

Sławomir NIEDBAŁA  
Politechnika Śląska, Instytut Informatyki

## ZASTOSOWANIE SYSTEMÓW ZARZĄDZANIA TREŚCIĄ W SERWISACH INTERNETOWYCH

**Streszczenie.** Utrzymanie rozbudowanego serwisu internetowego w porządku i dbanie o aktualność prezentowanych informacji należą do bardzo wymagających zadań. Ich wykonanie byłoby praktycznie niemożliwe, gdyby nie wsparcie oferowane przez specjalistyczne narzędzia określane mianem systemów zarządzania treścią. W artykule przeanalizowano podstawowe założenia istniejących systemów oraz przedstawiono autorskie rozwiązanie.

**Słowa kluczowe:** zarządzanie treścią, serwisy internetowe.

## APPLICATION OF CONTENT MANAGEMENT SYSTEMS IN INTERNET SERVICES

**Summary.** Maintenance and keeping up-to-date of complex Web sites are very difficult tasks. Successful execution of both tasks is impracticable without support by so-called content management systems. In this paper, the structure of those systems is discussed and one of them is presented.

**Keywords:** content management, Web sites.

### 1. Wstęp

Utrzymanie rozbudowanego serwisu internetowego w porządku i dbanie o aktualność prezentowanych informacji należą do bardzo wymagających zadań. Ich wykonanie byłoby praktycznie niemożliwe, gdyby nie wsparcie oferowane przez specjalistyczne narzędzia określane mianem systemów zarządzania treścią. Na rynku istnieje kilka uniwersalnych produktów oferujących bardzo duże możliwości konfiguracyjne i administracyjne, jednak rozwiązania te są bardzo drogie i przez to nieosiągalne dla wielu grup klientów. W efekcie

powstaje wiele tańszych systemów specjalizowanych pod konkretne architektury sprzętowe, ale nie ustępujących wiele droższym odmianom pod względem funkcjonalnym. W artykule omówiono podstawowe założenia istniejących systemów oraz przedstawiono utworzony system EIP.net.

## 2. Zarządzanie treścią

Zarządzanie treścią (ang. content management) to proces zbierania, zarządzania i udostępniania informacji w docelowych publikacjach. Dostarczona informacja jest przechowywana w oryginalnej postaci lub też konwertowana do określonego formatu. Następnie przechowywane informacje są przetwarzane i na końcu publikowane. Zakres typów tworzonych publikacji zależy od specyficznych wymagań firmy, jednak najczęściej spotykaną jest prezentacja na stronach internetowych lub intranetowych.

Problem zarządzania nie należy do łatwych zadań, gdyż jest mocno uzależniony od struktury i wymagań określonej firmy bądź instytucji. Sposób realizacji rozwiązania zależy w głównej mierze od metody uzyskiwania informacji i formatu dostarczanych danych, a także od przyjętego modelu sterowania. Proces zarządzania treścią bardzo często musi być zintegrowany z procesem zarządzania dokumentami, gdyż istotnym źródłem danych są istniejące informacje przechowywane w różnego typu dokumentach.

W przypadku małych ilości przetwarzanej informacji zdarza się, że proces zarządzania treścią odbywa się w tradycyjny sposób, tzn. przez samodzielną obróbkę dokumentów i tworzenie publikacji. Jednak wraz ze wzrostem ilości przetwarzanych dokumentów rozwiązanie to staje się wyraźnie nieekonomiczne. Poważnym problemem zaczyna być umiejętna selekcja potrzebnych informacji oraz niedeterministyczny czas tworzenia publikacji. W efekcie pojawiają się opóźnienia, które mogą się jeszcze bardziej kumulować, gdy publikacja jest zatwierdzana przez inną osobę. W przypadku negatywnej oceny wprowadzenie ewentualnych poprawek wymaga powtórzenia całego procesu tworzenia publikacji.

