLogic-Datenbank-Cluster während ihres gesamten Lebenszyklus erlauben. Mit einem umfassenderen Überblick über das System und einem optimierten Admin-Tools profitieren Datenbankadministratoren (DBA) von einer stärkeren Automatisierung. So bleibt mehr Zeit für Aufgaben, die zur Wertschöpfung für das gesamte Unternehmen beitragen.

### **ZUSÄTZLICHE ERWEITERUNG VON MARKLOGIC-FUNKTIONEN**

Neben Verbesserungen in den thematischen Bereichen Datenintegration, Management und Sicherheit bietet MarkLogic 9 viele weitere zusätzliche Funktionen und Erweiterungen. Compliance-Archive vereinfachen das Speichern, Verwalten und Sichern historischer Daten und ereignisbasierter Richtlinien. Tiered Storage – eine Funktion zum Verschieben von Daten auf günstige Speichermedien – ist jetzt noch schneller und flexibler. Geodaten liefern zusätzliche Aussagekraft und Präzision. Suchabfragen sind dank der erweiterten Sprachunterstützung und anderer Verbesserungen jetzt noch stabiler. Die Abfrage-Konsole funktioniert jetzt mit JavaScript-Profiling und automatischem Vervollständigen. Und wir unterstützen jetzt auch die Ausführung von MarkLogic in der Microsoft Azure Cloud. Ganz zu schweigen von Hunderten weiterer Verbesserungen, die die Messlatte für Enterprise NoSQL um einiges höher legt.



**ENTITY SERVICES:** Provides a model to design applications around real-world concepts, or entities, like Customers and Products.

## SCHNELLERE, EINFACHERE DATENINTEGRATION

Zwei der wichtigsten Alleinstellungsmerkmale von MarkLogic sind das flexible Datenmodell und die ausgefeilten Indexfunktionen. Gemeinsam sorgen sie dafür, dass die Integration von Daten aus Silos schneller erfolgt und weniger riskant ist. MarkLogic 9 bietet mehrere neue Funktionen, die genau diese Besonderheiten nutzen, um die Datenintegration noch schneller, einfacher und leistungstärker zu gestalten.

Zu diesen neuen Funktionen gehören die **Entity Services**, mit denen sich geschäftliche Entitäten (*Kunden und Aufträge, Handelsvorgänge und Gegenparteien, Anbieter und Ergebnisse* usw.) mit gewünschter Detailtiefe katalogisieren und verwalten lassen. Die neue **Optik-API** bietet eine einzigartige Flexibilität: Entwickler erhalten damit eine nahtlose Schnittstelle für Abfragen aller Daten mit JavaScript, XQuery, Java oder Node.js. Mit der Optik-API werden auch effiziente Verknüpfungen und Aggregationen über Dokumente hinweg möglich – damit ist MarkLogic die einzige Datenbank auf der Welt, die so etwas kann. Die Technologie, die dies ermöglicht, ist die **Template Driven Extraction (TDE)**. Hiermit können Sie eine relationale Linse über Ihre Dokumentdaten legen, um Teile Ihrer Daten mit Standard-SQL oder über die neue Optik-API

abzufragen – ohne die zugrunde liegenden Daten zu verändern. Sie können sogar Triples-Linsen für Ihre Dokumentdaten anwenden und Abfragen mit SPARQL oder der neuen Optik-API definieren.

Die SQL Engine in MarkLogic 9 wurde von Grund auf überarbeitet. SQL läuft damit noch schneller. **ODBC-Treiber** Auch der Zugriff funktioniert schneller, wodurch MarkLogic noch besser in Ihre Tools für Business Intelligence und Analytics integriert ist. Aber was wäre der Datenzugriff Wert, wenn Daten nicht leicht in MarkLogic geladen oder aus MarkLogic gezogen werden könnten? Mit der **Data Movement SDK** – eine neuartige Java-Bibliothek für Anwendungen – können Sie große Datenmengen innerhalb eines MarkLogic-Clusters verschieben.

## ENTITY SERVICES: MODELLBASIERTE DATENINTEGRATION FÜR ENTITÄTEN, MIT DENEN IHR UNTERNEHMEN ARBEITET

Mit seinem flexiblen Datenmodell und der Fähigkeit, mehrere Schemas gleichzeitig zu speichern, ist MarkLogic die ideale Datenbank für die Datenintegration. Aber Unternehmen brauchen nicht nur die Möglichkeit, gleichzeitig mehrere Schemas zu speichern. Sie müssen auch feste, vorhersehbare Aspekte der Daten abfragen können, die reale Dinge oder Entitäten repräsentieren. Beispiele für solche Entitäten sind *Kunden* und *Aufträge*,

Handelsvorgänge und Gegenparteien oder Anbieter und Ergebnisse.

