

dLibra – platforma do budowy repozytoriów cyfrowych

Adam Dudczak¹, Marcin Heliński¹, Cezary Mazurek¹, Tomasz Parkoła¹, Marcin Werla¹

Streszczenie: Niniejszy artykuł opisuje oprogramowanie dLibra – platformę do budowy bibliotek cyfrowych rozwijaną przez Poznańskie Centrum Superkomputerowo-Sieciowe. W artykule omówiono podstawowe funkcje systemu dLibra, jego architekturę oraz najważniejsze moduły. Przedstawiono również zaawansowane mechanizmy systemu dLibra takie jak: zarządzanie metadanymi przechowywanych obiektów, mechanizmy integracji oprogramowania dLibra z innymi systemami informatycznymi czy mechanizmy bazujące na otwartym protokole wymiany informacji między rozproszonymi instalacjami oprogramowania dLibra w Polsce.

Słowa kluczowe: metadane, obiekty cyfrowe, przetwarzanie zasobów rozproszonych, repozytoria cyfrowe, biblioteki cyfrowe, systemy rozproszone

1. Wstęp

Coraz większa ilość cyfrowych informacji tworzonych na całym świecie, wymaga zastosowania profesjonalnych narzędzi, umożliwiających ich efektywne gromadzenie, przechowywanie, zarządzanie i udostępnianie. W roku 1996 w Poznańskim Centrum Superkomputerowo-Sieciowym rozpoczęto prace badawcze w obszarze repozytoriów dokumentów cyfrowych. Efektem tych prac był prototyp oprogramowania do budowy bibliotek cyfrowych, na bazie którego w 1999 roku rozpoczęto tworzenie platformy do budowy bibliotek cyfrowych o nazwie „dLibra”. Platforma ta jest pierwszym tego typu polskim oprogramowaniem.

W roku 2001 uruchomiony został program PIONIER, zdefiniowany jako „Polski Internet Optyczny – Zaawansowane Aplikacje, Usługi i Technologie dla Społeczeństwa Informacyjnego”. Jednym z trzech kluczowych obszarów tego programu było „wytworzenie i przetestowanie pilotowych usług i aplikacji dla społeczeństwa informacyjnego, stanowiących podstawę dla wdrożeń w takich dziedzinach, jak: nauka, edukacja, opieka zdrowotna, środowisko naturalne, administracja rządowa i samorządowa, przemysł i usługi”. Oprogramowanie dLibra rozwijane dalej było jako jeden z projektów w obszarze e-content programu PIONIER.

Pod koniec roku 2002 w ramach współpracy Poznańskiego Centrum Superkomputerowo-Sieciowego z Poznańską Fundacją Bibliotek Naukowych uruchomiona została pierwsza publicznie dostępna biblioteka cyfrowa oparta na oprogramowaniu dLibra – Wielkopolska Biblioteka Cyfrowa. Jest to obecnie największa w Polsce

¹ Poznańskie Centrum Superkomputerowo-Sieciowe ul. Z. Noskowskiego 12/14, 61-704 Poznań
<http://www.man.poznan.pl/>

biblioteka cyfrowa. Śladem WBC poszły również inne instytucje, tworząc w Polsce kolejne biblioteki cyfrowe.

Obecnie system dLibra jest najpopularniejszym w Polsce oprogramowaniem do budowy bibliotek cyfrowych. System ten jest wykorzystywany przez ponad 150 instytucji do budowy regionalnych oraz instytucjonalnych repozytoriów. W czerwcu 2007 roku w sieci PIONIER było uruchomionych około 20 takich systemów. Tworzą one wspólnie sieć rozproszonych bibliotek cyfrowych, dającą dostęp do ponad 80 000 tysięcy obiektów cyfrowych wraz z ich metadanymi oraz do zaawansowanych usług takich jak wyszukiwanie rozproszone czy wirtualne kolekcje.

