

Projekt EuropeanaLocal – europejski wymiar agregowania zasobów cyfrowych

MARCIN WERLA

*Poznańskie Centrum Superkomputerowo-Sieciowe
mwerla@man.poznan.pl*

Streszczenie

W ramach niniejszego artykułu przedstawiono koncepcję wielowarstwowego agregowania zasobów rozproszonych bibliotek cyfrowych na potrzeby ich integracji na poziomie europejskim. Koncepcja ta wykorzystywana jest przy tworzeniu Europeany – europejskiej cyfrowej biblioteki, muzeum i archiwum. W referacie opisano poszczególne poziomy tego typu międzynarodowych agregacji oraz pokazano korzyści i nowe możliwości płynące z ich tworzenia. Doświadczenia opisane w referacie związane są przede wszystkim z agregowaniem polskich zasobów cyfrowych w Federacji Bibliotek Cyfrowych oraz z pracami projektu EuropeanaLocal, związanymi z włączaniem zasobów lokalnych i regionalnych bibliotek cyfrowych do Europeany.

Słowa kluczowe: biblioteki cyfrowe, agregacja metadanych, Europeana, Federacja Bibliotek Cyfrowych

Wstęp

Tożsamość kulturowa społeczeństwa jest niezwykle istotna dla jego wewnętrznej jedności. Jest ona kluczowa dla stabilnej identyfikacji poszczególnych jednostek ze społeczeństwem, a więc i dla trwałości tego społeczeństwa [1]. Dla wytworzenia się tego typu tożsamości niezbędne jest istnienie (a więc i zachowanie) dziedzictwa kulturowego oraz możliwość dostępu do tego dziedzictwa. Obecnie problemy związane z dostępem do dziedzictwa kulturowego są związane zarówno z geograficznym rozproszeniem tego dziedzictwa, jak i z rozproszeniem osób nim zainteresowanych. Jednym ze sposobów rozwiązywania problemów związanych z zachowaniem i udostępnianiem dziedzictwa kulturowego jest wykorzystanie nowoczesnych technologii informatycznych. Dzięki procesom digitalizacji możliwe jest przechowywanie dziedzictwa kulturowego w postaci cyfrowej. Taka postać pozwala nie tylko na długoterminowe zachowanie dziedzictwa, ale również jego udostępnienie wszystkim zainteresowanym poprzez Internet.

W kolejnym rozdziale referatu przedstawiono jedną z największych, realizowanych obecnie, inicjatyw współfinansowanych przez Komisję Europejską, związaną z szerokim dostępem do dziedzictwa kulturowego. Jest to Europeana (<http://europeana.eu>), europejska cyfrowa biblioteka, muzeum i archiwum. Omówiono tu cele, jakie Europeana ma realizować oraz sposób ich realizacji. W pracy skoncentrowano się na wielopoziomowym modelu agregacji i udostępniania dziedzictwa kulturowego rozproszonego w całej Europie, w szczególności w instytucjach o charakterze lokalnym i regionalnym. Wspomniano również doświadczenia i plany związane z włączaniem do tego modelu zasobów dostępnych w polskich bibliotekach cyfrowych.

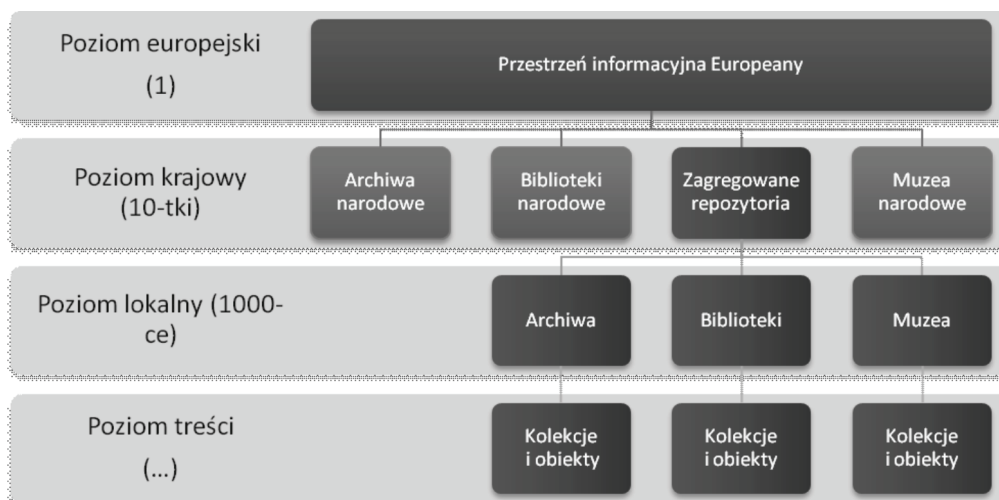
Europeana – Europejska cyfrowa biblioteka, muzeum i archiwum

