

Schwerpunktthema: Content Management und digitale Bibliotheken

In diesem Heft des Datenbank-Spektrums sollen die Themen *Content Management* und *digitale Bibliotheken* aufgegriffen und insbesondere die Bezüge zur Thematik »Datenbanksysteme« erläutert werden. Dazu werden in diesem Leitartikel die Begriffe *Content* und *Content-Management-Systeme* sowie *digitale Bibliotheken* und *digitale Bibliothekssysteme* kurz erklärt. Darüber hinaus werden die drei Artikel des Schwerpunkts in den thematischen Rahmen eingeordnet.

Es konnten noch nicht alle Aspekte des Content Managements und der digitalen Bibliotheken in diesem Heft in eigenen Artikeln aufgegriffen werden. Wir hoffen jedoch, dass in Zukunft beide Themen unter Datenbankern selbstverständlich werden – und dass auch in dieser Zeitschrift in regelmäßigem Abstand Beiträge zu diesen Themengebieten erscheinen.

Was ist Content?

Was ist zunächst der Content, der zu verwaltende Inhalt in einem Content-Management-System? Grob sind dies Informationen wie strukturierte Daten oder multimediale Dokumente, die in LANs oder WANs »veröffentlicht«, also bereitgestellt werden sollen.

Geht man bei WANs nun vom Web aus, spricht man auch spezieller von Web Content. Wichtig ist dabei, dass der Content sowohl aus einfachen Daten in strukturierten Datenbanken als auch aus multimedialen Dokumenten bestehen kann.

Was ist Content Management?

Content Management ist der Prozess der Verwaltung des bereitzustellenden Contents. Verwaltung heißt dabei, dass Daten und Dokumente erstellt, bearbeitet, freigegeben, wiedergefunden, gespeichert und langfristig archiviert werden können. Das Speichern und Wiederfinden von Daten geschieht dabei üblicherweise über ein relationales Datenbanksystem, das Speichern und Wiederfinden von Dokumenten dagegen über ein Dokumenten-

Management-System oder ein »Content Repository«.

Web Content Management ist der Prozess der Verwaltung von Daten und Dokumenten, die im Web bereitgestellt werden sollen. Oftmals wird dies reduziert auf die Verwaltung von Daten, aus denen sich dann dynamische Web-Dokumente erzeugen lassen. Ein einfaches Web-Content-Management-System unterstützt daher nur die Ablage strukturierter Daten in Datenbanksystemen, aber weniger die Verwaltung multimedialer Dokumente.

Web-Content-Management-Systeme mutieren daher oft zu datenbankgestützten Autorenwerkzeugen für Web-Anwendungen – was nur ein kleiner Teilaspekt des großen, ursprünglichen Gedankens von Content Management ist.

In diesem Schwerpunktheft wollen wir diese Schiefelage korrigieren. Zwar wird in einem Beitrag die Auswahl eines herkömmlichen Web-Content-Management-Systems behandelt, in weiteren Beiträgen werden aber sowohl das komplexe Problem des Enterprise Content Managements aus Herstellersicht als auch das Problem digitaler Bibliotheken als Anwendung von Content Management aus Anwendersicht beschrieben.

Phasen des Content Managements

Betrachtet man die einzelnen Arbeitsschritte bei der Verwaltung von Content genauer, so wird klar, dass es zwei Ebenen von zu unterstützenden Prozessen gibt:

- Die erste Ebene unterstützt das Zusammentragen und Gewinnen von Content sowie die redaktionelle Arbeit mit dem Content. Dazu gehören das Erzeugen von Dokumenten, das Sammeln und Integrieren auch fremder Dokumente, die Bearbeitung von Dokumenten, der Review und die Freigabe von Dokumenten sowie die Publikation der freigegebenen Dokumente.
- Die zweite Ebene unterstützt die Verwaltung, das Wiederfinden und die

Archivierung des Contents. Hier setzt die übliche Datenbankfunktionalität an. Besser: Die Datenbankfunktionalität muss hier von strukturierten Daten nun auch auf komplexe multimediale Dokumente erweitert werden.

Unterstützende Prozesse für diese zwei Ebenen sind einerseits das Workflow-Management, mit dem sich beispielsweise die Redaktionsprozesse steuern lassen, und weitere administrative Prozesse wie die Verwaltung von Nutzern und Personalisierungsinformationen.

Teilt man die erste Ebene noch einmal auf in das Sammeln von Content einerseits und das redaktionelle Bearbeiten von Content andererseits, so kann man insgesamt drei Arten von Content-Management-Systemen identifizieren:

- *Integrationssysteme* sammeln und integrieren Content und stellen diesen einheitlich zur Verfügung. Diese Funktionalität findet man oft in Portalsystemen.
- *Redaktionssysteme* unterstützen die redaktionelle Bearbeitung des Contents und die Ablage strukturierter Informationen über die Dokumente, die so genannten Metadaten, in einem Datenbanksystem.
- *Repository-Systeme* unterstützen die Verwaltung, das Wiederauffinden, die Verteilung und die langfristige Archivierung von Content. Dabei können strukturierte Informationen in einem Datenbanksystem gehalten werden, für die multimedialen Dokumente benötigt man entweder ein Multimediale Datenbanksystem oder ein spezielles Dokumenten-Verwaltungssystem.

