

## Wykorzystanie regionalnej biblioteki cyfrowej do tworzenia repozytorium instytucjonalnego

JAKUB BAJER<sup>1</sup>, KRZYSZTOF OBER<sup>2</sup>

<sup>1</sup> *Biblioteka Politechniki Poznańskiej, jakub.bajer@put.poznan.pl*

<sup>2</sup> *Poznańska Fundacja Bibliotek Naukowych, krzychu@pfsl.poznan.pl*

### Streszczenie

Ogólnoświatowy trend otwierania dostępu do publikacji naukowych mobilizuje coraz więcej instytucji naukowych do tworzenia własnych repozytoriów cyfrowych. Instytucje planujące utworzenie własnego repozytorium stają przed wyborem platformy systemowej, na bazie której zbudowane zostanie repozytorium. Jednym z istniejących rozwiązań umożliwiającym zarządzanie obiektami cyfrowymi jest dLibra – polska platforma do budowy profesjonalnych repozytoriów, rozwijana przez zespół z Poznańskiego Centrum Superkomputerowo-Sieciowego (PCSS). Celem niniejszego opracowania jest pokazanie możliwości wykorzystania regionalnej biblioteki cyfrowej opartej na platformie dLibra do utworzenia naukowego repozytorium instytucjonalnego na przykładzie projektu repozytorium Politechniki Poznańskiej.

**Słowa kluczowe:** dLibra, repozytorium instytucjonalne, biblioteka cyfrowa, obiekt cyfrowy

### Wstęp

Obecnie coraz więcej instytucji naukowych zaczyna dostrzegać potrzebę tworzenia własnych repozytoriów cyfrowych. Jest to związane z ogólnoświatowym trendem otwierania dostępu do publikacji naukowych. Wiele instytucji nakłada wręcz na swoich pracowników obowiązek publikowania artykułów i wyników badań na zasadach Open Access. Posiadanie repozytorium instytucjonalnego jest istotne zarówno ze względów naukowych (dostępność rezultatów pracy, komunikacja naukowa), jak i marketingowych (rozpoznawalność i prestiż instytucji).

Biblioteka Politechniki Poznańskiej, realizując swoje zadania w zakresie wspomagania badań naukowych, rozpoczęła prace nad utworzeniem własnego repozytorium naukowego. Punktem wyjścia jest wybór platformy systemowej, na bazie której zbudowane zostanie repozytorium. Istnieją gotowe rozwiązania umożliwiające zarządzanie obiektami cyfrowymi. Jednym z nich jest dLibra – polska platforma do budowy profesjonalnych repozytoriów obiektów cyfrowych. dLibra – rozwijana przez zespół z Poznańskiego Centrum Superkomputerowo-Sieciowego – wykorzystywana jest obecnie do budowy regionalnych i instytucjonalnych bibliotek cyfrowych w Polsce.

Wywodząca się z Poznania dLibra została wykorzystana w realizacji projektu Wielkopolskiej Biblioteki Cyfrowej (WBC). Od początku istnienia WBC, czyli od października 2002 roku, swój udział w pracach nad jej rozwojem oraz współtworzeniem zasobów ma również Biblioteka Politechniki Poznańskiej. Biblioteka udostępnia za pośrednictwem WBC m.in. rozprawy habilitacyjne oraz skrypty dydaktyczne. Możliwości udostępniania tego typu materiałów są jednak ograniczone z powodu konieczności przestrzegania prawa autorskiego.

### Projekt repozytorium instytucjonalnego Politechniki Poznańskiej

Od 2002 roku Biblioteka Politechniki Poznańskiej tworzy zasób e-materiałów, udostępniając zdigitalizowane materiały dydaktyczne w ramach Wielkopolskiej Biblioteki Cyfrowej. Zdigitalizowane książki i inne publikacje znajdują się w dwóch kolekcjach: materiały dydaktyczne oraz dziedzictwo kulturowe.

Od roku akademickiego 2006/2007 biblioteka udostępnia (również w ramach WBC) elektroniczne wersje rozpraw habilitacyjnych oraz innych materiałów drukowanych ze swoich zbiorów najczęściej wykorzystywanych przez studentów i pracowników Politechniki Poznańskiej [podstawa prawna art. 28.3 ustawy z dnia 1 kwietnia 2004 r. o zmianie ustawy o prawie autorskim i prawach pokrewnych – Dz. U. Nr 91 poz. 869]. Uwarunkowania

prawne wymuszają jednak ograniczenie dostępu do w/w publikacji tylko do komputerów znajdujących się w sieci lokalnej biblioteki.

