

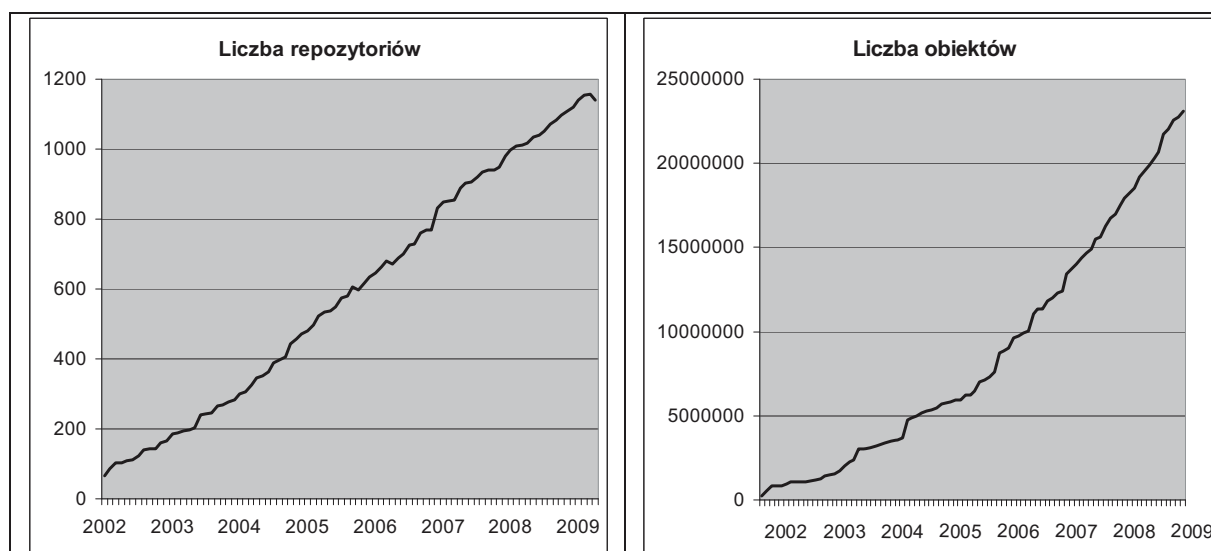
Władysław Marek Kolasa

Biblioteki cyfrowe na świecie – powstanie i rozwój

1. Wstęp

Biblioteki cyfrowe, jakkolwiek są zjawiskiem stosunkowo młodym, mają za sobą długą i rozbudowaną historią. Wprawdzie ich rozwój w obecnym kształcie liczy sobie dopiero niespełna dekadę, lecz wnikliwa analiza przekonuje, że poprzedziło ją szereg doświadczeń i eksperymentów, bez których obecna forma byłaby niemożliwa. W sensie porządkującym można wskazać kilka etapów, w których skryształizowały się kluczowe wizje i podjęto udane projekty. Pierwszy okres, który rozpoczął się od literackiej wizji G. H. Wellsa z 1937 roku trwał do początku lat 90. XX wieku; w czasie tym powstawały rozmaite koncepcje, budowano prototypy urządzeń i pojawiły się kolekcje eksperymentalne. Drugi okres trwał całą dziesiątą dekadę XX wieku i przesunął się o kilka lat na wiek XXI. W czasie tym powstały już pierwsze biblioteki cyfrowe, które szybko zaistniały w rodzącej się sieci Internet oraz podjęto badania podstawowe nad ich modelem i funkcją; w efekcie podjętych prac powstały pierwsze standardy i skryształizowały się grupy wsparcia. Trzeci okres związany immanentnie z ideą web 2.0 rozpoczął się u progu pierwszej dekady XXI wieku i trwa do dziś. Trudno wprawdzie wskazać konkretną datę, ale analiza danych czołowego największego światowego serwisu OAISTER (rys. 1) przekonuje, że proces ów rozpoczął się ok. 2001 roku i uległ przyspieszeniu w 2005.

Rys. 1, Biblioteki cyfrowe w serwisie OAISTER w latach 2002-2009 ¹



¹ Przeliczenia własne na podstawie OAIster Statistics – <http://www.lib.umich.edu/digital-library-production-service-dlps/oaister-statistics> [2011.05.19].

Ostatni okres można określić jako przejście do fazy dojrzałej; w sensie praktycznym charakteryzuje się on silnym trendem w kierunku budowy serwisów rozproszonych i orientacją w kierunku mediów społecznościowych.

2. Pierwsze próby i doświadczenia

Ideę biblioteki cyfrowej znaleźć można już w opublikowanych przed wojną esejach Herberta George Wellsa. W zbiorze *World brain* (Wells 1938) poświęconym rozważaniom na temat przyszłości organizacji wiedzy i edukacji zamieścił autor szkic *The Idea of a Permanent World Encyclopaedia*, w którym opisał ideę globalnej encyklopedii opartej na mikrofilmach. Miała ona wedle wizji autora obejmować uniwersum ludzkiej wiedzy i stać się motorem jednoczenia rasy ludzkiej. O krok dalej poszedł znany amerykański wynalazca Vannevar Bush, uważany za ojca idei sieci komputerowych i hipertekstu. Bush był zaniepokojony nadmiarem informacji, na jaki napotykał naukowcy i zaproponował kilka rozwiązań technologicznych, które w przyszłości mogłyby rozwiązać ten problem. W sławnym artykule *As We May Think* opublikowanym na łamach *Atlantic Monthly* w 1945 roku przedstawił ideę urządzenia Memex [Memory Extender], które składało się biurka (pulpitu) ze skośnymi ekranami, które umożliwiało wygodne czytanie. Mechanizm był swego rodzaju inteligentnym czytnikiem mikrofilmów, z których obrazy były inteligentnie wyszukiwane, dekodowane i transportowane na ekran z dużą prędkością, nade wszystko jednak były zdolne do automatycznego przechodzenia od zawartych w nich odniesień do innych prac, czyli hipertekstowe (Tedd & Large 2005). Kolejne ważne pomysły pojawiły się już w czasie, gdy istniały komputery; rozwijał je twórczo m.in. amerykański cybernetyk Joseph Licklider w systemem przetwarzania wiedzy *Procognitive* i uwiecznił m.in. w sławnej książce *Libraries of the future* (1965) (Hauben 2004). Zapewne także pod wpływem idei Busha był Douglas Engelbart, twórca licznych wynalazków (np. myszy komputerowej), uważany za pioniera w dziedzinie interfejsów. Uczony ten prowadził intensywne prace w zakresie hipertekstu i szukał rozwiązań w zakresie komunikacji naukowej dla zespołów geograficznie rozproszonych (Engelbart 1963); prace te uwieczniło pierwsze wdrożenie pod nazwą oN-Line System (NLS). W grudniu 1968 Engelbart z grupą 17 naukowców pracujących w Augmentation Research Center (Stanford Research Institute), przedstawił udaną 90-minutową prezentację NLS na żywo²; system był – warto dodać – wyposażony w interfejsy audio i

² Zapis filmowy eksperymentu – <http://sloan.stanford.edu/mousesite/1968Demo.html> [2011.05.20].

wideo. Niebagatelny wpływ na rozwój technologii miały także prace Theodora Nelsona prowadzone w ramach projektu Xanadu. Późniejszy autor terminu *hipertext* rozwijał w ramach projektu ideę *idealnej przestrzeni pamięci literackiej*, która była w swej istocie załączkiem obecnego systemu www. Koncepcje swe Nelson uwiecznił w dwu książkach: *Literary Machines* [Maszyny literackie] (1981) oraz *Computer Lib/Dream Machines* [Maszyny snów] (1974). Ukoronowaniem idei, którą rozpoczął Wells i Engelbart były podjęte w latach 1989-1991 prace Tima Bernersa-Lee dające podstawy teoretyczne i praktyczne współczesnych systemów hipertekstowych, czyli www. Przełomowe idee wyżej wspomnianych uczonych w sferze systemów informatycznych stworzyły grunt dla pierwszych doświadczeń w zakresie organizacji i manipulacji dużych ilości danych, które można uznać za prekursorów bibliotek cyfrowych (Tedd & Large 2005).

Odrębny tor rozważań obejmował rozwój technologii (głównie sprzętu) oraz realizowane w tym czasie pierwsze projekty biblioteczne. Warto zauważyć, że aż do końca lat 70. dostępna infrastruktura informatyczna była dopiero w fazie szkieletowej, co nie tworzyło sprzyjających warunków do jej wykorzystania. Przypomnijmy też, że aż do początku lat 80. dominowały komputery klasy mainframe, które były używane jedynie w dużych ośrodkach badawczych; niewielka była także liczba komputerów wpiętych do sieci (23 hosty w 15 węzłach w 1971, która wzrosła do 1000 dopiero w 1984), zaś przepustowość sieci niska. Ogromnym problemem było też przechowywanie danych; pierwsze dyski produkcyjne twarde (Segate) pojawiły się w 1980, a płyty CD dopiero na przełomie 1982 i 1983 roku. W efekcie potencjalna liczba użytkowników był niewielka. Mimo owej niesprzyjającej aury już w 1971 roku pojawił się pierwszy projekt biblioteki cyfrowej – student Uniwersytetu w Illinois Michael Hart rozpoczął tworzenie zbioru tekstów, które z czasem przerodziły się w *Gutenberg Project* [<http://www.gutenberg.org/>]. Z uwagi na brak uregulowań prawnych Hart postanowił umieszczać jedynie utwory z domeny publicznej, które aż do początku lat 90. rozpowszechniał pocztą elektroniczną w formacie tekstowym. Ręczne przepisywanie było pracochłonne, stad mimo udziału wolontariuszy zbiór powiększał się wolno: w 1989 kolekcja liczyła 10 pozycji i wzrosła do 100 w 1994. Okres intensywnego rozwoju rozpoczął się dopiero po 1995, gdy na szerszą skalę wykorzystano OCR; w kolejnych latach kolekcja rosła dynamiczniej i już w 1997 przekroczone 1000 pozycji (1999 – 2000, 2003 – 10 000), zaś w 2008 – 25000 (Lebert 2008). Pod koniec lat 90. zwiększono też liczbę dostępnych wersji (pojawiły się m.in. formaty multimedialne i mobilne), rozbudowano system metadanych oraz interfejs (wyszukiwarka od 1994). Już w latach 90. projekt znalazł pierwszych naśladowców; bliźniacze inicjatywy powołano m.in. w Niemczech (1994) [<http://projekt.gutenberg.de/>],

Australii (2001) [<http://gutenberg.net.au/>] i in. W 2004 pojawiła się również poszerzona wersja projektu *Project Gutenberg Consortia Center* [<http://www.gutenberg.us/>] (75 tys. poz.), która podjęła ścisłą współpracę z *WorldLibaryt.net*. Z dzisiejszej perspektywy *Gutenberg Project* reprezentuje nico przestarzałe podejście do zasobu i prymitywną technologię, niemniej sama idea, którą wdrożył Hart przez wiele lat inspirowała następców, zapisując się tym samym trwale w historii. Prócz opisywanego projektu do końca lat 80. pojawiło się jeszcze kilka innych inicjatyw, np. *Perseus Digital Library* [<http://www.perseus.tufts.edu>] założona w 1987 na Tufts University; zbiór ten początkowo deponowano na płytach CD, a od 1994 udostępniono w sieci. Z innych ważniejszych inicjatyw wczesnego etapu warto jeszcze wymienić: serwis skandynawski *Project Runeberg* [<http://runeberg.org/>] (od 1992) oraz *Bartleby* [<http://www.bartleby.com/>], *The On-Line Books Page* [<http://digital.library.upenn.edu/books/>], pierwsze pełne wydanie on-line dzieł Szekspira [<http://shakespeare.mit.edu/>], *MUSE* [<http://muse.jhu.edu/>] (od 1993), czy *The Humanities Text Initiative* [<http://www.hti.umich.edu/>] (1994)³.

Z zupełnie z innych powodów ogromną rolę odegrał także serwis preprintowy *arXiv* [<http://arxiv.org/>] uruchomiony w 1991 roku. System ten bowiem zapoczątkował pewien model naukowych zachowań publikacyjnych określanych potem jako *kultura reprintsu*, które następnie znalazły szeroki odzew w świecie nauki (szczególnie w naukach ścisłych). System *arXiv* – należy podkreślić – otworzył także drogę do idei Open Access, co znalazło finał m.in. w późniejszej Deklaracji Berlińskiej z 2003 roku i szeregu innych uregulowaniach z tego zakresu, intensywnie potem promowanych przez wiele urzędów i agencji np. Radę Unii Europejskiej. *arXiv* stał się zatem prototypem instytucjonalnych repozytorium, czyli systemów, które realizują funkcje *self-publishing* (składanie, przeglądanie, edytowanie, itd.) oraz rozpowszechniania dokumentów w formacie cyfrowym. Pierwsze tego typu systemy były nieskomplikowane i miały prostą architekturę; prócz *arXiv* były to m.in. baza dysertacji *Electronic Thesis and Dissertations Repositories* (od 1996) i ekonomiczny *RePEc* [<http://repec.org/>] (od 1997). Z kolei pierwszym systemem, który oferował usługi w zakresie *self-publishing* z wykorzystaniem PDF (od 1996) był istniejący od 1991 roku *Networked Digital Library of Theses and Dissertations* (NDLTD) [<http://www.ndltd.org/>]. Systemy te po licznych ewolucjach przetrwały do współczesności i obecnej perspektywy mogą być uważane za prototypy bibliotek cyfrowych (Candela, Castelli & Pagano 2011).

