

Projekt ENRICH – współpraca europejska wokół utworzenia wirtualnego środowiska badań i prezentacji historycznego dziedzictwa kulturowego

GRAŻYNA PIOTROWICZ

*Biblioteka Uniwersytecka we Wrocławiu
grazyna.piotrowicz@bu.uni.wroc.pl*

Streszczenie

W referacie wstępnie zaprezentowano szerszy europejski kontekst powstania potencjału bibliotek cyfrowych, w odniesieniu do inicjatyw finansowych podejmowanych w ramach Unii Europejskiej. Na tym tle szczegółowo przedstawiono główne cele i uczestników projektu ENRICH. Następnie dokonując opisu planowanych działań omówiono poszczególne pakiety zadań, plan implementacji projektu oraz stan zaawansowania prac. Na zakończenie opisano zasady współpracy Biblioteki Cyfrowej Uniwersytetu Wrocławskiego (BCUWr) z czeską biblioteką cyfrową Manuscriptorium, w ramach realizowanego projektu ENRICH.

Słowa kluczowe: digitalizacja, historyczne dziedzictwo kulturowe, biblioteki cyfrowe, projekt ENRICH, wirtualne środowisko badań.

Wprowadzenie

Od 1 czerwca 2005 roku, gdy Komisja Europejska zaakceptowała inicjatywę *i2010 – a European Information Society for Growth and Employment*, zaproponowaną przez Vivienne Reding, udostępnianie w Internecie zasobów europejskich bibliotek i archiwów zaczęło mieć kluczowe znaczenie dla negocjowania nowej europejskiej tożsamości [1].

Europejska Biblioteka Cyfrowa – portal Web, który zapewnia skonsolidowany dostęp cyfrowy do materiałów z pojedynczych bibliotek [2] – stał się technologiczną manifestacją idei globalnego dostępu do lokalnych kolekcji. Europejska przestrzeń dziedzictwa tworzona jest w oparciu o wizję uniwersalnego dostępu do europejskiej pamięci zbiorowej, natomiast administracyjnie „europejska przestrzeń dziedzictwa” definiowana jest poprzez nałożenie ram Unii Europejskiej na lokalny kontekst instytucjonalny.

Różnorodność jako istota dzisiejszej polityki kulturalnej opiera się na podstawie dziedzictwa poszczególnych narodów, które z kolei przywiązują coraz większe znaczenie do kolekcji narodowych i lokalnych [3]. Procesy, które spowodowały wzrost masy krytycznej zdigitalizowanego materiału dotyczącego dziedzictwa wiążą się z powstawaniem i rozwojem bibliotek cyfrowych.

Powstawanie bibliotek cyfrowych (elektronicznych, wirtualnych) było silnie związane z różnorodnymi programami dofinansowywania. We wczesnych latach dziewięćdziesiątych XX wieku potencjał bibliotek cyfrowych był wzmacniany przez szereg inicjatyw finansowych w Unii Europejskiej, Digital Library Initiatives (DLI) w USA, czy Electronic Libraries Programme (eLib) w Wielkiej Brytanii. Przez to planowe dofinansowywanie projekty powstające w ich ramach stały się bardziej zauważalne niż wiele następnych inicjatyw o podobnym charakterze, gdyż wokół nich skupiała się

dyskusja nt. bibliotek cyfrowych i oczekiwanych w nich zmian. Programy te miały różnie rozłożone akcenty i dostarczyły zarówno możliwości uczenia i doksztalcania się dla uczestniczącego w nich personelu, jak i ogólnych ram odniesienia do dyskusji toczącej się w środowisku. Wśród prac zrealizowanych w ramach wymienionych programów można było wyodrębnić dwa przewijające się i wzajemnie połączone nurty tematyczne: 1) odkrywanie by dostarczać (*discovery to delivery* – D2D) oraz 2) repozytoria i zarządzanie treścią.

W grupie prac należących do pierwszego nurtu przyglądano się usługom, pozwalającym na mediowanie dostępu pomiędzy użytkownikiem a rozproszonymi zasobami, usprawniającym łańcuchy dostawcze i skupiającym się na integracji procesów zorientowanych na użytkownika. Tu grupowały się zagadnienia dotyczące odkrywania zasobów, łączenia się z nimi i zamawiania. W czasie lat pojawiały się też tematy dotyczące wyszukiwania i metawyszukiwania, portali, pośredników, federacji, a potem też harvestingu i rozdzielczości. Było wiele prac nt. formatów metadanych oraz protokołów wyszukiwawczych.

