

# Wie kommt der vorhandene Content in das neue Oracle Universum?

**Gunther Thielemann**  
**Slix Gesellschaft für Computersystem mbH**  
**Au i. d. Hallertau**

**Tino Albrecht**  
**ORACLE Deutschland B.V. & Co. KG**  
**Potsdam**

## **Schlüsselworte:**

Enterprise Content Management, ECM, DMS, WebCenter Content, Universal Content Management, UCM, Stellent, Oracle Webcenter Content Management, CMSDK, Content Management SDK, Oracle Collaboration Suite, Content Services, Content DB, Migration Content Management

## **Einleitung**

Mit Oracle WebCenter Content lässt sich das ganze Spektrum von Anforderungen im Content Management Bereich abdecken. Bisher wurde das zugehörige Basisprodukt von Oracle unter „Universal Content Management“ (UCM) angeboten. Das kürzlich vorgenommene Rebranding schränkt die universelle Bedeutung jedoch keineswegs ein. Von Fileserver- Konsolidierung bis hin zu anspruchsvollem Web Content Management bietet Oracle hiermit eine robuste und skalierbare Plattform für Content-basierende Anwendungen auf einer leistungsfähigen Infrastruktur.

Wie der Titel schon andeutet, konzentriert sich dieser Vortrag auf die Core Capabilities (vorher Oracle Universal Content Management) und darauf, wie Sie den vorhandenen Content in dieses neue System transferieren können. Auf der Basis von UCM ein neues System aufzubauen ist einfach. Wenn man nicht von Null beginnt, erscheint eine Migration jedoch als große Herausforderung. Wir zeigen an Hand eines konkreten Beispiels, wie ein solcher Umzug geplant und erfolgreich durchgeführt werden kann.

## **Ausgangssituation**

Beim dem zu migrierenden System, „FIADAS“ (File Administration and Archiving System), handelte es sich um ein Dokumenten Management System zur vorgangsbezogenen Ablage, Verwaltung und Archivierung von Dokumenten einschließlich E-Mail, das auf der Basis des Oracle Content Management SDK (CMSDK) implementiert war. Mehrere Hundert Anwender von mehreren Standorten weltweit haben über die Jahre den Bestand auf etwa 1,5 Millionen Dokumente in ungefähr 300.000 Vorgangsakten wachsen lassen. Metadaten und Kategorien beziehen sich nicht auf einzelne Dokumente sondern auf die jeweilige Vorgangsakte, die als Container fungiert. Über die Protokolle SMB und IMAP wurden grundlegenden Funktionen zum Erstellen und Bearbeiten von Dokumenten unter Windows und Microsoft Office und MS Outlook zur Verfügung gestellt. Recherche, die Bearbeitung der Metadaten und der Zugang zu erweiterten Optionen waren im Altsystem nur über eine Web Oberfläche möglich.

## **Gründe für den Wechsel**

Da die Basistechnologie (CMSDK, Oracle Application Server 10.1.2, Java 1.4) das Ende des Lebenszyklus erreicht hatte, wurde im Ergebnis eines Evaluierungsprozesses Ende 2008 entschieden, das vorhandene System nicht mehr zu erweitern, sondern nach UCM zu migrieren und die sich damit bietenden neuen Möglichkeiten für den Ausbau der Funktionalität zu nutzen. Schlüsselrollen dabei spielten die umfangreiche Integration in den Windows Desktop mit der Desktop Integration Suite (DIS), die weitreichenden Konfigurationsmöglichkeiten des Systems und die zu erwartenden Steigerungen in der Performance. Ein weiterer gewichtiger Punkt war wie Eingangs erwähnt aber eben auch die Perspektive für Verfügbarkeit und Weiterentwicklung der Technologie durch den Produkthersteller.