Das Problem mit herkömmlichen relationalen Datenbanken ist ihr statisches Datenbankmodell und dass sie nur ein Schema speichern können. Dies gestaltet das Management von Entitäten extrem schwierig. Der Kontext und die Bedeutung der Daten "sich überall – in Datenbankabfragen, Anwendungscode, veralteten Anwendungsspezifikationen und Diagrammen über Entitäten-Beziehungen – *außer* in der Datenbank. Das bedeutet, dass es quasi unmöglich ist, Daten sinnvoll zu interpretieren:

*Welche Daten stellen einen Kunden zutreffend dar? Was sind die definierenden Eigenschaften? In welcher Beziehung stehen diese zu anderen Entitäten? Welche Systeme können Kunden generieren? Wie werden Kunden in Anwendungen dargestellt? Welche Kunden erfüllen nicht die Geschäftsregeln?*

Entity Services liefert eine bessere Methode, um Entitäten sowie die sich ständig verändernden Daten zu verwalten. Sie können sich dies als ein Verzeichnis vorstellen, das als Referenz für Ihre Entitäten und Beziehungen dient und das die Handhabung und Programmierung von Daten einfacher gestaltet.. DDateneigentümer können mit Entity Services ein Modell entwickeln, das Governance-RichtlinienS undDatenregeln in der Datenbank erfasst, zu der sie gehören. DEntwickler können Entity Services in modellbasierte Workflows einbinden, um automatisch Datentransformationen, Validierungsregeln, Indexkonfigurationen und SQL-Ansichten zu erstellen, die allesamt Fehler reduzieren und dazu beitragen,E Änderungen gerecht zu werden.

## MARKLOGIC OPTIK-API: VERKNÜPFUNGEN UND AGGREGATIONEN ÜBER DOKUMENTE HINWEG

Dokumentendatenbanken bieten eine natürliche, denormalisierte Repräsentation von Daten, bei denen Dokumente Entitäten in ihrer Gesamtheit ohne Verknüpfungen oder Komplexität darstellen. Der Vorteil dieses Datenmodells liegt darin, dass es den Kontext beibehält sowie Entwicklungs- und Suchfunktionen unterstützt. Dies ist möglich, weil sich alle Informationen über eine Entität an einer einzigen Stelle befinden und sich effizient abfragen

und aktualisieren lassen. Aber viele Fragen basieren auch auf den Beziehungen zwischen Entitäten. In solchen Fällen brauchen Sie Verknüpfungen und Aggregationen *über* Dokumente hinweg.

Die innovative Optik-API vereint die relationale Welt und die Welt der NoSQL-Dokumente mit dokumentübergreifenden Verknüpfungen und Aggregationen. MarkLogic ist die einzige Datenbank der Welt, die dazu in der Lage ist. Eine der Funktionen, die hierfür verwendet wird, ist die Template Driven Extraction (TDE). Damit werden relationale „Linsen“ über Dokumente gelegt. Diese Ansichten basieren auf Vorlagen, die die Teile eines Dokuments festlegen, die in einer Zeile angezeigt werden sollen. Eine Vorlage verändert Ihre Dokumente in keiner Weise. Sie verändert nur die Art und Weise, wie die Indizierung anhand der Werte erfolgt, die das Dokument enthält.

Sie können dann auf Ihre Daten zugreifen und die Daten durch Ihre Linsen betrachten. Beides funktioniert über SQL oder die API. Mit der Optik-API erhalten Sie eine Volltextsuche für Dokumente, die Sie als Filter, für Zeilenoperationen, zum Aggregieren oder Abfragen von Daten oder auch zum Zusammenstellen von Dokumenten bei der Ausgabe

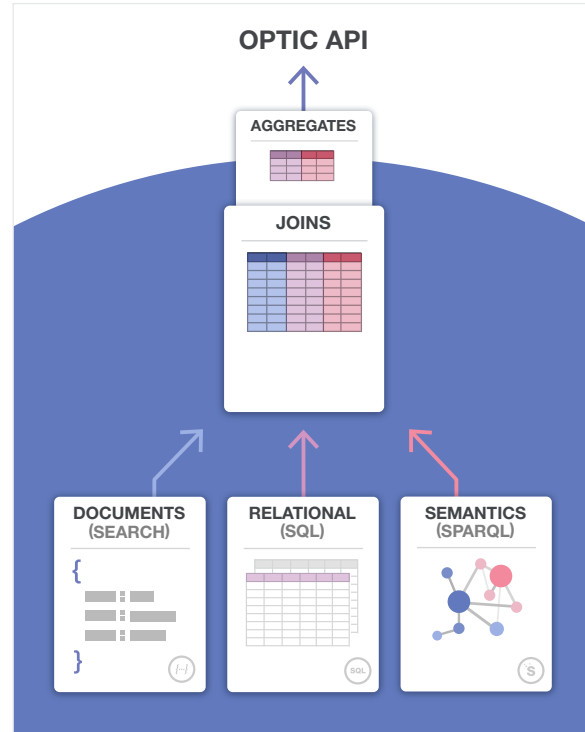


Abbildung 2: Die neue Optik-API bietet eine Schnittstelle für native Abfragesprachen, um Verknüpfungen und Aggregationen über Dokumente hinweg auszuführen.

verwenden können. Sie können bei der Optik-API sogar mehrere Linsen in einer einzigen Abfrage (einschließlich Triples) kombinieren.

