

**Adam Dudczak, Marcin Heliński, Cezary Mazurek, Tomasz Parkoła, Marcin Werla**

**Poznańskie Centrum Superkomputerowo-Sieciowe**  
ul. Z. Noskowskiego 12/14, 61-704 Poznań, <http://www.man.poznan.pl/>

## **ANALIZA FUNKCJONALNOŚCI WYBRANYCH MODELI I SYSTEMÓW ZARZĄDZANIA BIBLIOTEKAMI CYFROWYMI**

### **Streszczenie**

Zachodzące w ostatnich czasach zmiany technologiczne w połączeniu z popularyzacją Internetu dają w efekcie ogromny wzrost ilości powstających na świecie cyfrowych informacji. W celu efektywnego gromadzenia, przechowywania i udostępniania takich informacji tworzone są specjalne systemy informatyczne. Rodzajem takich systemów są m.in. biblioteki cyfrowe. Obecnie istnieje na świecie kilka modeli takich bibliotek. Jest też kilka najpopularniejszych, mających wiele wdrożeń, systemów do ich budowy. Każdy z tych modeli i systemów jest w jakiś sposób niepowtarzalny. Celem niniejszego artykułu jest próba określenia podstawowego zakresu funkcjonalnego systemu biblioteki cyfrowej. Podstawą jest tu przegląd i analiza funkcjonalności wybranych modeli i systemów bibliotek cyfrowych. Zdefiniowany podstawowy zakres funkcjonalny bibliotek cyfrowych powinien ułatwić tworzenie nowych bibliotek cyfrowych.

Badania finansowane z grantu badawczego MNiSW nr 3 T11C 023 30.

### **1. WSTĘP**

Wraz z rozwojem i popularyzacją technologii informatycznych, w szczególności Internetu, znacznie wzrasta liczba dostępnych w globalnej sieci informacji. Z porównania raportów [1] i [2] firmy Netcraft, śledzącej rozwój Internetu już od prawie 12 lat, widać, iż w okresie od 2002 do 2007 roku liczba serwerów WWW dostępnych w Internecie wzrosła dziesięciokrotnie i przekroczyła 110 milionów. Według opublikowanego niedawno raportu firmy Miniwatts Marketing Group na przestrzeni lat 2000 – 2007 na świecie nastąpił ponad 200% wzrost liczby użytkowników Internetu (w Polsce w tym okresie ponad 307%) [3].

Tak duży wzrost ilości dostępnych informacji i liczby użytkowników ma znaczący wpływ na sam Internet, sposób jego wykorzystania, a także sposób tworzenia oraz udostępniania treści cyfrowych. Poza samą formą prezentacji informacji, która ewoluowała od postaci tekstowej do w pełni multimedialnej, dużej zmianie uległa koncepcja i architektura sieciowych systemów gromadzenia, przetwarzania i udostępniania informacji. Rozwój technologii internetowych pozwolił, aby z prostych serwisów WWW będących

zbiorem statycznych plików HTML systemy te przekształciły się w zaawansowane, zróżnicowane aplikacje WWW takie jak portale, wiki, fora dyskusyjne, blogi czy systemy typu CMS<sup>1</sup>. Wraz ze wzrostem ilości i szybkości zmiany informacji zaczęto dostrzegać problem długoterminowego gromadzenia, przechowywania i udostępniania dużych ilości zróżnicowanych danych. Problem ten jest bardzo złożony i obejmuje wiele różnych aspektów, od kwestii transferu dużych ilości danych i ich składowania, poprzez metadane, aż po efektywny dostęp i przeszukiwanie gromadzonych informacji.

Odpowiedzią na ten problem są systemy bibliotek cyfrowych (BC). Jednym z pierwszych projektów, który można określić mianem BC był „Project Gutenberg” [4] zapoczątkowany w 1971 roku. Miał on na celu gromadzenie i udostępnianie cyfrowych postaci książek. Obecnie projekt ten udostępnia teksty ponad 20 000 publikacji. Jednak współczesne BC to nie tylko systemy, w których gromadzone są cyfrowe książki. Popularne jest m.in. stosowanie systemów BC do gromadzenia materiałów audiowizualnych, czego przykładem może być serwis BBC Motion Gallery (<http://www.bbcmotiongallery.com/>). Formą BC są również tak niezbędne obecnie naukowe repozytoria cyfrowe, w których składowane są np. dane z eksperymentów naukowych czy sekwencje DNA. Podobnym obszarem zastosowań BC są systemy telemedyczne, w których przechowywana jest dokumentacja dotycząca poszczególnych przypadków medycznych, wyników badań pacjentów itp.

