

# Open Source ECM auf dem Oracle-Stack

Frank Closheim, inxire GmbH

Die Content-Management-Strategie von Oracle ist seit einigen Jahren im Wandel. Das Portfolio wurde oft durch Zukäufe erweitert und ergänzt, doch nicht alle Kunden konnten von dieser Strategie profitieren. Häufig mussten bestehende und stabile Landschaften durch das Nachfolgeprodukt ersetzt werden. Methoden, die Analyse, Verständnis und Kommunikation von Daten für Menschen vereinfachen.

Ein Content-Management-Produkt aus der Oracle-eigenen Entwicklung sorgte bislang für Kontinuität: das CM SDK. Oracle hat dieses Juwel Anfang des Jahres an die Open-Source-Community übergeben, um die Investitionen der Kunden zu schützen und die Weiterentwicklung auf eine neue Basis zu stellen. Dieser Artikel zeigt die Entwicklung von Oracle CM SDK auf und stellt Kundenbeispiele, Einsatzbereiche sowie die Möglichkeiten der Open-Source-Lösung vor.

## Die Anfänge

Noch bevor der Begriff „ECM“ geprägt war, hatte Oracle im Mai 2000 mit dem Oracle Internet File System 1.0 („IFS“) ein Content Repository auf Basis der Datenbank veröffentlicht. Der Grundgedanke war, alle Inhalte und Metadaten vollständig in der Datenbank abzulegen und den Zugriff über ein Java SDK sowie diverse Dateizugriffsprotokolle wie SMB, NFS, IMAP, WebDAV oder HTTP zu ermöglichen. Dieser Ansatz war allen anderen überlegen und fand später auch Einzug in JSR-170 und JSR-283. Dass die Content-Management-Infrastruktur dabei auf der Oracle-Datenbank aufsetzt, zeigte sich als entscheidender Vorteil gegenüber den Wettbewerbern. Die Anwender profitierten von der Oracle-Datenbank-Technologie und nutzten das vorhandene Know-how für Hochverfügbarkeit, Skalierbarkeit, Datensicherheit sowie Performance-Tuning.

## Oracle begibt sich ins Lösungsgeschäft

Anfangs wurde iFS noch mit der Oracle-Datenbank ausgeliefert. Im Jahr 2002 entschied sich Oracle für eine zweigleisige Strategie, um noch größeres Interesse an

dem Produkt zu wecken. So wurde iFS als Entwicklungsplattform unter dem neuen Namen „Oracle Content Management SDK“ (CM SDK) weiterentwickelt und als Teil des Oracle Application Servers ausgeliefert. Kunden und Partner waren somit in der Lage, eigene Content-zentrierte Lösungen effizient zu entwickeln. Laut Oracle nutzten im Jahr 2004 weltweit mehr als 2.000 Kunden dieses Produkt [1]. In Deutschland setzten viele Oracle Partner CM SDK schon früh als ECM-Entwicklungsplattform ein.

Auch Oracle nutzte CM SDK, um eigene Lösungen wie Oracle Files zu entwickeln. Damit bot Oracle erstmals eine seinerzeit moderne Web-Anwendung für die Dokumenten-Verwaltung an. Oracle Files war Bestandteil der Collaboration Suite. Oracle hat somit zu einem erstaunlich frühen Zeitpunkt das Thema Konvergenz von Content Management und Collaboration berücksichtigt und integriert. Während sich die Collaboration Suite langfristig nicht durchsetzen konnte, stellte Oracle Files seine Leistungsfähigkeit unter anderem im Oracle-eigenen Intranet noch jahrelang unter Beweis. Mit dem Einsatz von Oracle Files konnten mehr als 1.200 Windows File-Server in den Datacentern konsolidiert werden. Rund 70.000 Benutzer arbeiteten täglich mit dem System und erzeugten darauf mehr als 30.000 neue Dokumente pro Tag.