Następny rozdział niniejszego artykułu przedstawia najważniejsze funkcje oprogramowania dLibra. Rozdział 3 opisuje architekturę tego systemu ze szczególnym uwzględnieniem podstawowego modułu systemu - serwera. Rozdział 4 przedstawia wybrane zaawansowane mechanizmy systemu dLibra takie jak zarządzanie metadanymi obiektów cyfrowych czy wymiana tych metadanych pomiędzy rozproszonymi repozytoriami.

2. Funkcjonalność oprogramowania dLibra

dLibra jest oprogramowaniem, które może być wykorzystywane do budowy bibliotek cyfrowych, instytucjonalnych repozytoriów dokumentów, cyfrowych archiwów oraz innych systemów zarządzania treścią. System ten pozwala na gromadzenie, długoterminowe przechowywanie i udostępnianie obiektów cyfrowych w dowolnym formacie (np. wieloplikowe dokumenty HTML, PDF, nagrania audio i video, itd.). Każdy z przechowywanych obiektów opisany może być przy pomocy zdefiniowanego w ramach danego repozytorium zestawu metadanych. Udostępnianie obiektów cyfrowych w Internecie odbywa się poprzez dedykowane strony WWW repozytorium. Repozytoria dLibra mogą także współpracować z innymi systemami internetowymi przy użyciu otwartych standardów komunikacyjnych takich jak: OAI-PMH (patrz Lagoze, 2004), kanały RSS. System dLibra składa się z trzech podstawowych elementów:

- Serwer biblioteki cyfrowej – odpowiedzialny za realizację wszystkich funkcji biblioteki, uruchomiony na dedykowanym komputerze (lub kilku komputerach), niedostępny bezpośrednio dla użytkowników,
- Aplikacja redaktora i administratora – pozwala użytkownikom tworzącym repozytorium cyfrowe (wprowadzającym obiekty cyfrowe, opisującym je itp.) oraz administratorom na korzystanie z funkcji repozytorium,
- Aplikacja czytelnika – dostępna poprzez strony WWW, pozwala użytkownikom („czytelnikom”) korzystać z zasobów zgromadzonych w repozytorium cyfrowym.

Jedną z podstawowych funkcji oprogramowania dLibra jest przechowywanie obiektów cyfrowych. W ramach tego zadania system dLibra udostępnia takie mechanizmy jak: planowanie wprowadzania obiektów cyfrowych, grupowanie powiązanych obiektów w większe struktury, zarządzanie wersjami obiektów cyfrowych

oraz tworzenie kolekcji powiązanych ze sobą obiektów (patrz Mazurek, 2005). Jak wcześniej wspomniano, obiekty cyfrowe mogą być opisane przy użyciu zdefiniowanego zestawu metadanych. Poza metadanymi opisowymi w systemie dLibra istnieją również metadane techniczne (struktura obiektu, informacje o częściach składowych obiektu) oraz administracyjne (uprawnienia użytkowników do obiektu, lokalizacja oraz stan obiektu w repozytorium).

Użytkownicy systemu dLibra, korzystający z zasobów repozytorium poprzez strony WWW, mają do dyspozycji wiele różnych funkcji takich jak: przeglądanie indeksów obiektów cyfrowych, analizę statystyk ich wykorzystania itp. Kluczową funkcją jest tu jednak funkcja wyszukiwania w lokalnych zasobach. Wyszukiwanie to odbywa się w metadanych opisowych obiektów cyfrowych oraz w treści tych obiektów, z których można ją wyodrębnić w postaci tekstowej (czyli np. dokumenty HTML czy PDF). Korzystając z zaawansowanego formularza wyszukiwania, możliwe jest określenie zakresu wyszukiwania oraz wykorzystanie wyrażeń logicznych do łączenia kilku zapytań prostych. Poza wyszukiwaniem w zasobach lokalnych system dLibra pozwala na przeszukiwanie metadanych opisowych obiektów, znajdujących się w innych (zdalnych) repozytoriach cyfrowych, wspierających otwarty standard wymiany metadanych OAI-PMH. Dzięki tej właściwości możliwe było utworzenie polskiej platformy bibliotek rozproszonych w sieci PIONIER, gdzie każde repozytorium udostępnia funkcję wyszukiwania zasobów rozproszonych z całej sieci bibliotek. Użytkownicy WWW mogą również przeglądać zasoby repozytorium, używając indeksów wartości z metadanych lub kolekcji obiektów cyfrowych.