Entwickler profitieren von der Einfachheit einer einheitlichen Abfrage-Schnittstelle und einer effizienten, datenbankinternen Verarbeitung. Die Optik-API ist für alle von MarkLogic unterstützten Schnittstellen verfügbar: Java, JavaScript und REST für den Client sowie JavaScript und XQuery für den Server. Jede Implementierung übernimmt sprachspezifische Muster, sodass Entwickler, die mit relationalen Datenbanken Erfahrung besitzen, das Konzept als vertraut empfinden und dank der Syntax bisherige Programmierkenntnisse leicht übertragen können. Beispielsweise muss nicht erst ein String programmiert und für die eigens für diesen Zweck vorgesehene Abfrage-Sprache (SQL oder SPARQL) angepasst werden. Stattdessen können Sie Ihre Abfrage mit der Programmiersprache durchführen, die Sie bevorzugen.

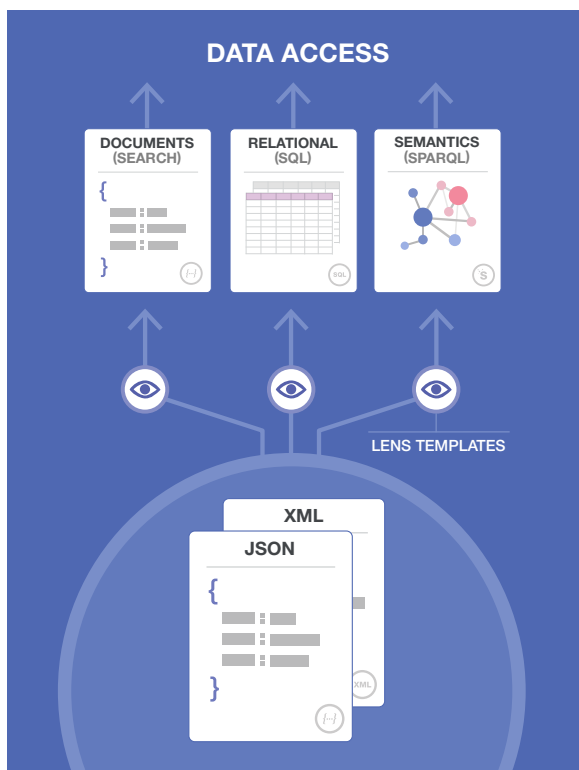


Abbildung 3: Mit Template Driven Extraction (TDE) können Sie sich Ihre Dokumente durch relationale und semantische Linsen ansehen.

## TEMPLATE DRIVEN EXTRACTION: SPEZIELLE ANSICHTEN FÜR IHRE DOKUMENTE

MarkLogic unterstützt das Dokumentmodell. Dahinter verbirgt sich ein flexibler, schema-agnostischer Ansatz, der Rich Data (wie JSON oder XML) nativ

handhabt. Aber vielleicht möchten Sie einige Teile Ihrer Daten aus relationaler Sicht ansehen und diese Daten wie gewohnt mit SQL abfragen.

Mit der Template Driven Extraction (TDE) können Sie Teile Ihrer Dokumentdaten mit einer relationalen Ansicht überlagern, ohne dass die Dokumente geändert werden. Wir nennen das die *relationale Linse*. Mit einer Vorlage können Sie relationale Linsen über ein Dokument legen und genau angeben, welche Teile eines Dokuments als Spalten in einer Ansicht dargestellt werden sollen. Wenn Sie diese Vorlage in die Datenbank einfügen, wird eine SQL-Ansicht erstellt und die Indexe werden entsprechend aufgefüllt. Sie können dann die Abfrage in SQL (beim Server mit SQL oder ODBC) oder über die Optik-API vornehmen. Dies ist besonders praktisch, wenn Sie Berichte und Visualisierungen mit Tools erstellen wollen, die sich SQL kommunizieren, oder wenn Sie Entitäten verknüpfen und dokumentübergreifende Aggregationen vornehmen wollen.