Poszczególne kraje Europy, w ramach prowadzonych przez siebie strategii, programów i projektów digitalizacyjnych, utworzyły i nadal tworzą wiele serwisów internetowych, umożliwiających

dostęp do wybranych wycinków dziedzictwa kulturowego. Są to zazwyczaj konkretne kolekcje tematyczne lub szczególnie cenne zasoby wybranych instytucji. W liście do Komisji Europejskiej z 25 sierpnia 2005 roku władze sześciu krajów (w tym również Polski) zasygnalizowały potrzebę utworzenia wirtualnej biblioteki europejskiej, dającej szeroki i otwarty dostęp do europejskiego dziedzictwa kulturowego i naukowego. List ten był wynikiem wcześniejszej wypowiedzi dyrektora francuskiej biblioteki narodowej, który zwrócił uwagę na postępującą monopolizację Internetu przez firmę Google oraz masowe skanowanie i udostępnianie przez tę firmę prawie wyłącznie anglojęzycznych (w większości amerykańskich) książek i innych obiektów dziedzictwa kulturowego. Wobec skali poczynił Google, brak intensywnych i skoordynowanych działań w celu udostępnienia i popularyzacji w Internecie dziedzictwa innych kultur, mógłby doprowadzić w niedalekiej przyszłości do zaniku tych kultur w świecie cyfrowym. Postulaty zawarte we wspomnianym liście zaowocowały ogłoszeniem w 2005 roku inicjatywy Komisji Europejskiej nazwanej *i2010 Digital Libraries* [2], która dała początek współfinansowaniu przez Komisję Europejską serii projektów realizujących kolejne fragmenty Europejskiej Biblioteki Cyfrowej. Pierwszy prototyp Europeany, bo tak nazywa się owa biblioteka, został udostępniony publicznie w listopadzie 2008 roku pod adresem <http://europeana.eu>.

W opracowaniu specyfikacji funkcjonalnej, tj. dokumentu określającego szczegółowo, jakie funkcje ma Europeana realizować [3], twórcy Europeany założyli, że użytkownicy tego serwisu będą tam najczęściej trafiali z czterech podstawowych powodów. Pierwszy i podstawowy cel to rozrywka. Użytkownik Europeany oczekujący rozrywki nie będzie najprawdopodobniej szukał niczego konkretnego, ale raczej wprowadzał przypadkowe zapytania i „surfował” po wynikach od jednego do drugiego obiektu cyfrowego. Drugi scenariusz użycia Europeany to poszukiwanie informacji na temat konkretnego obiektu, osoby, wydarzenia etc. Osoby realizujące ten scenariusz będą w przybliżeniu wiedziały czego poszukują, mogą jednak nie być w stanie sformułować zapytań wyszukiwawczych zgodnych ze sposobem opisu obiektów dostępnych w Europeanie. Trzeci cel przyświecający użytkownikom Europeany to poszukiwanie informacji w związku z planowaną podróżą. Dotyczy to zarówno sytuacji, w której użytkownik szuka informacji na temat ciekawych obiektów znajdujących się w miejscu, do którego się udaje, jak i sytuacji, w której takie obiekty są już znane, a użytkownik poszukuje dodatkowych informacji na ich temat lub na temat obiektów znajdujących się w pobliżu. W tym scenariuszu należy zadbać również o użytkowników korzystających z urządzeń mobilnych i poszukujących informacji w kontekście lokalizacji, w której się w danej chwili znajdują. I wreszcie czwarty, najbardziej zaawansowany scenariusz wykorzystania Europeany, który zakłada, że użytkownik jest zainteresowany wymianą wiedzy i współpracą z innymi użytkownikami o podobnych zainteresowaniach, powiązanych z materiałami widocznymi w Europeanie. Może tu chodzić o wymianę wiedzy i współpracę zarówno na poziomie profesjonalnym, jak i amatorskim czy hobbystycznym.