Die positiven Erfahrungen mit der eigenen Umsetzung, gepaart mit den Anforderungen von mehreren Hundert Kunden waren Triebfeder für die Entwicklung von Oracle Content Services, eine Weiterentwicklung von Oracle Files. Dies begünstigte auch die Neuerung der zugrun-

de liegenden Infrastruktur CM SDK, deren Version 9.0.4.2 Anfang 2005 herauskam. Zeitgleich versuchte man mit Oracle Content Services, sich im Content-Management-Lösungsgeschäft zu etablieren.

Im Jahr 2006 wurde Oracle Content Services in Oracle Content Database (CDB) umbenannt und fortan als Option der Datenbank vertrieben. Nach einem Exkurs über den Middleware- und Collaboration-Vertrieb gehörte das Thema „Content Management“ strategisch wieder zum Vertriebsbereich der Datenbank.

## Der Markt konsolidiert sich

Ebenfalls im Jahr 2006 setzte eine Konsolidierung unter den ECM-Anbietern ein. IBM kaufte FileNet und OpenText sicherte sich Hummingbird. Für Oracle war die Auswahl durch die strikt vorgegebenen Übernahmebedingungen eng gesteckt: Die Kriterien für Zukäufe von Oracle sind eine ausreichend große Kundenbasis, kompatible Technologie sowie positive wirtschaftliche Eckdaten. Da Oracle schnell handeln musste, war im Jahr 2007 die Übernahme von Stellent (damals im Leader-Quadranten) die folgerichtige Entscheidung. Diese Firmenzukäufe sorgten bei Kunden bisweilen für Verwirrung, da sich nun einige Produkte fachlich überlappten.

Im Jahr 2008 entstand, als weiteres Produkt auf Basis von CM SDK, das Universal Online Archive (UOA). Dieses neue Produkt hatte den Fokus auf Archivierung und lieferte verschiedene Archivierungsdienste für E-Mail, File Server, Daten aus Unternehmenssoftware oder jegliche andere Art von elektronischen Inhalten. Weitere Akquisitionen von Oracle auf dem Content-Management-