W obrębie drugiej grupy prac przywiązywano wagę do zarządzania dużymi repozytoriami treści cyfrowych i udostępniania ich różnymi metodami. Akcent kładziony na prowadzenie tego typu prac z czasem spowodował wzrost dostępności materiałów cyfrowych, zarówno tych powstałych w wersji cyfrowej (*digital born*), jak i takich, które do wersji cyfrowej zostały przekonwertowane z innych formatów (*digital reborn*). Prace publikowane ostatnio w tej grupie koncentrują swe zainteresowania na integrowaniu infrastruktur pojawiających się repozytoriów z podlegającym ciągłej ewolucji aparatem „odkrywania by dostarczać” (D2D) [4].

Analizując proces powstawania bibliotek cyfrowych z perspektywy czasowej i w kontekście europejskim – można wyodrębnić w nim kilka okresów.

Pierwszy, przypadający na lata 1998-2002, identyfikuje pierwszą generację bibliotek cyfrowych i pokrywa się z inicjatywą DELOS, sfinansowaną przez 5. Program Ramowy Komisji Europejskiej – *Technologie Społeczeństwa Informacyjnego* (IST-FP5). Celem tej inicjatywy było dostarczenie bodźca do prowadzenia działalności sprzyjającej rozwojowi „przyjaznego dla użytkownika społeczeństwa informacyjnego”, treści i narzędzi multimedialnych oraz niezbędnych technologii i interfejsów [5]. W tym czasie biblioteki definiowały relacje pomiędzy nowym medium, jakim był Internet, a swymi kolekcjami i zakończyły zakrojone na szeroką skalę projekty digitalizacyjne. Okres ten może być postrzegany jako moment definiujący adaptację Internetu do tworzenia bibliotek cyfrowych oraz jego instytucjonalizację. Internet jako medium jest bowiem wyjątkowo zdolny do integrowania trybów komunikacji oraz różnych form treści. Procesy zachodzące w tym pierwszym okresie rozwoju bibliotek cyfrowych spowodowały wzrost masy krytycznej zdigitalizowanego materiału dotyczącego dziedzictwa.

Druga generacja bibliotek cyfrowych mniej więcej pokrywa się z 6. Programem Ramowym *Badań i Rozwoju Technologicznego* (FP6), który obejmował lata 2002-2006. Koncentrował się on na koordynacji działalności badawczej europejskich zespołów pracujących w dziedzinach związanych z bibliotekami cyfrowymi, z wyznaczonym celem nadrzędnym dotyczącym rozwoju technologii dla bibliotek cyfrowych następnej generacji [6]. Zgodnie z propozycją Komisji Europejskiej z 2006 roku, dotyczącą zaleceń w sprawie zabezpieczania cyfrowego – digitalizacja dziedzictwa kulturowego jest działaniem kluczowym [1]. Dokonując przeglądu działań dotyczących europejskich bibliotek cyfrowych w ciągu ostatniej dekady (1995-2005) Jia Liu stwierdza, że przedmiotem ich zaintereso-

wania było dziedzictwo kulturowe i że ukazały one rzeczywiście wysoką świadomość istnienia kolekcji specjalnych i książek rzadkich, często wiązały się z wysiłkami kooperatywnymi i zaprezentowały przejście od projektu do infrastruktury informacyjnej na poziomie biblioteki europejskiej [7].

Z kolei celem 7. Programu Ramowego (FP7 – 2007) był wzrost badań nad digitalizacją, zabezpieczeniem cyfrowym i dostępem do dziedzictwa kulturowego poprzez „centra kompetencji” w digitalizacji i zabezpieczaniu. Podobne cele przyświecają programowi eContentPlus, który w latach 2005-2008 wniósł wkład 60 milionów euro do wykorzystania dla poprawienia zdolności wzajemnej współpracy narodowych kolekcji cyfrowych, ułatwiania wielojęzycznego dostępu i użytkowania materiałów kulturowych [1].

Programy mające stworzyć rzeczywiste biblioteki cyfrowe w Europie obejmują projekty sfinansowane w ramach FP6 (2005), FP7 (2007) i eContentPlus Programme (2005-2008).

Rozpoczęcie inicjatywy i2010 dla Europejskiej Biblioteki Cyfrowej jest częścią ostatniej fali innowacji bibliotek cyfrowych. Te prace rozwojowe są wielowymiarowym procesem ze składnikiem innowacji zanurzonym w dyskurs toczący się wokół dziedzictwa kulturowego i ról oraz kompetencji instytucji kultury gromadzących zbiory, w odniesieniu do (coraz bardziej cyfrowego) dziedzictwa [8]. Rozpoczęcie inicjatywy i2010 dla Europejskiej Biblioteki Cyfrowej położyło nowy akcent na problemy wyrastające z digitalizacji i transformacji historycznego dziedzictwa w treści cyfrowe.