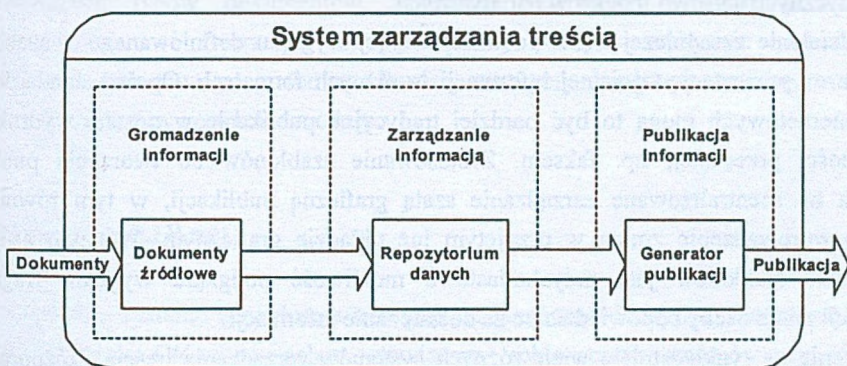
## 3. Systemy zarządzania treścią

Systemy zarządzania treścią (ang. content management system) umożliwiają rozwiązanie problemu efektywnej selekcji i przetwarzania posiadanych informacji. Pojęciem tym określa się oprogramowanie, które umożliwia zorganizowanie całego procesu zarządzania treścią i zautomatyzowanie poszczególnych jego faz. Ogólny model systemów CMS przedstawia



rys. 1. Trzy podstawowe zadania wyszczególnione na rysunku to gromadzenie informacji, zarządzanie nią i tworzenie publikacji. Każde z zadań jest realizowane przez osobny podsystem.

Moduł gromadzenia informacji odpowiada za kolekcjonowanie informacji pochodzących z różnych źródeł. Zebrane dane są rozkładane na składowe i następnie przesyłane do tzw. repozytorium danych. Przechowywane w repozytorium informacje dzielą się na zasadniczą treść dokumentów oraz na dodatkowe parametry umożliwiające np. wyszukiwanie, określenie daty ważności informacji, czasu planowanego zarchiwizowania czy też osoby odpowiedzialnej za zatwierdzenie fragmentu publikacji.



Rys. 1. Ogólny model systemu zarządzania treścią

Fig. 1. General model of content management system

Kolejnym elementem jest moduł zarządzania, który odpowiada za odpowiednie rozmieszczanie przechowywanych danych w repozytorium, przetwarzanie istniejących informacji oraz za udostępnianie wymaganych danych w czasie tworzenia publikacji. W zakres prac określanych mianem przetwarzania danych wchodzi generowanie list dokumentów i indeksów, automatyczne uaktualnianie zależności pomiędzy dokumentami, sprawdzanie terminów ważności i dopuszczanie do publikacji.

Ostatnim modulem systemu jest podsystem publikacji informacji. Głównym zadaniem tej części systemu jest automatyczne tworzenie publikacji zgodnie z wcześniej zdefiniowanymi szablonami, wykorzystując treści dostarczone przez moduł zarządzania informacją. Dodatkowo moduł publikacji odpowiada za zarządzanie szablonami, które są tworzone przez osoby odpowiedzialne za graficzny kształt tworzonych publikacji.

Przyjęcie przedstawionej architektury systemów zarządzania treścią umożliwia dokonanie podziału obowiązków wśród osób zaangażowanych w proces publikacji. Redaktorzy odpowiedzialni za dostarczenie treści samodzielnie umieszczają odpowiednie dokumenty w systemie. Dzięki możliwości podglądu mogą natychmiast obejrzyć wynik swojej pracy i ewentualnie wprowadzić żądane korekty. Osoby odpowiedzialne za wygląd publikacji nie muszą się zajmować konwertowaniem i wprowadzaniem treści do systemu. Ich obowiązkiem

jest tylko przygotowanie odpowiednich szablonów publikacji. Kadra kierownicza może natychmiast po wprowadzeniu informacji do systemu obejrzeć efekt i dokonać oceny. Proces tworzenia publikacji jest w pełni zautomatyzowany, a każdy z pracowników ma wyraźnie określony zakres obowiązków. Wprowadzenie automatycznego zarządzania treścią zapewnia, że w publikacji nigdy nie pojawi się przypadkowy, przez nikogo nie zatwierdzony dokument albo dokument, który jest już nieaktualny. Mechanizm dba również o spójność informacji przekazywanych do publikacji, weryfikuje powiązania pomiędzy elementami publikacji, eliminując odwołania do nieistniejących fragmentów oraz umożliwia generowanie automatycznych spisów i indeksów tematycznych.