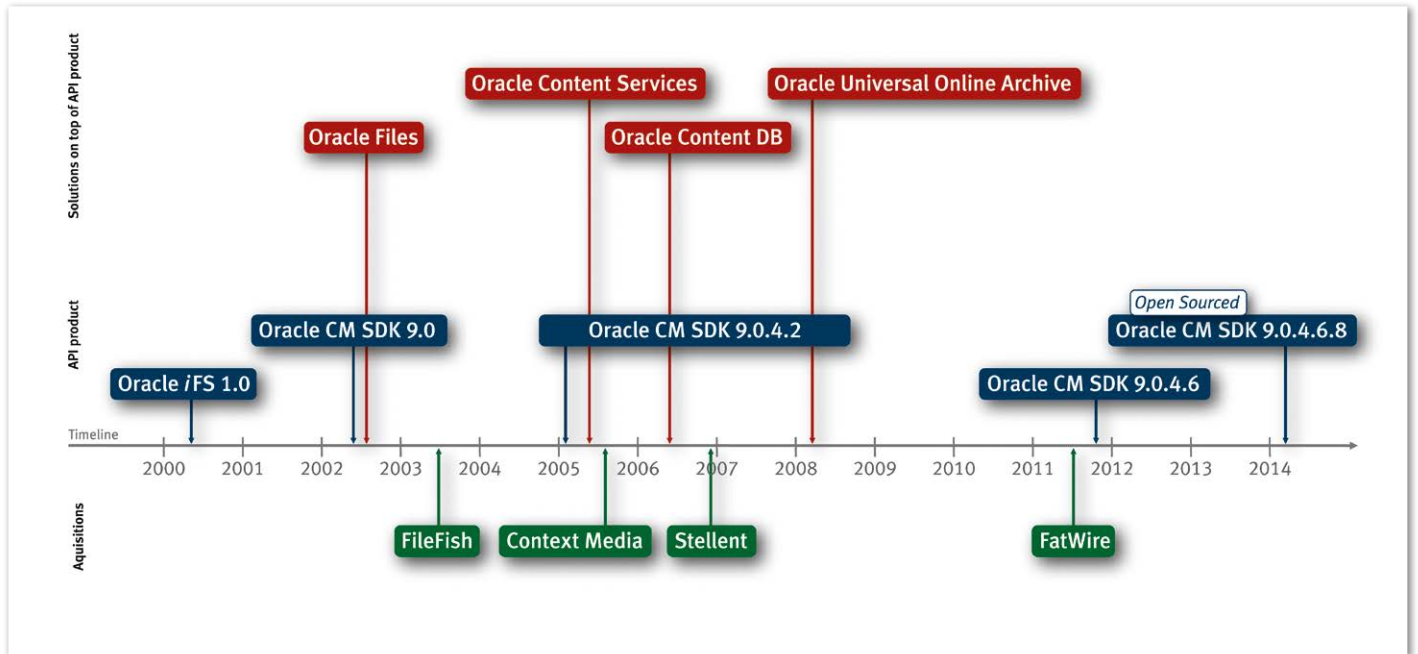


Abbildung 1: Evolution des Oracle CM SDK

Markt wie FatWire und zuletzt Compendium brachten zusätzliche Dynamik in die Produktstrategie. Betrachtet man die neuesten Akquisitionen sowie Oracle Documents Cloud Service, so ist eine strategische Neuausrichtung im ECM-Umfeld erkennbar.

### Investitionsschutz

Viele Kunden sind der Content-Management-Strategie von Oracle über viele Jahre gefolgt. Einige haben von einem auf das nächste Produkt migriert, während eine große Anzahl der Kunden ihre eigenen Lösungen mithilfe des Oracle-API-Produkts realisiert haben.

Um den zufriedenen CM-SDK-Kunden, die Oracle immer wieder bescheinigten, die skalierbarste und flexibelste Content-Management-Infrastruktur weltweit zu haben, auch weiterhin die Vorteile dieser Technologie anbieten zu können, entschied Oracle im Jahr 2011, CM SDK noch zukunftssicherer zu gestalten. Die Abhängigkeiten zum Application Server sollten aufgelöst und die aktuellste Java-Version unterstützt werden.

Im Dezember 2011 wurde eine neue Version von CM SDK 9.0.4.6 [2] veröffentlicht. Damit war das Produkt unabhängig vom Application Server für Java 1.6 freigegeben sowie für die Oracle-Datenbank 11g R2 zertifiziert. CM SDK kann mit dem Web-

Logic Server, aber auch unter GlassFish, Tomcat oder JBOSS betrieben werden.

Hinzu kamen einige Weiterentwicklungen wie die Unterstützung für Secure Files, automatisiertes Clustering und Failover, Policy-getriebenes Lifecycle Management, Anbindung an externe Archivsysteme, feingranulares und konfigurierbares Auditing, Event Handling sowie eine zusätzliche Skriptsprache zur Administration aus der Kommandozeile. Seitdem erscheinen regelmäßige Updates von Oracle CM SDK 9.0.4.6 [3].

### Oracle öffnet sich

Im Januar 2014 hat sich Oracle dann dazu entschlossen, den Quellcode des Oracle CM SDK als Open Source Code unter der BSD-Lizenz [4] der Öffentlichkeit zur Verfügung zu stellen [5]. Damit bestätigt Oracle nicht nur erneut sein Engagement für die Open-Source-Community, sondern folgt auch einer Strategie, die bereits mit der Freigabe von GlassFish und Java erfolgreich umgesetzt wurde (siehe Abbildung 1).

Die Open-Source-Strategie öffnet nicht nur für Bestandskunden ganz neue Möglichkeiten; auch Interessenten, die nach einer reifen, robusten, hoch skalierbaren und flexiblen ECM-Infrastruktur suchen, die sich nahtlos in eine bestehende Oracle-Umgebung integriert, finden hier eine Lösung.

