



itechWorks
the web professionals

cmsWorks

Content Import / Export

Juni 2010

Übersicht:

1. itechWorks – Partner der Agenturen.....	1
2. cmsWorks - Content Management	1
3. Content Import / Export mit cmsWorks	1
4. Beispielhafte Projekte.....	2
5. Jede Standardlösung ist einzigartig	5

1. itechWorks - Partner der Agenturen

itechWorks steht für Software und Service seit 1996 - Profitieren Sie von unserer Erfahrung bei Konzeption, Aufbau und Weiterentwicklung großer Internetportale wie sport1.de, spox.com und fussball.de.

Beim Aufsetzen der Datenstrukturen und der Entwicklung der Selektionskriterien für In- und Export konnten sich die Mitarbeiter von itechWorks im Laufe von 10 Jahren mit drei großen Content Management Systemen (CoreMedia, FirstSpirit, cmsWorks) einzigartige Erfahrung aneignen. Dieses Know-How stellen wir in der Konzeptions- und Implementierungsphase auch unseren Partnern gerne zur Verfügung.

2. cmsWorks - Content Management

cmsWorks ist ein von itechWorks entwickeltes Web Content Management System, das inzwischen in der Version 3.1 zur Verfügung steht. Die gesammelten Erfahrungen aus über acht Jahren, in denen sich itechWorks mit High Performance Content Solutions beschäftigt, sind in die Entwicklung eingeflossen.

cmsWorks ist besonders gut für reichweitenstarke Newsportale und internationale Unternehmensseiten geeignet, denn es ist schnell, bietet sehr flexible Strukturen und unterstützt Redaktion und Autoren mit einer intuitiven, webbasierten Oberfläche. Die technische Basis ist hoch verfügbar, modular erweiterbar und skaliert nahezu linear.

3. Content Import / Export mit cmsWorks

Jedes Content Management System (CMS) ermöglicht die Erfassung, strukturierte Ablage, Verarbeitung und Ausgabe von Daten. Die häufigste konkrete Ausprägung ist die manuelle Erfassung von Texten und multimedialen Inhalten durch eine Redaktion und die Ausgabe auf eine Webseite. Ebenso wichtig ist aber der automatisierte Import von Daten in das System und der gesteuerte Export an verschiedene Ziele in verschiedenen Formaten.

cmsWorks nutzt als Basis einen Application Server. Auf dieser serviceorientierten Architektur (SOA) können neben dem CMS beliebig viele und beliebig komplexe Import- und Export-Dienste aufgesetzt werden. Diese Systematik ist sicher, flexibel und effizient.

Das Standardvorgehen zur Implementierung eines Export-Dienstes:

- cmsWorks Generator erzeugt XML (oder ein anderes Ausgabeformat)
- ein Application Server Dienst sichert die Auslieferung

Das Standardvorgehen zur Implementierung eines Import-Dienstes:

- Ein Application Server Collector-Service nimmt Daten an (oder holt sie ab)
- cmsWorks nimmt die geprüften und aufbereiteten Daten in den Dokumentenpool auf

Abhängig von Aufgabenstellung werden die Import/Export-Daten nicht zwischengespeichert, sondern direkt weitergeleitet. In der Datenbank werden dann nur Konfigurationsdaten (pro Kunde/Quelle) und Logdaten (über alle Aktionen und deren Erfolg) abgelegt.

Es gibt zwei mögliche Ansätze, die Exportvorgänge auszulösen:

- Zeitgesteuert (z.B. alle Minute, einmal pro Stunde)
- Ereignisgesteuert (z.B. bei Erstellung, Änderung, Löschen, Freigabe oder Publikation einer Ressource)

In vielen Fällen kann alternativ zur Java-Programmierung cmsWorks ContentScript für Ein- und Ausgabeoperationen eingesetzt werden.

4. Beispielhafte Projekte

Content Syndication für mobile Endgeräte

Aufgabenstellung: Ein großes Sportportal bietet seine Inhalte nicht nur über die Webseite an, sondern auch über mobile Endgeräte. Eine Auswahl von Nachrichten werden an einen externen Dienstleister übergeben, der diese für eine Vielzahl gängiger mobiler Endgeräte umformatiert und bereitstellt.