Oddzielenie zasadniczej treści publikacji od jej wyglądu definiowanego w szablonach pozwala na prezentację tej samej informacji w różnych formatach. Oprócz standardowych stron internetowych mogą to być bardziej tradycyjne publikacje w postaci wydruku lub wiadomości przesłanej, np. faksem. Zastosowanie szablonów do tworzenia publikacji pozwala na scentralizowane zarządzanie szatą graficzną publikacji, w tym również na szybkie wprowadzenie zmian w przyjętym już układzie graficznym. Ważnym aspektem stosowania szablonów jest natychmiastowa możliwość podglądu wyglądu fragmentu publikacji przez osoby odpowiedzialne za dostarczanie informacji.

Obecnie na rynku istnieje wiele różnych systemów zarządzania treścią. Różnorodność oferowanych rozwiązań wynika z faktu, że różne firmy tworzące systemy CMS odmiennie interpretują termin „zarządzanie treścią”. Innym elementem wpływającym na mnogość stworzonych rozwiązań jest fakt, że każda instytucja posiada inną strukturę i stawia inne wymagania systemom zarządzania treścią. Dlatego systemy CMS różnią się między sobą pod względem ilości i rodzaju obsługiwanych formatów wejściowych i typów publikacji, architektury wewnętrznej systemu oraz wymaganiami sprzętowymi i programowymi. Ta różnorodność produktów powoduje, że nie ma jednoznacznej odpowiedzi, które z rozwiązań jest najlepsze.

#### **4. Zarządzanie treścią serwisów WWW**

Duża popularność serwisów internetowych spowodowała, że praktycznie każdy system zarządzania treścią oferuje możliwość publikacji dokumentów na stronach WWW, a duża ich ilość została stworzona wyłącznie w tym celu. Oferowane rozwiązania umożliwiają tworzenie mniej lub bardziej rozbudowanych portali lub wielojęzycznych serwisów korporacyjnych. Zakres obsługiwanych przez system formatów dokumentów źródłowych w każdym przypadku jest inny, ale zazwyczaj informacje dostarczane są za pomocą specjalnych programów lub w postaci plików HTML, rzadziej jako dokumenty Microsoft



Office lub StarOffice. Oprócz dokumentów tekstowych systemy te oferują również obsługę plików graficznych, muzycznych i audiowizualnych. Jak już wcześniej wspomniano, publikacje mogą nie ograniczać się tylko do stron prezentowanych w serwisach internetowych, ale mogą to być np.: dokumenty PDF, drukowane katalogi.

Bardzo często zdarza się jednak, że oferowane możliwości uniwersalnych systemów zarządzania treścią okazują się nadmiarowe, zaś cena produktu nie do zaakceptowania. Dlatego dużą popularnością cieszą się tańsze rozwiązania o uproszczonej architekturze i nieco mniejszych możliwościach. Uproszczenia uzyskuje się przez ograniczenie zakresu obsługiwanych typów dokumentów, mniejsze możliwości integracji z istniejącym oprogramowaniem i tworzenie rozwiązań dedykowanych wyłącznie pod określone konfiguracje sprzętowo-programowe. Bardzo ważnym elementem tych systemów jest możliwość rozbudowy o własne moduły.

## 5. System EIP.net

Przykładem systemu zarządzania treścią przeznaczonego dla małych i średnich firm jest system EIP.net. Głównym celem stawianym sobie przez twórców systemu było zautomatyzowanie procesu umieszczania informacji w obsługiwanych serwisach internetowych. Najistotniejszą cechą stworzonego systemu jest możliwość dostarczania informacji w postaci dokumentów pakietu Microsoft Office. System pomógł również w precyzyjnym określeniu odpowiedzialności za poszczególne elementy serwisu.

