

Content Management Systeme in der MPG

31.10.05 Peter Wittenburg

Die Verwaltung der zunehmenden Datenmengen und deren Verfügbarmachung stellt in der Wissenschaft eine der großen Herausforderungen dar. Abhängig von der Zielsetzung und dem zugrundeliegenden Datenmodell wird man verschiedene Arten von Datenverwaltungssystemen einsetzen. Traditionell bekannt sind Datei-Systeme und relationale Datenbanken mit ganz unterschiedlichen Vorzügen. In den letzten Jahren sind verstärkt Content Management Systeme (CMS) auf den Markt gekommen; sie haben auch bereits in der MPG in mehreren Instituten und der GV Einzug gehalten. Eine besondere Note bekommen die CMS dadurch, daß sie oft mit dem Web-Auftritt von Instituten wie auch der MPG und damit deren öffentlichem Erscheinungsbild in Verbindung stehen.

Wegen der unterschiedlichen Anforderungen bei den verschiedenen Einsatzgebieten, für die MPG insgesamt und für die einzelnen Institute, werden auch in Zukunft in der MPG verschiedene CMS zum Einsatz kommen. Ein Wildwuchs sollte jedoch vermieden werden. Um die Auswahl zu erleichtern, stellt der BAR hier eine kurze Einschätzung dieser Systeme für den Bereich der MPG vor und empfiehlt die folgenden Grundsätze:

- CMS sind Software Werkzeuge wie andere auch, die in der Wissenschaft sinnvoll für den Web-Auftritt aber auch vermehrt für das Management wissenschaftlicher Daten eingesetzt werden können.
- Abhängig von den intendierten Aufgaben wird man ein spezielles CMS auswählen. Benötigt ein MPI ein Werkzeug, um lediglich den Webauftritt zu gestalten, so reichen bereits kostengünstige Lösungen aus, die mit professionellen Benutzer-Schnittstellen ausgerüstet sind und durch IT Laien betreut werden können. Wählt man ein komplexeres und flexibleres System aus, um höheren Ansprüchen zu genügen bzw. disziplin-spezifische Eigenschaften zu erhalten, so ist im allgemeinen von einem höheren Betreuungsaufwand auszugehen und es muss auch höher qualifiziertes Personal eingesetzt werden.
- Es werden in der MPG verschiedene CMS zum Einsatz kommen, da die Zielsetzungen sehr unterschiedlich sind. Selbst innerhalb eines MPI können mehrere CMS zum Einsatz kommen, da sich die Personen, die z.B. die reine Webdarstellung pflegen, sehr stark von den Spezialisten unterscheiden, die wissenschaftliche Resultate und Ressourcen über das Web zur Verfügung stellen und mit Fachkollegen austauschen.
- Gleichwohl sollte man bei der Auswahl von CMS auf Homogenisierung achten und die in der MPG vorhandene Expertise nutzen und bündeln.
- Der BAR begrüßt und unterstützt daher die Initiative einiger Institute, ZOPE/PLONE als Basis zu wählen, um darauf aufbauend diverse Module zu erzeugen, die von breiterem Interesse sind, und Wissens-Netzwerke zu bilden, die sich gegenseitig unterstützen.
- Da die Kosten für die reine Beschaffung zumeist weit unter der 50.000 € Grenze liegen, geht der BAR im Allgemeinen davon aus, dass die Beschaffung derartiger Systeme von den Instituten finanziert werden. Wenn auch der BAR weiterhin gemeinsame Vorhaben mehrerer MPIe unterstützen wird, so muss doch darauf hingewiesen werden, daß die MPIe in jedem Falle die Folgekosten tragen müssen.

Erläuterungen:

Funktionalität

CMS gehen anders als z.B. relationale Datenbanken nicht von einem eng abgegrenzten Datenmodell aus, sondern können mit Daten unterschiedlicher Struktur umgehen. Sie haben Funktionen, die sie auch für schwach-strukturierte Daten wie z.B. textuelle Dokumente und Web-Seiten nutzbar machen. CMS sind Container, die eine Offenheit bezüglich der zu speichernden Dokument-Typen haben. Sie bieten ebenfalls eine Reihe von Basis-Funktionalitäten an wie z.B. zur Katalogisierung der Dokumente, zum Suchen sowohl über die Kataloge als auch über die Inhalte, wobei verschiedene Standard-Formate wie relationale Datenbanken, WORD, PDF und XML bei der Generierung der Indexe berücksichtigt werden, und zur Darstellung dieser Dokumente. Insbesondere bezüglich der freien Inhaltssuche, die zumeist auf statistischen Retrieval-Techniken basiert, bieten einige der CMS ausgezeichnete Systeme an. Dies ist der Bereich, in dem die CMS im Allgemeinen mit großen Vorteilen einherkommen.