Lösung:

Die Redaktion kann komfortabel über Standard-Verteilseiten festlegen, welche Rubriken und welche Inhalte syndiziert werden, und diese nach Bedarf umstellen (per drag-and-drop). Alles weitere erfolgt automatisch beim Anlegen von neuen Artikeln – auch dank eines leistungsfähigen Dokumentenmodells. So, kann beispielsweise in jedem einzelnen Dokument (z.B. Artikel oder Bild), die Syndizierung unterdrückt werden (z.B. wenn die entsprechenden Lizenzen fehlen). Soweit notwendig können den Artikeln zusätzliche Bildformate und ein so genannter Syndizierungs-Teaser vorangestellt werden, der einen von Markennamen gesäuberten Einführungstext enthält.

Ein Generator stellt dem externen Dienstleister XMLs für Index, referenzierte Dokumente (also weitere Verteilseiten und letztlich Artikel) und eingehängte Bilder zur Verfügung.

Entscheidend: Die Syndizierung der Inhalte erfolgt ohne nennenswerten Mehraufwand.



RSS Distribution

Aufgabenstellung: Ein Portal will seine Inhalte in Form von mehreren RSS-Feeds zur Verfügung stellen. Die Feeds sollen verschiedene Themen abdecken (z.B. Fußball-Bundesliga) und in verschiedenen Formaten angeboten werden (z.B. Media RSS)

Lösung:



Der Administrator kann komfortabel über die Oberfläche und in einfachen Config-Files festlegen, welche Kategorien und welche Themen als RSS-Feeds angeboten werden. Diese werden dann über einen XML-Generator zur Verfügung gestellt.

Für eine optimale Auslieferung des RSS-Daten, wurde dem RSS-Generator ein Caching-Dienst vorangestellt, der regelmäßig die RSS-Feeds abrufen und dann statisch anbieten.

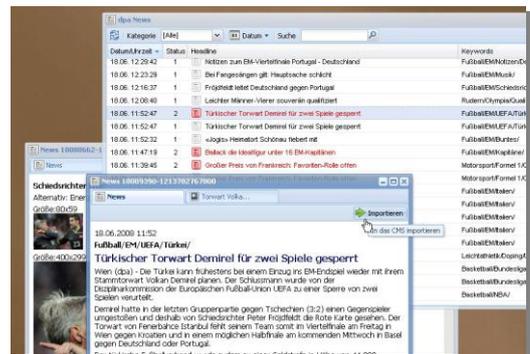
Presseagentur Tool: Selektion und Import aus externen Quellen

Aufgabenstellung: Ein Familienportal nutzt Agenturmeldungen (hier DPA) um sein redaktionelles Angebot zu ergänzen. Beiträge und anhängende Bilder sollen über eine integrierte Anwendung einfach in neue Artikel überführt werden.

Lösung:

Das „Presseagentur-Tool“ ist eine optionale Erweiterung für den cmsWorks Desktop. Es bietet eine komfortable Oberfläche für einen mehrstufigen Import und Selektionsprozess von Tickermeldungen, Tabellen und Bildmaterial aus Ticker-Quellen wie DPA oder SID.

- Die Presseagentur liefert die Daten (Artikel, Tabellen, Bilder) in XML-Format an.
- Der zwischengespeicherte Datenbestand wird abhängig von der Angebotsstrategie analysiert und für die weiteren Schritte aufbereitet.
- Ein Such-Service indexiert die eingegangenen Daten für eine Volltextsuche
- Über das Tool, das nahtlos in den cmsWorks Desktop integriert ist, können nun die eingegangenen Nachrichten ausgewählt und eingesehen werden. Verschiedene Filter- und Suchfunktionen helfen dem Redakteur das Richtige schnell zu finden.
- Mit nur einem Knopfdruck können die ausgewählten Inhalte in das CMS importiert werden. Dabei werden die benötigten Artikel, Medien- und Referenz-Dokumente angelegt, und die Inhalte gleich in die jeweiligen Felder übernommen. Ein Fenster bleibt offen, damit – soweit gewünscht – eine Nachbearbeitung erfolgen kann



- Mit einem weiteren Klick kann der Redakteur die Meldung sofort live stellen. Sie erscheint dann auch gleich in den Top-News Bausteinen der Webseite. Im Hintergrund laufen zusätzliche Archivierungsprozesse ab.