Podejmowane działania miały jak dotąd charakter doraźny i nie były w zadowalającym stopniu skoordynowane. Aby uporządkować organizacyjnie proces publikowania oraz umożliwić otwarty dostęp do zasobów gromadzonych w bibliotece cyfrowej, podjęto decyzję o utworzeniu repozytorium. W związku z tym dyrektor Biblioteki powołał zespół ds. repozytorium naukowego, którego zadaniem jest opracowanie projektu instytucjonalnego repozytorium naukowego Politechniki Poznańskiej (*Postanowienie Nr 1/2009 Dyrektora Biblioteki Głównej Politechniki Poznańskiej z dnia 5 stycznia 2009 w sprawie powołania w Bibliotece Politechniki Poznańskiej Zespołu ds. repozytorium naukowego*).

Głównym celem tego projektu jest bieżące dokumentowanie i archiwizowanie dorobku piśmienniczego i wydawniczego pracowników i doktorantów oraz prac dyplomowych studentów. Inne cele realizowane w ramach tego projektu to:

- zwiększenie dostępności najczęściej wykorzystywanych przez użytkowników materiałów dydaktycznych (podręczników, skryptów, wykładów, prezentacji i in.),
- ułatwienie dostępu do prac naukowych (rozpraw habilitacyjnych i doktorskich, prac dyplomowych i in.) naukowcom z kraju i zagranicy,
- zwiększenie efektywności pracy z materiałami przechowywanymi w bibliotece i archiwach Politechniki Poznańskiej, ale ze względów ochrony prawnej i bezpieczeństwa udostępnianych wąskiej grupie użytkowników,
- zintegrowanie udostępnionych materiałów dydaktycznych rozproszonych na serwerach jednostek organizacyjnych,
- stworzenie cyfrowych kopii najcenniejszych dzieł dziedzictwa kulturowego z zakresu dziejów techniki,
- obniżenie kosztów udostępniania źródeł informacji w bibliotece.

Aby osiągnąć powyższe cele, niezbędne jest określenie jednolitych zasad pozyskiwania, archiwizacji i udostępniania elektronicznych wersji prac naukowych (rozpraw doktorskich i prac dyplomowych) poprzez odpowiednie zarządzenie rektora Politechniki Poznańskiej.

Działania w zakresie publikowania e-materiałów podejmowane do tej pory przez Bibliotekę Politechniki Poznańskiej zakładały ich udostępnianie w WBC. W związku z tym zrodził się pomysł, aby naukowe repozytorium Politechniki Poznańskiej nadal funkcjonowało w WBC, ale jako wydzielona kolekcja.

### Wybór platformy systemowej

Pomysł wykorzystania WBC do tworzenia repozytorium instytucjonalnego implikuje wybór oprogramowania, na bazie którego powstanie repozytorium systemu dLibra. Wybór ten wiąże się z odrzuceniem innych platform open source przeznaczonych dla repozytoriów naukowych, w tym dwóch najbardziej popularnych: DSpace i Eprints (wykorzystywanych odpowiednio przez 35% i 16% repozytoriów na świecie – wg OpenDOAR, stan na dzień 20.09.2010). Mimo, że są to rozwiązania darmowe, dostarczane w pełnej zgodzie z założeniami *open source* i *open access*, to koszt przystosowania tego oprogramowania do potrzeb repozytorium instytucjonalnego Politechniki Poznańskiej oraz dodatkowych zasobów koniecznych do instalacji i utrzymania repozytorium byłby zdecydowanie wyższy. Czas potrzebny na uruchomienie repozytorium również uległby wyraźnemu wydłużeniu. Jednocześnie dLibra spełnia wymagania stawiane przed oprogramowaniem obsługującym repozytoria, m. in.:

- przechowywanie obiektów cyfrowych w dowolnych formatach,
- opisywanie obiektów cyfrowych za pomocą zestawów metadanych, które można definiować odpowiednio dla potrzeb repozytorium,
- wykorzystanie otwartych standardów komunikacji (OAI-PMH, RSS, SSO, LDAP, RDF, MARC, DublinCore), dzięki czemu możliwa jest współpraca z innymi systemami,
- zaawansowane funkcje wyszukiwania i przeglądania,
- udostępnianie zawartości repozytorium w sposób umożliwiający indeksowane i wyszukiwane obiektów cyfrowych za pomocą różnych wyszukiwarek internetowych (np. Google),

- rozbudowany mechanizm zarządzania prawami dostępu (hasła dostępu, autoryzacja IP),
- generowanie statystyk.