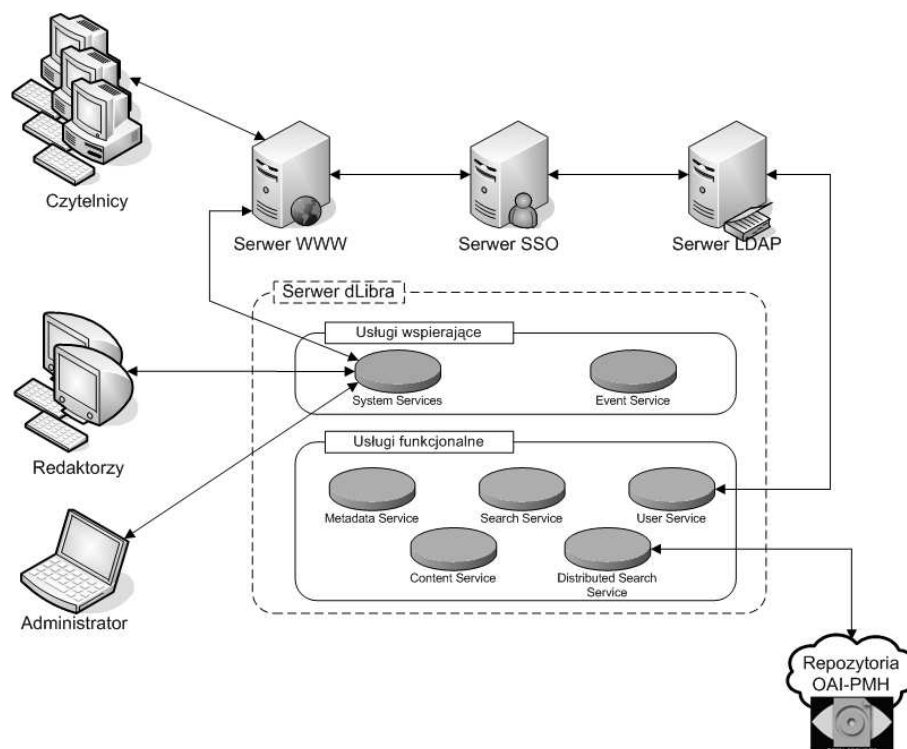
System dLibra opiera się na publicznie dostępnych standardach, umożliwiającą integrację tego środowiska z innymi systemami informatycznymi. Jest to przede wszystkim wsparcie dla protokołu OAI-PMH, dzięki czemu każde repozytorium dLibra może zarówno pobierać informacje z innych repozytoriów OAI-PMH, jak i udostępniać informacje o swoich lokalnych zasobach. Poza protokołem OAI-PMH system dLibra umożliwia przeprowadzanie procesu uwierzytelniania przy użyciu zewnętrznych serwerów LDAP lub systemów Single Sign-On. Wiele informacji dostępnych na stronach WWW repozytorium dLibra jest dostępnych również w formie kanałów RSS, a mechanizm wyszukiwawczy ma dodatek w formacie Open Search.

3. Architektura systemu dLibra

Jak wspomniano wcześniej, system dLibra jest systemem wielowarstwowym, w którym możemy wyróżnić trzy podstawowe moduły:

- Moduł dLibra Serwer,
- Moduł dLibra Redaktor,
- Moduł dLibra Czytelnik.