Projekt ENRICH

Projekt ENRICH (*European Networking Resources and Information concerning Cultural Heritage*) jest dwuletnim (1. 12. 2007 – 30. 11. 2009) projektem celowym typu agregacyjnego finansowanym w ramach Programu eContentPlus i realizowanym przez konsorcjum złożone z 18 partnerów europejskich. Wpisuje się on w pełni zarówno w założenia Programu eContentPlus, jak i wielowymiarowe działania innowacyjne wobec bibliotek cyfrowych przewidywane w inicjatywie i2010.

Główne cele

Głównymi celami projektu ENRICH są:

- utworzenie wirtualnego środowiska badawczego przeznaczonego do pracy nad historycznym dziedzictwem kulturowym, składającym się z: manuskryptów, inkunabułów, starodruków, historycznych materiałów archiwalnych, map, itp;
- adaptacja istniejącej platformy Manuscriptorium do udostępniania, wykorzystywania i harvestingu danych ze zdigitalizowanych manuskryptów i innych materiałów historycznych przechowywanych przez biblioteki, muzea i archiwa w krajach członkowskich Unii Europejskiej, krajach stowarzyszonych, kandydujących i innych krajach europejskich, która umożliwi wzajemną współpracę pomiędzy obiektami zdigitalizowanymi i kolekcjami;
- adaptacja istniejącej platformy Manuscriptorium, umożliwiająca wzajemną współpracę systemów istniejących u poszczególnych partnerów projektu dla autoryzowanego wykorzystywania treści cyfrowych;
- implementacja wielojęzycznych ontologii, umożliwiająca wyszukiwanie danych w językach lokalnych, ich pobieranie w językach źródłowych, indywidualne indeksowanie oraz tworzenie osobistych kolekcji i prezentacji.

Uczestnicy

Konsorcjum ENRICH składa się z 18 partnerów, pochodzących z 12 państw Europy. Wyróżnia się wśród nich partnerów odpowiedzialnych za aspekty techniczne oraz partnerów odpowiedzialnych za dostarczenie treści, przy czym część z nich występuje w podwójnej roli. Koordynatorem całego projektu jest Biblioteka Narodowa Republiki Czeskiej, koordynatorem technicznym jest AiP Beroun Ltd., a sprawy administracyjne wspiera Cross Czech Prague, a.s. Pierwsze dwie instytucje są także liderami pakietów zadań, podobnie jak Oxford University Computing Services, Università degli Studi di Firenze – Centro per la comunicazione e l'integrazione dei media, Instytut Matematyki i Informatyki w Wilnie, SYSTRAN Paris i Biblioteka Narodowa Hiszpanii. Pozostali partnerzy to: Københavns Universitet – Nordisk Forskningsinstitut, Biblioteca Nazionale Centrale di Firenze, Biblioteka Uniwersytecka w Wilnie, Biblioteka Uniwersytecka we Wrocławiu, Stofnun Árna Magnússonar í íslenskum fræðum in Reykjavík, Computer Science for the Humanities – Cologne University, St. Pölten Diocese Archive (Monasterium project, Austria), Narodowa i Uniwersytecka Biblioteka Islandii, Budapest University of Technology and Economics i Poznańskie Centrum Superkomputerowo-Sieciowe.

Ponadto, projekt wspierany jest przez szereg instytucji, wśród których znajduje się wiele posiadających cenne materiały dziedzictwa kulturowego, a zainteresowanie współpracą wyraziły między innymi biblioteki narodowe Węgier, Kazachstanu, Mołdawii, Polski, Rumunii, Serbii, Szwecji i Turcji, jak również biblioteki uniwersyteckie w Bratysławie, Bukareszcie czy Heidelbergu. Lista partnerów stowarzyszonych jest poszerzana podczas trwania projektu.

Konsorcjum ENRICH będzie również ściśle współpracować z Biblioteką Europejską (The European Library, TEL), a w momencie uruchomienia Europejskiej Biblioteki Cyfrowej (The European Digital Library, EDL) zasoby projektu ENRICH dostępne poprzez Manuscriptorium staną się jej integralną częścią.