Die Beiträge dieses Schwerpunkthefts werden sich auf die Redaktionssysteme und Repository-Systeme konzentrieren. Im Beitrag über das Open-Source-System MyCoRe zur Verwaltung digitaler Bibliotheken etwa wird auf einem kommerziellen Repository-System ein an Belange von Universitätsbibliotheken adaptiertes Redaktionssystem aufgebaut.

Was ist eine digitale Bibliothek?

Eine digitale Bibliothek ist

- eine Sammlung von Dokumenten mit bleibendem Wert,
- deren Dokumente mit Metadaten beschrieben (erschlossen) werden,

- die langfristig zitierbar bleiben müssen,
- von denen es verschiedene Versionen geben kann, wobei eine spezielle Version aber nach der Publikation nicht mehr geändert werden kann,
- die verkauft und gekauft werden können (Dokumente kann man also »besitzen«).

Einige der Aspekte einer digitalen Bibliothek stellen also höhere Anforderungen an ein Content-Management-System. So legen Web-Content-Management-Systeme, die zur Darstellung von kurzfristig gültigem Web Content ausgelegt sind, nicht unbedingt Wert auf die langfristige und unveränderte Zitierbarkeit bereitgestellter Dokumente. Im Web ist vieles flüchtig, die Dokumente werden oft nur aus Datenbanken dynamisch generiert und sind am nächsten Tag schon nicht mehr (in der alten Form) verfügbar. Ein System für digitale Bibliotheken muss nun diesen Anspruch zusätzlich erfüllen.

Was ist ein digitales Bibliothekssystem?

Was ist nun ein System zur Verwaltung von digitalen Bibliotheken? Ein unglücklicher deutscher Begriff wäre digitales Bibliothekssystem, da er eine falsche Betonung suggeriert und damit auch einen falschen Zusammenhang herstellt. Ein solches System ist zunächst einmal nicht nur ein System für die organisatorische Einheit »Bibliothek«, sondern ein System zur Verwaltung digitaler Dokumente, die obige Anforderungen haben.

Unter den digitalen Dokumenten »mit Wert« kann man dann die wissenschaftlichen Artikel einer Universitätsbibliothek,

ebenso aber auch die digitalen Audio- und Videodateien eines privaten Musik- oder Film-Freundes verstehen.

Ein digitales Bibliothekssystem unterstützt alle Prozesse auf einer digitalen Bibliothek, die auch das Content-Management-System anbietet, nur mit einer anderen Gewichtung, da Archivierung etwa ein sehr wichtiger Teilprozess wird. Somit kann man sagen, dass ein Content-Management-System auch digitale Bibliotheken verwalten kann.

Überblick über die Beiträge dieses Heftes

In diesem Heft erscheinen drei Beiträge zum Schwerpunktthema Content Management und digitale Bibliotheken:

- James A. Reimer (IBM, San Jose) beschreibt das Thema »Enterprise Content Management« aus einer allgemeinen Sicht. Als Herausforderungen sieht er sowohl die Seite der Redaktionssysteme als auch der Repository-Systeme. Der Artikel beschreibt zwar die Herstellersicht, bleibt aber herstellerunabhängig.
- Frank Lützenkirchen (Universität Essen) schildert die Umsetzung eines Systems für digitale Bibliotheken: »MyCoRe – Ein Open-Source-System zum Aufbau digitaler Bibliotheken«. Ein Verbund von Universitäten nutzt ein kommerzielles Repository-System, um die speziellen Redaktionsprozesse einer Universitätsbibliothek darauf abzubilden.
- Marlis Brunk und Heiko Armin Schneider (T-Systems, Berlin) beschreiben die Auswahl eines Web-Content-Management-Systems, also

eines klassischen Redaktionssystems für die Verwaltung von Web-Auftritten. Dabei werden weniger kommerzielle Produkte in den Vordergrund gestellt, sondern die Art und Weise des Auswahlprozesses bei einem großen Anwender.

Die drei Beiträge decken schon ein großes Spektrum im Themengebiet Content Management und digitale Bibliotheken ab. Wir hoffen aber, dass in Zukunft noch weitere Beiträge im Datenbank-Spektrum erscheinen werden, etwa

- zum Forschungsgebiet »Digitale Bibliotheken«, auf dem es gerade in Deutschland einige erfolgreiche Projekte im Bereich der Grundlagenforschung gibt,
- zu konkreten Redaktionssystemen und Repository-Systemen, sowohl aus Herstellersicht als auch aus Anwendersicht,
- zu konkreten (weiteren) Anwendungen außerhalb klassischer Bibliotheken.

Wir wünschen dem Leser viel Spaß bei der Lektüre der Schwerpunktbeiträge.

Prof. Dr. Andreas Heuer
Lehrstuhl Datenbank- und Informationssysteme
Fachbereich Informatik
Universität Rostock
18051 Rostock
heuer@informatik.uni-rostock.de
<http://www.db.informatik.uni-rostock.de/~ah>

Prof. Dr. Klaus R. Dittrich
Database Technology Research Group
Institut für Informatik
Universität Zürich
Winterthurerstr. 190
CH-8057 Zürich
dittrich@ifi.unizh.ch
<http://www.ifi.unizh.ch>