Na Rys. 1 przedstawiono schemat architektury systemu dLibra. Moduł dLibra Serwer (oznaczony linią przerywaną) jest podstawową jednostką systemu dLibra,



Rysunek 1. Schemat architektury systemu dLibra

w skład której wchodzi szereg powiązanych ze sobą usług, odpowiedzialnych za poszczególne funkcje całego systemu. Usługi modułu dLibra Serwer można podzielić na dwa rodzaje: funkcjonalne oraz wspierające. Usługi funkcjonalne (Metadata, Content, User, Search i Distributed Search) odpowiedzialne są za wszystkie funkcje wykonywane przez serwer. Usługi wspierające (Event i System Services) udostępniają mechanizmy odkrywania, uwierzytelniania usług oraz komunikacji między nimi. Każda z usług systemu dLibra może być uruchomiona na osobnym komputerze (serwerze) lub też usługi te mogą być łączone w grupy. Dodatkowo każda z usług systemu dLibra wymaga do swojego działania relacyjnej bazy danych. Usługi mogą współdzielić między sobą jedno konto bazy danych, ale mogą też korzystać z odrębnych baz danych, uruchomionych na osobnych serwerach (patrz Mazurek, 2005).

Dwa pozostałe moduły stanowią interfejs użytkownika systemu dLibra. Moduł dLibra Czytelnik (aplikacja WWW) uruchamiany jest na dedykowanym komputerze, a jego zadaniem jest obsługa czytelników korzystających z zasobów repozytorium poprzez strony WWW. W celu realizacji swoich zadań moduł dLibra Czytelnik komunikuje się z modułem dLibra Serwer oraz z systemem Single Sign-On, jeśli repozytorium zintegrowane jest z takim systemem. Moduł dLibra Redaktor jest

programem okienkowym, który uruchamiany jest na komputerach użytkowników wprowadzających zasoby do repozytorium lub zarządzających całym repozytorium. Moduł ten, w celu wykonywania wszelkich operacji na repozytorium, komunikuje się z modułem dLibra Serwer.

4. Zaawansowane mechanizmy systemu dLibra

System dLibra posiada zaawansowane mechanizmy zarządzania obiektami cyfrowymi. Wśród najważniejszych mechanizmów można wyróżnić: mechanizm zarządzania metadanymi obiektów cyfrowych wraz z obsługą wielojęzyczności oraz wyrazów bliskoznacznych, mechanizmy pozwalające na łatwą integrację systemu dLibra z innymi systemami informatycznymi oraz mechanizm wymiany metadanych opisowych.

4.1. Mechanizm zarządzania metadanymi

Obiekty cyfrowe w repozytorium dLibra mogą być opisane przy użyciu zdefiniowanego w ramach konkretnej instalacji schematu metadanych. Schemat metadanych w systemie dLibra jest hierarchiczny. Każdy element schematu nazywany jest atrybutem. Atrybutami mogą być przykładowo: tytuł, podtytuł, autor, data wydania itd. Jako że schemat metadanych zorganizowany jest hierarchicznie, każdy atrybut jest węzłem, który może posiadać atrybuty podrzędne (np. atrybut 'identyfikator' może mieć podatrybuty 'URI', 'ISBN', 'ISSN' i 'DOI', wykorzystywane zależnie od kontekstu danego obiektu cyfrowego). Atrybuty podrzędne rozdzielają zakres atrybutu nadrzędnego na mniejsze obszary, dlatego wszystkie wartości przypisane do atrybutów podrzędnych są automatycznie dodawane do wartości atrybutu nadrzędnego. Schemat metadanych (drzewo atrybutów) może być dowolnie modyfikowany przez administratora systemu, przy czym domyślnie wprowadzonym schematem jest schemat Dublin Core w wersji 1.1 (patrz NISO, 2003).