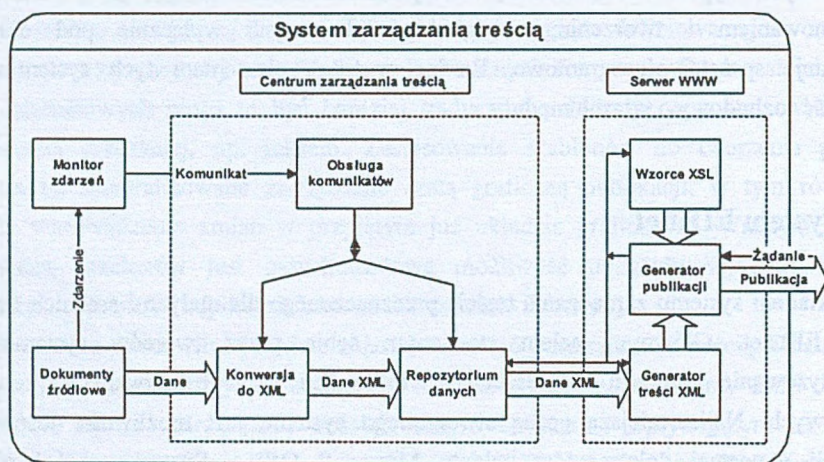
Przedstawiany system zarządzania treścią działa na systemie operacyjnym Windows2000 Server, na którym zainstalowano platformę .NET. Jako serwer WWW wykorzystano serwer IIS w wersji 5. Repozytorium danych zostało zorganizowane w postaci odpowiednich plików XML przechowywanych na dysku, dzięki czemu nie jest obowiązkowe posiadanie bazy danych.

System umożliwia publikowanie treści dostarczonych w jednym z trzech formatów: dokumenty Word, arkusze Excel oraz strony w języku HTML lub XHTML. Wszystkie informacje są konwertowane do postaci XML i następnie publikowane za pomocą odpowiedniej transformacji XSL. System umożliwia również przechowywanie innych plików, w tym plików graficznych i dokumentów PDF.

Architektura systemu jest przedstawiona na rys. 2.

Wszystkie dokumenty źródłowe są umieszczane w specjalnie do tego celu przygotowanej przestrzeni dyskowej. Zawartość tej przestrzeni jest ciągle nadzorowana przez specjalny monitor, który o każdym zaistniałym zdarzeniu informuje system zarządzania treścią. Jeśli zachodzi potrzeba umieszczenia nowych danych w repozytorium, to system tworzy kopię

dokumentu źródłowego, przekształca ją do postaci XML i rozkłada zawartą w nim informację na zasadniczą treść dokumentu oraz dodatkowe parametry przekazane przez autora dokumentu. Do określania parametrów został wykorzystany mechanizm właściwości dokumentów pakietu Office, który umożliwia również tworzenie własnych parametrów. Równoległe dokumenty źródłowe są indeksowane za pomocą odpowiedniej usługi systemowej, którą system wykorzystuje do wyszukiwania określonych informacji. System nie wprowadza własnego mechanizmu tworzenia zapasowych kopii dokumentów, gdyż postanowiono w tym celu skorzystać ze standardowych narzędzi oferowanych wraz z systemem operacyjnym.



Rys. 2. Architektura systemu EIP.net  
Fig. 2. Architecture of EIP.net system

Po umieszczeniu danych w repozytorium można je wykorzystać w aplikacji internetowej działającej na serwerze WWW. Proces prezentacji jest inicjowany przez aplikację podczas każdego żądania odczytu strony. Pierwszym zadaniem modułu publikacji jest sprawdzenie, czy poszukiwany dokument istnieje, oraz zweryfikowanie uprawnień do publikacji. Kolejny etap to skompletowanie dokumentu XML na podstawie przechowywanych w repozytorium informacji. System umożliwia podłączenie zewnętrznych źródeł danych, dzięki czemu możliwe jest zintegrowanie systemu z niezależnymi, już istniejącymi aplikacjami, a także tworzenie nowych własnych modułów. Ostatnią fazą jest wykonanie odpowiedniej transformacji XSL generującej stronę serwisu WWW. Wykorzystanie różnych szablonów transformacji umożliwia tworzenie serwisów w kilku wersjach przeznaczonych dla różnych grup odbiorców. Dzięki zastosowaniu technologii XML i XSL możliwe jest uporządkowanie formatu dostarczonych informacji i zastosowanie jednolitej szaty graficznej dla wszystkich generowanych stron. Osoba odpowiedzialna za wygląd stron w serwisie określa zestaw dopuszczalnych stylów formatowania dokumentów, które mogą zostać wykorzystane



przez autorów dokumentów. Wszelkie inne style i niedopuszczone elementy są filtrowane podczas transformacji.