Przesłanki dla realizacji

Inicjatywa taka jak TEL stanowiła ważne badania w dziedzinie tworzenia wspólnego dostępu do europejskich repozytoriów kultury. Jednakże TEL skupiała się raczej na repozytoriach wielkich organizacji narodowych, takich jak biblioteki narodowe, podczas gdy repozytoria lokalne nie mogły być jeszcze dostępne poprzez ten portal. W projekcie ENRICH biorą udział partnerzy odpowiedzialni za aspekty techniczne oraz za dostarczanie treści, a wielu z nich jest członkami inicjatyw: TEL i TELplus. Zarówno TEL, jak i TELplus są inicjatywami horyzontalnymi, dodającymi nowe funkcjonalności do tego, co w przyszłości powinno się stać Europejską Biblioteką Cyfrową (European Digital Library).

Projekt ENRICH jest natomiast przedsięwzięciem o podejściu wertykalnym, które wielu typom instytucji kultury pozwoli zintegrować swe repozytoria w ramach TEL. ENRICH umożliwi również ciągłe dołączanie metadanych i danych obrazowych pochodzących z rozproszonych zasobów do wspólnej platformy z ujednoliconym interfejsem, który w czasie rzeczywistym będzie współdziałał ze zdalnymi źródłami danych, włączając w to różnorodne banki obrazów zgromadzonych w oryginalnych lokalizacjach właścicieli treści.

Ważnym punktem odniesienia dla działań ENRICH-a były rezultaty projektu TEL-ME-MOR, który dostarczył raport na temat Aktywności Badawczej Europejskich Bibliotek Narodowych w Dziedzinie Dziedzictwa Kulturowego oraz Technologii Komputerowych i Informatycznych (lipiec

2006). Raport ten podaje ważne dane liczbowe związane z digitalizacją, bibliotekami cyfrowymi oraz manuskryptami. Stwierdza się w nim m.in., że europejskie biblioteki narodowe uczyniły niewielki postęp w reformatowaniu swych kolekcji analogowych, co oznacza, że znaczna część naszego dziedzictwa dokumentalnego w najbliższej przyszłości może zostać utracona. Ponadto, podkreśla się pilną konieczność mikrofilmowania i masowej digitalizacji materiałów znajdujących się na kwaśnym papierze oraz konieczność prowadzenia digitalizacji zabezpieczającej. Zauważa się też, że portale stanowią ciągle nowość dla bibliotek narodowych, a biblioteki cyfrowe znajdują się w nich w początkowej fazie rozwoju. Zwraca się uwagę na fakt, iż biblioteki narodowe zbyt mocno koncentrują się na swej tradycyjnej działalności (jak np. katalogowanie), a ich cyfrowe biblioteki oraz usługi, ani nie są w stanie zapewnić użytkownikom dobrej jakości, ani też nie są zorientowane na użytkownika.

Opis planowanych działań

Uczestnicy projektu na bazie czeskiej biblioteki cyfrowej Manuscriptorium (<http://www.manuscriptorium.com>) planują zapewnienie ciągłego dostępu do zdigitalizowanych dokumentów historycznego dziedzictwa kulturowego, znajdujących się w różnych europejskich instytucjach kultury oraz utworzenie wirtualnego środowiska badawczego przeznaczonego do pracy nad manuskryptami, inkunabułami, starodrukami oraz innymi dokumentami historycznymi, które zostanie bezpośrednio podłączone do TEL (The European Library). Ponadto, planowana jest praktyczna walidacja możliwości i określenie warunków dla zintegrowania pod interfejsem biblioteki cyfrowej Manuscriptorium istniejących, lecz rozproszonych, treści cyfrowych poprzez wzbogacanie metadanych oraz koordynację pomiędzy heterogenicznymi metadanymi i standardami danych. Główną innowacją ENRICH-a jest wspólny, łatwy w użytkowaniu interfejs, który umożliwi koncentrację rozproszonych zasobów w jednym w swoim rodzaju środowisku badawczym i wyszukiwanie danych ze zdalnych serwerów. Właściciele repozytoriów (biblioteki, archiwa) mogą wykorzystywać swe własne istniejące już platformy. Platforma Manuscriptorium będzie praktycznie weryfikować możliwość wykorzystywania istniejących danych ze zdalnych serwerów dla ciągłego dostępu. Manuscriptorium jest środowiskiem posiadającym wiele rodzajów wartości dodanej; prezentuje ono bezpośrednio materiał badawczy, podczas gdy istniejące rozwiązania portalowe (jak TEL) dostarczają jedynie odnośników do kolekcji, bez możliwości nawigowania poprzez wirtualne reprezentacje ich pozycji. ENRICH i platforma Manuscriptorium umożliwią dzielenie się repozytoriami danych i szczegółowymi metadanymi oraz wyszukiwanie poprzez pojedynczy punkt dostępu, bez potrzeby przełączania się na inne strony domowe z różnymi interfejsami, wykonanymi w różnych językach. Ponadto, platforma Manuscriptorium umożliwi tworzenie spersonalizowanych metadanych i pakietów danych, np. dla celów badawczych i edukacyjnych (tworzenie nowych dokumentów cyfrowych przez połączenie istniejących części danych i metadanych).