### Vielfältige Einsatzbereiche

CM SDK bietet zahlreiche Funktionen, die über Java-APIs und -Services aufgerufen werden können. Neben den klassischen Dokumenten-Management-Funktionen wie die Arbeit mit Ordnern, Check-in/Check-out von Dokumenten, Versionierung, erweiterbare Metadaten, feingranulare Zugriffsteuerung und Volltextsuche (via Oracle Text), bietet es auch Content-Management-Funktionalitäten, die bei anderen Anbietern kaum zu finden sind. Ein Beispiel ist der „LegalHold“-Service: Im Rahmen von Auditing-Verfahren müssen oft bestimmte Dokumente eingefroren werden. Ein LegalHold pausiert einen vorhandenen Lebenszyklus eines Dokuments und verhindert somit, dass unter einem Audit stehende Dokumente verändert oder nach definierten Aufbewahrungsfristen vom System gelöscht werden. Zudem lässt sich CM SDK als Archivsystem für SAP verwenden. Dabei können Regeln für eine einheitliche Ablage („AutoFoldering“) definiert werden.

Für das Thema „Compliance“ stellt CM SDK neben der feingranularen Zugriffsteuerung auch Auditing-Funktionen zur Verfügung, die jede gewünschte Aktion im System protokollieren und auf Wunsch in einem Report anzeigen kann.

Für hohe Verfügbarkeit und Skalierbarkeit sorgen das Clustering- und Failover-

System. Damit erfolgt eine automatische und dynamische Lastverteilung über mehrere Knoten (Dienste) hinweg.

Als weiteres Highlight bietet CM SDK die Möglichkeit, nahezu beliebige externe Speichersysteme anzubinden. Über den sogenannten „CustomMediaDriver“ können die Binärdaten völlig transparent auf externe Filesysteme, Archivsysteme oder Bänder geschrieben und von dort gelesen werden. Weiter lassen sich auch bequem Oracle-Datenbank-Features wie Secure Files, Deduplication, Compression, Transparent Data Encryption oder auch Multimedia nutzen.

### Kundenbeispiele

Die Anwendungsbereiche des Oracle CM SDK sind so vielfältig wie die Kunden, die es einsetzen. Neben international bekannten Unternehmen wie Cisco, TOYS“R“US, Quantas Airways, General Motors, Volkswagen oder Deutsche Bahn nutzt es auch die National Ignition Facility (NIF). Diese ist Teil der Lawrence Livermore National Laboratory (LLNL) und für den größten und kraftvollsten Laser weltweit bekannt. In der Forschung zur thermonuklearen Fusionszündung finden diverse datenintensive Prozesse statt, die Unmengen an Daten (ca. 27 TB pro Jahr) produzieren. Diese wandern zunächst in ihrem Rohformat in CM SDK und werden dort in mehreren Stufen automatisiert nachbearbeitet, um sie so anderen Wissenschaftlern global zur weiteren Analyse und Zusammenarbeit für einen Zeitraum von dreißig Jahren zur Verfügung zu stellen. CM SDK wird bei NIF als Teil einer Big-Data-Lösung unter anderem als Datenbank-zentriertes hierarchisches Speichersystem und Wissensdatenbank genutzt.

Im Bereich „Anspruchsvolle Sicherheitsarchitekturen“ steht CM SDK auch im Dienst der Bundeswehr und des Bundeskriminalamtes. Die Bundeswehr nutzt das Produkt unter anderem für die sichere Aufbewahrung und Verteilung von schutzwürdigen Dokumenten. Auch das polizeiliche Informationssystem des Bundes und der Länder wurde als Eigenentwicklung des BKA auf Basis von CM SDK umgesetzt. Weitere Anwendungsbereiche sind neben Big Data und Datensicherheit auch Telematik, SAP-Archivierung, B2B oder Transaktions-basierte Dateisysteme.