## **Umfang der Migration**

Eine solche Migration beschränkt sich nicht darauf, die Daten in das neue System zu importieren, sondern betrifft alle Anwendungsschichten:

- Präsentationsschicht,
- Geschäftslogik und
- Datenschicht

## **Anwendungsmigration**

Auch wenn Webcenter Content ein ganzes Arsenal von Tools zur Anwendungs-Integration zur Verfügung stellt, mit der Content Integration Suite sogar ein eigenes Produkt, hatten wir uns nach gründlicher Analyse entschieden, die vorhandene kundespezifische Geschäftslogik nach UCM zu portieren. Um die Desktop Integration Suite optimal nutzen zu können, war es ohnehin erforderlich, die Präsentationsschicht in UCM zu implementieren.

Die modulare, serviceorientierte Architektur des Content Server bietet einen optimalen Rahmen für Anpassungen und Erweiterungen. Dieses grundlegende Konzept gilt nicht nur für die Java Komponenten, sondern findet sich auch in der Gestaltung der HTML Oberfläche wieder und reduziert den Aufwand bei Anpassungen enorm. Oracle verwendet hier eine eigene, sehr mächtige und erweiterbare Scriptsprache, die den Zugriff auf Content und Services bietet. Neben dieser stehen zusätzlich auch noch Java und Webservices Schnittstellen zur Verfügung um das System auf Basis von Standards zu erweitern.

Das gewählte Vorgehen hat sich im konkreten Fall bewährt und führte zügig zum Erfolg. Im geplanten Zeitraum konnten alle vorhandenen Funktionen re-implementiert und die geplanten Erweiterungen anforderungsgemäß umgesetzt werden.

Dieses Vorgehen kann jedoch nicht in allen Fällen empfohlen werden. Eine zweckmäßige Entscheidung bedarf der gründlichen Analyse der bestehenden Anwendung einschließlich der nichtfunktionalen Anforderungen.

## Datenmigration

Auch wenn es keine genaue Entsprechung der einzelnen Objekte in beiden Systemen gibt und sich die Sicherheitskonzepte grundlegend unterscheiden, konnte das Modell vollständig und weitestgehend automatisiert übernommen werden. Für notwendige Transformationen ließen sich Regeln definieren.

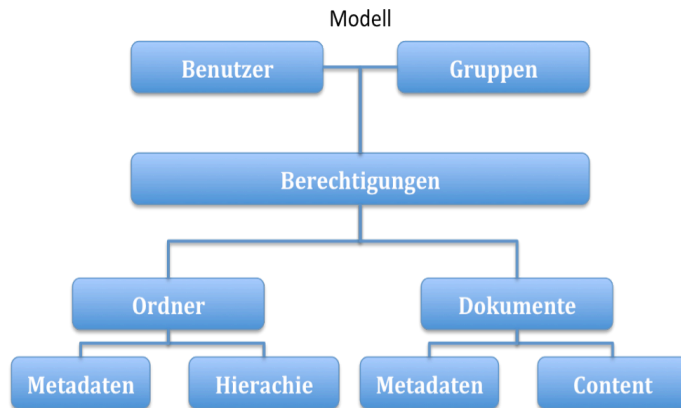


Abbildung 1 Modell DMS

Die Dokumente mit Metadaten zu übernehmen, erwies sich als unproblematisch. Um die Hierarchie und Containerobjekte übertragen zu können, sind in UCM die Komponenten für Folders und WebDav erforderlich. Die im Content Management SDK vorhandenen Kategorien lassen sich in UCM über Metadaten abbilden.

Prädestiniertes Tool für den Import von Dokumenten in den Content Server ist der Batchloader. Dieses Programm ist einfach zu verwenden, lässt sich über Text Dateien skripten und hat sich auch für Migrationen als sehr mächtig und nutzbringend erwiesen, erfordert jedoch den Zugriff auf ein Dateisystem.

Das Lesen der Metadaten aus CMSDK, Content DB oder Content Services muss unter Verwendung der vorhandenen API's erfolgen. Von Oracle stehen dafür keine fertigen Tools zur Verfügung. Die von der Slix GmbH entwickelten Migrationskomponenten können diese Aufgabe übernehmen und haben sich im aktuellen und weiteren Projekten bewährt. Dazu später mehr.