Ponieważ Manuscriptorium stanowi platformę, która jest już wykorzystywana – zostanie utworzona nowa, prototypowa i równoległa platforma Manuscriptorium (tzw. klon Manuscriptorium), by tworzyć i testować na niej nowe narzędzia do harvestingu zdalnych danych. Platforma ta pozwoli na testowanie nowo przetworzonych danych (np. przetłumaczonych automatycznie). Narzędzia i własności podlegające sukcesywnemu testowaniu i walidacji będą stopniowo implementowane na istniejącej i działającej już platformie, czyli powszechnie udostępniane.

Możliwości wykorzystania nowoczesnych narzędzi do automatycznego tłumaczenia metadanych i wielojęzycznych wyszukiwarek będą testowane na istniejących danych i metadanych dostarczanych

przez wszystkich, uczestniczących w projekcie, właścicieli treści. W ramach rozwoju i implementacji zasad wyszukiwania używanych w Manuscriptorium – zostaną poddane ocenie zasady oparte na wielojęzycznych ontologiach dostarczonych przez SYSTRAN Paris, który opracował je podczas realizacji unijnego projektu VICODI. Ta ocena pozwoli na wybór najbardziej odpowiednich technologii wyszukiwawczych spełniających potrzeby użytkowników końcowych. Takie rozwiązanie pozwoli użytkownikom prowadzić wyszukiwanie w swych lokalnych językach.

Bardzo ważną cechą projektu ENRICH jest integracja różnych typów reprezentacji dziedzictwa kulturowego (rekordy z opisami katalogowymi, obrazy cyfrowe, ustrukturyzowane pełne teksty, dokumenty audialne itp.). Zasada integracji polega na centralizacji opisowych metadanych w bibliotece cyfrowej Manuscriptorium i dystrybucji danych wśród innych zasobów w wirtualnym środowisku sieciowym. Rezultatem pracy będzie ciągły dostęp do europejskiego dziedzictwa kulturowego, w obszarze historycznych kolekcji książek i dokumentów archiwalnych.

Obecna postać Manuscriptorium to rezultat 15 lat prac prowadzonych wspólnie przez AIP Beoun Ltd. oraz Bibliotekę Narodową Republiki Czeskiej. Aktualnie serwis ten udostępnia ponad milion zdigitalizowanych stron, co czyni go najbogatszym źródłem cyfrowych postaci manuskryptów w Europie. Podstawą funkcjonowania Manuscriptorium są dane przechowywane w formacie XML. Ich istotną częścią są opisy manuskryptów przygotowane w formacie MASTER. Do obecnie zgromadzonych danych zostaną dołączone kolejne rekordy pobrane z systemów uczestników projektu przy pomocy protokołu OAI-PMH. Rekordy te będą zawierały zarówno opisy manuskryptów, jak i odnośniki do ich cyfrowych reprezentacji. W celu umożliwienia automatycznego przepływu informacji opracowane również zostaną specjalne procedury integracyjne, a dla instytucji, które obecnie nie mają technicznych możliwości do samodzielnego udostępniania swoich zasobów – specjalne narzędzia online.

Docelową grupę użytkowników projektu ENRICH stanowią biblioteki, muzea, archiwa, pracownicy naukowcy i studenci oraz wszyscy inni zainteresowani. Projekt pozwoli im na przeszukiwanie i dostęp do dokumentów, które są trudno osiągalne w inny sposób. Poza obrazami zdigitalizowanych dokumentów możliwy będzie również dostęp do innych typów obiektów, takich jak zapisane w formacie TEI transkrypcje czy pliki audio oraz wideo. Użytkownicy będą mieli również do dyspozycji narzędzia, które pomogą im tworzyć własne dokumenty i osobiste biblioteki cyfrowe. Narzędzia te będą również wspierały wielojęzyczność, w tym także wielojęzyczne ontologie, umożliwiające wyszukiwanie danych w języku lokalnym, a ich pobieranie w języku źródłowym.

Pakiety zadań

Projekt ENRICH realizowany jest w ramach 8 pakietów zadań.

Pakiet zadaniowy nr1 (WP1) – *Zarządzanie projektem:*

- 1.1 Kierowanie techniczne;
- 1.2 Administrowanie projektem i monitorowanie zasobów;
- 1.3 Koordynacja finansowa;
- 1.4 Koordynacja administracyjna;
- 1.5 Zapewnienie jakości.