³ Por. P. Sauber: *Timeline of the Open Access Movement* - <http://www.earlham.edu/~peters/fos/timeline.htm> [2011.02.13].

Obraz wczesnej fazy rozwoju bibliotek cyfrowych byłby niepełny, gdyby pominąć pierwsze serwisy czasopism elektronicznych i inicjatywy z tym związane. Istotną rolę odegrały tu w szczególności projekty *Mercury*, *CORE*, *TULIP* i *Red Sage*. Projekt *Mercury* (1989-1992) prowadzony w Mellon University w Pittsburgu był jednym z pierwszych prób udostępnienia na terenie kampusu artykułów z czasopism naukowych z zakresu informatyki⁴. Kolejnym dużym przedsięwzięciem stał się prowadzony w latach 1991-1995 projekt *Chemistry Online Retrieval Experiment* (CORE), w który zaangażowane był m.in. Cornell University, Chemical Abstracts Service i OCLC; owocem prac była digitalizacja i udostępnienie 20 czołowych czasopism chemicznych. Teksty te przygotowano w dwu wersjach: jako obrazy rastrowe oraz jako pliki tekstowe ze znacznikami SGML (razem 425 tys. stron tekstu). Wielkie znaczenie miał także koordynowany przez Elsevier Science projekt *TULIP* (1991-1995)⁵, w czasie którego testowano rozmaite modele dystrybucji i rozwiązania techniczne. Na użytek testów i badań zeskanowano i poddano OCR ponad 500 tys. stron tekstu. Projekt ów prowadzony we współpracy w 15 czołowymi amerykańskimi uniwersytetami zakończył się spektakularnym sukcesem i z czasem stał się załącznikiem obecnego serwisu *ScienceDirect* (Nahotko 2007). Wyliczone projekty, jakkolwiek nie były bibliotekami cyfrowymi *sensu stricto* pomogły wskazać korzyści, jakie można osiągnąć poprzez cyfrową dystrybucję wydawnictw oraz przyczyniły się do postępu technologicznego w zakresie techniki skanowania i OCR (Tedd & Large 2005).

W pierwszej połowie lat 90. podjęto także pierwsze próby refleksji teoretycznej nad bibliotekami cyfrowymi. Duże znaczenie w tym zakresie miały zorganizowane w 1991 roku warsztaty *Future Directions in Text Analysis, Retrieval and Understanding*. W latach 1992-1993 odbyło się kilka kolejnych konferencji i spotkań, na których kontynuowano i rozwijano zainicjowane w 1991 wątki. Pokłosiem wspomnianych imprez prac wydana sumptem National Science Foundation obszerna elektroniczna książka pt. *Source Book on Digital Libraries* (Fox 1993). Pracę tę można postrzegać jako swoistą syntezę stanu wiedzy na temat bibliotek cyfrowych u schyłku 1993 roku; zawiera ona zarówno rozważania teoretyczne, jak i szczegółowe opisy projektów i kalendaria oraz bibliografię przedmiotu (28 pozycji). Najważniejszym jednak efektem wspomnianych wyżej spotkań było przygotowanie gruntu pod duże projekty; ukoronowaniem tych starań był rozpoczęty w 1994 *Digital Library Initiative*, który zapoczątkował nowy okres rozwoju bibliotek cyfrowych.

⁴ *The Mercury Electronic Library and Library Information System II*. "Mercury Technical Reports Series", No 6 (1992). Dostęp: <http://www.cs.cornell.edu/wya/papers/Mercury6.doc> [2011.05.21].

⁵ Informacje o projekcie – <http://www.elsevier.com/wps/find/librariansinfo.librarians/tulip> [2011.05.21].

3. Prace badawczo-rozwojowe i ewolucja bibliotek cyfrowych w latach 1994-2001

Od początku lat 90. XX wieku, czyli wraz z narodzinami www rozpoczyna się okres dynamicznego rozwoju bibliotek cyfrowych i ich stopniowa ewolucja. Pierwsze z nich w sensie technicznym były nieskomplikowane; zwykle tworzono je jako proste witryny HTML zawierające wykazy hiperłączy do publikacji (tzw. model *strona www*), które z czasem, w miarę powiększania się zbiorów, wyposażono w mechanizmy wyszukiwawcze i system metadanych (tzw. model *katalog obiektów cyfrowych*). Wadami tych rozwiązań był jednak brak mechanizmów konserwacyjnych oraz brak kontroli dostępu, zarządzania zasobem i metadanymi. Obydwa modele odzwierciedlały więc filozofię pierwszych bibliotek cyfrowych (np. *Gutenberg Project*); charakterystycznym ich rysem było programowe ukierunkowanie na udostępnienie kanonu kulturowego (klasyka literacka, dziedzictwo historyczne) oraz założenie, że dostęp do zbiorów winien być nielimitowany. W efekcie biblioteki te realizowały przede wszystkim cel archiwalny (nie biblioteczny); z drugiej zaś strony udostępniane zasoby nie wymagały stosowania żadnych mechanizmów zabezpieczeń i kontroli dostępu, gdyż były z natury swej wolne od ograniczeń wynikających z ochrony praw autorskich. Warto dodać, że nie interesowano się też szerzej problematyką użytkownika (interfejs) ani mechanizmami wyszukiwawczymi. Takie usytuowanie celów nie sprzyjało rozwojowi bibliotek cyfrowych. Było bowiem niewystarczające do realizacji funkcji bibliotecznych, gdyż stało w sprzeczności z jedną z zasadniczych cech biblioteki, zakładającej, że proces gromadzenia jest permanentny (stały, ciągły, powtarzalny), a także otwarty na nowe formy dokumentów. Z drugiej zaś strony istniała uzasadniona niechęć wydawców do nielimitowanego udostępniania, gdyż nie zabezpieczało to ich interesu ekonomicznego (Kolasa 2009).

Czynniki te stały się motorem podjęcia licznych prac zmierzających do zmiany istniejącej sytuacji i rozpoczęła się stopniowa ewolucja systemów do form bardziej zaawansowanych. Prace te nabrały tempa w drugiej połowie ostatniej dekady XX wieku, gdy podjęto wiele ważnych inicjatyw mających charakter badań podstawowych. W ich wyniku powstało szereg standardów i rozwiązań, które – jak czas pokazał – okazały się trwałe. Rozmach w tym obszarze najlepiej ilustruje krótki wykaz wybranych dokonań tego okresu: uruchomiono duże projekty badawczo-rozwojowe ukierunkowane na biblioteki cyfrowe (*Digital Library Initiative* [1994-1998 i 1999-2004]; *eLib* [1998-2001]); narodowe i

międzynarodowe strategie digitalizacji (np. amerykański *National Digital Library Program*⁶ [1995], czy *Memory of the World Programme* [1992]); powstały pierwsze stowarzyszenia bibliotek cyfrowych (*Digital Library Federation* [1995]) i organizacje koordynujące (np. *Joint Information Systems Committee – JISC* [1993], *Electronic Information for Libraries Direct - eIFL* [1999]); tematyczne konferencje (*Joint Conference on Digital Libraries* [1994]; *European Conference on Digital Libraries (ECDL)* [1997]; *International Conference on Asian Digital Libraries* [1998]); czasopisma fachowe (*D-Lib Magazine* [1995]; *Ariadne* [1996]) oraz standardy i specyfikacje (Dublin Core [1995]; DIENST [1995]; RDF [1999]; OAI-PMH [2001]). Atmosferę charakteryzowanego okresu wiernie obrazują materiały zbierane przez profesora Edwarda A. Foxa cyklicznie udostępniane na stronie Virginia Tech w latach 1998-2001 [<http://ei.cs.vt.edu/~dlib/>]. Dokumenty te zebrane w formie *silva rerum* zawierają wiele informacji historycznych, w tym zrzuty nieistniejących już serwisów. Interesujący jest też wachlarz rozważanych wówczas problemów. Ich analiza przekonuje, że już u progu XXI wieku wiedza na temat bibliotek cyfrowych osiągnęła wysoki poziom dojrzałości. Udowadnia to krótki wykaz tematów z opracowań E. A. Foxa (2001):

1. Podstawy (definicje, publikacje, projekty); 2. Wyszukiwanie, przeglądanie (metody, indeksowanie, wyszukiwanie rozproszone, roboty/pajaki, integracja linków); 3. Multimedia (tekst, audio, obraz, film, grafika, animacja, JPEG, MPEG, wyszukiwanie oparte na treści dokumentu); 4. Architektura (komponenty, protokoły, InfoBus, protokoły pośrednie, ramki, OAI); 5. Interfejs (psychologia, środowisko, taksonomia elementów, wizualizacja, design, użyteczność); 6. Metadane (ontologie, RDF, MARC, DublinCore, IMS, mapowanie, konwersja); 7. Publikacje elektroniczne (autorstwo, prezentacja, obiektowy model DOM, równoległe publikowanie, style XLS, e-książki, struktura, adiuścacja, tagowanie); 8. Problemy baz danych (technologia, struktura informacji, dane multimedialne i hipermedialne, replikacja, magazynowanie); 9. Oprogramowanie agentowe do zarządzania informacją (problemy, protokoły, automatyczne indeksowanie); 10. Handel, ekonomia, wydawcy (zabezpieczenia, ochrona, problemy Open Access, modele płatności); 11. Prawa własności intelektualnej (ochrona praw autorskich, copyright, zarządzanie prawami); 12. Problemy społeczne (współpraca, ocena, użytek w edukacji, dziedzictwo kulturowe, organizacje).

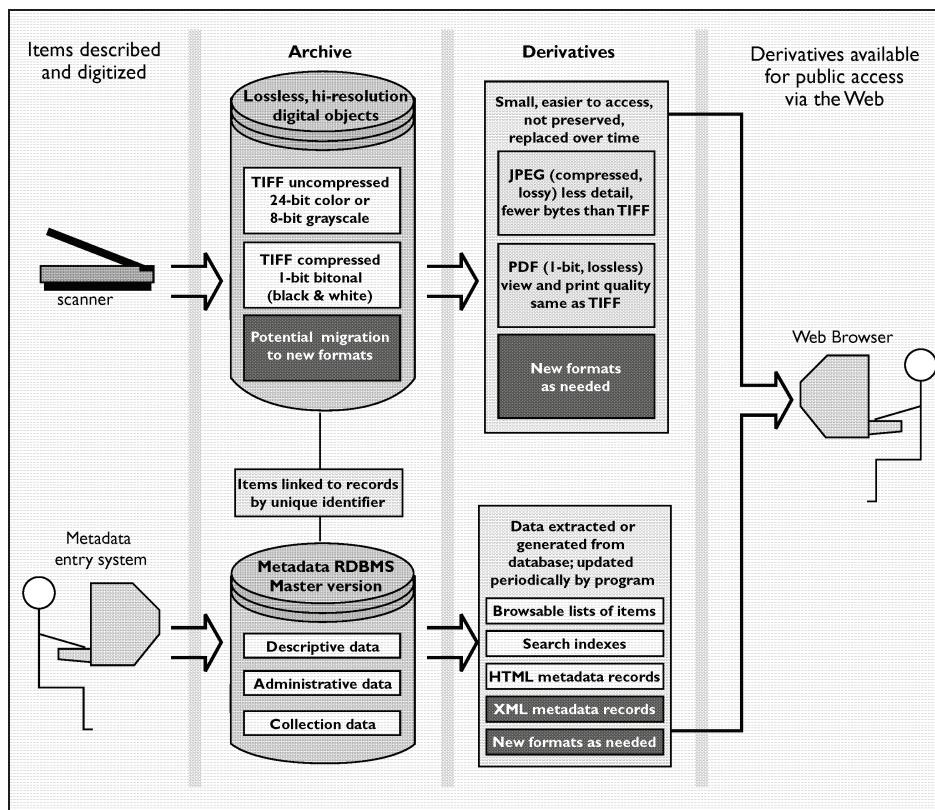
Bliższa analiza wyliczonych tematów pozwala stwierdzić, że już w II połowie lat 90. podjęto niemal wszystkie wątki rozważne wspólnie.