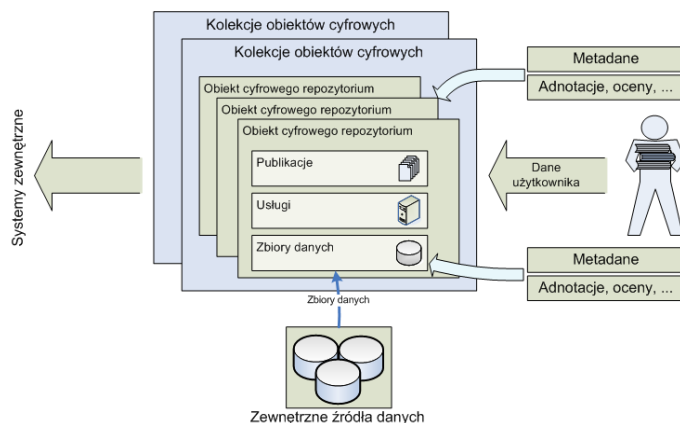
Współczesne systemy BC projektowane są tak, aby móc długotrwale przechować dane cyfrowe w dowolnym formacie, opisane przy pomocy zróżnicowanych schematów metadanych (patrz rys. 1), a ich architektura coraz częściej zorientowana jest na usługi (SOA)<sup>2</sup> [5]. Takie podejście do tworzenia BC umożliwia integrowanie ich z innymi systemami i wykorzystywanie ich funkcjonalności oraz gromadzonych w nich danych.

Z zestawienia obszarów zastosowań BC wynika jasno, iż konieczne w BC są funkcje pozwalające na gromadzenie, przechowywanie i udostępnianie zróżnicowanych treści cyfrowych. Jednak do zaistnienia spójnej BC konieczne są jeszcze dodatkowe specjalizowane funkcje umożliwiające długoterminowe przechowywanie, zaawansowaną prezentację gromadzonych treści, ułatwiające dostęp do nich i pozwalające zmaksymalizować możliwości ich wykorzystania. W dalszej części pracy podjęto próbę określenia podstawowego zakresu funkcjonalnego systemów BC – określenia, jakie cechy funkcjonalne są niezbędne do ich efektywnego działania i wykorzystania przez użytkowników i zewnętrzne usługi.

---

<sup>1</sup> CMS (Content Management System) – oprogramowanie mające wspomagać użytkowników w procesie tworzenia, zarządzania i publikowania treści cyfrowych, często z uwzględnieniem mechanizmów pracy grupowej. Najpopularniejszymi tego typu systemami są systemy do tworzenia serwisów internetowych, realizowane w technologiach PHP+MySQL, takie jak Joomla czy Drupal.

<sup>2</sup> SOA (Service-Oriented Architecture) – Architektura systemów informatycznych, w której podstawową zasadą jest wykorzystanie luźno powiązanych usług do realizacji logiki systemu. Poszczególne usługi udostępniane są jako niezależne od siebie, a ich użycie nie wymaga wiedzy o szczegółach ich realizacji.



Rys. 1. Struktura biblioteki cyfrowej.

Określenie zakresu funkcjonalnego systemu BC rozpoczęte zostało od próby zdefiniowania systemu BC zawartego w następnym punkcie artykułu. W kontekście wybranych definicji, w punktach 3 i 4 przedstawiono wybrane modele BC oraz funkcjonalność najczęściej używanego na świecie oprogramowania do budowy BC. Informacje te posłużyły do opracowania zestawu podstawowych wymagań funkcjonalnych dla systemu BC, który przedstawiony został w punkcie 5. Punkt 6 zawiera podsumowanie i prezentuje kierunki dalszych prac.