Pakiet zadaniowy nr 2 (WP2) – *Przygotowanie do implementacji systemu i rozszerzania treści:*

- 2.1 Badanie dostępności zdigitalizowanych manuskryptów i innych materiałów historycznych;
- 2.2 Warsztaty partnerów technicznych oraz partnerów odpowiedzialnych za treści;

2.3 Dialog w celu przygotowania technicznych warunków do komunikacji wewnątrz systemu;

2.4 Definicja standardów i rozwój metadanych.

Pakiet zadaniowy nr 3 (WP3) – *Standaryzacja wspólnie użytkowanych metadanych*:

3.1 Konwersja między platformami do opisu manuskryptów TEI P4 i TEI P5;

3.2 Implementacja harwestera OAI-PMH do Manuscriptorium;

3.3 Rozszerzenie wewnętrznego środowiska Manuscriptorium przez implementację konteneryzacji METS dla systemu Manuscriptorium;

3.4 Poprawa i uogólnienie wykorzystania UNICOD-u w Manuscriptorium.

Pakiet zadaniowy nr 4 (WP4) – *Personalizacja użytkownika*:

4.1 Analiza potrzeb typowych użytkowników końcowych, w odniesieniu do implementacji wspólnych kolekcji do Manuscriptorium;

4.2 Tworzenie indywidualnych kolekcji dla użytkowników końcowych;

4.3 Tworzenie wirtualnych dokumentów dla celów badawczych i edukacyjnych;

4.4 Ocena zdolności wyszukiwania bibliograficznego w Manuscriptorium;

4.5 Implementacja możliwości głębokiego wyszukiwania we wszystkich metadanych oraz danych tekstowych: poprawa jakości wyszukiwania zarówno w metadanych, jak i w pełnych tekstach.

Pakiet zadaniowy nr 5 (WP5) – *Personalizacja dla dostawców*:

5.1 Narzędzia online do strukturyzacji istniejących metadanych i danych związanych z manuskryptami;

5.2 Analiza i rozwój narzędzi do wykorzystania wielkich zewnętrznych zbiorów danych;

5.3 Pilotażowa implementacja wielkich zbiorów danych dostarczonych do struktur Manuscriptorium przez wybranych partnerów;

5.4 Integracja zewnętrznych danych do Manuscriptorium.

Pakiet zadaniowy nr 6 (WP6) – *Wielojęzyczny, przyjazny dla użytkownika dostęp*:

6.1 Rozwój dostępu wielojęzycznego;

6.2 Zaprojektowanie i wykorzystanie Arkusza Stylów Przekładu;

6.3 Implementacja VICODI.

Pakiet zadaniowy nr 7 (WP7) – *Ocena, testowanie i walidacja*:

7.1 Zdefiniowanie strategii oceny;

7.2 Testowanie i ocenianie dostępności, użyteczności oraz zdolności adaptacyjnych rozwiniętych aplikacji;

Pakiet zadaniowy nr 8 (WP8) – *Rozpowszechnianie i eksploatacja*:

8.1 Przygotowanie planu rozpowszechniania;

8.2 Materiały reklamowe;

8.3 Aktualizacja i utrzymywanie strony domowej;

8.4 Przygotowanie planu eksploatacji;

8.5 Zorganizowanie konferencji końcowej.

Plan implementacji projektu oraz stan zaawansowania prac

Plan prac podzielony jest na cztery główne obszary:

- 1) przygotowanie do implementacji systemu i poszerzania treści,
- 2) opracowanie rozwiązań technicznych lub modyfikacja i implementacja,

- 3) walidacja i ocena,
- 4) rozpowszechnianie i eksploatacja.