Aby użytkownicy Europeany mogli realizować opisane powyżej scenariusze, niezbędne jest oczywiście utworzenie dla nich stosownych narzędzi. W tym celu Komisja Europejska finansuje projekty takie jak EDLnet (projekt już zakończony), Europeana v1.0 (<http://version1.europeana.eu/>) czy EuropeanaConnect (<http://www.europeanaconnect.eu/>). Jednak same narzędzia to nie wszystko. Aby usatysfakcjonować użytkowników, serwis Europeana musi oferować dostęp do dużej ilości odpowiednio ustandaryzowanych opisów obiektów rozproszonych w bibliotekach, archiwach i muzeach cyfrowych w całej Europie. Ogólną koncepcję realizacji tego celu, a więc i koncepcję architektury przepływu informacji w Europeanie przedstawiono na rycinie 1.



Ryc. 1. Schemat struktury agregacji metadanych na potrzeby Europeany

Europeana zakłada połączenie w jednej przestrzeni informacyjnej metadanych obiektów pochodzących z archiwów, bibliotek i muzeów, zarówno tych o charakterze narodowym, jak i regionalnym czy lokalnym. Współpraca na poziomie technicznym oraz organizacyjnym z tak dużą liczbą instytucji pochodzących z wielu różnych krajów europejskich jest bardzo trudnym zadaniem. Aby zminimalizować ryzyko wystąpienia problemów z tym związanych, zdecydowano się gromadzić informację w Europeanie za pomocą szeregu serwisów pośredniczących, nazwanych agregatorami. Usługi te odpowiadają za gromadzenie, uporządkowanie, znormalizowanie i wzbogacenie metadanych znajdujących się w rozproszonych bibliotekach, archiwach czy muzeach cyfrowych, a następnie przekazanie tak przetworzonych metadanych na wyższy poziom agregacji. W strukturze przedstawionej na rycinie 1 najwyższym poziomie agregacji jest Europeana, należy się jednak liczyć z tym, że w przyszłości Europeana będzie udostępniać gromadzone metadane innym, nadrzędnym czy równorzędnym serwisom.

W celu realizacji struktury współpracujących ze sobą serwisów internetowych powołano szereg projektów współfinansowanych przez Komisję Europejską. I tak obiekty z archiwów narodowych będą agregowane w projekcie APENet (<http://www.apenet.eu/>), a muzea narodowe połączy projekt ATHENA (<http://www.athenaeurope.org/>). Biblioteki narodowe już od kilku lat współpracują w ramach inicjatywy The European Library (<http://www.theeuropeanlibrary.org/>), początkowo w zakresie katalogów bibliotecznych, później również w obszarze obiektów cyfrowych. Przestrzeń informacyjną Europeany zdefiniowaną przez trzy wspomniane wyżej projekty dopełnia EuropeanaLocal, którą opisano w następnej części referatu.

EuropeanaLocal

Projekt EuropeanaLocal (<http://europeanlocal.eu>) rozpoczęto w czerwcu 2008 roku jako trzyletni projekt współfinansowany przez Komisję Europejską w ramach programu eContentPlus. Projekt ten ma formę „sieci najlepszych praktyk” (ang. *best practice network*). Jego celem jest opraco-