## 2. BIBLIOTEKA CYFROWA JAKO ZBIÓR USŁUG

Trudno podać precyzyjną definicję terminu „biblioteka cyfrowa”. Jego dwuczłonowa budowa sugeruje, iż możliwe są dwie perspektywy jego interpretacji: perspektywa dziedzin takich jak bibliotekoznawstwo i informacja naukowa oraz perspektywa szeroko pojętej informatyki, rozumianej jako dziedzina nauka i techniki zajmująca się przetwarzaniem informacji. W kontekście bibliotekoznawstwa i informacji naukowej BC definiuje się często w sposób zbliżony do poniższego:

„Biblioteka cyfrowa – biblioteka, w której znacząca większość zgromadzonych zasobów dostępna jest w postaci interpretowalnej maszynowo (w przeciwieństwie do postaci drukowanej czy mikrofilmu), udostępnianej przy pomocy komputerów.” [6]

Szerzej BC definiuje firma Sun Microsystems Inc. w raporcie zatytułowanym „Digital Library Trends”:

„[BC to] elektroniczne rozszerzenie funkcji, które zazwyczaj dostarcza użytkownikom tradycyjna biblioteka i zasobów, z których użytkownicy w takiej bibliotece korzystają.” [7]

Mamy tu nadal odwołanie do tradycyjnych bibliotek, jednak określenie „elektroniczne rozszerzenie funkcji” jest na tyle szerokie, iż pozwala zakładać, że BC może dawać swoim użytkownikom znacznie więcej możliwości niż jej tradycyjne odpowiedniki. Z kolei Michael Lesk, światowej klasy specjalista w zakresie BC, w swojej definicji odchodzi od terminologii książek czy publikacji, w zamian stosuje szerszy termin, bardziej odpowiadający współczesnej roli BC – pisze o cyfrowej informacji:

*„Biblioteki cyfrowe są zorganizowanymi kolekcjami cyfrowych informacji. Łączą one gromadzenie i organizowanie informacji, które od dawna prowadzone jest przez biblioteki i archiwa z cyfrową reprezentacją tych informacji możliwą dzięki komputerom.” [8]*

W powyższych definicjach mamy do czynienia z porównaniem BC i gromadzonych w nich zasobów do tradycyjnych bibliotek czy archiwów. Grupa robocza do spraw metryk bibliotek cyfrowych działająca przy czasopiśmie D-Lib zaproponowała następującą, nie nawiązującą wprost do tradycyjnych bibliotek, definicję BC:

*„Biblioteka cyfrowa to:*

- *kolekcja usług*
- *oraz kolekcja obiektów informacyjnych,*
- *która wspiera użytkowników w wykorzystaniu obiektów informacyjnych*
- *i organizację oraz prezentację tych obiektów*
- *dostępnych bezpośrednio lub pośrednio*
- *przy użyciu elektronicznych/cyfrowych środków.” [9]*

Podobnie jak w przytoczonej wcześniej definicji podanej przez Michaela Leska, również w powyższej definicji również mowa jest o pewnej kolekcji cyfrowych informacji. Ponadto wyodrębniona jest tu kolekcja usług pełniąca niezwykle istotną rolę zawartą w dwóch środkowych punktach definicji. Ta rola to w ogólności wsparcie w zakresie organizacji, prezentacji i wykorzystania obiektów informacyjnych. Bez istnienia kolekcji usług BC pozostaje tylko kolekcją cyfrowych obiektów informacyjnych. Szczegółowa funkcjonalność usług BC może być zróżnicowana w zależności od charakteru danej biblioteki, gromadzonych w niej informacji czy korzystających z niej użytkowników i innych systemów. Pewien podstawowy zakres powinien jednak pozostać taki sam, aby spełniać podstawowe założenia zawarte w powyższych definicjach.

### **3. WYBRANE MODELE BIBLIOTEK CYFROWYCH**

#### **3.1. Wprowadzenie**

W ramach prac badawczo-rozwojowych prowadzonych na całym świecie w zakresie BC opracowano wiele modeli tych bibliotek. Poniżej zaprezentowano skrócone opisy dwóch wybranych modeli. Pierwszy z nich to opracowany w USA model OAIS. Opisuje on system otwartej informacji archiwalnej zarówno na poziomie technicznym jak organizacyjnym. Drugi model to zaproponowany niedawno przez europejską organizację DELOS referencyjny model BC.