Oprogramowanie dLibra wywodzi się ze środowiska lokalnego i jest rozwijane z uwzględnieniem potrzeb użytkowników – w tym dodawanie funkcjonalności niezbędnych w nowych obszarach zastosowań. Rozwój ten i rozwiązywanie ewentualnych problemów technicznych nie generują dodatkowych kosztów.

### **Naukowe repozytorium instytucjonalne Politechniki Poznańskiej jako wydzielona kolekcja Wielkopolskiej Biblioteki Cyfrowej**

Naukowe repozytorium Politechniki Poznańskiej będzie funkcjonowało jako wydzielona kolekcja w działającym już projekcie Wielkopolskiej Biblioteki Cyfrowej. Za takim rozwiązaniem przemawia wiele argumentów:

- ośmioletnie doświadczenie w eksploatacji platformy;
- oprogramowanie dLibra jest rozwijane przez zespół PCSS w ramach współpracy z poznańskimi bibliotekami (biblioteki mają wgląd w prace rozwojowe i mogą wpływać na ich kierunek);
- brak dodatkowych kosztów; zakup sprzętu i oprogramowania realizowany jest przez Poznańską Fundację Bibliotek Naukowych (PFBN) ze składek członków Fundacji, licencja systemu dLibra jest nieodpłatna dla bibliotek zrzeszonych w PFBN;
- opisy bibliograficzne w WBC są indeksowane przez wyszukiwarki internetowe;
- minimalny okres wdrażania – uruchomienie repozytorium może odbyć się praktycznie natychmiast (część publikacji w WBC została już tam umieszczona na zasadach repozytorium – np. doktoraty Uniwersytetu Medycznego, czasopisma AWF i UAM, skrypty UAM, AWF, a także Politechniki Poznańskiej);
- WBC wraz z innymi bibliotekami cyfrowymi współtworzy Federację Bibliotek Cyfrowych – polską platformę rozproszonych bibliotek cyfrowych w sieci PIONIER, dającą dostęp do kilkudziesięciu tysięcy obiektów cyfrowych wraz z ich metadanymi oraz do zaawansowanych usług, takich jak wyszukiwanie rozproszone;
- obiekty cyfrowe w WBC są opisywane za pomocą metadanych w formacie Dublin Core; zestaw atrybutów opisu może być dowolnie modyfikowany dla potrzeb tworzonego repozytorium (np. dodanie atrybutu określającego dziedzinę, której dotyczy obiekt cyfrowy);
- dLibra współpracuje z systemami Single Sign-On (system pojedynczego logowania), dzięki czemu repozytorium będzie mogło być zintegrowane z używanym w Politechnice Poznańskiej systemem centralnego uwierzytelniania;
- dla potrzeb WBC został opracowany mechanizm wspomagający proces publikowania, obejmujący: konwersję dokumentów (Document Express), rozpoznawanie tekstu (ABBYY Recognition Server) i automatyczne umieszczanie publikacji na serwerze WBC; mechanizm ten zostanie wykorzystany przy tworzeniu naukowego repozytorium Politechniki Poznańskiej.

Utworzenie naukowego repozytorium Politechniki Poznańskiej jako wydzielonej kolekcji WBC nie spowoduje utraty jego autonomii. Repozytorium będzie można przeszukiwać za pomocą interfejsu WWW WBC, ograniczając zakres wyszukiwania do wybranej kolekcji (repozytorium). Stworzony zostanie również oddzielny interfejs wyszukiwawczy, który zapewni funkcjonalne wydzielenie repozytorium. Repozytorium, które fizycznie będzie stanowiło integralną część WBC, dla użytkownika końcowego będzie widziane jako odrębny system z własnymi mechanizmami przeszukiwania i przeglądania. Dzięki takiemu podejściu repozytorium będzie funkcjonalnie wydzielonym systemem, chociaż fizycznie będzie korzystało z zasobów sprzętowo-programowych wykorzystywanych przez WBC.