## Migrationsschritte

Es wird empfohlen, in folgender Reihenfolge vorzugehen:

- Erstellen der Objekte für die Zugriffsteuerung (Security Groups, Accounts und Roles)
- Erstellen der Metadaten
- Anlegen der Ablagestruktur (Bei Verwendung von Folders).
- Import der Content Objekte und der Metadaten
- Aufbau des Search Index

Die Benutzerkonten können entweder vorher erstellt oder von einem Directory Service übernommen werden. Im letzteren Fall kann die Zuweisung der Rollen und Konten (Accounts) an die Benutzer auf Basis der Organisation oder Gruppen über Regeln konfiguriert werden. Während der Übernahme von Content und Metadaten sollte der Index Service ausgeschaltet werden um den Durchsatz zu steigern und ein unnötiges Fragmentieren des Volltext Index zu vermeiden.

## Migrationskomponenten

Die von Slix entwickelten Komponenten ermöglichen die Automatisierung der folgenden Massenoperationen:

- Erstellen von Benutzern,
- Erstellen der Metadaten,
- Anlegen Ablagestruktur im Content Server<sup>1</sup> auf Basis der im Quellsystem vorhandenen Hierarchie,
- Import von Content und Metadaten oder
- Erstellen von Skripten für den Batchloader.

Für die Verwendung sind keine detaillierten Kenntnisse über Aufbau und Funktion beider Systeme erforderlich. Darüber hinaus bieten die Komponenten Optimierungen zur Steigerung des Durchsatzes bei Massenimporten. Die Konfiguration erfolgt über Mapping Dateien.

Der Import von Content und Metadaten kann entweder direkt durch eine Komponente oder in einem zweistufigen Prozess erfolgen. In der zweiten Variante werden Eingabedateien für den Batchloader generiert. Der eigentliche Import erfolgt dann durch das von Oracle bereitgestellte Tool. Der kombinierte Einsatz bietet folgende Vorteile:

Es besteht die Möglichkeit, die generierten Dateien vor dem Import zu prüfen und gegebenenfalls ohne erneutes Einlesen der Daten anzupassen. Da die vom Batchloader generierte Fehlerausgabe dem Eingabeformat entspricht, können diese generierten Dateien nach Beseitigung der Fehlerursache für einen erneuten Lauf verwendet werden.

## Konzeption und Planung

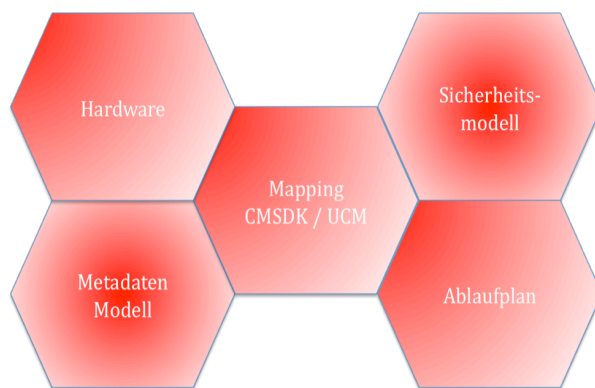


Abbildung 2 Konzeptbausteine

Für UCM ist das Berechtigungskonzept mit anderen Mechanismen als im CMSDK umzusetzen, da es keine direkte Entsprechung für die im Quellsystem vorhandenen ACLs gibt. In diesem Zusammenhang ist zu definieren, ob und wie die Integration in die vorhandenen Directory Services erfolgen soll. Mit der mächtigen Weblogic Infrastruktur gestaltet sich das in der Regel weitaus einfacher als über das bisherige Single-Sign-On mit OID (Oracle Internet Directory).

Es ist zu beurteilen, ob für die Migration zusätzliche Hardware erforderlich ist. Unter anderem ist zu berücksichtigen, dass für den Aufbau des Search Indexes ausreichend Plattenplatz im Verzeichnis des Content Servers vorhanden sein muss. Da der Bedarf im laufenden Betrieb wesentlich geringer ist, hat sich eine zwischenzeitliche Erweiterung der Kapazität als ausreichend erwiesen.