#### **3.2. Open Archival Information System (OAIS)**

W styczniu 2002 roku Consultative Committee for Space Data Systems (CCSDS) opublikowała rekomendację [10] zawierającą opis modelu otwartego systemu informacji archiwalnej (Open Archival Information System, OAIS). Opisany w tej rekomendacji model stanowić ma podstawę dla dalszej standaryzacji cyfrowych archiwów informacji. OAIS zdefiniowane jest jako archiwum składające się z osób i systemów, które przyjęły na siebie odpowiedzialność długoterminowego przechowywania informacji i udostępniania ich określonym społecznościom.

Model OAIS wyróżnia trzy podstawowe rodzaje aktorów występujących w otoczeniu archiwum:

- producentów – dostarczają informacje, które mają być przechowane w archiwum,
- konsumentów – wyszukują i pobierają wyszukane informacje z archiwum, w szczególności powinni być w stanie zrozumieć pobraną informację,
- zarządzających – odpowiadają za ustalenie zasad działania archiwum w szerszym jego kontekście, nie są związani z codziennymi działaniami archiwum – za to odpowiedzialne są jednostki będące częścią archiwum.

Model OAIS nie jest modelem ściśle technicznym, mówi on również jak powinno być zorganizowane archiwum jako instytucja. Stąd też wyraźne wyróżnienie wśród podstawowych rodzajów aktorów osób zarządzających. Z punktu widzenia analizy funkcjonalnej OAIS jako systemu technicznego możemy jednak skupić się na producentach i konsumentach. Według modelu OAIS między archiwum, a producentami i konsumentami zachodzą następujące podstawowe interakcje:

- producenci przesyłają dane do archiwum, a archiwum przechowuje te dane;
- konsumenci pobierają dane udostępnione z archiwum – przy czym wyróżnia się dwa podstawowe rodzaje „zamówienia danych”:
  - zamówienie „ad hoc” – nie poprzedzone żadnym z góry znanym zdarzeniem, konsument może nie wiedzieć jakiego typu dane są w archiwum i które z tych danych są dla niego interesujące;
  - zamówienie wywołane zdarzeniem – przy czym zdarzenie to może być zdarzeniem okresowym lub unikalnym i jednorazowym.

Interakcje te określają podstawowy zakres funkcjonalny wg modelu OAIS, jaki powinno oferować archiwum zewnętrznym systemom. Ponadto w szczegółowej analizie funkcjonalnej modelu OAIS można znaleźć jeszcze następujące elementy odnoszące się do aspektów technicznych: generowanie raportów dotyczących gromadzonych zasobów, utrzymywanie kopii bezpieczeństwa i obsługa sytuacji awaryjnych, kontrola spójności treści przechowywanych zasobów, budowanie hierarchii gromadzonych obiektów i tworzenie kolekcji oraz kontrola dostępu.

### 3.3. DELOS Digital Library Reference Model

We wrześniu 2006 roku instytucje członkowskie DELOS udostępniły dokument zatytułowany „Reference Model for Digital Library Management Systems” [11]. Przedstawiono w nim szczegółową specyfikację trzech warstw modelu BC opisanego dokładnie w „Digital Libraries Manifesto” [12] oraz podjęto próbę zdefiniowania funkcjonalności BC na poziomie każdej z tych warstw modelu. Warstwy te to:

- Biblioteka cyfrowa (Digital Library)
- System biblioteki cyfrowej (Digital Library System)
- System zarządzania biblioteką cyfrową (Digital Library Management System)

Pierwsze dwie warstwy mają taką samą funkcjonalność – taką, przez pryzmat której zewnętrzni, końcowi użytkownicy postrzegają BC. Zakres funkcjonalny BC został tu podzielony na pięć podstawowych grup, w ramach których wyróżniono konkretne funkcje:

- Uzyskaj dostęp do informacji - przeglądaj, wizualizuj, zachowaj (*persist*), przetłumacz, przeszukuj, dostarcz.
- Zarządzaj informacją – stwórz, uaktualnij, przekształć, nanieś adnotacje, złóż (*compose*), prześlij, waliduj.

- Zarządzaj biblioteką cyfrową – rozpowszechnij (*disseminate*), publikuj, wycofaj, przechowaj, zarządzaj zasadami, zarządzaj kolekcjami, zarządzaj użytkownikami (w szczególności: rejestruj użytkowników, zarządzaj rolami).
- Personalizuj – wyraż zainteresowanie, zaloguj się, zapisz się, zastosuj profil, dostarcz informację zwrotną.
- Współpracuj.