Przez całe lata 90. równoległe ze wzrostem liczby repozytoriów (było ich wówczas ok. 300-500) następowała ewolucja systemów i doskonalenie standardów związanych z wspieraniem zapisu cyfrowego, następował rozwój formatów i nośników danych. Z czasem podjęto także próby udostępnienia modeli i narzędzi oferujących zakres usług zbliżonych do tradycyjnej biblioteki (gromadzenie, opracowanie, udostępnianie) oraz zainteresowano się interfejsem. Typowy model biblioteki cyfrowej tego okresu obrazuje rys. 2. Wyróżniono w nim najczęściej trzy poziomy: 1. poziom tworzenia (skanowanie i edycja metadanych), 2.

⁶ National Digital Library Program – <http://international.loc.gov/ammem/dli2/html/lcndlp.html> [2011.05.21].

poziom bazy danych: system archiwalny i prezyencyjny oraz 3. poziom użytkownika (przeglądarka). Na styku poszczególnym poziomów wskazywano na rozmaite problemy, np.: jakość kopii archiwalnej i użytkowej, typy metadanych, mechanizmy wyszukiwawcze. Warto zwrócić uwagę, że w modelu tym zupełnie marginalne znaczenie miała współpraca z innymi systemami (XML, new formats), w efekcie system działał niejako w pojedynkę. Zmiana tej perspektywy dokonywała się jednak stopniowo od 1999 roku, zaś po 2002 weszła na trwałe.

Rys. 2, Model biblioteki cyfrowej w 2001 roku (McCray, Gallagher 2001)



Jakkolwiek prace badawczo-wdrożeniowe nad bibliotekami cyfrowymi podjęto w wielu ośrodkach, jedynie trzy inicjatywy pozostawiły trwałe ślady w historii informatyki: *Digital Library Initiative* (DLI1 [1994-1998] i DLI2 [1999-2004]) oraz *eLib*. Bez wątpienia największy wpływ na liczbę wdrożeń i postęp w zakresie badań nad bibliotekami cyfrowymi miały projekty amerykańskie. Rozpoczął je w 1994 roku pierwszy etap *Digital Library Initiative* [DLI1], którego głównym inicjatorem była agencja rządowa National Science Foundation wspierana przez Department of Defense Advanced Research Projects Agency (DARPA) oraz National Aeronautics and Space Administration (NASA). Projekt trwał 4 lata i obejmował finansowanie prac realizowanych na sześciu uczelniach. Głównym celem DLI1 były i rozwój podstawowych technologii na rzecz bibliotek cyfrowych. Rozwijano w jej ramach trzech grup problemowych, które obejmowały: 1. utrwalanie, kategoryzacja i

porządkowanie informacji; 2. wyszukiwanie, przeglądanie, filtrowanie, prezentacja i wizualizacja; 3. protokoły i standardy sieciowe. Wykonawcami były czołowe amerykańskie uniwersytety: University of California w Berkeley, University of California w Santa Barbara, Carnegie Mellon University, University of Illinois w Urbana, University of Michigan oraz Stanford University. DLI1 okazała się sporym osiągnięciem, stąd w roku 1999 ogłoszono jej drugi etap [DLI2], trwający kolejne sześć lat. W fazie drugiej skoncentrowano się na wdrożeniach wcześniej badanych technologii i testowaniem poszczególnych rozwiązań. Czas pokazał, że szczególne znaczenie miały prace z zakresu architektury i problemów użyteczności. Najistotniejsze zaś wydaje się przyjęcie koncepcji, w której architektura systemu jest zorientowana na użytkownika. W trakcie realizacji DLI2 do grona uczestników dołączyli nowi wykonawcy, m.in.: Library of Congress, National Library of Medicine, National Endowment for the Humanities i wielu innych (Fox 1999; Madalli 2003). Po obu odsłonach projektu pozostały trwałe osiągnięcia. Sześć z nich – jak podkreślają znawcy technologii (Candela, Castelli & Pagano 2011) – było szczególnie istotnych: 1. *Projekt California Environmental Digital Library*⁷, który wniósł wiele w zakresie technologii obsługi dużych, rozproszonych zbiorów ikonograficznych (fotografii, zdjęć satelitarnych, wideo, map); 2. *Alexandria Digital Library* [<http://www.alexandria.ucsb.edu/>], gdzie rozwijano problemy bibliotek rozproszonych oraz zagadnienia tworzenia kolekcji zawierających obiekty geograficzne (mapy, zdjęcia lotnicze, obrazy satelitarne) oraz badano zapytania zdefiniowane geograficznie; 3. *Projekt Digital Video Library* [<http://www.informedia.cs.cmu.edu/>], który zrealizował udaną koncepcję dużej cyfrowej kolekcji wideo obsługującej wyszukiwanie oparte na wiedzy; 4. Projekt *The Interspace*, koncentrujący się na budowie dużych kolekcji z zakresu techniki inżynierskiej i fizyki, które były jednocześnie skutecznie przeszukiwane z poziomu jednolitego interfejsu; 5. *The University of Michigan Digital Library* [<http://www.lib.umich.edu/digital-library-production-service-dlps>], specjalizujący się w tworzeniu architektury systemów bibliotecznych, wykorzystujących oprogramowanie agentów; 6. *The Stanford Digital Library Project* [<http://diglib.stanford.edu:8091/>], badający zastosowanie zasad interoperacyjności w środowisku heterogenicznych usług i zbiorów, z użyciem specjalnego protokołu [InfoBus], który zapewniał dostęp do usług i informacji z różnych źródeł poprzez agentów. Mimo że większość wspomnianych projektów nie działa współcześnie w pierwotnej formie stanowiły one w ówczesnym okresie ogromny wkład do wiedzy informatycznej.

⁷ Aktualnie: California Environmental Digital Library Network (CalEDLN) – <http://caledln.projects.atlas.ca.gov/> [2011.05.21].

Na marginesie warto podkreślić, że jednym z najważniejszych sukcesów wynikające z historii tych projektów jest Google. Późniejsi twórcy wyszukiwarki Lawrence Page i Sergey Brin rozpoczęli prace nad jej silnikiem w czasie, gdy byli doktorantami na Uniwersytecie Stanforda w trakcie realizacji Stanford Digital Library Project. DLI wykracza daleko poza konkretne prace i wynalazki; ze współczesnej perspektywy można ocenić, że projekty te dały podstawę do współczesnego oblicza bibliotekarstwa cyfrowego, jako nowej dyscypliny badawczej, która stała się sumą wysiłków teoretycznych kilku specjalności: informatyki, informacji naukowej, językoznawstwa, psychologii i in. (Candela, Castelli & Pagano 2011).

Równoległe poza projektem DLI podjęto w USA szereg innych inicjatyw i prac z interesującego nas zakresu, m.in. na University of Arizona badano wydajne systemy klasyfikacyjne – realizujące postulat *od wyszukiwania informacji do zarządzania wiedzą*; na Carnegie Mellon University rozpoczęto prace nad koncepcją późniejszego serwisu *Universal Library* [<http://www.ulib.org/>]; na University of South Carolina testowano oprogramowanie do bibliotek; na Stanford University badano technologie na rzecz bibliotek cyfrowych zaś w Tufts University kontynuowano rozwój *The Perseus Digital Library Project*. Dużą rolę w konsolidacji wysiłków odegrało założone w 1995 konsorcjum *Digital Library Federation* (DLF) [<http://old.diglib.org/>]. Inicjatorami przedsięwzięcia były m.in. duże biblioteki realizujące własne projekty cyfrowe, m.in. New York Public Library [<http://www.nypl.org/collections>], Library of Congress (*American Memory* [<http://memory.loc.gov/ammem/>]) oraz instytucje nadzorujące dostęp do dokumentów, m.in. National Archives and Records Administration i Commission on Preservation and Access (CPA); ostatnia z nich zapewniała finansowanie pozyskane od IBM. W krótkim czasie liczba członków DLF urosła do 37, zrzeszając czołowe amerykańskie biblioteki oraz kilka księżnic zagranicznych. Cele strategiczne DLF były jasno sformułowane: rozwój standardów, promocja najlepszych praktyk, stymulacja rozwoju poprzez wpływ na instytucje finansujące, ochrona własności intelektualnej, a nade wszystko budowa wspólnej infrastruktury⁸. Ostatni z celów dalece wybiegał w przyszłość i zaowocował pracami ukierunkowanymi na rzecz tworzenia bibliotek rozproszonych (współpraca przy tworzeniu OAI-PMH oraz jego impementacji w serwisie *OAI-STER* afiliowanym w University of Michigan od 2002 roku) (Greenstein, Thorin 2002).

Nie sposób przecenić także prac prowadzonych w ramach *Open Archive Initiative*, organizacji finansowanej przez Fundację Mellona, powołanej wspólnie przez DLF i

⁸ Szczegółowy wykaz projektów – <http://old.diglib.org/produce.htm> [2011.05.21].

partnerów społecznych. Owocem prowadzonych prac był wspomniany wyżej protokół OAI-PMH, którego pierwszą wersję opublikowano w 1999. Rozwiązanie to wychodziło naprzeciw idei interoperacyjności, czyli zdolności sieci do efektywnej współpracy w celu zapewnienia wzajemnego dostępu użytkowników do usług poprzez dedykowany model wymiany danych. Warto dodać, że podjęcie prac nad OAI-PMH wiązało się z kontrowersyjną decyzją zarzucenia eksperymentów z użyciem wcześniejszych standardów, m.in. z39.50, InfoBus [1995] oraz Dienst Protocol [1995]⁹. Istotną rolę odegrały też prace grupy zaangażowanej w rozwój metadanych. Główną rolę odegrała w tym zakresie wspólna grupa OCLC i NCSA, która w 1995 roku zainicjowała powstanie Dublin Core Metadata Initiative (DCMI). Rozwijany w ramach DCMI standard Dublin Core, który prezentowano na cyklicznych warsztatach szybko wzbudził zainteresowanie społeczności międzynarodowej i dużą popularność wśród twórców bibliotek cyfrowych. Fakt ów sprawił, że Dublin Core zaczął w zastosowaniach cyfrowych stopniowo wypierać wcześniejsze specyfikacje (np. MARC21 i jego odmiany) jeszcze na długo przed jego oficjalną premierą (2001). Istotną rolę w rozwoju bibliotek cyfrowych spełniły projekty realizowane m.in. przez Virginia Tech i partnerów, szczególnie dwa repozytoria: baza reprintów z zakresu informatyki *Networked Computer Science Technical Reference Library* [<http://www.ncstrl.org/>] oraz międzynarodowa baza dysertacji *The Networked Digital Library of Theses and Dissertation* [<http://www.ndltd.org/>], działająca od 1996 roku. Ogromny wpływ na kierunki rozwoju bibliotekarstwa cyfrowego wywarł też projekt edukacyjny *The National Science Digital Library* [<http://nsdl.org/>] realizowany przez kilka organizacji i udostępniony w 2000 roku (Madalli 2003).

Prace wdrożeniowe na wyspach brytyjskich koncentrowały się wokół projektów koordynowanych przez *Joint Information Systems Committee* (JISC) [<http://www.jisc.ac.uk/>] od 1993 roku. Pierwsza faza prac (do 1998 roku) pod nazwą *eLib - The Electronic Libraries Programme* [<http://www.ukoln.ac.uk/services/elib/>], obejmowała kilkanaście projektów dotyczących szeroko pojętej digitalizacji, m. in. bibliotek hybrydowych (np. *Agora*, *Malibu*), systemów zabezpieczeń (CEDRAS) oraz szeroko pojętej technologii dokumentów cyfrowych (czasopisma elektroniczne, druk na zadanie, reprints). W 1996 roku powołano też specjalny dwumiesięcznik *Ariadne* [<http://www.ariadne.ac.uk/>]. Inne prace ogniskowały się także w

⁹ DIENST funkcjonuje w trzech znaczeniach: 1. koncepcja architektoniczna rozproszonych bibliotek cyfrowych; 2. protokół komunikacyjny w tej architekturze; 3. oprogramowanie, które realizuje funkcje tego protokołu. Zasadniczo Dienst to system konfiguracji szeregu indywidualnych usług na serwerach rozproszonych do współpracy między bibliotekami cyfrowymi – <http://www.cs.cornell.edu/cdlrg/dienst/DienstOverview.htm> [2011.05.21].

ramach powołanej przez *JISC Resource Discovery Network* [<http://www.rdn.ac.uk/>]¹⁰. Szerszą działalność JISC rozwinęła jednak dopiero po 2003 roku. W Europie w latach 1994-1998 najważniejsze prace koncentrowały się w ramach II etapu projektu UE *Telematics for Libraries Program* [109 projektów]¹¹ oraz grupy roboczej DELOS (Raitt 2000). Wysiłki tej grupy skupiły się początkowo nad pracami koncepcyjnymi, które twórczo rozwinięto w latach późniejszych, dopiero po 2001 roku, czyli po przekształceniu w sieć DELOS Thematic Network (Candela, Castelli & Pagano 2011).