ad 1) *Przygotowanie do implementacji systemu i poszerzania treści.*

Ta część projektu jest realizowana w ramach pakietu zadaniowego nr 2 (WP2). Jej głównym celem jest zidentyfikowanie różnych wymagań użytkowników dotyczących pracy z piśmienniczym dziedzictwem kulturowym (średniowieczne manuskrypty lub inne dokumenty historyczne) w zintegrowanym środowisku i ustalenie zasad praktycznej współpracy dotyczącej zbierania heterogenicznych danych. W końcu – określenie formalnych zasad gromadzenia danych. Pierwszym krokiem w tych działaniach było przeprowadzenie wśród partnerów projektu odpowiedzialnych za dostarczanie treści – ankiety nt. dostępności zdigitalizowanych materiałów historycznych oraz określenie przez nich stopnia, w jakim chcą dzielić się danymi z Manuscriptorium (np. harvesting i odnośniki do oryginalnych źródeł z możliwością wyświetlania oryginalnych danych poprzez ten sam interfejs, utworzenie struktur kompatybilnych z Manuscriptorium wraz z narzędziami Manuscriptorium i rozproszone przechowywanie danych, wymagania dotyczące transformacji metadanych, reżimy dotyczące dostępności, wymagania lingwistyczne, itd.). Po ocenie i analizie rezultatów ankiety w trzecim miesiącu trwania projektu przeprowadzono warsztaty partnerów technicznych oraz partnerów odpowiedzialnych za dostarczanie treści i określono początkowy pakiet wymagań dla projektu ENRICH, który miał dotyczyć też partnerów stowarzyszonych. Dalszy dialog między grupą partnerów technicznych i odpowiedzialnych za dostarczanie treści ma na celu przygotowanie technicznych warunków dla określonego sposobu komunikacji wewnątrzsystemowej takich, jak harvesting w oparciu o protokół OAI-PMH, rozproszone przechowywanie danych, wytwarzanie dokumentów kompatybilnych z Manuscriptorium. Planuje się też określenie typu i zakresu metadanych opisowych, administracyjnych, zabezpieczających i strukturalnych, które będą rozwijać partnerzy odpowiedzialni za dostarczanie treści.

ad 2) *Opracowanie rozwiązań technicznych lub modyfikacja i implementacja.*

W oparciu o rezultaty WP2 grupa techniczna ma przeprowadzić szereg modyfikacji mających spełnić wymagania partnerów dostarczających treści oraz pozwolić na pełne działanie i wzajemną współpracę wszystkich zaprojektowanych składników systemu. Techniczna część projektu składa się z WP3, WP4, WP5, WP6. W ramach WP3 została przygotowana formalna specyfikacja TEI (w oparciu o TEI Guidelines release P5) dla schematu używanego w Manuscriptorium. Powstają narzędzia systemowe, które pomogą w konwersji rekordów powstałych dla wcześniejszej specyfikacji MASTER do nowego schematu TEI P5. Obecnie trwają prace nad zaimplementowaniem do Manuscriptorium harvestera protokołu OAI-PMH, który zapewni harvesting metadanych z zasobów zarządzanych w ramach Manuscriptorium. W ramach WP4 przygotowano ankietę online w językach partnerów projektu, złożoną z 23 pytań, która badała potrzeby i oczekiwania użytkowników bibliotek cyfrowych w kwestii tworzenia indywidualnych kolekcji oraz indywidualnych wirtualnych dokumentów. Przeanalizowane wyniki ankiety są wykorzystywane do realizacji zadań przewidzianych w ramach WP4. W obrębie WP5 trwają prace związane z przygotowaniem narzędzi online do strukturyzacji istniejących metadanych oraz analiza i rozwój narzędzi do wykorzystania wielkich zewnętrznych zbiorów danych. Jeśli chodzi o WP6 – SYSTRAN pracuje nad rozwojem dostępu wielojęzycznego oraz zaprojektowaniem i wykorzystaniem Arkusza Stylów Przekładu.

ad 3) *Walidacja i ocena.*

Prace dotyczące oceny, testowania i walidacji prowadzone są w ramach WP7. W oparciu o podstawowe zasady i kryteria oceny opracowane przez znane w świecie zespoły (np. W3C, Minerva Technical Guidelines, Abo Academy – Finland) – została opracowana Metodologia testowania i walidacji e-aplikacji, która zostanie wykorzystana do testowania i oceny dostępności, użyteczności i zdolności adaptacyjnych – aplikacji opracowanych w ramach projektu.

ad 4) *Rozpowszechnianie i eksploatacja.*

Prace tego typu prowadzone są w obrębie WP8. Do chwili obecnej opracowane zostały: Plan rozpowszechniania informacji o projekcie, Projekt wizualizacji i prezentacji projektu, witryna internetowa poświęcona projektowi (<http://enrich.manuscriptorium.com>). Ponadto, opracowano podstawowe informacje nt. projektu, z przeznaczeniem do opublikowania w prasie i rozpowszechniania innymi kanałami przekazu. Przygotowano także ulotkę w języku angielskim nt. Projektu ENRICH, która po przetłumaczeniu na języki lokalne może być wykorzystywana do celów promocyjnych. Część partnerów zorganizowała już warsztaty i/lub konferencje poświęcone projektowi.