Schemat metadanych definiowany jest na poziomie całego repozytorium, dlatego każdy obiekt cyfrowy opisany jest przy użyciu tego samego zestawu metadanych (zestawu atrybutów). Każdy atrybut zdefiniowany w schemacie metadanych posiada odrębny słownik wartości. Słownik ten tworzony jest automatycznie na podstawie wartości tego atrybutu, występujących w metadanych zgromadzonych obiektów cyfrowych, ale może być również uzupełniony dodatkowymi wartościami przez administratorów systemu. Wartości występujące w słowniku mogą być grupowane, tworząc tzw. grupy wartości. Każda grupa wartości może zawierać jedną lub więcej wartości z słownika, przy czym w każdej grupie musi znajdować się dokładnie jedna wartość bazowa, która identyfikuje grupę. W słowniku wartość bazowa może wystąpić tylko raz. Grupy wartości służą do definiowania grup wyrazów bliskoznacznych, które wykorzystywane są podczas procesu wyszukiwania uruchamianego przez użytkownika repozytorium. Zapytanie użytkownika rozbudowywane jest poprzez dodanie wszystkich wartości z grupy, do której należy wartość z zapytania, jako alternatywne do wartości z zapytania. Aby ułatwić tworzenie słowników atrybutów, podczas wprowadzania opisu obiektu cyfrowego, system dLibra automatycznie dodaje nowe wartości do słownika, tworząc dla każdej nowej wartości

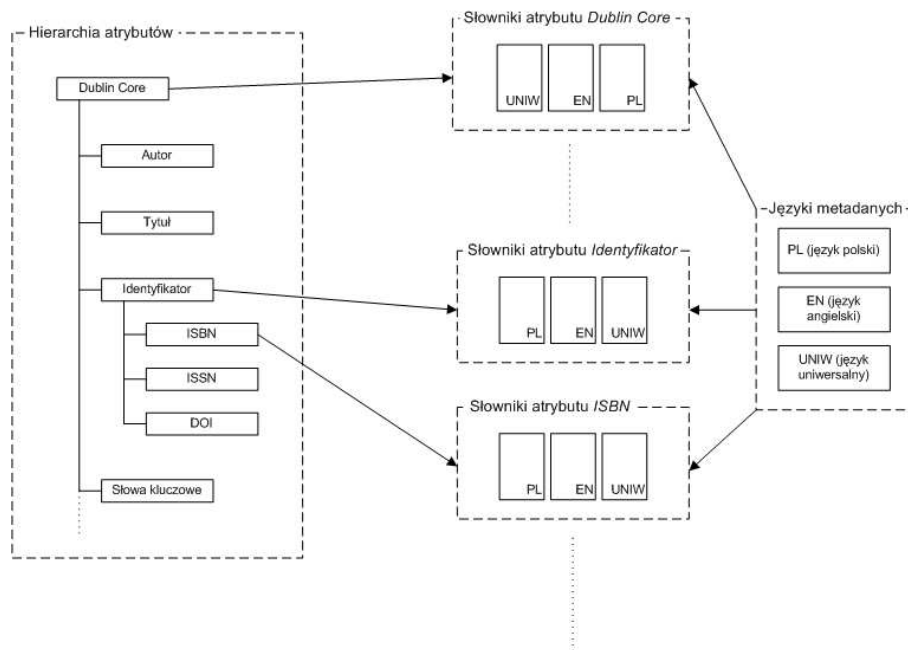
grupę wartości, w której znajduje się tylko ta nowa wartość. Słownikami wartości atrybutów można zarządzać wykorzystując moduł dLibra Redaktor. Zarządzanie słownikami obejmuje między innymi: tworzenie nowych grup wartości, tworzenie nowych wartości, przenoszenie wartości z jednej grupy do innej, modyfikację wartości. Przez to że opisy obiektów powiązane są z wartościami w słowniku, każda modyfikacja wykonana na słowniku wpływa na opisy obiektów, które wykorzystują modyfikowaną wartość lub wartości. Takie podejście pozwala zarządzać i kontrolować wartości wykorzystywane w opisach obiektów na poziomie całego repozytorium.