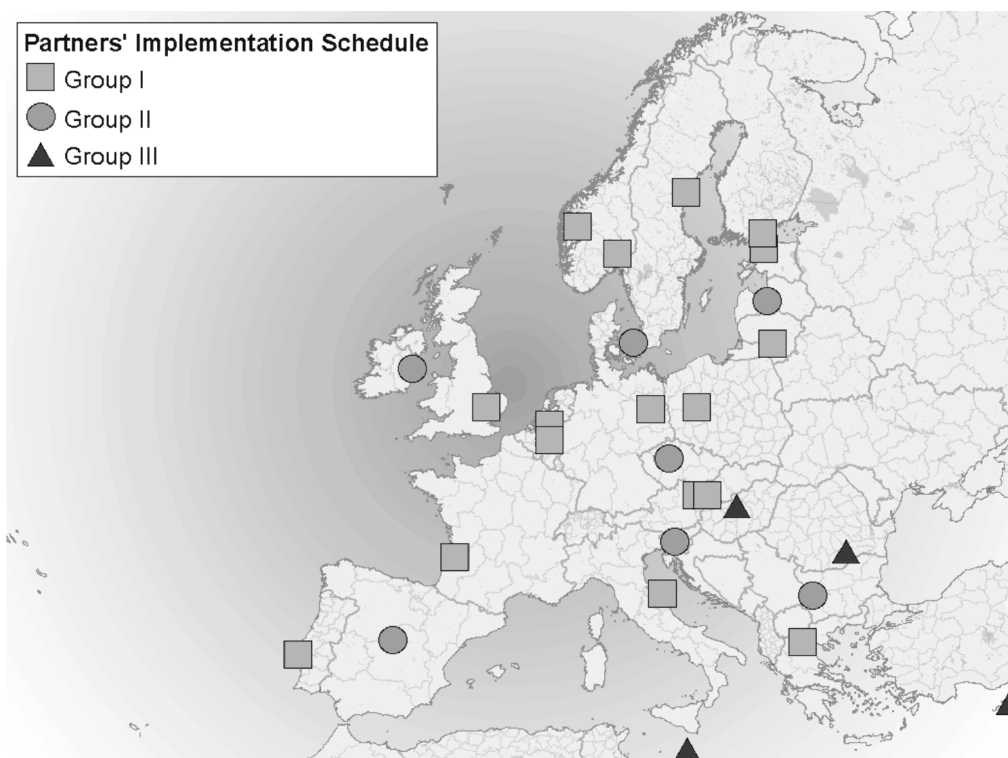
wywanie i propagowanie najlepszych praktyk związanych z tematyką projektu oraz praktyczne działania związane ze wspieraniem wdrażania tych praktyk. Projekt dotyczy poprawy interoperacyjności treści cyfrowych pochodzących z lokalnych i regionalnych bibliotek cyfrowych poprzez zwiększenie możliwości zautomatyzowanego użycia tych treści w zewnętrznych usługach sieciowych. Kolejnym celem jest wspieranie tworzenia usług agregujących obiekty cyfrowe z wielu różnych instytucji i doprowadzenie do utworzenia sieci regionalnych repozytoriów cyfrowych i usług agregujących, będących w stanie komunikować się z Europeaną.

Projekt EuropeanaLocal jest realizowany przez 32 partnerów z 27 krajów europejskich. Koordynatorem projektu jest instytucja Sogn og Fjordane County Municipality z Norwegii, a firma MDR Partners z Wielkiej Brytanii odpowiada za koordynację naukową i wsparcie zarządzania projektem. W projekcie uczestniczy ponadto czterech partnerów technicznych (w tym Fundacja EDL) oraz 27 partnerów pełniących rolę koordynatorów projektu w poszczególnych krajach. W Polsce koordynatorem jest Poznańskie Centrum Superkomputerowo-Sieciowe afiliowane przy Instytucie Chemii Bioorganicznej PAN.

Pierwsze zadania krajowych koordynatorów projektu dotyczą przeprowadzenia w poszczególnych krajach ankiety na temat wykorzystywanych schematów metadanych i słowników kontrolowanych oraz zakresu tematycznego obiektów, które mogą zostać w perspektywie kilkunastu miesięcy udostępnione w Europeanie. Kolejnym elementem będzie opracowanie planu utworzenia w danym kraju infrastruktury zgodnej z Europeaną, w tym określenie poziomu agregacji treści cyfrowych, który jest najważniejszy dla danego kraju (np. jeden agregator na cały kraj lub sieć agregatorów regionalnych). Inne zadania to m.in. opracowanie mapowania pomiędzy schematami metadanych stosowanymi w danym kraju a schematem wymaganym przez Europeaną, a także opracowanie lub wskazanie słowników, które ułatwią osiągnięcie zgodności metadanych obiektów cyfrowych opracowywanych w danym kraju z mechanizmami semantycznymi Europeany. W celu zapewnienia trwałości rezultatów projektu konieczne będzie również opracowanie zestawu wytycznych i/lub procedur pozwalających na przyłączenie się do wypracowanej infrastruktury bibliotekom cyfrowym, które powstaną już po zakończeniu projektu.