Jak widać, model opracowany przez DELOS jest od strony technicznej zdecydowanie bardziej rozbudowany i szczegółowy niż model OAIS. Operuje on na poziomie systemów informatycznych, ich funkcjonalności, architektury, użytkowników itp., podczas gdy model OAIS operuje na poziomie organizacji czy instytucji. Takie podejście w modelu DELOS ułatwia praktyczne wykorzystanie go przy budowie systemów BC, czy też przy tworzeniu oprogramowania do ich budowy.

## 4. SYSTEMY ZARZĄDZANIA BIBLIOTEKAMI CYFROWYMI

### 4.1. Wprowadzenie

Obecnie na świecie dostępnych jest wiele systemów do budowy BC. W większości są to jednak systemy budowane na potrzebę jednego, konkretnego projektu i zazwyczaj nie są one po zakończeniu tego projektu dalej utrzymywane. Poniżej przedstawiono opis trzech wybranych systemów do budowy BC, które mają więcej niż jedną instalację i są utrzymywane i rozwijane już od dłuższego czasu. Dwa pierwsze (Fedora oraz DSpace) są to bardzo popularne projekty amerykańskie, mające wiele wdrożeń na całym świecie. Trzeci z opisanych projektów (Open DLib) to włoski projekt o podłożu badawczym mający kilka wdrożeń w ramach prac organizacji DELOS.

### 4.2. Fedora<sup>3</sup>

Fedora to oprogramowanie do budowy repozytoriów cyfrowych rozwijane przez Cornell University Information Science i University of Virginia Library. Podstawowym celem tego oprogramowania jest dostarczanie podstawowych podsystemów niezbędnych do budowy zróżnicowanych systemów zarządzania informacją. Oprogramowanie to cechuje się architekturą zorientowaną na usługi. Całość oparta jest na modelu powiązanych z sobą obiektów cyfrowych, które udostępniane są użytkownikom na różne sposoby (różne widoki tego samego obiektu). Podstawowe cechy oprogramowania Fedora to:

- obiekty cyfrowe mogą przechowywać dowolną liczbę i dowolne formaty danych,
- metadane opisujące obiekt cyfrowy mogą występować w dowolnej liczbie formatów,
- obiekty cyfrowe zawierają metadane reprezentujące różnorodne powiązania między obiektami, metadane te są indeksowane,
- dla każdego obiektu przechowywana jest historia jego zmian,

Obiekty cyfrowe w repozytorium Fedora składają się z unikalnego trwałego identyfikatora, metadanych niezbędnych do zarządzania obiektem, własności opisujących obiekt, strumieni danych (lokalnych bądź zdalnych) oraz metod rozpowszechniania obiektu

---

<sup>3</sup> Fedora Project, <http://www.fedora.info/>

(*disseminators*). Obiekty cyfrowe w repozytorium Fedora mogą być ze sobą powiązane na wiele sposobów, a zależności te zapisywane są jako metadane obiektów. Metadane każdego obiektu cyfrowego są indeksowane podczas jego tworzenia. Indeksowaniu podlegają zarówno metadane systemowe, jak i metadane zapisywane w schemacie Dublin Core [13].

Użytkownicy pracujący z obiektami cyfrowymi przechowywanymi w repozytorium Fedora mają do nich dostęp poprzez aplikacje klienckie komunikujące się z serwerem Fedora wykorzystując cztery rodzaje interfejsów: zarządzania, dostępu, wyszukiwania oraz komunikacji przez OAI-PMH<sup>4</sup>.

### 4.3. DSpace<sup>5</sup>

DSpace to oprogramowanie rozwijane wspólnie przez MIT Libraries i Hewlett-Packard Labs wykorzystywane do budowy różnego rodzaju repozytoriów cyfrowych. W odróżnieniu od systemu Fedora, DSpace nie ogranicza się tylko do dostarczania podstawowej platformy repozytorium, ale oferuje kompleksowe rozwiązanie włącznie z interfejsem WWW przeznaczonym dla użytkowników. Podstawowym elementem struktury repozytorium opartego na DSpace są wspólnoty, które mogą zawierać wspólnoty podrzędne. W ramach każdej ze wspólnot występować mogą kolekcje obiektów cyfrowych. Obiekty te mogą mieć wiele różnych reprezentacji, a każda z reprezentacji może składać się z wielu różnych plików z treścią.