Obiekty cyfrowe w systemie dLibra mogą być opisywane metadanymi w różnych językach. Efektem jest to, że dany atrybut posiada tak naprawdę kilka słowników wartości – jeden słownik dedykowany jednemu językowi. Zestaw wykorzystywanych języków definiowany jest na poziomie całego systemu, przez co każdy obiekt cyfrowy w repozytorium może być opisany we wszystkich zdefiniowanych językach. Słownikiem danego języka zarządza się niezależnie od słowników dla innych języków. Dodatkowo w systemie zdefiniowany jest specjalny język „uniwersalny”. Język ten widoczny jest tylko dla użytkowników wprowadzających zasoby i ma na celu przyspieszenie pracy przy wprowadzaniu metadanych obiektu. Wszystkie wartości wprowadzone w języku „uniwersalnym” dla konkretnego obiektu, dodawane są automatycznie do metadanych tego obiektu w innych językach. Jeśli zatem istnieją wartości metadanych, które powinny być wprowadzone we wszystkich językach, np. identyfikator w systemie zewnętrznym to wystarczy wprowadzić je w języku uniwersalnym.

4.2. Integracja systemu dLibra z innymi systemami informatycznymi

Podstawowym mechanizmem pozwalającym na integrację repozytorium dLibra z innymi systemami informatycznymi jest mechanizm wymiany metadanych, wykorzystujący protokół OAI-PMH. System dLibra w pełni wspiera ten protokół, działając zarówno jako repozytorium (ang. repository) oraz klient (ang. harvester) OAI-PMH. Jako repozytorium OAI-PMH system dLibra udostępnia metadane opisowe obiektów cyfrowych, które przechowuje. Metadane mogą być prezentowane w schemacie Dublin Core 1.1 oraz w schemacie konkretnej instalacji systemu dLibra. Dodatkowo każda kolekcja obiektów cyfrowych utworzona w konkretnej instalacji systemu dLibra prezentowana jest poprzez protokół OAI-PMH jako zdefiniowany zbiór obiektów. Dostęp do tych metadanych jest otwarty, co oznacza że każdy inny system informatyczny może z nich korzystać bez żadnych ograniczeń. Jako klient OAI-PMH system dLibra może pobierać metadane obiektów cyfrowych z innych repozytoriów OAI-PMH. Lista repozytoriów OAI-PMH z których mają być pobierane metadane definiowana jest przez administratora systemu. Wszystkie metadane obiektów pobierane przez system dLibra z innych repozytoriów są przetwarzane (indeksowane), przez co mogą być przeszukiwane przez czytelników stron WWW repozytorium dLibra.

Użytkownicy, którzy wprowadzają obiekty cyfrowe do repozytorium, mogą korzystać z mechanizmu automatycznego importu metadanych, zapisanych w innym formacie niż ten zdefiniowany w systemie dLibra. Import metadanych możliwy



Rysunek 2. Zależności między metadany, językami oraz atrybutami

jest z następujących formatów: format komunikacyjny MARC 21, BibTeX, XML. W ramach XML możliwe jest importowanie opisów z szeregu konkretnych formatów, które wykorzystują standard XML jako formę prezentacji danych. Mechanizm importu metadanych bazuje na wtyczkach (ang. plugins), dzięki czemu w łatwy sposób można dołączyć do systemu wsparcie importu z innych formatów metadanych. Wtyczki udostępniane domyślnie w repozytorium dLibra są w pełni konfigurowalne, przez co mogą być w łatwy sposób zaadaptowane do konkretnych potrzeb określonej instytucji.

Użytkownicy, którzy korzystają z zasobów repozytorium dLibra, mają do dyspozycji (poza stronami WWW) szereg informacji udostępnianych w formie kanałów RSS. Kanały RSS dostępne w repozytoriach dLibra zawierają informacje takie jak: lista ostatnio dodanych obiektów cyfrowych do repozytorium, lista planowanych obiektów cyfrowych (obektów które mają być lub są w trakcie digitalizacji) oraz ostatnio dodane wiadomości na stronach repozytorium. Informacje znajdujące się w kanałach RSS mogą być także wykorzystywane przez inne systemy informatyczne (np. przez systemy agregujące informacje z różnych źródeł).