Aktualny sposób współpracy Biblioteki Cyfrowej UW r z Manuscriptorium

W trakcie prac nad projektem ENRICH wynikała potrzeba pobrania do repozytorium projektu opisów stworzonych i przechowywanych w zasobie Biblioteki Cyfrowej UW r (BCUWr.). Z powodu różnicy pomiędzy używanymi formatami metadanych (ENRICH – MASTER, dLibra – Dublin Core) oraz formatem samych obiektów w bibliotece (ENRICH – pliki graficzne w formacie *jpg*, dLibra pliki w formacie *DjVu z ocr*) zespół programistów z PCSS stworzył narzędzia umożliwiające współpracę pomiędzy wspomnianymi systemami.

Proces integracji z Manuscriptorium wykorzystuje rozszerzenie protokołu OAI-PMH poprzez dynamiczne zbiory. Dynamiczne – a więc ustalane na podstawie określonych kryteriów, takich jak np.: typ publikacji, rok wydania, format, które powodują, że w miarę dodawania kolejnych publikacji do Biblioteki Cyfrowej UW r – zbiór „napęlnia się sam”, jeśli tylko któraś z nowo dodanych publikacji spełnia kryteria przynależności do niego.

Po stworzeniu mechanizmu do generacji dynamicznych zbiorów rekordów – kolejnym problemem do rozwiązania stał się format, w jakim rekordy należy zaprezentować drugiej stronie. W tym przypadku programiści projektu dLibra „opakowali” oryginalny schemat Dublin Core w schemat METS. W formacie tym oprócz samych metadanych typu tytuł, autor, rok wydania itd. zawarte są też informacje opisujące poszczególne pliki publikacji. Dodatkowy problem związany był z formatem samych obiektów w bibliotece cyfrowej, gdyż jak już wspomniano wcześniej – Manuscriptorium do prezentacji używa wyłącznie plików graficznych *jpg*. Rozwiązano go tworząc mechanizm konwersji, który przekształca pliki *DjVu*, używane w BCUWr., na pliki w formacie *jpg*.

Proces ten jest na tyle szybki, że nie ma potrzeby wcześniejszego tworzenia (na zapas) plików *jpg* do danej publikacji. Zamiast tego, tuż po otrzymaniu żądania danej publikacji – konwerter tworzy niejako „w locie” potrzebne pliki graficzne, a odnośniki do nich umieszcza w przygotowanym pliku METS. Wszystkie informacje o publikacji są widoczne po wybraniu podglądu źródła strony.

Całość działa na tyle płynnie, iż użytkownik Manuscriptorium nie odczuwa istotnej różnicy w przeglądaniu publikacji z BCUWr – choć ta jest tworzona „na jego żądanie”.

W analogiczny sposób mogą być w przyszłości prezentowane obiekty cyfrowe pochodzące z innych polskich bibliotek cyfrowych opartych na oprogramowaniu dLibra, które zechciałyby skorzy-

stać z zaproszenia do współpracy w charakterze partnerów stowarzyszonych, jakimi obecnie (listopad 2008) są już Biblioteka Uniwersytetu M. Kopernika w Toruniu i Biblioteka Uniwersytetu Zielonogórskiego.

Piśmiennictwo

- [1] Komisja Europejska. „Comission Unveils Plans for European Digital Libraries”. Document no. IP/05/1202. September 30, (2005). [dostęp 20 listopada 2008], Dostępny w World Wide Web: <http://europa.eu/rapid/pressReleasesAction.do?reference=IP/05/1202&format=HTML&aged=0&language=EN&guiLanguage=en>.
- [2] M. Rogers, *European Digital Library in 2010?*, Library Journal 131(6), (2006).
- [3] A. Prochaska, *National Collections, Global Collecting: The Responsibilities of Librarians as Collectors*, Libraries & Culture 37(1), 72-76 (2002).
- [4] T. Davenport, *The coming commodization of processes*. Harvard Business Review, 101-108. (June 2005).
- [5] European Commission, Fifth framework Programme of the European Community for Research, Technological development and Demonstration Activities (1998-2002), Europa, [dostęp 19 listopada 2008], Dostępny w World Wide Web: <http://europa.eu.int/comm/research/fp5.html>
- [6] DELOS Network of Excellence on Digital Libraries, 2005, [dostęp 19 listopada 2008], Dostępny w World Wide Web: <http://www.delos.info>
- [7] J. Liu, *Digital Library Activities in Europe: Brief Overview*. Journal of Educational Media and Library Sciences 42(4), 455-469 (June 2005); [dostęp 16 listopada 2008], Dostępny w World Wide Web: http://research.dils.tku.edu.tw/joemls/index_e.htm
- [8] M., E. Phillips, *What should we preserve? The Question for Heritage Libraries in a Digital World*. Library Trends 54(1), 57-71 (Summer 2005).