Sie können mit TDE auch eine semantische Linse definieren. Dafür geben Sie einfach an, welche Daten aus einem Dokumentenbestand RDF-Triples bilden, die dann im Triple-Index indiziert werden. Auf diese Weise werden die Daten mit SPARQL oder der Optik-API abfragbar.<sup>1</sup>

Mit TDE werden Vorlagen vollkommen unabhängig von den Dokumenten – sie verändern Ihre Dokumente in keiner Weise. Sie verändert nur die Art und Weise, wie die Indexe anhand der Werte aufgefüllt werden, die die Dokumente enthalten.

## ODBC-TREIBER: BESSERE INTEGRATION MIT BI UND ANALYTICS TOOLS

Dank der besseren SQL Engine von MarkLogic 9 laufen Business Intelligence und Analytics Tools stabiler und problemloser. Die neue SQL Engine lässt sich einfacher einrichten, für mehr Spalten (ohne Memory-Mapped-Index) skalieren und führt Abfragen schneller aus. Die neue SQL Engine trägt zur besseren Nutzung von BI-Investitionen bei, die viele MarkLogic-Nutzer wegen der Standardisierung für ODBC-Verbindungen getätigt haben. MarkLogic unterstützt ODBC-Verbindungen seit Version 6. Mit der überarbeiteten SQL Engine profitiert der ODBC-Treiber von Funktions- und Leistungsverbesserungen innerhalb der MarkLogic-Architektur.

<sup>1</sup> Laden Sie sich das [Datenblatt zum flexiblen Datenmodell und das Semantik-Datenblatt](#) herunter, um sich genauer über die Multi-Modell-Funktionen von MarkLogic zu informieren.

## DATA MOVEMENT SDK: MASSENDATENBEWEGUNG MIT JAVA

Mit MarkLogic lassen sich einfach Daten in eine Datenbank verschieben, aus einer Datenbank verlagern oder innerhalb einer Datenbank bewegen. Sie können z. B. das Laden großer Datenmengen aus Altsystemen, das Erstellen umfassender Berichte oder die Umwandlung von Millionen Dokumenten in einem Arbeitsschritt einfach in die Nachtstunden verlegen.

MarkLogic 9 bietet zudem ein neues SDK-Kit für die Softwareentwicklung als Teil der vorhandenen Java API. Dieses Data Movement SDK ermöglicht Entwicklern die einfache Einbeziehung von Massendatenbewegungen in Java-Anwendungen. Das Data Movement SDK übernimmt die gesamte komplizierte Abstimmung, die ein paralleler Datenbankzugriff und eine ständige Verfügbarkeit erfordern – selbst wenn sich der MarkLogic-Cluster ändert (z. B. beim Hinzufügen von Knoten).

Das Data Movement SDK ergänzt die bisherige Java Client API um eine asynchrone Schnittstelle zum Lesen, Schreiben und Umwandeln von Daten in einem MarkLogic-Cluster. Damit sind Integrationen in vorhandene ETL-Abläufe möglich, wenn z. B. Datenströme aus einer Nachrichtenwarteschlange geschrieben oder relationale Daten über JDBC oder ORM (Object-Relational Mapping) übertragen werden sollen.

Zu den wichtigsten Einsatzbereichen für das Data Movement SDK gehören:

- Erfassen von großen Datenmengen oder -strömen aus einer beliebigen Java I/O-Quelle, die Ihren gesamten Cluster zur Leistungserweiterung nutzen
- Massenverarbeitung von Dokumenten direkt in der Datenbank mit JavaScript oder XQuery-Code, der per Remote-Zugriff über Java gestartet und innerhalb der Datenbank – nahe den Daten – ausgeführt wird.
- Exportieren von Dokumenten basierend auf einer Abfrage, optional mit Aggregation zu einem einzigen Artefakt wie z. B. CSV-Zeilen

## MARKTFÜHRENDE DATENSICHERHEIT

Datensicherheit und Datenschutz besitzen in jedem Unternehmen Priorität. MarkLogic 9 bietet deshalb leistungsstarke neue Funktionen für beide Bereiche. Daten, Konfigurationsinformationen und Protokolle lassen sich per **Verschlüsselung** im Ruhezustand vollständig schützen. Zudem gibt es eine noch granularere **Sicherheit auf Elementebene**. Diese bietet Zugriffskontrolle auf Ebene der XML-*Elemente* oder JSON-*Eigenschaften* innerhalb von Dokumenten und zusätzlich zur vorhandenen Sicherheit auf Dokumentenebene. Zudem lassen sich Adressen aus Datenschutzgründen **schwärzen**, wobei beim Importieren, Exportieren oder Kopieren von Daten Informationen entfernt oder maskiert werden können. Dies verhindert unbefugte Zugriffe auf sensible Informationen.