W innych krajach wysiłki koordynacyjne podejmowały najczęściej biblioteki narodowe, np. w Kanadzie rolę tę pełnił afiliowany przy bibliotece narodowej spis - *Canadian Inventory of Digital Initiatives* (1999) [<http://publications.gc.ca/>]; w Australii - *Australian digitisation Project* [<http://www.nla.gov.au/libraries/digitisation/>] (od 1998); w Azji i Pacyfiku – doroczna konferencja ICADL [<http://www.icadl.org/>] (1998), zrzeszająca rozwinięte państwa regionu (Australia, Japonia, Chiny, Korea, Indie, Tajlandia, Wietnam, Indonezja). Osobno wysiłki Indii i Chin wspierało kilka projektów amerykańskich, szczególnie inicjatywa *Universal Digital Library* [<http://www.ulib.org/>] rozpoczęta przez Carnegie Mellon University oraz wspomniany wcześniej NTDL [<http://www.ndltd.org/>] wspierany przez Science and Technology Collaboration. Inne projekty azjatyckie to m.in. pilotażowy program malezyjski prowadzony przez Malaysian National Digital Library [<http://www.mylib.com.my/>] (od 2000) oraz koreański rozwijany od 1996 przez National Digital Library of Korea [<http://www.dlibrary.go.kr/>].

Duże znaczenie w rozwoju wiedzy o bibliotekach cyfrowych miały także międzynarodowe konferencje, które odbywały się cyklicznie już od połowy lat 90. Najaktywniejsze było środowisko badaczy amerykańskich, kanadyjskich i australijskich, które od 1994 roku cyklicznie urządzało branżowe imprezy. Najpierw organizowano je spontanicznie (1994-1995), potem osobno przez IEEE Computer Society (1995-2000) oraz Association for Computing Machinery (1996-2000), a następnie wspólne pt. *Joint Conference on Digital Libraries* (JCDL) [<http://www.jcdl.org/>] (razem 24 konferencje). Równie bogate tradycje ma związana z DELOS impreza pt. *European Conference on Digital Libraries* (ECDL) [<http://ecdconference.isti.cnr.it/>], która jest głównym europejskim forum z zakresie bibliotek cyfrowych; od 1997 roku odbyło się 14 edycji, których efektem były poważne prace teoretyczne zdeponowane w Springer Verlag. Region Azji i Pacyfiku reprezentuje natomiast

¹⁰ Hiom, D. (2006): Retrospective on the RDN. *Ariadne*. Iss. 47 (April). Dostęp: <http://www.ariadne.ac.uk/issue47/hiom/> [2011.05.21].

¹¹ Telematics for Libraries / Project – <http://cordis.europa.eu/libraries/en/projects.html> [2011.06.01].

doroczna *International Conference on Asian Digital Libraries* [<http://www.icadl.org/>], która odbywa się od 1998 (razem 13 edycji).

Prócz wyliczonych uprzednio projektów zbiorowych (zwykle inspirowanych przez rządy lub organizacje naukowe) warto wspomnieć osobno o niektórych dużych realizacjach. Wśród ok. 300-500 bibliotek cyfrowych działających w latach 1995-2001 najwyższy poziom reprezentowały amerykańskie projekty członków DLF¹². Biblioteki te bowiem jako pierwsze testowały i wdrażały najnowsze rozwiązania oraz szybko powiększały zbiory. Spośród dużych przedsięwzięć, które nieznacznie zmienione działają współcześnie na uwagę zasługują m.in. projekty prowadzone przez Library of Congress (głównie *American Memory*), część serwisu Internet Archive oraz francuska *Gallica*. Prace pilotażowe nad projektem *American Memory* (AM) [<http://memory.loc.gov/ammem/>] prowadzono w Library of Congress już w latach 1990-1994, zaś oficjalny debiut nastąpił w 1994. Portal gromadzi zasób, który odzwierciedla kanon kulturowy USA (głównie teksty i ikonografia) i składa się ze 140 części. Charakterystycznym rysem AM jest heterogeniczność kolekcji. Ponieważ poszczególne części AM realizowały różne zespoły w czasach, gdy nie istniały jeszcze stabilne standardy, prawie każda kolekcja nosi na sobie indywidualne piętno twórców. Fakt ten utrudnia nawigację i jest uciążliwy dla czytelnika. Aby nieco złagodzić tę ułomność z czasem całość wyposażono w dedykowaną wyszukiwarkę. Drugą wadą AM jest archaiczne podejście do formatów dokumentów, gdyż nawet w zbiorach typowo tekstowych są one prezentowane albo jako grafika rastrowa, albo HTML. Mimo pewnych ułomności projekt zyskał wysoką renomę i szybko się rozwijał (ponad 9 milionów jednostek), dając przykład innym. Podobne podejście legło u podstaw narodowego projektu francuskiego *Gallica* [<http://gallica.bnf.fr/>] zainicjowanego w 1997 roku przez Bibliothèque nationale de France (BnF). Serwis prezentuje podejście typowe dla księżnicy narodowej i gromadzi szeroko pojętą literaturę oraz multimedia francuskie i Francji dotyczące (razem ponad 1,4 mln jednostek). W odróżnieniu od AM *Gallica* jest jednak serwisem nowocześniejszym, który od niemal samego początku jest blisko związany z innymi serwisami BnF, oferuje doskonałe mechanizmy

¹² Wykaz członków: Bibliotheca Alexandrina; British Library; California Digital Library; Carnegie Mellon University; Columbia University; Cornell University; Council on Library and Information Resources; Dartmouth College; Emory University; Harvard University; Indiana University; Johns Hopkins University; Library of Congress; Massachusetts Institute of Technology; New York Public Library; New York University; North Carolina State University; Oxford University; Pennsylvania State University; Princeton University; Rice University; Stanford University; University of California, Berkeley; University of California, Los Angeles; University of Chicago; University of Illinois - Urbana-Champaign; University of Michigan; University of Minnesota; University of Pennsylvania; University of Southern California; University of Tennessee; University of Texas - Austin; University of Virginia; University of Washington; U.S. National Archives and Records Administration; U.S. National Library of Medicine; Yale University – źródło: <http://old.diglib.org/about.htm> [2011.05.21].

wyszukiwawcze oraz dostęp do zasobu w kilku wariantach i wspiera idee współpracy z użyciem protokołu OAI. Zupełnie inny model reprezentuje działające od 1996 roku serwis *Interent Archive* (IA) [<http://www.archive.org/>], którego znaczącą częścią jest kolekcja książek (IA/Texts). Serwis ten nie wytwarza własnych zasobów, lecz gromadzi kopie lustrzane z bibliotek współpracujących, które są replikowane w wielu formatach. Właściwy rozwój IA nastąpił dopiero od 2005, gdy założyciel przystąpił do *Open Content Alliance*, od tego czasu jego zasoby wzrosły z nieco ponad 150 tys. do ok. 2,8 mln. i nastąpiła integracja z *Open Library* [<http://openlibrary.org/>]. Warto też dodać projekty, które inspirowała inicjatywa UNESCO *Memory of the World*; wśród nielicznych owoców tego przedsięwzięcia pozytywnym wyjątkiem jest m.in. czeski serwis udostępniający książki historyczne *Manuscriptorum* [<http://manuscriptorium.com/>] (od 1992) oraz kolekcja historyczna *Memory of the Netherlands* [<http://www.geheugenvannederland.nl/>], która może uchodzić za wzór narodowego repozytorium historycznego (od 1999). Spośród innych serwisów o zasięgu globalnym warto wymienić przedsięwzięcie komercyjne *World Public Library* [<http://worldlibrary.net/>], działające od 1996 roku (2 mln jednostek); serwis ten – co warto podkreślić – wyprzedził podobne przedsięwzięcia o niemal pełną dekadę.

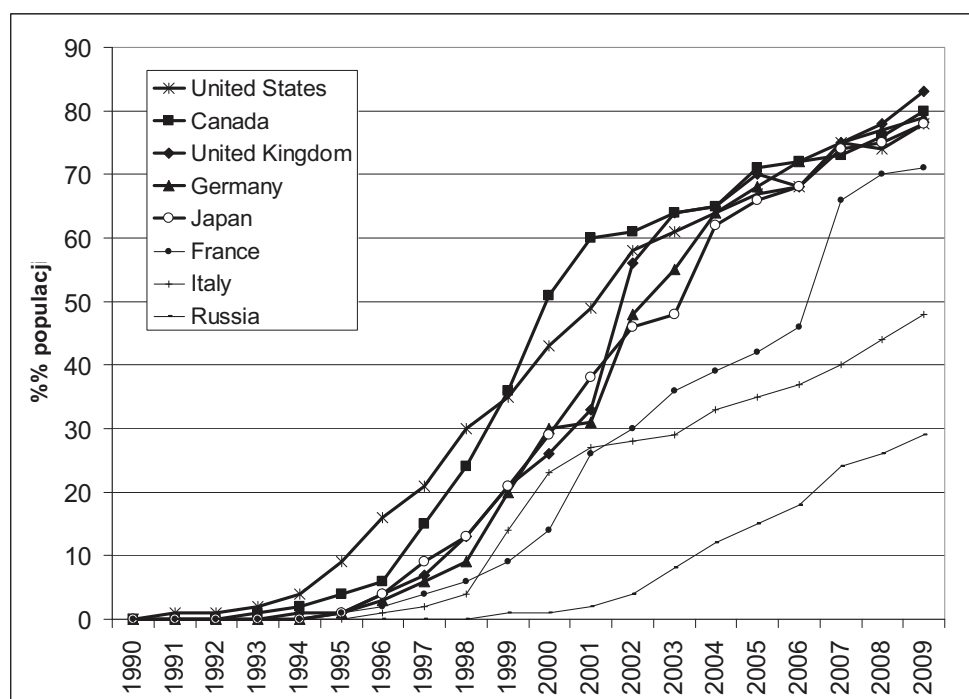
4. Biblioteki cyfrowe po 2002 roku. Współczesność

U progu XXI wieku w bibliotekarstwo cyfrowe wkroczyło w nowy etap rozwoju. Tendencja ta staje się łatwo uchwytna jeśli za miernik postępu przyjąć liczbę funkcjonujących serwisów i wzrost dostępnych obiektów (rys. 1). Źródeł tego zjawiska można dopatrywać się w kilku płaszczyznach. Z jednej strony zadziałał tu pozytywnie rozwój technologii i osiągnięcia organizacyjne w sferze bibliotek cyfrowych, czyli tendencje, które przedstawiono wyżej. Z drugiej jednak – nałożył się na to ogólny trend rozwoju usług on-line, którego obiektywnym miernikiem wydaje się być wzrost liczby użytkowników. Na przestrzeni 2000-2003 nastąpił bowiem proces, który można nazwać przekroczeniem progu umasowienia, gdyż z usług sieciowych korzystało już ponad 50% populacji w najbogatszych krajach świata (rys. 3). Fakt ów był niewątpliwie motorem dalszych zmian, w szczególności przeobrażeń sposobu komunikacji w internecie i mechanizmów generowania treści. *Signum temporis* tego przełomu jest symboliczna data narodzin internetu określanego jako web 2.0 (2001).

Z technicznego punktu widzenia idea web 2.0 oznaczała stosowanie nowych technologii zorientowanych na użytkownika, choć – jak czas pokazał – najistotniejszy okazał się aspekt społeczny, który oznaczał zrównanie użytkowników i twórców. W środowisku

web 2.0 pojawiły się też inne zjawiska, istotne dla bibliotek cyfrowych, jak: tworzenie społeczności sieciowych, trend do używania folksonomii, nowy rodzaj komunikacji i zachowań; od strony technicznej zaś – powszechną drożenie do integracji wszelkich usług sieciowych oraz komunikacja maszynowa. Nie ulega wątpliwości, że idea web 2.0 postawiła bibliotekarzy w nowej sytuacji. Niekontrolowany zalew łatwo dostępnych informacji generowanych przez użytkowników stał się realnym zagrożeniem dla tradycyjnego modelu biblioteki. Aby temu przeciwdziałać twórcy serwisów bibliotecznych zaczęli aktywnie uczestniczyć w środowisku web 2.0 oferując informację kwalifikowaną. Pierwszym i najłatwiej uchwytnym aspektem przemian w tym zakresie był powszechny trend do tworzenia serwisów rozproszonych.