Do ciekawszych funkcji systemu Dspace zaliczyć można ściśle zdefiniowany przepływ pracy dotyczący obiektu cyfrowego, możliwość przechowywania obiektów cyfrowych w zaawansowanych systemach składowania danych, zbieranie statystyk na temat gromadzonych obiektów oraz możliwość zapisu informacji o zmianach w systemie przy użyciu ontologii ABC [14] opracowanej w ramach projektu Harmony<sup>6</sup>.

### 4.4. Open DLib<sup>7</sup>

System BC Open DLib rozwijany jest przez grupę DLib pracującą w instytucie ISTI – części Italian National Research Council. Projekt ten jest efektem badań nad aplikacjami wymagającymi obsługi dużych kolekcji treści multimedialnych dostępnych przez Internet. Ze względu na zakres funkcjonalny wyróżniono tu następujące typy usług:

- Repository, Multimedia Storage – przechowują odpowiednio dokumenty w formacie DoMDL [15] oraz dokumenty wideo w postaci strumieni.
- Library Management – obsługuje dodawanie, usuwanie i wymianę dokumentów.
- Index, Query Mediator, Browse – usługi odpowiedzialne za obsługę wyszukiwania, generowanie indeksów dokumentów i przeglądanie ich.

---

<sup>4</sup> OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) – protokół sieciowy oparty na protokole HTTP, opracowany przez Open Archives Initiative (<http://www.openarchives.org/>). Daje możliwość pobierania rekordów metadanych z repozytoriów cyfrowych, umożliwiając wykorzystanie tych metadanych w innych usługach.

<sup>5</sup> DSpace Project, <http://www.dspace.org/>

<sup>6</sup> Harmony Project, <http://www.metadatalib.net/harmony/>

<sup>7</sup> Open DLib System, <http://www.opendlib.com/>

- Registry – przechowuje informacje na temat użytkowników i uprawnień.
- Manager Service – zarządza innymi usługami.
- Collection – umożliwia tworzenie wirtualnych kolekcji dokumentów.
- User Interface – odpowiedzialny za interakcje użytkownika z systemem.

Open DLib charakteryzuje się otwartą architekturą zorientowaną na usługi, umożliwiającą elastyczną konstrukcję BC ze zbioru usług oraz ich dynamiczną rekonfigurację czy rozbudowę.

## **5. PODSTAWOWE WYMAGANIA FUNKCJONALNE SYSTEMU BIBLIOTEKI CYFROWEJ**

Zestawienie i analiza informacji zawartych w poprzednich punktach doprowadziła do opracowania zestawu podstawowych wymagań funkcjonalnych jakie powinien spełniać system BC. Podstawowym kryterium doboru funkcji było założenie, iż wybrane będą funkcje niezbędne do utworzenia BC, a pominięta będzie rozszerzona funkcjonalność związana z możliwościami występującymi tylko w niektórych BC, często wyłącznie na poziomie interfejsu użytkownika. Przykładem może być tutaj usługa powiadamiania o nowych obiektach cyfrowych. Występuje ona w każdej z omówionych wcześniej platform do budowy BC. Nie ma jednak potrzeby uznawać ją za funkcję podstawową, gdyż może być ona łatwo zrealizowana przez twórców interfejsu użytkownika konkretnej BC chociażby jako wyszukiwanie obiektów spełniających określony warunek daty utworzenia (np. ostatnie 24 godziny). Tak więc założono, że twórcy BC będą wykorzystywali wybrane funkcje podstawowe do stworzenia funkcji złożonych, wymaganych w kontekście danej biblioteki.

Kolejnym kryterium była niezależność funkcji – podstawowe funkcje powinny być od siebie wzajemnie niezależne, a ewentualne powiązania między nimi powinny być tworzone poprzez obiekty cyfrowe, na których poszczególne funkcje będą wykonywane. Przykładem może być tutaj funkcjonalność przechowywania obiektów cyfrowych i funkcjonalność przechowywania metadanych dotyczących tych obiektów. Obydwie te funkcje są od siebie niezależne – można przechowywać treść obiektu cyfrowego w oderwaniu od jego metadanych, i odwrotnie – można przechowywać metadane niezależnie od treści. Funkcje te mają rozłączny charakter i żadna z nich nie wymaga drugiej. Powiązanie między tymi funkcjami następuje w momencie, gdy w ramach danej BC przechowujemy pewien obiekt i jego metadane. Wtedy BC korzysta z funkcji przechowywania treści i metadanych jako z dwóch komplementarnych funkcji umożliwiających podstawową działalność systemu BC, a powiązanie między tymi funkcjami następuje właśnie na poziomie BC poprzez przetwarzany obiekt cyfrowy. Dzięki takiemu podejściu możliwe jest konstruowanie bibliotek cyfrowych zgodnie z założeniami SOA, a komponenty realizujące poszczególne funkcje mogą być wydzielonymi usługami współdzielonymi pomiędzy wiele systemów.