### VERSCHLÜSSELUNG: ERWEITERTE VERSCHLÜSSELUNG IM RUHEZUSTAND

Zunehmende interne und externe Sicherheitsbedrohungen, umfassendere Compliance-Anforderungen und Cloud Computing sind nur einige Gründe, warum die Verschlüsselung wichtig ist. Nehmen wir beispielsweise folgende Szenarien, die in der Praxis häufig vorkommen: ein Systemadministrator mit uneingeschränktem Zugriff auf die Dateien auf einem Server, unberechenbare Attacken, mit denen sich Hacker Zugang verschaffen, oder eine Datenbank, die bei einem Cloud-Anbieter gehostet wird, über den Sie keine Kontrolle haben. Ohne Verschlüsselung – und auch trotz Dateiverschlüsselung – können der Systemadministrator, der Cloud-Anbieter oder der Hacker auf Dateien zugreifen und diese verändern, einschließlich der Dateien in Ihrer Datenbank.

Die Verschlüsselung ist eine neue Funktion von MarkLogic 9, mit der Daten, Konfigurationen und Protokolle transparent verschlüsselt werden können. Das bedeutet, dass MarkLogic-Dateien *auf Festplatten* im Ruhezustand verschlüsselt werden, sodass kein Zugriff außerhalb von MarkLogic möglich ist. Diese Funktion erfordert keine Anpassung von Anwendungen, die auf Grundlage von MarkLogic-Datenbanken entwickelt wurden. Die Verschlüsselung in MarkLogic ist umfassend und flexibel. Systemadministratoren können entscheiden, welche Datenbanken verschlüsselt werden sollen sowie ob die Konfiguration oder auch Protokolle ebenfalls durch eine Verschlüsselung zu schützen sind.

Die Verschlüsselung funktioniert mit Schlüsseln, die mit MarkLogic oder einem externen Key Management System (KMS) generiert werden. Die Verwendung eines externen KMS bietet eine weitere Sicherheitsstufe, da die Kontrolle zwischen dem MarkLogic-Administrator und dem Sicherheitsadministrator des KMS geteilt wird. Unabhängig davon verhindert die Verschlüsselung von Protokollen und Konfigurationen, dass keine unberechtigten Benutzer – selbst nicht MarkLogic-Administratoren – Ereignisse in einer Log-Datei löschen oder ändern können. Daher ist bei einem Fehler keine Verschleierung der Spuren möglich.

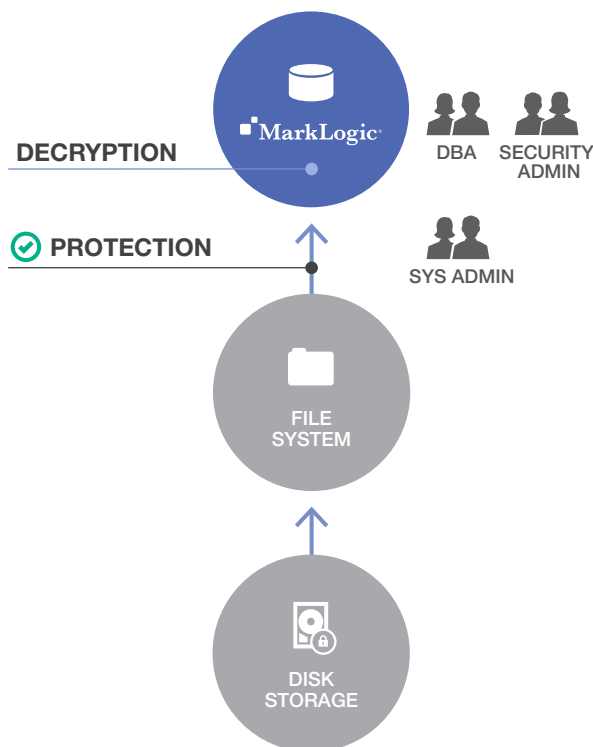


Figure 4: Advanced Encryption prevents access from outside MarkLogic by ensuring files resting on disk are fully encrypted.



Abbildung 5: Die Sicherheit auf Elementebene bietet in Echtzeit Kontrolle darüber, wer den Inhalt von Dokumenten anzeigen darf.