Rys. 3, Użytkownicy internetu w krajach G8 (1990-2009)¹³

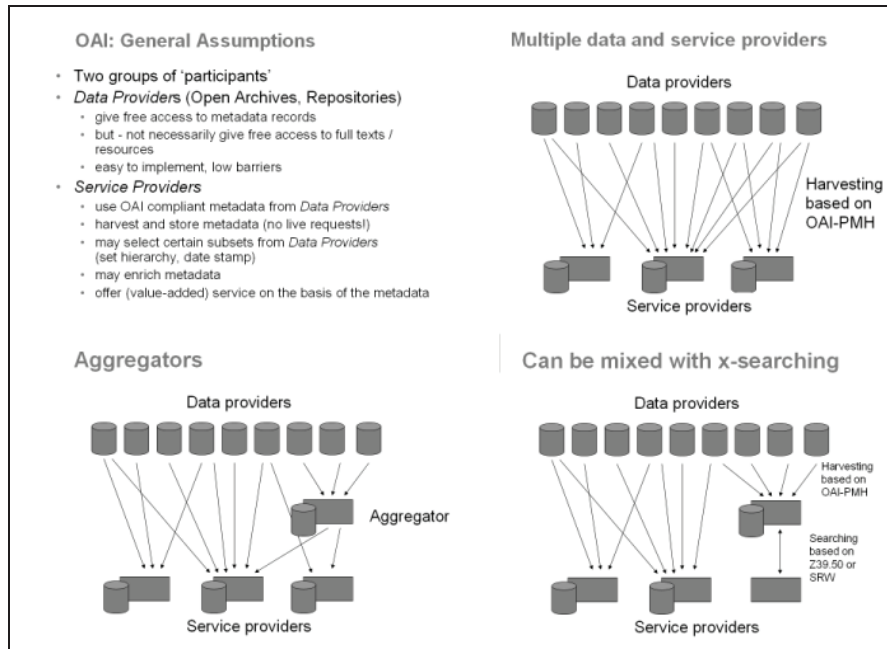


Nad ideą komunikacji pomiędzy bibliotekami cyfrowymi podjęto prace już w połowie lat 90., zaś ich ukoronowaniem był opracowany w ramach Open Archive Initiative dedykowany protokół OAI-PMH (1999). Protokół ten [Open Archive Initiative Protocol for Metadata Harvesting] służy do selektywnego pobierania metadanych opisujących obiekty cyfrowe; umożliwia on automatyczną wymianę informacji pomiędzy systemami gromadzącymi takie metadane, a systemami chcącymi te dane uzyskać. Jest przy tym bardzo elastyczny, gdyż dane są zapisane w otwartym standardzie XML (zwykle RDF). Idea jego

¹³ Przeliczenia na podstawie danych: The United Nations Statistics Division – <http://data.un.org/> [2011.05.28].

działania jest zatem zupełnie inna niż *z39.50* czy *SRU*, gdyż nie działa on w trybie rzeczywistym.

Rys. 4, Model biblioteki rozproszonej OAI-PMH (Hunter, 2005)

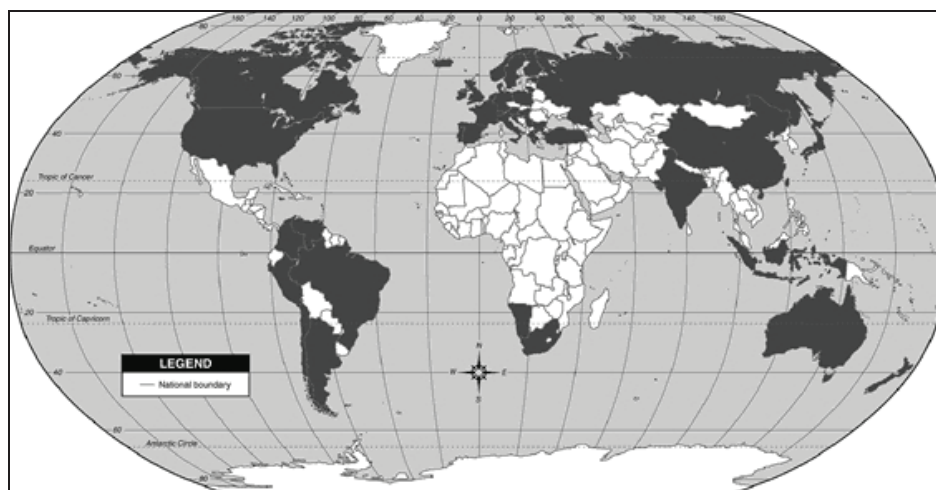


W praktyce centrum serwisu rozproszonego [service provider] pełni w tym rozwiązaniu dedykowany serwer [harvester], który okresowo pobiera metadane z określonych bibliotek [data provider] (rys. 4). Serwis providera indeksuje własnymi mechanizmami metadane i pełni rolę bramki dostępowej dla użytkownika. Warto dodać, że mogą istnieć dwa rodzaje dostawców: pojedynczy serwis lub tzw. agregator – czyli lokalny pośrednik. Drugi typ dostawcy stwarza wielkie możliwości, gdyż biblioteki mogą się łączyć w dowolne grupy i elastycznie współpracować.

Największą udaną realizacją serwisu opartego na OAI-PMH był portal *OAISTER*, uruchomiony w lipcu 2002 roku na University of Michigan. Do października 2009 działał samodzielnie, następnie stał się częścią serwisu OCLC *Worldcat* [<http://www.worldcat.org/>], istnieje też odrębna bramka dostępowa [<http://oaister.worldcat.org/>]. Rozwój portalu następował szybko (rys. 1). W grudniu 2006 roku zrzeszał już 726 dostawców oferujących prawie 10 mln jednostek (Garrison 2010), najwięcej dostawców pochodziło z USA (272), Wielkiej Brytanii (78), Niemiec (72), Francji (32), Kanady (27), Brazylii (23), Włoch i Australii (po 21), Holandii (19), Hiszpanii (18), Japonii (16), Szwecji (14), Polski (12) oraz

innych państw Europy (55)¹⁴. Obecnie (maj 2011) zrzesza ponad 1100 dostawców, oferując ponad 25 mln rekordów. Zaprezentowany rozkład, ze względu na reprezentatywność może stanowić dobrą podstawę do uogólnień geograficznych (rys. 5). Z przedstawionych danych wynika w szczególności lapidarna teza, że biblioteki cyfrowe funkcjonują jedynie w państwach rozwiniętych i rozwijających się oraz daleko od ognisk konfliktów zbrojnych.

Rys. 5, Biblioteki cyfrowe na świecie w rozkładzie terytorialnym (2006)¹⁵



Prócz bazy *OAISTER* powstało wiele innych serwisów rozproszonych o różnej wielkości. Oficjalny wykaz na stronach OAI¹⁶ wlicza ich ponad 30, lecz z całą pewnością jest ich znacznie więcej, niektóre zaś pełną wyłącznie funkcję agregatorów dla innych portali. Wśród samodzielnych serwisów na szczególną uwagę zasługują portale naukowe, m.in.: serwis *Scirus* [<http://www.scirus.com>] należący do Elseviera (410 mln jednostek, głównie artykułów); *ScientificCommons* [<http://www.scientificcommons.org/>], założony w 2007 roku na szwajcarskim Universität St. Gallen (38 mln jednostek, 1269 dostawców), repozytorium dysertacji *DART-Europe* [<http://www.dart-europe.eu/>] (338 dostawców) czy rozbudowana w 2006 wersja serwisu *Perseus* [<http://www.perseus.tufts.edu/>], która scala zasoby filologiczne. Nadto liczne serwisy narodowe: holenderski *DARE* (od 2004), norweska *NORA* [<http://www.ub.uio.no/nora/>] (2004); hiszpańska sieć muzealna *Hispana*

¹⁴ OAister Repositories: Argentina (3), Australia (21), Austria (6), Belgium (9), Brazil (23), Bulgaria (1), Canada (27), Chile (5), China (2), Columbia (2), Croatia (1), Denmark (3), England (63), Estonia (1), Finland (7), France (32), Germany (72), Greece (2), Hong Kong (3), Hungary (3), Iceland (1), India (11), Indonesia (1), Ireland (3), Italy (21), Japan (16), Lithuania (1), Namibia (1), Netherlands (19), New Zealand (3), Northern Ireland (1), Norway (1), Peru (3), Poland (12), Portugal (5), Russia (1), Scotland (13), Serbia (1), Slovenia (1), South Africa (4), Spain (18), Sweden (14), Switzerland (6), Taiwan (2), Turkey (1), United States (272), Venezuela (2), Wales (2), West Indies (1) - <http://www.lib.umich.edu/digital-library-production-service-dlps/oaister-statistics> [2011.05.21].

¹⁵ Źródło: OAISTER – <http://www.oaister.org/stats.html> [2009.04.13].

¹⁶ Registered Service Providers – <http://www.openarchives.org/service/listproviders.html> [2011.05.21].

[<http://hispana.mcu.es/>] (2006; 148 dostawców, 3 mln jednostek) czy polska sieć *Federacja Bibliotek Cyfrowych* [<http://fbc.pionier.net.pl/>] (2007; 69 dostawców, 648 tys. jednostek). Z kolei spisy dostawców, których aktualna liczba sięga zapewne 2-3 tysięcy gromadzi kilka dedykowanych serwisów, m.in.: *OpenDOAR* [<http://www.opendoar.org/>] (1966 dostawców) czy *Openarchives.eu*: [<http://www.openarchives.eu/>] (1200 dostawców).

Na osobną uwagę zasługuje jedna z największych realizacji tego typu – serwis *Europeana* [<http://www.europeana.eu/>], który uruchomiony w listopadzie 2008. Projekt stał się realizacją inicjatywy *i2010* finansowanej przez Komisję Europejską w ramach w ramach programu *eContentplus* (2005-2008) i ma na celu zgromadzenie całego dziedzictwa kulturowego Europy (zbiory biblioteczne, muzealne i archiwalne). Warto dodać, że projekt funkcjonował w postaci pilotażowej już od 2006 roku jako *EDLnet* [<http://edlproject.eu/>] (Ayris 2009). Podstawą *Europeany* były zasoby nieudanego przedsięwzięcia UE *The European Library*, które realizowano w latach 2001-2004 [<http://www.theeuropeanlibrary.org>]. Aktualnie *Europeana* zrzesza ok. 1500 instytucji (ok. 200 dostawców oferujących ponad 15 mln obiektów). Znaczną grupę partnerów (16) stanowią narodowi agregatorzy, m.in.: francuska *Culture.fr/Collections* (480 instytucji, w tym np. Luwr), brytyjska *Collections Scran* (380); polska *Federacja Bibliotek Cyfrowych* (250); niemiecka *Bayerische Staatsbibliothek* (158); rumuńska *cIMeC - Institutul de Memorie Culturala* (59); holenderska *Het Geheugen van Nederland* (57) i in.¹⁷.

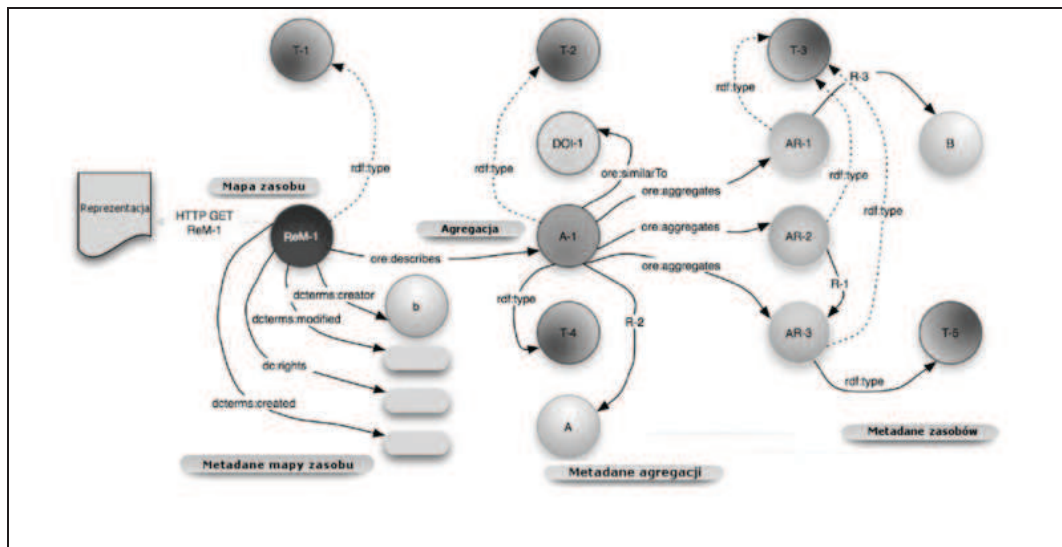
Na marginesie warto dodać, że omawiany protokół OAI-PMH był szeroko wykorzystywany do lokalizacji zasobów przez popularne wyszukiwarki (*Google, Yahoo*) oraz do aktualizacji danych w portalach społecznościowych (np. *Wikipedii*).

Realizacje projektów rozproszonych opartych na OAI-PMH mają jednak pewną słabość. Protokół ten nie potrafi poprawnie reprezentować struktury obiektu wielopoziomowego i powiązań z innymi obiektami oraz wizualizować tych zależności. Wyjściem naprzeciw temu problemowi jest nowa specyfikacja OAI-ORE (ang. Open Archive Initiative – Object Reuse and Exchange)¹⁸ ogłoszona w 2008 roku. Specyfikacja ta (rys. 6) wprowadza nowy typ zasobu – agregację, która identyfikuje oraz reprezentuje zbiór powiązanych obiektów. Ma ona własny URI, lecz nie zgodnie z postulatami sieci semantycznej, nie posiada swojej reprezentacji. W tym zakresie agregację uzupełnia mapa zasobu utrzymana w postaci czytelnej do przetwarzania maszynowego czyli RDF.