4.3. Sieć bibliotek cyfrowych w Polsce

Zrealizowanie zaawansowanych mechanizmów udostępniających obiekty cyfrowe znajdujące się w repozytoriach sieci PIONIER możliwe było w oprogramowaniu

dLibra dzięki wykorzystaniu protokołu OAI-PMH. Obecnie w sieci PIONIER uruchomionych jest około 20 repozytoriów cyfrowych opartych na oprogramowaniu dLibra, które tworzą polską platformę repozytoriów cyfrowych. Łącznie platforma ta zawiera obecnie ponad 80 000 obiektów cyfrowych.

Każde repozytorium oparte o oprogramowanie dLibra aktualizuje okresowo oraz przyrostowo metadane obiektów znajdujących się w pozostałych repozytoriach. W efekcie każde repozytorium dLibra posiada informacje o metadanych wszystkich obiektów cyfrowych znajdujących się w całej sieci repozytoriów. Bazując na tym procesie, każde repozytorium dLibra udostępnia wyszukiwanie w opisach obiektów cyfrowych rozproszonych w całej sieci połączonych repozytoriów. Użytkownicy końcowi repozytorium otrzymują jedną listę wyników, na której zasoby zdalne są specjalnie oznaczonymi hiperłączami i wskazują na konkretne obiekty w ich oryginalnych repozytoriach (patrz Mazurek, 2006).

W oparciu o wyszukiwanie zasobów rozproszonych zrealizowany został mechanizm dynamicznych kolekcji, który uzupełnia statycznie definiowane kolekcje. Statyczne kolekcje definiowane są przez administratorów systemu dLibra. Aby dany obiekt należał do kolekcji statycznej, użytkownik tworzący go, musi takie powiązanie wprost zdefiniować. Kolekcje dynamiczne bazują na zupełnie innym podejściu. Kolekcja dynamiczna definiowana jest indywidualnie przez użytkownika stron WWW dla jego potrzeb. Jest to zbiór obiektów cyfrowych, których metadane pasują do zdefiniowanego przez użytkownika zapytania wyszukiwawczego. Wyniki takiego zapytania stanowią kolekcję dynamiczną. Jeżeli w platformie repozytoriów pojawi się nowy obiekt cyfrowy, propaguje się on przez wszystkie repozytoria i automatycznie trafia do tych kolekcji dynamicznych, których zapytanie pasuje do jego metadanych. Kolekcje dynamiczne mogą być subskrybowane jako kanały RSS lub wyświetlane na stronie WWW repozytorium jako wyniki wyszukiwania pod specjalnie przygotowanym adresem WWW.

Obydwa wyżej opisane mechanizmy bazują na protokole wymiany metadanych OAI-PMH. Aby taka wymiana mogła nastąpić, wymagany był mechanizm unikalnej i trwałej w czasie identyfikacji obiektu. W systemie dLibra mechanizm ten zrealizowany został poprzez wykorzystanie formatu OAI identyfikator (patrz Lagoze, 2006). Format OAI identyfikator specyfikuje sposób konstruowania unikalnych identyfikatorów zasobów, znajdujących się w repozytoriach OAI-PMH. Wymaganiem tego formatu jest to, aby identyfikatory w ramach konkretnego repozytorium były unikalne. Składnia formatu OAI identyfikator jest ograniczeniem składni URI (patrz Berners-Lee, 1998). Identyfikator OAI składa się z schematu „oai” wskazującego na to, iż jest to identyfikator OAI, identyfikatora w przestrzeni nazw repozytoriów (będącego najczęściej nazwą domenową adresu WWW repozytorium) oraz identyfikatora lokalnego obiektu w konkretnym repozytorium. Przykładowy identyfikator OAI to „oai:www.wbc.poznan.pl:234”, gdzie „oai” to schemat identyfikatora, ciąg znaków „www.wbc.poznan.pl” to identyfikator w przestrzeni nazw repozytoriów, a „234” to identyfikator lokalny obiektu w repozytorium. Identyfikatory OAI mogą być również używane przez czytelników, jako identyfikatory do obiektów cyfrowych.