## SICHERHEIT AUF ELEMENTEBENE: GRANULARE SICHERHEIT AUF DOKUMENTEN-UNTEREBENE

Bei vielen Datenbanken funktioniert der Zugriff nach dem „Alles-oder-nichts“-Prinzip. Das bietet jedoch keinen ausreichenden Schutz vor heutigen Cybergefahren für die Sicherheit. Heutzutage müssen Unternehmen definieren können, wer welche Daten auf granularster Ebene sehen kann. Nehmen wir z. B. eine Anwendung, die die Kreditkartendaten und personenbezogenen Informationen für einen Callcenter-Betreiber anzeigt. Die Offenlegung dieser Informationen – selbst gegenüber ordnungsgemäßen Anwendern – kann gegen Datenschutzgesetze verstoßen und die Daten Ihres Unternehmens unnötigen Risiken aussetzen. Die Kontrolle über den Datenzugriff auf Datenbankebene eliminiert solche Risiken und verhindert, dass sensible Daten über eine Anwendung oder ein Netzwerk offengelegt werden und in die falschen Hände gelangen. Die Sicherheitskontrollen erfolgen auf Ebene der Datenbank. Es handelt sich also nicht um irgendeine Ergänzung, die Anwendungsentwickler erst in den Code oben im Stack integrieren müssen.

MarkLogic bietet seit jeher eine rollenbasierte Zugriffskontrolle auf Ebene des einzelnen Dokuments. Der Schutz bei MarkLogic 9 reicht noch weiter: Hiermit können Sicherheitsadministratoren die gleichen Kontrollen auf einzelne Teile eines Dokuments





Figure 5: Redaction is similar to element level security, but is designed to control who sees what data when the data is exported from MarkLogic.

anwenden. Die Sicherheit auf Elementebene ermöglicht eine Zugriffskontrolle auf Ebene von JSON-Eigenschaften oder XML-Elementen innerhalb von Dokumenten. Das bedeutet, dass bestimmte Informationen in einem Dokument vor einem bestimmten Benutzer (abhängig von seiner Benutzerrolle) verborgen werden können, während der gleiche Benutzer andere Informationen in dem Dokument sehen kann.

Die Sicherheit auf Elementebene ist äußerst flexibel, weil sie beschreibt, welche Elemente geschützt werden sollten. Beispielsweise kann ein Administrator eine Sozialversicherungsnummer schützen, wenn diese zur Dokumentstruktur gehört (d. h. unabhängig vom Schema). Der Administrator verwendet dafür einen Rich-Path-Standardausdruck statt einer konkreten Spezifikation.

Diese Ausdrücke können Attribute eines Elements enthalten – nicht nur die Elementbezeichnung. Beispielsweise kann ein Element mit einem Klassifizierungsattribut als geheim eingestuft werden. Eine restriktivere Sicherheitsregel kann für alle Elemente angewandt werden, die dieses Attribut haben.

### **SCHWÄRZEN: EXPORTKONTROLLE VERHINDERT DAS TEILEN SENSIBLER DATEN**

Manchmal sind Cyberattacken äußerst raffiniert. Ein anderes Mal werden Daten versehentlich einem Risiko ausgesetzt. Nehmen wir z. B. den häufigen Fall, wenn Anwendungen mit Musterdaten aus der Produktion getestet werden. Ein Entwickler oder Testleiter erhält dann vielleicht Zugriff auf echte Kreditkartendaten oder personenbezogene Informationen. Heutzutage kann selbst das versehentliche Offenlegen solcher Informationen

gegen Datenschutzvorschriften verstoßen und Ihr Unternehmen einem Risiko aussetzen.

Durch ein solches Schwärzen wird eine Bekanntgabe sensibler Informationen verhindert. Dafür werden vorhandene Informationen entfernt oder durch andere Werte ersetzt, die für die Arbeit der Endbenutzer keine Rolle spielen.

Mit Regeln wird beim Schwärzen definiert, welche Informationen beim Exportieren von Daten entfernt werden müssen. Der Prozess ist einfach, flexibel und sicher. Zuerst erstellt der MarkLogic-Sicherheitsadministrator Richtlinien mit Regeln, die definieren, welche sensiblen Informationen auf welche Weise geschwärzt werden sollen. Dann entscheidet der Administrator, welche Regel beim Exportieren anzuwenden ist. Mit diesem Prozess kann der Administrator Datensätze für bestimmte Zwecke generieren, die sensible Daten unterschiedlich behandeln (z. B. für die Entwicklung oder die Datenanalyse). Beispielsweise kann ein Datensatz für Entwickler eine zufällig generierte Kreditkartennummer enthalten, während ein Datensatz für Analysten überhaupt keine Kreditkarteninformationen aufweist. Zudem werden die Regeln und Aktionen im Protokoll erfasst, damit Sie später sämtliche Exportaktivitäten überprüfen können.

### **VERBESSERTE HANDHABBARKEIT**

MarkLogic 9 kommt mit zahlreichen neuen Funktionen, mit denen sich MarkLogic noch einfacher und effizienter verwalten lässt. **Ops Director** bietet eine zentrale Ansicht für die Visualisierung und das Management mehrerer Cluster und Produktversionen

übergreifend über Umgebungen und verschiedene Benutzergruppen. **Telemetrie** ermöglicht stärker automatisierte Support-Abläufe: Diagnose- und Systeminformationen über Ihren MarkLogic-Cluster können direkt an den Support von MarkLogic gesendet werden, wodurch Support-Cases nicht nur schneller abgeschlossen werden, sondern auch viele manuelle Aufgaben entfallen.