¹⁷ Europeana / partnerzy – <http://www.europeana.eu/portal/partners.html> [2011.05.21].

¹⁸ OAI-ORE – <http://www.openarchives.org/ore/> [2011.05.23].

Rys. 6, Schemat komunikacji przy użyciu OAI-ORE (Lewandowska, Werla 2009)



W praktyce więc mapa zasobu kieruje do agregacji, zaś ta na wskazuje na konkretny zasób. W mapie zasobu można zawrzeć: 1. opis relacji między obiektami; 2. informacje o własnościach konkretnych zasobów; 3. informacje o własnościach samej mapy zasobu (Lewandowska, Werla 2009). Protokół nie doczekał się wprowadzić jeszcze większych wdrożeń, lecz jego możliwości są obiecujące i szeroko dyskutowane w kręgach specjalistów, np. *DELOS* i *Google*. Innym przykładem dużego systemu rozproszonego jest projekt *DRIVER* [<http://www.driver-community.eu/>] scalający zasoby naukowe. Jest to wynik prac dwóch projektów finansowanych przez Komisję Europejską w latach 2006-2009. Z technicznego punktu widzenia jest on najbardziej innowacyjny, gdyż opiera się na tzw. D-Net technology, czyli architekturze zorientowanej na usługi. Aktualnie serwis ten daje dostęp do ok. 3,5 mln z 38 krajów (295 repozytoriów).

Poza systemami rozproszonymi czołową rolę odgrywają współcześnie duże serwisy międzynarodowe, które można określić, jako megaprojekty. Do grupy tej należy zaliczyć: *Google Books* (2004), *Internet Archive/OCA* (2005), *Universal Digital Library – Million Book Collection* (2001) oraz omówioną wcześniej *Europeanę*. Istnieje nadto znaczna liczba projektów narodowych, regionalnych, specjalistycznych oraz trudna do oszacowania liczba projektów instytucjonalnych.

Jedną z największych współczesnych inicjatyw digitalizacyjnych był projekt *Google Print*, rozpoczęty formalnie jesienią 2004 roku (od 2005 pod nazwą *Google Books*). Na wstępie warto przypomnieć, że zainteresowanie bibliotekami twórców *Googla* Lawrence'a Page i Sergeya Brina nie było przypadkowe. Obaj rozpoczynali bowiem swą karierę naukową jako doktoranci w Stanford pracując przy projekcie *Stanford Digital Library Project* (część

Digital Library Initiative). Google tuż po ogłoszeniu planu uruchomienia nowej usługi intensywnie zabiegał o współpracowników posiadających zasoby i w przeciągu trzech lat (2005-2007) pozyskał jako partnerów 19 dużych bibliotek amerykańskich i europejskich¹⁹. Szczegóły umów nie są znane, lecz wiadomo, że biblioteka partnerska otrzymywała cyfrową kopię każdej skanowanej przez Google książki z jej zbiorów. Prócz wyliczonych bibliotek do projektu w kolejnych latach przyłączyło się wiele innych podmiotów, w szczególności wydawców, którym Google oferował korzystne warunki współpracy (Toobin 2007). Udostępniany zasób rósł szybko: w listopadzie 2008 osiągnięto 7 mln, zaś w czerwcu 2010 – 12 mln; wynik ten ośmielił twórców do karkołomnej deklaracji, wg której serwis zobowiązał się do zeskanowania w ciągu następnej dekady (2010-2020) wszystkich znanych książek na świecie – czyli ok. 130 mln. Serwis *Google Books* od początku swego działania wzbudzał liczne kontrowersje natury prawnej (Franke 2007). Już rok po debiucie skierowano przeciw portalowi pierwsze pozwy zbiorowe w Stanach Zjednoczonych (w 2008 zakończyły się one ugodą); w następnych latach kolejne procesy (instytucje w Niemczech, Francji i duże korporacje, np. Microsoft). Część z nich została zakończona ugodami, inne są w toku. System *Google Books* wydaje się być jednak skutecznie zabezpieczony przed łamaniem prawa, gdyż udostępnia książki w trzech trybach: pełnym – tylko pozycje z domeny publicznej; oraz ograniczonym – widoczne są fragmenty lub cytaty z zawartości. Abstrahując od innych cech serwisu warto podkreślić, że *Google Books* nie jest biblioteką cyfrową *sensu stricto*, lecz hybrydą (księgarnia-biblioteka), gdyż książki inne niż z domeny publicznej mogą być na wniosek wydawcy w każdej chwili usunięte z serwisu. Niemniej liczba pozycji z domeny publicznej jest znaczna (ok. 1 mln w 2008 roku i ok. 1,7 mln w 2010), co sprawia że *Google Books* jest liczącym się graczem w omawianym obszarze. Niezaprzeczalnym walorem serwisu jest przytłaczająca popularność samego serwisu matki – wyszukiwarki *Googe* (ok. 90-93% rynku)²⁰, jak i innych usług specjalistycznych, wśród których wschodzący *Google Scholar* zaczyna pełnić coraz donioślejszą rolę w świecie nauki.

W październiku 2005 roku w odpowiedzi na inicjatywę Google podjęto konkurencyjny projekt pod auspicjami konsorcjum *Open Content Alliance* (OCA), które zostało utworzone w wyniku porozumienia pomiędzy serwisem *Internet Archive*, *Yahoo!* oraz Adobe System, European Archive, University of Kalifornia, HP Labs, brytyjskim The

¹⁹ M.in. Bavarian State Library, Princeton University, University of California, The National Library of Catalonia, University Complutense of Madrid, Harvard University, University of Michigan, The New York Public Library, Oxford University, Stanford University, University of Texas at Austin, University of Virginia, University of Wisconsin–Madison – <http://books.google.pl/googlebooks/partners.html> [2011.05.21].

²⁰ Search Engine Barometer – <http://en.atinternet.com/Resources/Surveys/Search-engine-barometer/index-1-2-6-0.aspx> [2011.05.21].

National Archive) i wieloma innymi. Konsorcjum działa w oparciu o współpracę organizacji rządowych, non-profit, kulturalnych i technicznych z całego świata i ma zadanie trwałe zabezpieczenie szeroko pojętego dziedzictwa kulturowego (Urbaniec 2006). OCA nie tworzył jednak odrębnego portalu, lecz rozbudował istniejący wcześniej serwis należący do *Internet Archive* (IA) [<http://www.archive.org/>]. Portal ten założył w 1996 roku Brewster Kahle dzięki dotacji Alexa Internet i miał początkowo na celu archiwizację szeroko pojętych zasobów internetowych (stron web, plików programów, wideo i audio oraz tekstów). Z punktu widzenia bibliotek interesujący jest jedynie serwis tekstowy (działający od 1999), w którym udostępniane są zbiory książkowe. Serwis ten w sensie funkcjonalnym jest biblioteką cyfrową w ścisłym znaczeniu (jest nawet członkiem American Library Association). W 2005 roku, po powstaniu OCA utworzono przy IA kilka grup roboczych (m.in. do straw metadanych, ochrony, formatu, skanowania i protokołów komunikacyjnych), nade wszystko jednak przystąpiono do rozbudowy zasobów. Zrąb główny portalu tworzą dzieła zdigitalizowane przez instytucje partnerskie (23 biblioteki i archiwa)²¹ oraz zasoby z domeny publicznej pozyskane legalnie z innych portali (np. z *Google Books*). Aktualnie dominują zasoby z domeny, lecz udostępniana jest też pewna liczba zbiorów współczesnych objętych prawami, dla których autorzy udzielili wyraźną czytelną zgodę na publiczne wykorzystanie (dla niewielkiej części dostęp jest limitowany). Na początku 2005 roku zbiór książkowy IA był niewielki i liczył nieco ponad 25 tys., jednak po utworzeniu OCA szybko się powiększył (ok. 1 mln w 1998 oraz 2,8 mln w 2011); warto przy tym dodać, że do zbioru IA od 2007 roku włączone się książki z domeny publicznej dostępne w *Google Books* (ok. 900 tys.). Warto też podkreślić, że IA przyjęło specyficzne rozwiązanie techniczne, polegające na tworzeniu kopii zasobów pobranych od instytucji partnerskich i ich deponowania na własnej platformie w wielu formatach. Ostatnia cecha bardzo korzystnie wyróżnia IA na tle innych serwisów. Książki są tu bowiem dostępne w wielu równoległych wersjach: on-line, czyli renderowanie przez flasha, PDF, DJVu, czysty tekst, HTML, formaty mobilne (ePub, kindle, mobi) oraz przeznaczonym dla niewidomych i niedowidzących format DAISY (od 2009 dostępnych jest nieco ponad 1 mln pozycji w tym formacie). W miarę rozwoju projektu funkcjonalność kolekcji IA została poszerzona, gdyż zbiory zostały zintegrowane z *Open Library*, projektem rozwijanym przez OCA od 2006. *Open Library* jest uniwersalnym serwisem rozproszonym (funkcjonalnie zbliżonym do katalogu *WorldCat*), dla którego dostawcami danych stały się

²¹ OCA Contributors –<http://www.opencontentalliance.org/contributors/> [2011.05.21].

katalogi bibliotek amerykańskich i kanadyjskich (przede wszystkim Library of Congress) oraz dane Amazon (O'Leary 2009).

The Universal Digital Library - Million Book Collection (UDL) [<http://www.ulib.org/>] – to inicjatywa, nad którą prace rozpoczęły się w 2001 roku, a oficjalny debiut nastąpił we wrześniu 2003. Projekt ów jest wspólnym przedsięwzięciem amerykańsko-chińsko-indyjskim. Ze strony amerykańskiej prace koordynuje Carnegie Mellon University, zaś technicznie jest realizowany głównie w Zhejiang University (Chiny) i Indian Institute of Science (Indie) oraz siedmiu chińskich i ośmiu indyjskich uczelniach partnerskich; w projekcie uczestniczą też egipska Bibliotheca Alexandrina. UDL koncentruje się na książkach z domeny publicznej oraz dziełach, dla których pozyskano stosowne umowy od autorów i wydawców. Do grudnia 2007 roku udostępniono ponad 1,5 mln dzieł, wśród których dominują publikacje chińskie – 970 tys., nieco mniej udostępniono w języku angielskim – 360 tys., kilku językach indyjskich (telungu – 50 tys.; urdu – 17 tys.; sanskryt – 7 tys. i in) oraz arabskim – 40 tys., perskim – 7 tys. i kilku innych. W UDL zastosowano kilka interesujących osobliwości technicznych: baza danych istnieje na trzech serwerach lustrzanych (USA, Chiny, Indie), serwis jest wyposażony w wielojęzyczny interfejs (w tym języki wschodnie), zaś publikacje są poddawane OCR – co dla języków azjatyckich jest dużym wyzwaniem, a zasób serwowany jest w kilku formatach, m.in. djvu i TIFF; do obsługi ostatniego typu służy dedykowana wtyczka AlternatIFF [<http://www.alternatiff.com/>]. W listopadzie 2005 r. Indie, Chiny i USA uzgodniły, że UDL do przyłączy się do *Open Content Alliance*; w efekcie zbiory UDL stały równolegle dostępne w *Internet Archive* oraz katalogu *Open Library*.

Wśród licznych inicjatyw o charakterze międzynarodowym należy wspomnieć o kilku mniejszych projektach, m.in. inicjatywie Library of Congress i UNESCO – *World Digital Library* (WDL) [<http://www.wdl.org/>] (od 2009, ok. 1300 poz.); *International Children's Digital Library Foundation* (ICDL) [<http://en.childrenslibrary.org/>], tworzonej od 2002 roku na University of Maryland (ok. 2 tys. pozycji), która istotnie wzbogaciła zasób po roku 2006, gdy uzyskała finansowanie m.in. z Library of Congress i Microsoftu.

Liczba znaczących projektów o charakterze regionalnym i narodowym jest trudna do ustalenia. Punktem wyjściowym do oszacowania ich liczby może być baza UNESCO i IFLA *Directory of Digitized Collections* [<http://www.unesco.org/webworld/digicol/>], związana z inicjatywą *Memory of the World*. Źródło to jest wprawdzie dalekie od kompletności (238 bibliotek), ale z uwagi na staranny dobór w dobrym przybliżeniu odzwierciedla proporcje regionalne. Według stanu z 2007 roku na Europę i Amerykę Północną przypadało 83% bibliotek cyfrowych, Azję i Pacyfik – 15%, Amerykę Południową – 1,6% oraz Afrykę –

0,4%. Dane te potwierdza omawiany wyżej *OAISTER* – gdzie na Europę i Amerykę Północną przypadało 88% bibliotek cyfrowych (w tym 47% – Europa 41% i Ameryka)²².