Auf diesem Weg ist die Übernahme und Aufbereitung von Agenturmeldungen mit minimalem Aufwand in wenigen Sekunden möglich. Das gibt der Redaktion einen entscheidenden Zeitvorteil bei gleichzeitiger Arbeitsentlastung. Das Tool ist leicht an neue Datenquellen anpassbar.

Intelligente, simultane Aktualisierung mehrerer Websites

Aufgabenstellung: Ein großes Sportportal ging mit einer großen Jugendzeitschrift eine Kooperation für ein Jugendsportportal ein. Teil dieser Kooperation sollte die Anlieferung top-aktueller Sport-News und Bilder sein, sodass die Sportredaktion des Jugendmagazins wenig Arbeit hat und sich auf zielgruppenspezifischen Specials konzentrieren kann. Entscheidend ist, dass die ausgetauschten Inhalte nahtlos in das Zielsystem eingefügt werden und dort entweder gleich auf die Seiten verteilt werden (mit korrektem Ranking) oder zuvor noch von der Jugendredaktion überarbeitet werden können.

Lösung:

Das Jugendportal wurde im Zuge des Relaunches mit cmsWorks ausgestattet, was für die folgende Content-Verknüpfung vorteilhaft, jedoch nicht zwingend erforderlich war. Hierzu wurde nämlich die cmsWorks Content Bridge genutzt, welcher einem erweiterten XML-Import/Export entspricht. Der Mechanismus besteht aus folgenden Komponenten

- Ein XML-Exporter im Quellsystem, der unmittelbar nach „online-stellen“ eines neuen oder aktualisierten Dokuments (Publikationsevent) aktiv wird und dieses als XML verschickt. Das geschieht nur für Inhalte und alle verknüpften Elemente, die aus bestimmten thematischen Kategorien stammen und von den Redakteuren als „exportierbar“ gekennzeichnet werden.
- Ein XML-Importer im Zielsystem, der ...
 - Dokumenten-IDs des Quell- und des Zielsystems verwaltet und so eindeutige Zuordnungen (auch von knüpften Elementen) ermöglicht,
 - Dokumententypen des Quell-Systems erkennt, die Daten für das Dokumentenmodell des Ziel-Systems übersetzt (von einzelnen Artikeln bis zu ganzen Bildergalerien) und gegebenenfalls schon vorhandene Dokumente ersetzt,
 - Eingebettete Bilder in die Formate des Zielsystems umformatiert,
 - Ranking-Informationen in Dokumentendaten umwandelt und das Ranking selbst speichert
- Dynamische Elemente auf den Zielseiten, die frisch importierte Daten abhängig von der Konfiguration gleich an den vorgesehenen Stellen (und in der richtigen Ranking-Reihenfolge) auf die Zielseite stellen.

Auch hier ist eine ausgeklügelte Erfolgskontrolle wieder ein wichtiger Faktor. Zudem wurde sichergestellt, dass die importierten Inhalte die interne referenzielle Integrität des Zielsystems nicht gefährden. Denn Teil der cmsWorks Philosophie ist, dass „tote Links“, fehlende Bilder oder gar fehlende Seiten (404) nicht akzeptabel sind.

5. Jede Standardlösung ist einzigartig

Jeder Kunde, jede Kombination aus Quell- und Zielsystemen ist einzigartig. Anforderungen verändern sich mit den Geschäftsmodellen und den technischen Möglichkeiten. Selbst Standards wie XML sind immer nur ein gemeinsamer Nenner, der auf die konkreten Anforderungen angepasst werden muss. cmsWorks ist dafür besonders gut geeignet, da es selbst auf Standards aufsetzt, offen strukturiert ist und über eine gut dokumentierte API verfügt.

Kontakt:

itechWorks, Schwanthalerstraße 3, 80336 München

Ihr Ansprechpartner: Jürgen Melchhammer, 089/ 5486279-0, info@itechworks.de