Aby doprowadzić do bezpośredniego połączenia z Europeaną, konieczne będzie spełnienie narzuconych przez nią wymagań technicznych. Jest to przede wszystkim wykorzystanie protokołu OAI-PMH [4] oraz specjalnego schematu metadanych o nazwie Europeana Semantic Elements (ESE) [5]. W podstawowym zakresie nie są to szczególnie wygórowane wymagania. Protokół OAI-PMH jest uznanym na świecie, otwartym standardem wymiany metadanych i wiele systemów do budowy bibliotek cyfrowych wspiera go automatycznie. Schemat ESE jest w zasadzie nieznacznie rozszerzoną wersją schematu Dublin Core [6], bardzo powszechnie stosowanego w świecie bibliotek cyfrowych. Metadane publicznie dostępnych obiektów cyfrowych wyrażone w schemacie ESE i przesłane za pomocą protokołu OAI-PMH będą dla Europeany podstawą do prezentacji informacji o tych obiektach użytkownikom końcowym. Gdy użytkownicy odnajdą interesujące ich obiekty w Europeanie i będą chcieli zapoznać się z ich treścią, zostaną automatycznie skierowani do odpowiednich stron biblioteki cyfrowej, z której pochodzi dany obiekt. To właśnie już na stronach tej biblioteki realizowany będzie dostęp do treści obiektu. Takie podejście spowoduje, że Europeana staje się serwisem reklamującym zasoby europejskich bibliotek cyfrowych, a nie serwisem, który te biblioteki ma zastąpić.

Harmonogram budowy opisanej powyżej infrastruktury oraz przyłączenia zasobów poszczególnych krajów do Europeany w ramach projektu EuropeanaLocal przedstawiono schematycznie na rycinie 2.



Ryc. 2. Plany przyłączenia zasobów cyfrowych poszczególnych krajów w ramach projektu EuropeanaLocal

Dzięki tworzonej konsekwentnie od 2002 roku infrastrukturze bibliotek cyfrowych, opartej na protokole OAI-PMH, zdecydowano o przyłączeniu Polski w pierwszej grupie państw (najpóźniej pierwszy kwartał 2010 roku). Kluczowym elementem tej decyzji było istnienie funkcjonującego już agregatora metadanych, jakim jest Federacja Bibliotek Cyfrowych (<http://fbc.pionier.net.pl/>) dostępna w Polsce od czerwca 2007 roku. Obecnie w Europeanie dostępne są tylko zasoby Cyfrowej Biblioteki Narodowej POLONA, przyłączonej poprzez serwis TEL. Zasoby te liczą około 10 000 obiektów, co stanowi mniej niż 0,5% wszystkich zasobów Europeany.

Zakończenie

W referacie przedstawiono koncepcję gromadzenia na poziomie europejskim informacji o cyfrowych reprezentacjach zasobów dziedzictwa naukowego i kulturowego, znajdujących się w tysiącach instytucji rozproszonych po całym kontynencie. Koncepcja ta opiera się na wielopoziomowej agregacji metadanych opisujących obiekty cyfrowe, organizowanej zarówno wg lokalizacji geograficznej (np. agregacja repozytoriów z danego kraju czy regionu), jak i wg rodzaju instytucji (np. agregacja

zasobów z bibliotek narodowych). Element znajdujący się na najwyższym poziomie opisanej struktury to Europeana. Dzięki bardzo dużemu wsparciu ze strony Komisji Europejskiej, ma ona szansę stać się wkrótce głównym punktem dostępu do cyfrowych zasobów kulturowych Europy. Ważne jest, aby podjąć intensywne działania, dzięki którym polska kultura będzie tam również w sposób znaczący reprezentowana.

Piśmiennictwo

- [1] M.S. Szczepański, *Społeczności lokalne i regionalne a ład kontynentalny i globalny*. [w:] *Kręgi integracji i rodzaje tożsamości. Polska Europa Świat*, red. W. Wesołowski, J. Włodarek, Wydawnictwo Naukowe Scholar, PTS, XII Ogólnopolski Zjazd Socjologiczny. Warszawa 2005: 121-140.
- [2] *Inicjatywa i2010 Digital Libraries*.
http://ec.europa.eu/information_society/activities/digital_libraries/doc/brochures/dl_brochure_2006.pdf
- [3] Dekkers Makx, Gradmann Stefan, Meghini Carlo, *Europeana Outline Functional Specification ver1.7*,
http://version1.europeana.eu/c/document_library/get_file?uuid=a9e29cb4-a9b3-462a-a43d-0b480c677088&groupId=10602
- [4] Lagoze Carl, Van de Sompel Herbert, *The Open Archives Initiative Protocol for Metadata Harvesting*,
<http://www.openarchives.org/OAI/openarchivesprotocol.html>
- [5] *Europeana Semantic Elements specifications v3.2.1*, http://group.europeana.eu/c/document_library/get_file?uuid=c56f82a4-8191-42fa-9379-4d5ff8c4ff75&groupId=10602
- [6] DCMI Usage Board, *DCMI Metadata Terms*, <http://dublincore.org/documents/dcmi-terms/>