W największym pod względem inicjatyw regionie – czyli Europie zdecydowana większość istniejących bibliotek uczestniczy w roli partnerów omówionej wcześniej *Europeany*. W przedsięwzięciu uczestniczy ok. 200 bibliotek cyfrowych, które udostępniają zasoby samodzielnie oraz ponad tysiąc, które czynią to za pośrednictwem sieci agregujących (16 sieci)²³. W grupie tej można wyróżnić kilka typów dostawców: repozytoria i agregatorzy narodowi, np. *Gallica* (Francja) [<http://gallica.bnf.fr/>]; *Digital Library of Slovenia* (Słowenia) [<http://www.dlib.si/v2/>]; *The British Library* (Wielka Brytania) [<http://www.bl.uk/eresources/main.shtml>]; *Federacja Bibliotek Cyfrowych* (Polska) [<http://fbc.pionier.net.pl/>]; *Biblioteca Nacional Digital* (Portugalia) [<http://bnd.bn.pt/>]; *NLL Digital Library* (Łotwa) [http://www.lnb.lv/digitala_biblioteka/index_anglu.htm]; *Zentrales Verzeichnis Digitalisierter Drucke* (Niemcy) [<http://www.zvdd.de/startseite/>]; *Электронная Ленинка РГБ* (Moskwa) [<http://www.rsl.ru/>] oraz *Электронная библиотека РНБ* (St. Petersburg) [<http://www.nlr.ru/eng/>] (obie w Rosji); *Magyar Elektronikus Könyvtár* (Węgry) [<http://mek.oszk.hu/indexeng.phtml>]; *Helmi Digitised* (Finlandia) [<http://www-db.helsinki.fi/helmi/en/>]. Inny, dość liczny typ biblioteki to narodowe repozytoria czasopism, np.: *ANNO AustriaN Newspapers Online* (Austria) [<http://anno.onb.ac.at/>]; *Deutsche Exilpresse 1933-1945 Digital* (Niemcy) [<http://deposit.ddb.de/online/exil/exil.htm>]; *EPA Elektronikus Periodika Adatbázis* (Węgry) [<http://epa.oszk.hu/>]; *Internet Library of Early Journals* (Wielka Brytania) [<http://www.bodley.ox.ac.uk/ilej/>]. Wreszcie typy specjalne; najczęściej występują narodowe repozytoria historyczne – np. holenderski *Memory of the Netherlands* [<http://www.geheugenvannederland.nl/gvnmnl/>] (Bakker 1999) lub narodowe repozytoria literackie – np. *Poesie Française* [<http://poesie.webnet.fr/home/>]²⁴. Mozaikę tę uzupełnia duża liczba mniejszych portali narodowych, specjalnych, regionalnych i instytucjonalnych.

Obraz bibliotek cyfrowych w Stanach Zjednoczonych nie jest tak skonsolidowany, jak w Europie, gdyż zabrakło narodowego reprezentanta. Niemniej większość ważnych podmiotów o charakterze narodowym skupiło się wokół Library of Congress lub *Interent Archive*. Największy zasób skupia opisany szczegółowo wcześniej projekt *American Memory* (AM) [<http://memory.loc.gov/ammem/>], którego uzupełnieniem jest kilka innych serwisów

²² Por. przyp. 14.

²³ Europeana Partners – <http://www.europeana.eu/portal/partners.html> [2011.05.14].

²⁴ Pełniejszy wykaz inicjatyw europejskich i amerykańskich. W. M. Kolasa: *Największe światowe biblioteki cyfrowe i repozytoria OA* – http://mbc.malopolska.pl/dlibra/text?id=dlib_world [2011.05.21].

powstałych w czasie realizacji *National Digital Library Program* (NDLP)²⁵; warto dodać że poszczególne kolekcje AM²⁶ są tworamami na tyle samodzielnymi, że są rozróżniane w literaturze, np. *Votes for Women 1848-1921*. Bardziej klarowną konstrukcję ma literatura amerykańska w serwisie *Internet Archive*; tworzy ją subkolekcja *American Libraries* [<http://www.archive.org/details/americana>] – która scala w jednolitym interfejsie ogromny zasób *americanów* (ok. 1,5 mln) udostępnionych przez najważniejsze księżnice i muzea amerykańskie. Nie jest to bynajmniej komplet amerykańskich inicjatyw. Wiele starszych projektów funkcjonuje niezależnie, np. *The Berkeley Digital Library SunSITE* [<http://sunsite.berkeley.edu/>]; *The Humanities Text Initiative* [<http://www.hti.umich.edu/>]; *Historic American Sheet Music* [<http://scriptorium.lib.duke.edu/sheetmusic/>]. Istnieją też odrębne projekty tematyczne, np. kolekcja gazet *Historic American Newspapers* [<http://chroniclingamerica.loc.gov/>] czy serwis historyczny *Making of America* [<http://quod.lib.umich.edu/m/moagrp/>]. Obraz ten dopełnia ogromna liczba dużych bibliotek stanowych lub uniwersyteckich, których zbiory są niemal w komplecie dostępne w IA, wyróżniają się na tym tle: *California Digital Library* [<http://www.cdlib.org/>]; *The University of Pittsburgh Digital Research Library Collections* [<http://digital.library.pitt.edu/>]; *The Indiana University Digital Library Program* [<http://www.dlib.indiana.edu/>]; *Kentuckiana Digital Library* [<http://kdl.kyvl.org/>]; *LOUISiana Digital Library* [<http://louisdl.louislibraries.org/>]; *Michigan State University Libraries Digital and Multimedia Center* [<http://digital.lib.msu.edu/>]; *NYPL Digital Library Collections* [<http://www.nypl.org/digital/>]; *Oregon State University Library Digital Collections* [<http://digitalcollections.library.oregonstate.edu/>]; *The University of Michigan Digital Library Production Service Collection* [<http://www.umdl.umich.edu/>] czy *The University of Chicago Library Digital Activities and Collections* [<http://www1.lib.uchicago.edu/cgi-bin/nand/search/diglist>]²⁷. Podobną strukturę mają zasoby Kanady, które skupione są w dwu serwisach: IA oraz przy księżnicy narodowej. Warto podkreślić, że ogromne znaczenie w tym obszarze miał narodowy program *The Canadian Initiative on Digital Libraries*²⁸, który realizowano w latach 2000-2006. Aktualnie 38 największych kanadyjskich repozytoriów (łącznie 338 tys. jednostek) tworzy zwartą kolekcję wydzieloną w IA

²⁵ Por. 6.

²⁶ Wykaz kolekcji AM – <http://international.loc.gov/ammem/dli2/html/list.html> [2011.05.21].

²⁷ Pełniejszy wykaz dostawców do IA – <http://www.archive.org/details/americana> [2011.05.21].

²⁸ CIDL – <http://epe.lac-bac.gc.ca/100/206/301/lac-bac/cidl-ef/2007-09-28/cidl/040021-300-e.html> [2011.06.01].

[<http://www.archive.org/details/toronto>]; pozostałe inicjatywy (w tym 40 dużych kolekcji)²⁹ wyczerpująco prezentuje serwis [<http://www.collectionscanada.gc.ca/a-z-index/index-e.html>] prowadzony przez Library and Archives Canada.

Liczba bibliotek cyfrowych na kontynencie afrykańskim jest symboliczna. W zasadzie jedynym dużym przedsięwzięciem jest udział egipskiej Bibliotheca Alexandrina w opisanym uprzednio projekcie UDL [<http://www.ulib.org/>] oraz kilka egipskich kolekcji zlokalizowanych na jej serwerach [<http://www.bibalex.org>], m.in. *Modern Egypt* [<http://modernegypt.bibalex.org>]. Inne inicjatywy afrykańskie pojawiają się incydentalnie, np. powstały na University of Natal w RPA serwis *DISA* [<http://disa.nu.ac.za/>]. Stan ten nieco łagodzą portale zagraniczne interesujące się Afryką. Najważniejszym projektem tego typu jest powstały w 2008 w USA dzięki inicjatywie serwisu JSTOR serwis *ALUKa* [<http://www.aluka.org/>] – portal w całości poświęcony Afryce i gromadzący *africana*. Do końca 2007 roku serwis zgromadził pokaźny zasób: 76 kolekcji, 323 tys. jednostek pochodzących z 31 krajów (w tym 14 państw afrykańskich).

Proces tworzenia bibliotek cyfrowych w Ameryce Łacińskiej znajduje się dopiero na etapie wstępnym. W większości państw pojawiły się już kolekcje eksperymentalne, lecz nie istnieje ani jedno duże repozytorium. Według bazy *OAISER* najwięcej serwisów istnieje Brazylii (23) – są to jednak głównie repozytoria naukowe, zaś uniwersalnych kolekcji działa niewiele, np. *Biblioteca Digital de Obras Raras e Especiais* [<http://www.obrasraras.usp.br/>]. Stan ten jednak stopniowo ewoluuje i od niedawna pojawiają się inicjatywy skoordynowane, np. *Brazilian Government Document Digitization Project* [<http://www.crl.edu/brazil>]. W pozostałych państwach regionu (Chile, Boliwia, Kuba, Nikaragua, Peru, Urugwaj, Argentyna, Kolumbia, Wenezuela) istnieje po 1-5 bibliotek cyfrowych, najczęściej są to małe kolekcje eksperymentalne afiliowane przy bibliotekach narodowych lub ministerstwach, np. chilijska *Memoria Chliena* [<http://www.memoriachilena.cl/elibros/e-libros.asp>] lub argentyńska *Biblioteca Digital de Tratados* [<http://tratados.cancilleria.gob.ar/index.php>]. Obraz dopełniają inicjatywy zagraniczne, np. działający przy University of Miami portal poświęcony Kubie – *Cuban Heritage Collection* [<http://www.library.miami.edu/chc/>].

Główną inicjatywą azjatycką, która od 2003 roku zogniskowała wysiłki chińskie i indyjskie jest omówiona wcześniej *Universal Digital Library* (UDL) [<http://www.ulib.org/>]. Podkreślić jednak należy, że główną siłą koordynującą prace w tym zakresie były cykliczne konferencje ICADL, które odbywały się w regionie od 1998 roku. Dynamikę przemian

²⁹ Canada's Digital Collections: Sharing the Canadian Identity on the Internet - <http://www.collectionscanada.gc.ca/publications/002/015002-2170-e.html> [2011.06.01].

odzwierciedla liczba członków ICADL na przestrzeni 1998-2004. W okresie tym wzrosła ich liczba z 10 do ok. 40. Projekty działające w poszczególnych krajach są bardzo zróżnicowane, lecz przeważają projekty scentralizowane. Model taki reprezentują Chiny, gdzie prace nad głównymi inicjatywami narodowymi koordynuje rząd, wynikiem tych działań jest narodowy projekt *Digital Library of China* [<http://www.d-library.com.cn/>] oraz jej dopełnienie *China Digital Library for Visual Impairment* [<http://www.cdli.cn/>]. Prace przygotowawcze nad ich realizacją trwały od 2000 roku, zaś oficjalne otwarcie w 2005. W okresie tym na potrzeby pierwszego projektu przygotowano imponujący zasób: 1,1 mln książek, 450 tys. obrazów, 12 tys. tyt. czasopism i 600 tyt. gazet (Leye& Ping 2009); zasób ten rośnie w imponującym tempie i do końca 2009 roku urósł do ponad 27 mln (Zhen 2010). Główna inicjatywa w Indiach to *The Digital Library of India* [<http://www.dli.cdacnoida.in/>], wspólny projekt 21 narodowych i kilku zagranicznych instytucji, nad którymi prace koordynują Indian Institute of Science oraz International Institute of Information Technology (Hyderabad). Istnieją jednak przesłanki, by sądzić, że zawartość DLI pokrywa się z kolekcjami UDL (Swanepoel 2008). Znaczne osiągnięcia w zakresie budowy infrastruktury cyfrowej ma także Korea Południowa, gdzie istnieje 6 dużych projektów kierowanych przez National Digital Library of Korea [<http://www.dlibrary.go.kr/>] (od 2006). W pozostałych krajach regionu istnieją jedynie repozytoria w formie załączkowej, np. *Filipiniana* [<http://www.filipiniana.net/>] czy malezyjska *myLibrary* [<http://www.mylib.com.my/>] (od 2000). Istotnym dopełnieniem tego obrazu są tematyczne serwisy zagraniczne; wyróżnia się w tym zakresie *Portal to Asian Internet Resources* [<http://digicoll.library.wisc.edu/PAIR/>] – amerykańska inicjatywa prowadzona od 2005 pod kierunkiem University of Wisconsin-Madison.

