



Open-Source-Systeme für Content Management

Dr. Ulrich Wiesner

Services für Enterprise Content Management verbreiten sich zunehmend in den IT-Infrastrukturen moderner Unternehmen. Neben den Systemen etablierter Infrastrukturanbieter setzen sich zunehmend Open-Source-basierte Lösungen durch. Deren Vor- und Nachteile sowie typische Merkmale werden in diesem Dokument kurz beleuchtet.

Elektronische Dokumente in der Praxis

Der geordnete Umgang mit elektronischen Dokumenten ist mehr denn je eine Herausforderung für jedes Unternehmen: Sowohl im Kerngeschäft als auch in internen Prozessen spielen elektronische Dokumente, E-Mails und Portalinhalte eine entscheidende Rolle und haben Papier als Trägermedium weitgehend verdrängt.

Mehr noch: Der effektive Umgang mit elektronischen Dokumenten spart Kosten, hilft Risiken zu vermeiden und wird so zu einem relevanten Wettbewerbsvorteil.

Wie kann Technik dabei unterstützen?

Um diese Wettbewerbsvorteile zu erreichen, werden Systeme benötigt, die bei der systematischen Erstellung, Bearbeitung, Verwaltung, Verteilung, Publikation und Aufbewahrung von Inhalten jeglicher Art unterstützen.

Dies betrifft das Kerngeschäft, wo Dokumente im Prozesskontext eng mit den fachlichen Transaktionen zu verbinden sind, die interne Organisation (z. B. Prozessmanagement, QS-Verfahren), Team- und Projektarbeit (z. B. Abteilungsablage, Projektakte), den Außenauftritt eines Unternehmens im Web und die Kommunikation unter Management und Mitarbeitern (Intranet).

Da diese Inhalte i. d. R. nicht nur für einen Bereich relevant sind, sollten sie zentral verwaltet und den jeweiligen Personen gezielt bereitgestellt werden. So sind z. B. Dokumente über Zahlungen zur Schadenregulierung relevant für die Schadenakte, aber auch für die Buchhaltung. Ähnlich ist die Entwicklung von Organisationshandbüchern und Prozessbeschreibungen eine Kernaufgabe der Betriebsorganisation, aber alle Mitarbeiter müssen die freigegebenen Dokumente im Intranet abrufen können. Dieser Ansatz, Content organisationsübergreifend

bereitzustellen, hat zu dem Begriff Enterprise Content Management (ECM) geführt.

Wie genau helfen diese Systeme?

- Unstrukturierte Informationen, etwa digitalisierte Eingangspost, Akten, E-Mails und Web-Content sind auf Knopfdruck parallel an allen relevanten Stellen zugreifbar, können logisch gruppiert und in den Arbeitsablauf integriert werden.
- Einmal erfasste Inhalte werden automatisch regelbasiert für verschiedene Kanäle (Print, Fax, E-Mail, Web) angepasst, verlinkt und publiziert und damit den jeweiligen Zielgruppen aktiv personalisiert zur Verfügung gestellt.
- Mitarbeiter werden automatisch über anstehende Aufgaben informiert, sobald Post eintrifft oder vorangegangene Arbeitsschritte abgeschlossen sind (weniger Liegezeiten, weniger Prozessfehler).
- Arbeitsabläufe werden nachvollziehbar (Basis für Prozessverbesserung).

Welche Trends sind heute erkennbar?

Um elektronische Dokumente organisationsübergreifend als Informationsträger nutzen zu können, werden heute zentrale Services zur Verwaltung angeboten. ECM wird also Bestandteil der IT-Infrastruktur.

Moderne ECM-Systeme besitzen ein einheitliches Repository für alle Informationstypen in unterschiedlichen Formaten, um so Redundanz von Informationen zu vermeiden und konsistente Daten bereitstellen zu können. Darauf aufgesetzt ist oft eine Middleware für die Verwaltung der Daten, meistens in serviceorientierten Architekturen. Auf der Anwendungsseite bieten ECM-Systeme Komponenten als unabhängige Dienste, die in verschiedenen Anwendungen genutzt werden können.