### Professionelle Weiterentwicklung ist gesichert

Der Kern des ehemaligen CM-SDK-Entwicklungsteams hat sich mittlerweile mit dem Expertenwissen innerhalb der Community neu formiert. Unter dem neuen Namen „Enterprise Content Management SDK“ (ECMSDK) ist diese Technologie nun bei inxire GmbH aus Frankfurt beheimatet. Damit ist eine professionelle Weiterentwicklung dieses Produkts langfristig gesichert. Der weitere Fahrplan ist bereits gesteckt. Dazu gehört neben der Unterstützung der Oracle Datenbank 12c und Java 1.7 auch die Adaption von Oracle Engineered Systems. Darüber hinaus wird ECMSDK soziale Interaktionen an Inhalten ermöglichen. Hierzu stehen unter anderem kontextsensitive Konversationen und Benachrichtigungen bereit, die es beispielsweise ermöglichen, mit Kollegen über Inhalte zu diskutieren. Mit ECMSDK steht damit erstmals ein Open-Source-Produkt für Social ECM zur Verfügung.

### Weitere Informationen

- [1] <http://www.authorstream.com/Presentation/Christo-27668-1031-Oracle-Content-Management-Rest-Us-Agenda-Collaboration-Suite-Background-Files-as-Entertainment-ppt-powerpoint>
- [2] <http://www.oracle.com/technetwork/middleware/cmsdk/documentation/oracle-cmsdk-9046-data-sheet-2005978.pdf>
- [3] [https://updates.oracle.com/Orion/Advanced-Search/process\\_form?product=11803](https://updates.oracle.com/Orion/Advanced-Search/process_form?product=11803)
- [4] <http://www.oracle.com/technetwork/licenses/bsd-license-1835287.html>
- [5] <http://www.oracle.com/technetwork/middleware/cmsdk/documentation/oracle-cmsdk-sod-2112642.pdf>



Frank Closheim  
frank.closheim@inxire.com

## Prof. Dr. Harm Knolle ist DOAG-Botschafter 2014

Als DOAG-Botschafter werden jährlich diejenigen Personen gewürdigt, die sich in der Arbeit der DOAG durch hohes Engagement und Kompetenz auszeichnen.

Ein Komitee der DOAG, bestehend aus fünf Vorständen und Leitungskräften aus allen Communities, hat sich in diesem Jahr entschieden, den Preis „DOAG Botschafter 2014“ in der Kategorie „Technologie“ einer Person zuzusprechen, die der DOAG schon seit sehr langer Zeit eng verbunden ist: Prof. Dr. Harm Knolle, Professor für das Fachgebiet „Datenbanksysteme“ am Fachbereich Informatik der Hochschule Bonn-Rhein-Sieg.

Sein erster Kontakt mit der DOAG geht bereits in seine Zeit als Oracle-Daten- und Datenbank-Administrator Mitte der 1990er-Jahre bei der LTU in Düsseldorf zurück. Damals hat er häufig sehr wertvolle Tipps aus den DOAG-Publikationen erhalten. Auf die DOAG Konferenz + Ausstellung ist er im Jahr 2007 in seiner Rolle als Hochschullehrer in Schmalkalden aufmerksam geworden. Seitdem ist er in Begleitung zahlreicher Studierender regelmäßig dabei.

An der DOAG gefällt ihm am besten, dass der Verein ein Spiegelbild der aktuellen Situation vieler Bereiche der Informatik bietet. Gerade für Lehrer einer Hochschule für angewandte Wissenschaften sind die Themen der DOAG, die Konferenzen und die Publikationen stets eine gute und kompakte Quelle, um ihr Wissen auf dem Laufenden zu halten.

Am Rande der Konferenz erzählt Prof. Dr. Knolle eine kleine Anekdote im Zusammenhang mit der DOAG: Eine Studentin der Schmalkalder Delegation und ein Student der Bonner Delegation sind sich im letzten Jahr auf der DOAG Konferenz näher gekommen und heute ein Paar. So trägt die DOAG nicht nur zur Sicherung des Fachkräftebedarfs bei, sondern leistet auch ihren Beitrag in der Partnervermittlung.



Von links: Dr. Dietmar Neugebauer (DOAG), Prof. Dr. Harm Knolle, Urban Lankes (DOAG)