Ostatecznie zaproponowano następujący zestaw podstawowych wymagań funkcjonalnych dla systemu BC:

- Przechowywanie obiektów cyfrowych (w wielu różnych postaciach, w wielu dowolnych formatach)
- Przechowywanie metadanych i adnotacji dotyczących obiektów cyfrowych (w wielu dowolnych schematach)
- Przeszukiwanie metadanych i adnotacji obiektów cyfrowych



- Tworzenie obiektów cyfrowych składających się z innych obiektów cyfrowych (kolekcje, obiekty złożone) oraz powiązań pomiędzy obiektami cyfrowymi

Powyższe funkcje można podzielić na dwie grupy. Do pierwszej z nich należy zaliczyć funkcję przechowywania obiektów i funkcję przechowywania metadanych i adnotacji. Funkcje te są niezbędne do działania BC, umożliwiają realizację podstawowych zadań każdego z opisanych wcześniej systemów i modeli, i trudno sobie wyobrazić przykład BC, która tych usług by nie wymagała. Tak więc są to funkcje konieczne, lecz nie wystarczające do skonstruowania BC. Druga grupa to funkcje przeszukiwania i tworzenia złożonych obiektów cyfrowych. Funkcja przeszukiwania znacznie ułatwia dostęp do gromadzonych danych i ich wykorzystanie. Funkcja tworzenia powiązań między obiektami i budowania złożonych obiektów cyfrowych jest niezwykle istotna w systemach, których rola nie ogranicza się do składowania danych, lecz są one dla swoich użytkowników podstawą środowiska pracy nad gromadzonymi obiektami cyfrowymi. A tym właśnie powinny być BC. Wszystkie powyższe funkcje występują w każdym ze wspomnianych modeli i systemów BC. Jednak istnieje wiele przykładów prostych BC, w których nie ma funkcji wyszukiwania czy organizowania obiektów cyfrowych, a dostęp do konkretnego obiektu cyfrowego możliwy jest tylko po podaniu konkretnego identyfikatora.

W powyższej analizie pominięto wymagania niefunkcjonalne, takie jak skalowalność, wydajność, środowisko działania czy kontrola dostępu. Są one pewnymi cechami opisanych powyżej podstawowych funkcji i z pewnością w zdecydowanej większości zastosowań BC są niezbędne. Mechanizm kontroli dostępu występuje w każdym z opisanych w punkcie 4 systemów BC i zazwyczaj egzekwowany jest w przypadku użycia każdej z podstawowych funkcji BC. Mimo wszystko jednak wymagania niefunkcjonalne są zdecydowanie bardziej zależne od konkretnego zastosowania BC. Przykładem może tu być kwestia skalowalności czy rozproszenia systemu BC na kilka odrębnych serwerów. Możliwość taka jest rozwinięta w systemie FEDORA i Open DLib, a w przypadku systemu DSpace jest znacznie ograniczona. Wynika to z obszarów zastosowań tych systemów – DSpace służy zazwyczaj do budowy repozytoriów instytucjonalnych, a jego instalacje nie są narażone na bardzo duże obciążenia. Dodatkowa kwestia występująca przy wymaganiach niefunkcjonalnych, to sposób ich realizacji w architekturze systemu. Przykładem może być tutaj kwestia autoryzacji przy dostępie do systemu. Może ona być realizowana jako zupełnie odrębna usługa, jak np. usługa Registry w systemie Open DLib. Z drugiej strony mechanizmy autoryzacji mogą nie być wyróżnione jako osobny komponent systemu, tylko występować jako element realizacji wszystkich innych jego funkcji – tak jak to ma miejsce np. w systemie Fedora.