Viele der CMS beinhalten spezielle Funktionalitäten, um insbesondere Web-Seiten einfach generieren und verwalten zu können. Einfache Umgebungen machen es möglich, daß der Web-

Seiten-Gestalter sich nicht mehr um das Layout zu kümmern braucht, sondern lediglich seine Texte bzw. Menü-Optionen eingeben muß. Damit kann auch das Sekretariat bestimmte Teile des Web-Auftritts verwalten. Bei den besseren Systemen werden darüberhinaus Programmier-Umgebungen angeboten, die es z.B. erlauben, Daten aus wissenschaftlichen Tabellen in Web-Seiten einzublenden. Ebenso sind die besseren CMS multimedial ausgerichtet, d.h. sie haben Komponenten, die es ihnen erlauben, Bilder, Audio und Video Dokumente zu integrieren und zu präsentieren.

Einsatzgebiete

Es ist daher naheliegend, daß Institute und vor allem auch das Pressereferat der GV für die Öffentlichkeitsarbeit mit CMS ausgestattet sind, wobei der Grad an erforderlicher Professionalität vom Umfeld und den Anforderungen abhängig ist. Allerdings können insbesondere in den Geisteswissenschaften, deren Forschung ja zu einem wesentlichen Teil auf schwach-strukturierten Dokumenten und multimedialen Objekten basiert, CMS auch wissenschaftsnahe Aufgaben übernehmen. Hier gilt es, die wissenschaftlichen Anforderungen mit der gebotenen Funktionalität zu vergleichen. Generell gilt, daß die CMS eine Basis-Funktionalität anbieten und daß die spezielle Anwendungslogik selbst entwickelt werden muß. Hierbei muß die jeweils angebotene Programmierumgebung nüchtern mit den bekannten offenen Umgebungen, wie sie z.B. durch Java gegeben sind, verglichen werden. CMS könnten sich hier sogar sehr schnell als ein ernstes Hindernis erweisen.

Während die meisten CMS gegenwärtig keine spezielle Disziplin-bezogene Anwendungslogik aufweisen, kann davon ausgegangen werden, daß sich das in den kommenden Jahren ändern wird. Die Verfügbarkeit spezieller Komponenten wird die Attraktivität der CMS noch erhöhen. Hinzu kommt der wichtige und von anderer Software her bekannte Aspekt, daß sich bei disziplinbezogenen Lösungen ein Kenntnis-Netzwerk in der jeweiligen Disziplin herausbildet, von dem alle Teilnehmer in verschiedener Hinsicht profitieren können. Wie sich diese Software-Systeme im Detail entwickeln werden, ist offen.

Offenheit

Eine Frage von zunehmender Bedeutung in der Wissenschaft ist die nach der Offenheit von Systemen. Damit werden vor allem drei Aspekte angesprochen: (1) Offenheit der Programmierung und des Codes, (2) Offenheit der Format-Schnittstellen und (3) Offenheit der Daten. Von primärer Bedeutung ist die Verfügbarkeit der Daten in offenen Standard-Formaten wie z.B. XML, d.h. Institute werden es zumeist nicht akzeptieren, wenn offene Dokumente beim Einlesen in proprietäre Binärformate umgewandelt und gekapselt werden. Die entstandene Abhängigkeit von einem Produkt wäre nicht akzeptabel. Wichtig ist auch die Verfügbarkeit von Schnittstellen, so daß man z.B. interne Tabellen als XML Dateien ausgeben kann. Ebenfalls ist es wichtig, flexible Import/Export-Schnittstellen zur Verfügung zu haben, um in einfacher Weise Daten und Dokumente austauschen zu können. Dies ist insbesondere eine der wichtigen Funktionen für das CMS des Pressereferat der GV.

Immer mehr in den Fokus der Diskussion in der Wissenschaft kommen CMS, deren Source Code offenliegt. Dies hat vor allem zwei Gründe: (1) Man möchte andere zum Entwickeln von Modulen motivieren und von der Entwicklung anderer profitieren, ohne langfristig Mittel größeren Umfangs binden zu wollen und sich in Abhängigkeiten von Firmen zu begeben. (2) Man möchte auf bestehende Basis-Funktionalität aufbauen und seine eigene wissenschaftsbezogene Anwendungslogik in flexibler Weise hinzufügen können. Mit dem Open Source Anwendungs-Server ZOPE und dem darauf aufbauenden CMS Modul PLONE steht eine exzellente und weit-verbreitete Umgebung bereit, um diesen Ansprüchen gerecht zu werden. Um die Flexibilität von Open Source Lösungen voll nutzen zu können, ist ein gewisser Programmieraufwand erforderlich.

Kosten

Für die Verwaltung von Web-Auftritten bekommt man heute schon ausgezeichnete kommerzielle Systeme für weniger als 5.000 €. Für professionelle Systeme, die mit verschiedenen zusätzlichen Möglichkeiten ausgestattet sind, muß man entsprechend mehr Geld bezahlen. Die größten Ausgaben bei der Gestaltung zum Beispiel eines Web-Auftritts liegen jedoch im Anfertigen von Layouts, im Editieren der Inhalte und bei der Schulung von Mitarbeitern. Auch die Einführung eines Open Source Tools wie ZOPE/PLONE bringt Kosten mit sich. Dies gilt insbesondere dann, wenn man IT-Personal bereitstellen will, das mit den Internas eines solchen flexiblen Tools vertraut sein soll.