---

<sup>1</sup> Der Content Server ist das Basismodul von WebCenter Content und ist für die Kernfunktionalitäten, wie Datenverwaltung, Metadaten, Versionierung, Workflows, etc. zuständig

Zur Abbildung der vorhandenen Objekte im neuen System sind Beziehungen und Regeln zu definieren und Matrizen zur Steuerung der Migrationskomponenten zu erstellen. Für die Planung des konkreten Ablaufes ist unter Berücksichtigung aller Rahmenbedingungen zu entscheiden, ob die Migration in einem Schritt vollständig oder in Etappen erfolgen soll. Letzteres kann durchaus eine pragmatische Option sein, da es mitunter nicht erforderlich ist, den über mehrere Jahre gewachsenen Bestand in einem „Big Bang“ ins „neue Universum“ zu übertragen. In unserem Projekt sind wir in zwei Stufen vorgegangen. Nach der Übernahme des Gesamtbestandes nach UCM wurde ein zweiter Lauf zur Aktualisierung der zwischenzeitlichen Änderungen ausgeführt. Somit konnten die Anwender bis auf die endgültige Umstellung an einem Wochenende praktisch ohne Unterbrechungen weiter arbeiten.

### Spezialfall: E-Mail

Um E-Mails als solche unter UCM mit Outlook bearbeiten zu können, müssen sie im MSG-Format vorliegen. Da die Ablage im CMSDK nach RFC 822 erfolgt, ist eine Konvertierung erforderlich. Auch hier bietet Slix ein Tool an.

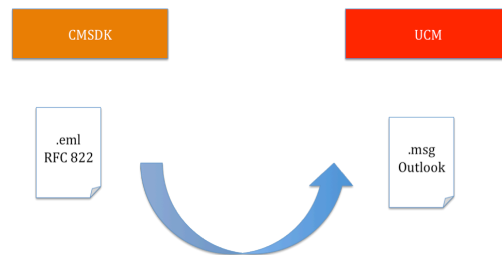


Abbildung 3 Konvertierung E-Mail

### Fazit

Insbesondere Oracle Kunden, die nach einer Ablösung für bestehende Content Management Systeme auf der Basis von Oracle's Content DB, Content Services oder Content Management SDK suchen, erhalten eine Orientierung für die Planung und Durchführung solcher Projekte. Die Slix GmbH vermittelt als Oracle Partner Erfahrungen, wie sich solche Systeme erfolgreich nach WebCenter Content portieren und migrieren lassen und beschreibt die Verwendung der verfügbaren Komponenten zur Automatisierung der Massenoperationen. Im konkreten Projekt wurden alle Daten komplett übernommen. Unter UCM konnten die vorhandenen Funktionen die geplanten Erweiterungen in vollem Umfang zur Verfügung gestellt werden. Die gewünschte erweiterte Integration in den Windows Desktop einschließlich der Recherche und der Bearbeitung der Metadaten wurde erreicht. Trotz höherer Benutzeranzahl und ausgebauter Funktionalität konnten die Wartungskosten deutlich gesenkt werden. Zu erwartende Projektlaufzeiten sind in hohem Maße von Umfang und Komplexität der vorhandenen individuellen Anpassungen und Erweiterungen des Altsystems abhängig und liegen zwischen mehreren Tagen und wenigen Monaten.

Kontaktadressen:

Gunther Thielemann  
Slix Gesellschaft für Computersysteme mbH  
Nandlstädter Weg 6  
D-84072 Au i. d. Hallertau

Tino Albrecht  
ORACLE Deutschland B.V. & Co. KG  
Schiffbauergasse 14  
14467 Potsdam

Telefon: +49 8752 219308  
Fax: +49 8752 85034  
E-Mail: [gunther.thielemann@slix.de](mailto:gunther.thielemann@slix.de)  
Internet: [www.slix.de](http://www.slix.de)

Telefon: +49 331 2007 126  
Fax: +49 331 2007 559  
E-Mail: [tino.albrecht@oracle.com](mailto:tino.albrecht@oracle.com)  
Internet: [www.oracle.de](http://www.oracle.de)