---

## **OPS DIRECTOR: VERSTÄNDLICHER ÜBERBLICK ÜBER DAS MARKLOGIC-SYSTEM<sup>2</sup>**

Während die Infrastrukturkosten fallen, steigen die Personalkosten. Die Folge: Heutzutage müssen Datenbankadministratoren (DBA) in Unternehmen mehr Daten und Server als je zuvor verwalten. Entscheidend ist, dass Datenbankadministratoren ihre Arbeit noch effizienter erledigen können. Dafür muss aber die Pflege moderner Datenbanken so einfach wie möglich sein. Aus diesem Grund wurde die Benutzeroberfläche für die Cluster-Administration in MarkLogic 9 komplett überarbeitet.

MarkLogic 9 bietet mit Ops Director eine zentrale Ansicht für die Administration. Dies erleichtert sowohl erfahrenen Datenbankadministratoren als auch Einsteigern in die Materie das Management von MarkLogic. Ops Director bietet Datenbankadministratoren eine konsolidierte Ansicht mit aussagekräftigen Diagrammen und Dashboards. Monitoring und Fehlerbehebung bei Clustern sind dank besserer Alarmfunktionen, Protokollsuche und Performance-Berichten deutlich einfacher. Zudem profitieren Sie mit der rollenbasierten Zugriffskontrolle und Informationssicherheit von höchsten Sicherheitszertifizierungen für Unternehmen. Mit Ops Director lassen sich Probleme leicht proaktiv erkennen und in Ruhe lösen. Dank der zentralisierten Erfassung, Bereitstellung und Speicherung der Daten können IT-Teams aus diesen Erfahrungen lernen und Analysen durchführen.

Hier eine kurze Liste mit einigen der vielen Möglichkeiten, die Ops Director bietet:

- Dashboard mit Übersicht über Ereignisse in Ihrem Cluster
- Alarmer in Echtzeit bei wichtigen Ereignissen
- Überwacht Cluster, Hosts, Datenbanken und Applikationsserver

- Zeigt Fehlerprotokolle an und reagiert entsprechend
- Analysiert problematische Hosts
- Kommuniziert mit dem MarkLogic-Support
- Management-Dashboard zur Auslastung von Ressourcen
- Analysiert die Performance von Festplatten, CPU, Speicher, Netzwerk, Datenbanken und Servern
- Nachverfolgen aktueller Aufgabenstatus für Wartungsroutinen
- Verwaltet die Konsolen-Einstellungen für Sicherheit, Lizenzierung, Konnektivität und Telemetrie

---

## **TELEMETRIE: BESSERER, SCHNELLERER SUPPORT**

Die Telemetrie gehört zu den Aspekten, an deren Verbesserung wir ständig arbeiten. Da Diagnosedaten über Ihre MarkLogic-Cluster automatisch gesammelt werden, genießen Sie einen besseren, schnelleren Support. Die Erfolgsbilanz des erstklassigen Kundensupports von MarkLogic kann sich sehen lassen und dank der Telemetriedaten verkürzt sich die Problemlösung erheblich.

Telemetrie ist eine Opt-in-Funktion. Ist diese Option aktiviert, werden Diagnose- und Systeminformationen über einen MarkLogic-Cluster gesammelt, verschlüsselt und an ein sicheres MarkLogic-Ziel übermittelt. Aus Datenschutz- und Sicherheitsgründen werden keine Benutzerdaten, Anwendungsprotokolle und Zugriffsprotokolle gesendet. Sie behalten vollständige Kontrolle darüber, welche Daten übermittelt werden sollen: Protokolldaten, Messdaten, Konfigurationsdateien und Support-Anfragen. Auch die Datengranularität beim Senden legen Sie fest, um Umgebungen mit geringer oder instabiler Bandbreite gerecht zu werden.

## **NOCH MEHR VERBESSERUNGEN**

Neben Verbesserungen in den Bereichen Datenintegration, Management und Sicherheit bietet MarkLogic 9 viele weitere zusätzliche Funktionen und Erweiterungen, die Sie bei Ihrer täglichen Arbeit zu schätzen wissen dürften. Geodaten werden immer aussagekräftiger und präziser. Die Suchfunktion wird zunehmend stabiler. Tiered Storage und

<sup>2</sup> Hinweis: Das erste Release von Ops Director folgt nach dem 9.0-1 Server-Release.

bitemporale Datenhaltung lassen sich immer einfacher implementieren, um Compliance-Vorgaben zu erfüllen. Dazu kommen Hunderte weiterer Verbesserungen, die im Bereich Enterprise NoSQL neue Maßstäbe setzen.