Protokół OAI-PMH jest publicznie dostępnym, otwartym standardem, który wykorzystywany jest przez wiele różnych systemów repozytoriów cyfrowych. Me-

tadane dostępne poprzez protokół OAI-PMH w repozytoriach dLibra mogą zatem być wykorzystywane przez inne systemy wspierające protokół OAI-PMH, co więcej, systemy te mogą być również dołączone do platformy bibliotek cyfrowych w sieci PIONIER.

5. Zakończenie

Oprogramowanie dLibra wykorzystywane jest przez wiele instytucji naukowych oraz publicznych w Polsce. Pozwala ono na budowanie skalowalnych, elastycznych oraz wygodnych w użyciu repozytoriów cyfrowych. Repozytoria dLibra posiadają zaawansowane mechanizmy zarządzania przechowywanymi obiektami oraz informacjami, które są z nimi powiązane. W Polsce repozytoria utworzone na bazie oprogramowania dLibra tworzą polską platformę repozytoriów cyfrowych, która udostępnia zaawansowane usługi związane z obiektami w platformie takie jak wyszukiwanie zasobów rozproszonych czy dynamiczne kolekcje obiektów.

Pomimo rozbudowanych możliwości, oprogramowanie dLibra jest stale udoskonalane na podstawie zebranych doświadczeń użytkowników oraz przeprowadzanych badań. Obecnie prowadzone są badania, które mają na celu identyfikację podstawowych usług repozytoriów cyfrowych oraz zdefiniowanie mechanizmów ich współdziałania oraz kompozycji. W najbliższym czasie rozwój oprogramowania dLibra będzie również koncentrował się na technologiach Web 2.0, pozwalając czytelnikom wspomagać tworzenie czy opisywanie zasobów repozytorium. W planach rozwoju oprogramowania dLibra uwzględnione są również prace związane z problemem długoterminowego przechowywania obiektów.

Literatura

- BERNERS-LEE, T. (1998) *Uniform Resource Identifiers (URI): Generic Syntax, RFC 2396*. The Internet Society. <http://www.ietf.org/rfc/rfc2396.txt>
- LAGOZE, C. and VAN DE SOMPEL, H. (2004)
- LAGOZE, C. and VAN DE SOMPEL, H. (2006) *Specification and XML Schema for the OAI Identifier Format*. Open Archives Initiative. <http://www.openarchives.org/OAI/2.0/guidelines-oai-identifier.htm>
- MAZUREK, C. and WERLA, M. (2005) Digital Object Lifecycle in dLibra Digital Library Framework. *Proceedings of DELOS 9th Thematic Workshop: Digital Repositories*. Kreta.
- MAZUREK, C. and WERLA, M. (2005) Distributed Services Architecture in dLibra Digital Library Framework. *Proceedings of 8th International Workshop of the DELOS Network of Excellence on Digital Libraries on Future Digital Library Management Systems*. DELOS.

MAZUREK, C., STROINSKI, M., WERLA, M. and WĘGLARZ J. (2006) Metadata harvesting in regional digital libraries in PIONIER Network. *Proceedings of TERENA Networking Conference 2006*. Włochy.

NISO (2003) *Information and documentation - The Dublin Core metadata element set, ISO Standard 15836-2003*. NISO.

The Open Archives Initiative Protocol for Metadata Harvesting. Open Archives Initiative.

<http://www.openarchives.org/OAI/openarchivesprotocol.html>.

dLibra – platform for building digital libraries

This article describes the dLibra software – platform for building digital libraries developed by Poznan Supercomputing and Networking Center. This article includes description of the most important functionality of the dLibra system, its architecture and core modules. Article describes advanced mechanisms which dLibra software provides such as management of digital objects metadata, integration of the dLibra system with other systems and mechanisms based on open metadata exchange protocol which is used by dLibra software installations in Poland.