Dla potrzeb repozytorium w systemie dLibra zaimplementowane zostaną dodatkowe funkcjonalności, m.in.:

- integracja z bazą bibliografii *Biblio*, zawierającą opisy bibliograficzne dorobku naukowego pracowników Politechniki Poznańskiej – możliwość uzyskania w wyniku wyszukiwania w repozytorium kompleksowej informacji o dorobku piśmienniczym wybranego autora (pełne teksty publikacji oraz opisy bibliograficzne); rozważany jest również scenariusz umieszczania w repozytorium opisów bibliograficznych (bez zawartości

pełnotekstowej) oraz zaimportowania rekordów z bazy *Biblio* do repozytorium; takie opisy mogłyby zostać uzupełnione o odnośniki do pełnych tekstów w serwisach zewnętrznych oraz do usług pozwalających pozyskać teksty drogą tradycyjną, np. zamawianie kopii, zamawianie z magazynu zamkniętego (agregacja zasobów z różnych źródeł);

- mechanizm zgłaszania i przesyłania publikacji do repozytorium przez autorów (*self-archiving*) – możliwość dodawania i modyfikowania własnych publikacji przez uprawnionych użytkowników z poziomu aplikacji WWW; utworzone w ten sposób publikacje zostaną upublicznione dopiero po zatwierdzeniu przez moderatora (redaktora z odpowiednimi uprawnieniami); mechanizm ten umożliwi autorowi i moderatorowi wymianę informacji dotyczących statusu publikacji oraz ewentualnych nieprawidłowości;
- eksport do menedżerów bibliografii – możliwość przekazywania metadanych opisujących obiekt cyfrowy do najczęściej stosowanych programów służących do gromadzenia, porządkowania i analizowania danych bibliograficznych (np.: Zotero, EndNote, BibTex, RefMan, RefWorks);
- alternatywne schematy Dublin Core – możliwość definiowania zestawów pól metadanych dla różnych obszarów zastosowań oraz mechanizmy elastycznego zarządzania tymi zestawami (np. zestaw zawierający dziedzinę wiedzy, która jest elementem wyszukiwawczym specyficznym dla repozytorium);
- integracja z uczelnianym systemem uwierzytelniania użytkowników – możliwość pojedynczego logowania się (Single Sign-On) za pomocą danych użytkownika znajdujących się w systemie centralnego uwierzytelniania Politechniki Poznańskiej (eKonto);
- zapisywanie i wysyłanie wyników wyszukiwania za pomocą e-mail.

### Podsumowanie

dLibra w wersji 5.0 dostarcza nowe funkcjonalności, które częściowo pokrywają się z opisanymi wcześniej potrzebami (np. *self-archiving*). Ponadto w wersji 5.0 wprowadzono wiele usprawnień dodatkowo przemawiających za wyborem tej platformy do utworzenia repozytorium, np.:

- możliwość ustalenia kolejności wartości atrybutów opisu,
- zaawansowane zarządzanie kontrolowanymi słownikami atrybutów,
- automatyczne tworzenie miniatur obiektów cyfrowych,
- łatwiejsze przeglądanie danych w aplikacji administratora,
- optymalizacja wyszukiwania i udoskonalenie listy wyników wyszukiwania,
- ulepszone mechanizmy prezentacji opisów i treści obiektów cyfrowych w aplikacji czytelnika.

Realizacja pozostałych funkcji przydatnych w tworzeniu repozytorium znajduje się obecnie w różnych stadiach zaawansowania: od prac koncepcyjnych, po implementację i testowanie. Ostateczny wynik tych działań zależy od dalszych, szczegółowych uzgodnień pomiędzy zespołami z PCSS i Politechniki Poznańskiej.

Dotychczasowa współpraca z zespołem programistów platformy dLibra oraz pozytywne doświadczenia w eksploatacji systemu pozwalają założyć, że przedstawione w niniejszym opracowaniu rozwiązanie sprawdzi się w praktyce i znajdzie naśladowców wśród instytucji planujących utworzenie własnego repozytorium.