Wenn ECM Bestandteil der Infrastruktur wird, ist es nachvollziehbar, dass die großen Infrastrukturanbieter hochwertige ECM-Applikationen in ihr Portfolio aufnehmen. Sie haben deshalb die leistungsfähigsten Spezialisten vom Markt übernommen (Documentum zu EMC², FileNet zu IBM, Stellent zu Oracle) und integrieren deren marktführende Produkte schrittweise als Services in ihre Infrastrukturen. Dieser Aufwand ist hoch, aber der Prozess ist aktuell in vollem Gange.

Parallel dazu haben diverse Standardisierungsinitiativen den Markt verändert: Besonders die aus der Java Community getriebenen JSR-Standards der OASIS¹ haben sich dabei prägend ausgewirkt. Die Spezifikationen JSR 170: Content Repository for Java Technology API und deren Nachfolger JSR 283: Content Repository for Java Technology API Version 2.0 haben wesentliche Eckpunkte zu einer systematischen Ablage der unstrukturierten Informationen vorgeschlagen.

Am 1. Mai 2010 hat die OASIS den Standard Content Management Interoperability Services (CMIS)² Version 1.0 verabschiedet. Damit existiert erstmalig ein Standard zum Zugriff auf verschiedene Content Repositories über einen Client. Egal, welches oder welche Content-Verwaltungssysteme einzubinden sind: CMIS stellt eine allgemeine Programmierschnittstelle auf REST³- oder Webservice-Basis bereit.

Vorreiter waren die Open-Source-Initiativen, u. a. Alfresco und das Apache-Jackrabbit⁴.

¹ OASIS: Organization for the Advancement of Structured Information Standards

² siehe auch <http://docs.oasis-open.org/cmisis/CMIS/v1.0/os/cmisis-spec-v1.0.html>

³ Representational State Transfer: Softwarearchitekturstil für verteilte Hypermedia-Informationssysteme wie das World Wide Web

⁴ siehe <http://jackrabbit.apache.org/>

Projekt. Auch die kommerziellen Anbieter wie EMC², IBM, Microsoft, Open Text, Oracle und SAP haben aktiv am Standard mitgearbeitet.

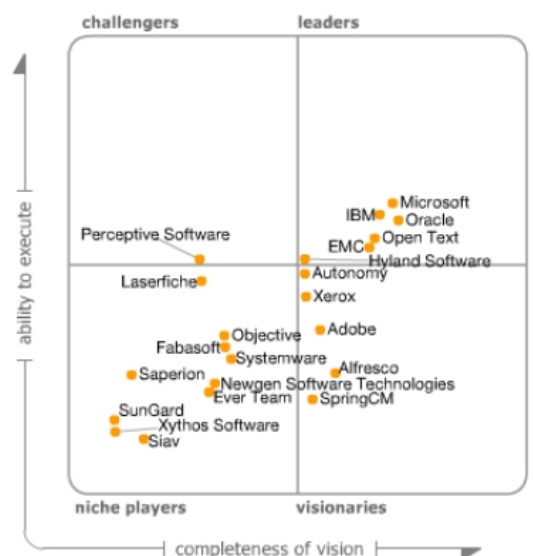
Auch zur Gestaltung von Oberflächen für Bedienung und Anwendungsintegration sind Richtlinien entwickelt (z. B. JSR 168: Portlet Specification) und von vielen Herstellern adaptiert worden.

Neuerdings tendieren Unternehmen zunehmend zu Social-Business-Systemen. Techniken, bekannt von Plattformen wie XING, Salesforce.com und Twitter, bieten Mitarbeitern schnelle unkomplizierte Informationszugänge und Kommunikationswege und beschleunigen damit innerbetriebliche Prozesse.

In Folge dieses Trends werden jedoch große Mengen unorganisierter Inhalte erzeugt, für die geeignete Kontrollstrukturen geschaffen werden müssen.

Wer sind die führenden Hersteller heute?

Der Trend zur Integration contentbezogener Services in Infrastrukturen ist am Anbietermarkt sehr gut nachvollziehbar. Der Gartner Magic Quadrant for Enterprise Content Management zeigt, dass sich EMC², IBM, Microsoft, und Oracle unter den Marktführern befinden:



Source: Gartner (November 2010)

Auf der anderen Seite ist in den letzten Jahren eine leistungsfähige Community von Open-Source-Anbietern entstanden, die deutlich spürbar den Markt beeinflusst.