- **Compliance-Archiv** <sup>NEU</sup> – Zum Schützen temporärer Dokumente vor Löschen, Updates und Bereinigungen mit zeit- oder ereignisbasierten Richtlinien und Speichern von Dokumenten im WORM Storage. Funktioniert mit der Verschlüsselung, damit kein Systemadministrator Daten auf einer Festplatte verändern kann.
- **Azure Support** <sup>NEU</sup> – Azure-Image-Vorlage und Empfehlungen für Implementierungen auf der Cloud-Plattform von Microsoft.
- **Geodaten** – leistungsstärkere, genauere Antworten dank geobasierter regionaler Suchen und doppelter Präzision, um Anwendungen mit strengen Geo-Anforderungen zu unterstützen.
- **Suche** – neue Plugin-API für die Tokensierung und das Stemming. Damit können Sie jeden externen Tokenizer oder Stemmer für jede beliebige Sprache verwenden und bessere sprachübergreifende Suchfunktionen anlegen. Auch das Suchen in benachbarten Datensätzen (*Near-Suche*) und mit *Platzhaltern* funktionieren in MarkLogic 9 noch besser.
- **Erweiterter Tiered Storage** – höhere Flexibilität und bessere Performance für mehrstufige Speicherungen. Dank der Indizierung und dem Load Balancing von MarkLogic können Dokumente mit abgestuften Abfragen durchsucht werden, die auf Regeln basieren. Solche stufenübergreifenden Abfragen sind effizienter, da die Abfrage in mehreren Teilen (partitioniert) erfolgt.
- **JavaScript** – Das serverseitige JavaScript wurde auf die neueste Version von JavaScript (ES2015) aktualisiert, welches neue Syntax- und Standardbibliotheken zur Verfügung stellt. Dadurch wird das Schreiben, Debuggen und Lesen von Code vereinfacht.
- **Abfrage-Konsole** – Das automatische Vervollständigen dafür, dass Funktionen, Namespaces, Scope-Variablen und Keywords je nach aktuellem Kontext richtig eingegeben werden.

Zudem gibt es einen neuen Profil-Modus, um die Ausführung der Abfrage besser nachvollziehen zu können. Sie können damit einen Bericht über die Prozesse erstellen, die der Input-Code „im Verborgenen“ ausführt, und was wie lange dauert.

- **Node.js** – unterstützt jetzt die neuen Sicherheits- und Geodaten-Funktionen von MarkLogic 9.

## ERSTE SCHRITTE MIT MARKLOGIC 9

MarkLogic 9 legt die Messlatte bei Enterprise NoSQL-Datenbank erneut höher und bietet jetzt noch mehr Funktionen, um Daten aus Silos in einer gemeinsamen Datenbank zu integrieren.

- [Laden Sie sich MarkLogic 9](#) von der MarkLogic-Website für Entwickler herunter.
- [Lesen Sie die Release Notes](#) in der Dokumentation, um mehr über MarkLogic 9 zu erfahren
- [Machen Sie einen Kurs an der MarkLogic University](#), um sich über alle Neuerungen oder nur über neue Funktionen für die Datenintegration zu informieren.

Für Kunden mit bestehendem Supportvertrag wird MarkLogic 9 als kostenloses Update angeboten. Weitere Informationen für ein Upgrade von älteren Versionen finden Sie in den Release Notes. In der Dokumentation wird auch genau beschrieben, wie Sie MarkLogic 9 in jeder beliebigen Umgebung ausführen können.

Als weitere Unterstützung für das Upgrade auf MarkLogic 9 bietet MarkLogic Consulting einen [Upgrade Accelerator](#) an. Dieses Tool wurde eigens entwickelt, damit bestehende Kunden schnell und einfach von den leistungsstarken Funktionen von MarkLogic 9 profitieren können. Wie Sie den MarkLogic 9 Upgrade Accelerator verwenden, erklärt Ihnen gern Ihr MarkLogic-Kundenbetreuer oder Consulting Director. Oder wenden Sie sich einfach direkt an uns unter +1 877 992 8885 oder per E-Mail an [consulting@marklogic.com](mailto:consulting@marklogic.com).

---

© 2017 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 | +49 897 10 42 2151  
Skyper Villa, Taunusanlage 1, Frankfurt 60329 | +49 695 05 06 0588  
[de.marklogic.com](http://de.marklogic.com) | [germany@marklogic.com](mailto:germany@marklogic.com)



Theatinerstr. 11, 8. Etage, München 80333 | +49 897 10 42 2151  
Skyper Villa, Taunusanlage 1, Frankfurt 60329 | +49 695 05 06 0588  
de.marklogic.com | [germany@marklogic.com](mailto:germany@marklogic.com)