Wśród podstawowych zalet stworzonego systemu można wymienić, podobnie jak dla innych tego typu rozwiązań: zautomatyzowanie i przyspieszenie procesu publikacji, redukcję związanych z tym kosztów, umożliwienie autorom dokumentów na bezpośrednie kontrolowanie treści publikacji, scentralizowanie systemu zarządzania wyglądem publikacji. Do cech charakterystycznych tego produktu należy możliwość przetwarzania dokumentów pakietu Office i stron HTML, integracja systemu bezpieczeństwa z mechanizmami zabezpieczeń systemu operacyjnego oraz stworzenie repozytorium danych bez wykorzystania bazy danych.

## 6. Podsumowanie

Z problemem efektywnego zarządzania treścią serwisów internetowych boryka się wiele firm i instytucji. Na ich potrzeby stworzono wiele różnorodnych systemów, umożliwiających zautomatyzowanie całego procesu i określenie obowiązków zainteresowanych osób. Wśród dostępnych rozwiązań istnieje wiele różnic wynikłych ze specyficznych wymagań określonych grup klientów, a także z ich możliwości finansowych. Mimo iż prognozy dotyczące rynku systemów CMS zapowiadają ograniczenie liczby oferowanych produktów do tylko tych najlepszych i najpopularniejszych, to wciąż powstają nowe autorskie rozwiązania próbujące nawiązać konkurencję z istniejącymi produktami.

W omawianym systemie EIP.net użytkownicy mogą dostarczać informacji w dokumentach pakietu Microsoft Office, dzięki czemu mogą bardzo szybko opanować sposób prezentowania informacji. Innym ciekawym rozwiązaniem jest stworzenie repozytorium danych w przestrzeni dyskowej zamiast tradycyjnej bazy danych. Mimo pewnych ograniczeń funkcjonalnych przedstawiony produkt cieszy się dość dobrą opinią klientów.

## LITERATURA

1. Browning P., Lowndes M.: Content Management Systems. Techwatch report TSW 01-02, 2001. [http://www.jisc.ac.uk/index.cfm?name=techwatch\\_report\\_0102](http://www.jisc.ac.uk/index.cfm?name=techwatch_report_0102).
2. Clyman J.: From Chaos to Control. PC Magazine, 2002. [http://www.pcmag.com/print\\_article/0,3048,a=30213,00.asp](http://www.pcmag.com/print_article/0,3048,a=30213,00.asp).
3. Guenther K.: Web Site Management - What is a Web Content Management Solution? ONLINE, 2001, t. 25, nr 4, p. 81-84.

4. Rzewuski M.: Systemy zarządzania treścią - redakcyjne kombajny. PCKurier 3/2001, s. 40-47. <http://www.pckurier.pl/archiwum/art0.asp?ID=4566> .
5. Rzewuski M.: Zarządzanie treścią - raport IDC - Czas oszczędności w przedsiębiorstwach. PCKurier 20/2001, s. 13. <http://www.pckurier.pl/archiwum/art0.asp?ID=5086> .
6. Arnold S. E.: Content Management's New Realities. ONLINE, 2003, t. 27, nr 1, p. 36-40.
7. Choo G.: Buy, build, or rent that CMS? 2002. <http://www.techrepublic.com/article.jhtml?id=r00520020820goo01.htm> .
8. Choo G.: Defining CMS technical requirements. 2002. <http://www.techrepublic.com/article.jhtml?id=r00520020308goo02.htm> .

Recenzent: Dr inż. Arkadiusz Sochan

Wpłynęło do Redakcji 31 marca 2003 r.

## Abstract

Maintenance and keeping up-to-date of complex Web sites are very difficult tasks. Processing a large amount of information, difficulties in doing design and navigation, and a lack of authority control make those tasks difficulty for responsible people. The solution to these problems is using so-called content management system. But there is no one perfect system for every company or institution. Each of them has different structure and other needs. For that reason, various different systems are developed and new ones are formed.

In this paper, the main idea of those systems is discussed. The general model consists of three essential parts responsible for extracting information from source documents, processing it, and creating publication. The CMSs usually are designed to put information on Web site, but the form of them relies on many factors. Depending of the form of source documents and abilities of processing them, systems can differ each from the others in many respects. Development of technologies of creating Web sites causes that the new systems can be cheaper, more flexible or have more possibilities. One of such systems, created by the author, is presented in this paper.

## Adres

Sławomir NIEDBAŁA: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16, 44-101 Gliwice, Polska, [niedbyk@star.iinf.polsl.gliwice.pl](mailto:niedbyk@star.iinf.polsl.gliwice.pl) .