ECM und Open Source

Was bedeutet Open Source Software?

Open Source Software bedeutet, dass mit Auslieferung der Software der Quellcode übergeben wird. Die Vorteile liegen darin, dass geschickte Programmierer nach entsprechender Einarbeitung in der Lage sind, die Software zu verändern und zu ergänzen. Für den Risikofall, dass der Hersteller der Software ausfällt, besteht die Möglichkeit, das System u. U. selbst zu pflegen und weiterzuentwickeln. Andererseits ist zu berücksichtigen, dass typische Enterprise-Funktionalitäten (z. B. LDAP-Integration) in kommerziellen Produkten heute noch besser unterstützt werden, als dies im Open-Source-Umfeld der Fall ist. Der Nachbau fehlender Funktionalität bedeutet Mehraufwand bei der Einführung.

Open Source Software ist jedoch nicht zwingend kostenfreie Software (Freeware). Lizenzkosten fallen zwar i. d. R. nicht an, es ist aber zu berücksichtigen, dass einige Open-Source-Anbieter für so genannte Enterprise Editions eine nicht unerhebliche jährliche Wartungsgebühr erheben. Diese sichert den Support und garantiert wohlgetestete Releases.

Der Einzug von Open-Source-Systemen in den ECM-Markt

Die von Gartner genannten führenden Anbieter von Open-Source-Systemen für Content Management sind Alfresco und Liferay. Beide verfolgen jedoch unterschiedliche Schwerpunkte.

Alfresco bietet nicht nur eine robuste Plattform für Entwickler und Unternehmen, die für das Erstellen jeder Art von Rich-Content-Applikatio-

nen konzipiert ist, sondern ist das am besten skalierbare Java-Content-Repository auf einfacher Hardware⁵.

Das System ist mit einer modernen Benutzerschnittstelle (Alfresco Share) ausgestattet, die stark auf Social Collaboration und Dokumentenmanagement durch viele Beteiligte ausgerichtet ist. Mit dem Einsatz offener Standards wie CMIS & JSR 168 ist Alfresco Enterprise eine Content-Plattform, die sowohl das Management und die Aufbewahrung von Content aus Business Systemen als auch moderne Social-Collaboration-Systeme unterstützt. Es bietet Anwendern schon im Standard die Einfachheit und Vertrautheit der Tools, die sie auch sonst verwenden: in Google suchen, in Wikipedia lesen, in einem Blog schreiben, Leute in sozialen Netzen finden, Benachrichtigungen in einem RSS-Reader erhalten und vieles mehr.

Ein weiterer Vorteil ist die schnelle Anwendungsentwicklung aus Komponenten, die mit standardbasierten Lightweight-Skripten erstellt wurden. Entwickler benötigen dazu keine Herstellerunterstützung, sondern nur gutes Programmier-Know-how.

Alfresco Software Ltd. wurde 2005 von John Newton, Mitbegründer von Documentum, und John Powell, früher COO bei Business Objects, gegründet und ist heute ein stabiles Unternehmen. Zu den Kapitalgebern gehören führende Investmentgesellschaften wie Accel Partners, Mayfield Fund und SAP Ventures. Namhafte Unternehmen wie SAP, Merck und andere setzen Alfresco produktiv sowohl im transaktionalen (Kerngeschäft, Rechnungswesen) als auch im kollaborativen Umfeld (z. B. Intranet, Mitarbeiterportal) ein.

⁵ Ein unabhängiger, zertifizierter Benchmark-Test der Unisys Corporation hat eine lineare Skalierbarkeit auf bis zu 100 Millionen Objekte bewiesen.

Das Open-Source-System Liferay⁶ ist als Portal für Publizierung, Inhalte und Zusammenarbeit ebenfalls weit verbreitet. Als eher leichtgewichtiges, Java-basiertes Portal orientiert es sich streng an bestehenden Standards. Nach eigenen Angaben ist Liferay skalierbar und performant, einfacher als WebSphere und flexibler als SharePoint⁷.