## 6. ZAKOŃCZENIE

Nieustannie rosnąca ilość informacji elektronicznej pociąga za sobą konieczność opracowywania nowych systemów gromadzenia, długoterminowego przechowywania i efektywnego udostępniania dużych ilości zróżnicowanych danych. Systemami takimi są między innymi BC. W ramach prac opisywanych w niniejszym artykule, przedstawiono wybrane modele i systemy takich bibliotek. Analiza tych modeli i systemów umożliwiła zaproponowanie zestawu podstawowych wymagań funkcjonalnych, które spełniać powinien system BC, aby sprostać stawianym mu wyzwaniom. System spełniający te wymagania pozwala na przechowanie obiektów cyfrowych w dowolnym formacie wraz z metadanymi i adnotacjami. Dostęp do tych obiektów cyfrowych jest ułatwiony poprzez

funkcje wyszukiwawcze, a funkcje wiązania obiektów pomiędzy sobą i tworzenia obiektów złożonych znacznie zwiększają możliwości organizowania i wykorzystywania gromadzonych obiektów.

W ramach dalszych prac konieczne będzie uszczegółowienie i doprecyzowanie zaproponowanych wymagań, łącznie z przykładowymi scenariuszami użycia. Tak opracowane wymagania będą stanowić podstawę do zdefiniowania zestawu usług sieciowych i protokołów komunikacyjnych pozwalających stworzyć funkcjonalną i spójną BC.

## BIBLIOGRAFIA

- [1] *February 2007 Web Server Survey*, Netcraft, 2007.  
[http://news.netcraft.com/archives/2007/02/02/february\\_2007\\_web\\_server\\_survey.html](http://news.netcraft.com/archives/2007/02/02/february_2007_web_server_survey.html)
- [2] *February 2002 Web Server Survey*, Netcraft, 2002. <http://survey.netcraft.com/index-200002.html>
- [3] *Internet World Stats: Usage and Population Statistics*, Miniwatts Marketing Group, 2007.  
<http://www.internetworldstats.com/stats.htm>
- [4] Hart M.: *The History and Philosophy of Project Gutenberg*, Project Gutenberg, 1992.
- [5] *SOA Practitioners' Guide*, BEA Systems, 2006.  
<http://dev2dev.bea.com/pub/a/2006/09/soa-practitioners-guide.html>
- [6] Reitz Joan M.: *Dictionary for Library and Information Science*, Libraries Unlimited, 2004.
- [7] *Digital Library Technology Trends*, Sun Microsystems Inc., 2002.  
[http://www.sun.com/products-n-solutions/edu/whitepapers/pdf/digital\\_library\\_trends.pdf](http://www.sun.com/products-n-solutions/edu/whitepapers/pdf/digital_library_trends.pdf)
- [8] Lesk M.: *Practical digital libraries: Books, bytes and bucks*, Morgan Kaufmann, 1997.
- [9] Leiner B. M.: *The Scope of the Digital Library*, D-Lib, 1998.
- [10] *Reference Model for an Open Archival Information System (OAIS)*, CCSDS, 2002.
- [11] Candela L., Castelli D.: *Reference Model for Digital Library Management Systems*, DELOS, 2006.
- [12] Candela L., Castelli D.: *Digital Libraries Manifesto*, DELOS, 2006.
- [13] *Information and documentation - The Dublin Core metadata element set*, ISO Standard 15836-2003, NISO, 2003.
- [14] Lagoze C., Hunter J.: *The ABC Ontology and Model*, Journal of Digital Information, Special Issue - selected papers from Dublin Core 2001 Conference.
- [15] *The Open DLib Document Model*, [http://www.opendlib.com/area4/info\\_domdl.html](http://www.opendlib.com/area4/info_domdl.html)

## FUNCTIONALITY ANALYSIS FOR CHOSEN DIGITAL LIBRARY MODELS AND MANAGEMENT SYSTEMS

### Summary

Recent technological changes together with growing popularization of the Internet results in vast increase of created digital information. In order to store, preserve and give access to large amount of information, various systems are created, including digital libraries. Currently several models of such systems exists. Additionally there is several digital library software packages. Each of those models and software packages is somehow unique. This paper presents an attempt to define basic functional scope of digital library systems. It is based on the functionality analysis of chosen digital library models and software packages. Defined scope should be very useful when building new digital libraries.

Works under MNiSW research grant no 3 T11C 023 30.