Nach unseren Erfahrungen ist Liferay sehr gut einsetzbar, wenn der Schwerpunkt auf einer flexiblen und modernen Portal-Oberfläche liegt. Sein eher proprietäres Dokumenten-Repository ist für einfache Strukturen ausgelegt und erfüllt derzeit noch nicht alle Anforderungen an eine reversionssichere Verwaltung des Content. Durch geschickte Kombination mit anderen Open-Source-Komponenten (z. B. Jackrabbit) kann diese Schwachstelle jedoch behoben werden.

Weitere Open-Source-Systeme für die Verwaltung elektronischer unstrukturierter Informationen sind z. B. Typo3⁸, Nuxeo⁹ oder Open-CMS¹⁰. Diese haben sich vielfältig bewährt, ihre Stärken liegen aber jeweils in unterschiedlichen Einsatzfällen.

Auch etablierte kommerzielle Anbieter wie Adobe setzen stark auf Open Source. Mit der Übernahme der Firma Day hat Adobe die JSR-283-konforme Open-Source-Repository-Schnittstelle des Apache-Jackrabbit-Projekts in sein Produktspektrum integriert.

⁶ siehe <http://www.liferay.com/>

⁷ siehe <http://www.liferay.com/de/products/liferay-portal/overview>

⁸ siehe <http://typo3.org/home/>

⁹ siehe <http://www.nuxeo.com/de/>

¹⁰ siehe <http://www.opencms.org/de/>

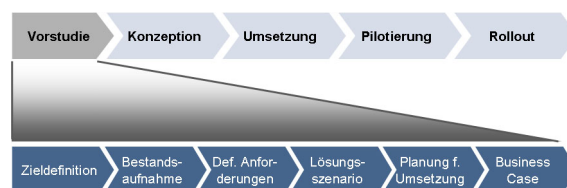
Fazit zu Open Source ECM

Open-Source-basierte Systeme sind ein etablierter Faktor für Enterprise Content Management. Die Systeme sind komplex und bieten einen Leistungsumfang, der denen der kommerziellen Systeme durchaus entspricht und alle typischen Funktionsbereiche abdeckt. Es ist also kein größeres Risiko mehr, unternehmensweite ECM-Lösungen auf Basis von Open-Source-Systemen zu errichten. Positiv zu bewerten sind die eingespielten Communities, in denen Entwickler und Anwender weltweit und vom Hersteller unbeeinflusst ihre Erfahrungen austauschen und sich gegenseitig unterstützen.

PPI unterstützt Sie

Auch die Einführungsprojekte für ECM-Systeme sind kein Hexenwerk mehr. Es gibt viele gute Produkte und leistungsfähige Dienstleister. PPI unterstützt Sie gern, beginnend bei strategischen Entscheidungen (über Zielfindung, Lösungsszenarien, Planung, Business Case) über die Systemauswahl bis zur Umsetzung und Einführung (Change Management).

Im ersten Schritt zur Entscheidungsfindung hat sich eine Vorstudie nach folgendem erprobten Verfahren bewährt:



Systematik und Nachvollziehbarkeit der Entscheidungen sind dabei Erfolgsfaktoren, an denen wir unsere Beratung messen lassen. Alle Ergebnisse erarbeiten wir mit Ihnen gemeinsam, so dass die Entscheidungen sicher und realistisch ausfallen.

Unsere Referenzen sprechen dafür, dass wir gemeinsam mit unseren Kunden messbaren Nutzen erzeugen. Wenn wir Ihr Interesse geweckt haben, sprechen Sie uns bitte an.

PPI AG Informationstechnologie
Dr. Ulrich Wiesner
Partner
+49 151 54375013
Ulrich.Wiesner@ppi.de

PPI AG Informationstechnologie

Die PPI AG ist seit 1984 erfolgreich für die Finanzbranche tätig. PPI stärkt die Marktposition von Versicherungen durch fachlich wie methodisch exzellente Lösungen für alle Kernprozesse der Assekuranz, basierend auf langjähriger Praxiserfahrung. In der Software-Entwicklung steht PPI für hohe Qualität und absolute Budgettreue. Das Consulting-Angebot erstreckt sich von strategischer über fachliche bis zur IT-Beratung.

Kontaktinformationen

PPI AG Informationstechnologie
Moorfuhrweg 13
D-22301 Hamburg
Tel.: +49 40 227433-0
Fax: +49 40 227433-1333
E-Mail: info@ppi.de
www.ppi.de