Odrębnie należy rozpatrywać projekty japońskie, gdyż mają one własną specyfikę. W Japonii dominują bowiem rozwiązania zbliżone do Europy lub USA, tj. działają tu w szczególności sieci informacji specjalistycznej i istnieje ogromna liczba repozytoriów instytucjonalnych. Serwisy tego typu działają przy większości ośrodków badawczych i uniwersytetów, np. *Kyoto University RIR* [<http://repository.kulib.kyoto-u.ac.jp/dspace/>]; *Nagoya University Academic Knowledge Factory* [<http://akf.nul.nagoya-u.ac.jp/>]; *Hiroshima University IR* [http://ir.lib.hiroshima-u.ac.jp/portal/index_e.html]. Istnieje nadto kilka serwisów uniwersalnych, które można uznać za narodowe; taki charakter na sześć kolekcji prowadzony przez National Diet Library [<http://www.ndl.go.jp/en/data/endl.html>], wśród których wyróżnia się serwis dawnych książek *Digital Library from the Meiji Era* [<http://kindai.ndl.go.jp/>] (150 tys. jednostek) oraz centralny portal cyfrowy archiwów japońskich: *PORTA* [<http://porta.ndl.go.jp/portal/dt>].

Zainteresowanie bibliotekami cyfrowymi w regionie Australii i Oceanii ma długą tradycję. Już od połowy lat 90. z inicjatywy bibliotek narodowych Australii i Nowej Zelandii uruchomiono wiele lokalnych projektów realizowanych przede wszystkim na uniwersytetach (Ianella 1996). Inicjatywy te zorientowane były zarówno na ochronę dziedzictwa kulturowego, np. *The Australian Museums On Line* czy *Preserving Oral History Recordings Project*, jak i dotyczyły dystrybucji wydawnictw współczesnych np. *Electronic Journal Project*. W efekcie infrastruktura cyfrowa regionu prezentuje poziom zbliżony do europejskiego (Swanepoel 2008). Wśród znaczących inicjatyw australijskich należy wyliczyć m.in.: *Project Gutenberg of Australia* [<http://gutenberg.net.au/>] realizowany od 2001; *Picture Australia* uruchomiony w 1998 [<http://www.pictureaustralia.org/>] (1,8 mln obrazów), *Australian Periodical Publications 1840-1845* [<http://www.nla.gov.au/ferg/>] oraz ponad 90 innych zestawionych w specjalnym przewodniku *Australian Digitisation Project* [<http://www.nla.gov.au/libraries/digitisation/projects.html>]. Znaczące osiągnięcia ma także Nowa Zelandia, która w ramach narodowego projektu, który rozpoczęto w 1994 roku zbudowała wiele kolekcji, w szczególności portal narodowy *New Zealand Digital Library* [<http://www.nzdl.org/>], realizowany w University of Waikato. Warto dodać, że na potrzeby projektu zaprojektowano oprogramowanie Greenstone [<http://www.greenstone.org/>], rozwijane później we współpracy z UNESCO, które jest dystrybuowane bezpłatnie za licencji GNU/GPL. W oparciu o system Greenstone powstało kilkaset bibliotek cyfrowych na całym świecie, np. *Biblioteca Nacional de Uruguay* [<http://coleccionedigitales.bibna.gub.uy>], *Armenian Rare Books* [<http://greenstone.flib.sci.am/gsd/cgi-bin/library.cgi>], czy *Oxford Digital Library* [<http://www2.odl.ox.ac.uk/>]. Wśród wielu innych przedsięwzięć prowadzonych w Nowej Zelandii warto też wspomnieć o kilku kolekcjach utrzymywanej przez bibliotekę narodową, m.in. serwisach dawnych czasopism: *AtoJs* [<http://atojs.natlib.govt.nz>] i *Papers Past* [<http://paperspast.natlib.govt.nz/>].

5. Perspektywy

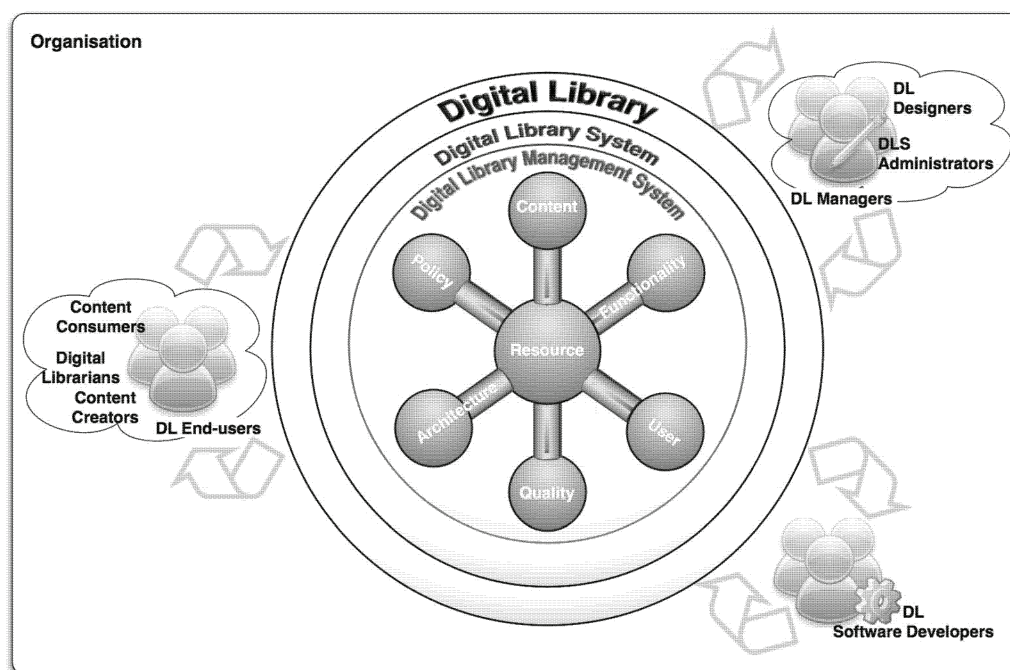
Rozwój prac w zakresie digitalizacji i budowy bibliotek cyfrowych, który rozpoczął się w latach 90. ubiegłego wieku i przyspieszył w pierwszej dekadzie bieżącego stulecia zaowocował znaczącym przyrostem projektów (por. rys. 1). Warto jednak zauważyć, że są one mocno zróżnicowane. Zdecydowana większość z nich to serwisy ukierunkowane na treści udostępniania (a nie funkcjonalność), zaś ich model organizacyjny opiera się na współpracy dostawców treści. Takie podejście było przez wiele lat dominujące i najczęściej uzasadniane jako strategia w celu zmniejszenia kosztów utrzymania biblioteki cyfrowej. Pierwszym

krokiem ku integracji serwisów (spójność na poziomie danych) było upowszechnienie się protokołów OAI-PMH i budowa sieci rozproszonych. Wnikliwa analiza przekonuje jednak, że znacznie trudniej osiągnąć spójność na poziomie funkcjonalności ze względu na wielką różnorodność modeli i ontologii przyjętych przez istniejące systemy. Innymi słowy jest to negatywny skutek braku systematycznego podejścia do interoperacyjności przez projektantów. Bariere tę od niedawna próbują przełamać niektóre projekty najnowszej generacji, np. *DL.org* [<http://www.dlorg.eu/>] (społeczność w ramach DELOS), której prace torują drogę dla pełnej realizacji interoperacyjności w systemach, co jest warunkiem koniecznym dla realizacji spójnej infrastruktury globalnej (Candela, Castelli & Pagano 2011).

Dyskusja nad budową takiej infrastruktury, nazwanej często globalną biblioteką cyfrową (Nahotko 2010) toczy się w dyskursie naukowym już od kilku lat. Warto jednak zaznaczyć, że w odróżnieniu od poprzedniej dekady, gdy prym wiodły ośrodki amerykańskie, w XXI wieku najbardziej wartościowe prace w tym zakresie powstały w Europie (głównie w ramach DELOS). Szczególnie interesująca jest dyskusja dotycząca modelowania bibliotek cyfrowych w zmieniających się warunkach. Jeszcze na początku XXI wieku najpopularniejszy był model OASIS (2002), który dzielił aktorów występujących w otoczeniu biblioteki na producentów, konsumentów i administratorów (Dudczak i in., 2007); konkurowały z nim inne rozwiązania, m.in. model CKESS dedykowany do współpracy społecznościowej (2002), model Bessera dotyczący współpracy rozproszonej z wykorzystaniem agentów (2002) i model 5S (2005), nazwany od pierwszych liter elementów modelu - streams (strumienie), structures (struktury), spaces (przestrzenie), scenarios (scenariusze) i societies (społeczności) oraz inne (Nahotko, 2010: 22-23). Warto podkreślić, że żaden z nich nie operował na odpowiednio wysokim poziomie abstrakcji, co sprawiało, że w rozmaitych okolicznościach modele okazywały się ułomne. Zmianę przyniósł dopiero rok 2006, wraz z pierwszą wersją modelu referencyjnego opracowanego w DELOS. Początkowo głównym celem DELOS (jako grupy ekspertów) było podnoszenie poziomu wiedzy w dziedzinie bibliotek cyfrowych poprzez koordynację prac europejskich zespołów badawczych. Jednym z pierwszych osiągnięć było nawiązanie współpracy z amerykańską National Science Foundation i powołanie wspólnych grup roboczych. Zespoły interesowały się m.in. techniką oraz kwestiami społecznymi i ekonomicznymi oraz – co najważniejsze – opublikowano zestaw zaleceń w odniesieniu do interoperacyjności, m.in. metadanych, własności intelektualnej i ekonomii, zasobów oraz informacji wielojęzycznych; zalecenia te są opublikowane w specjalnym czasopiśmie *International Journal of Digital Libraries* [<http://www.dljournal.org/>] (Griffin, Peters, Thanos 2005). Po przekształceniu DELOS z

grupy w sieć (DELOS Network), jej główną misją stała się integracja i koordynacja działań badawczych europejskich zespołów badawczych w zakresie bibliotek cyfrowych. Głównym zaś jej osiągnięciem stał się opracowany w 2006 model referencyjny biblioteki cyfrowej *Digital Library Reference Model* [DLRM], opisujący jej ramy koncepcyjne na bardzo wysokim poziomie abstrakcji (Candela, Castelli & Pagano 2007; Candela, Nardi 2011).

Rys. 7, Model referencyjny biblioteki cyfrowej (Candela, Nardi 2011)



Według DLRM biblioteka cyfrowa to – wirtualna organizacja, który kompleksowo gromadzi, zarządza i długotrwale zachowuje treści cyfrowe oraz oferuje docelowym społecznościom wyspecjalizowane funkcje, dbając o ich jakość i zgodność z przyjętymi kierunkami polityki (Candela, Nardi 2011). W dokumencie przedstawiono specyfikację trzech warstw modelu biblioteki opisanego wcześniej w *Digital Libraries Manifesto* (Candela, Castelli 2006) oraz podjęto próbę zdefiniowania funkcjonalności na poziomie każdej z tych warstw. W DLRM bibliotekę cyfrową tworzą trzy odrębne warstwy (rys. 7): 1. Biblioteka cyfrowa (Digital Library); System biblioteki cyfrowej (Digital Library System) i System zarządzania biblioteką cyfrową (Digital Library Management System). Rozróżnienie to jest istotne z poziomu administratora, zaś z punktu widzenia użytkownika warstwy te są tożsame. Należy zwrócić uwagę, że komponenty: content, policy, architecture, quality, user, functionality [odpowiednio: zawartość, polityka, architektura, jakość, funkcjonalność] znajdują się wewnątrz systemu i podlegają zarządzaniu. Zakres funkcjonalny został z kolei podzielony na pięć podstawowych grup, w ramach których wyróżniono konkretne funkcje: 1.

Uzyskaj dostęp do informacji - przeglądaj, wizualizuj, zachowaj, przetłumacz, przeszukuj, dostarcz; 2. Zarządzaj informacją – stwórz, uaktualnij, przekształć, nanieś adnotacje, złóż, prześlij, waliduj; 3. Zarządzaj biblioteką cyfrową – rozpowszechnij, publikuj, wycofaj, przechowaj, zarządzaj zasadami, zarządzaj kolekcjami, zarządzaj użytkownikami (rejestruj użytkowników, zarządzaj rolami); 4. Personalizuj – wyraż zainteresowanie, zaloguj się, zapisz się, zastosuj profil, dostarcz informację zwrotną; 5. Współpracuj (Dudczak 2007). Dalej idące wnioski można wyścignąć analizując zdefiniowane klasy aktorów (content consumers, digital librarians, content creator, end user, designers, administrators, managers, software developers [odpowiednio: konsumentów treści cyfrowych bibliotekarzy, twórcy treści, użytkowników końcowych, projektantów, administratorów, menedżerów, twórców oprogramowania]). Każda z tych grup definiuje wobec biblioteki cyfrowej własne oczekiwania i interakcje, np. użytkownik końcowy – język zapytań. Zaprezentowany model jest od strony technicznej zdecydowanie bardziej rozbudowany i pozwala na precyzyjne projektowanie biblioteki cyfrowej. DLRM jest także wyjściem naprzeciw heterogeniczności współczesnych repozytoriów, które z definicji są tak projektowane, aby móc długotrwale przechować dane cyfrowe w dowolnym formacie, opisane przy pomocy różnych schematów metadanych, a ich architektura coraz częściej zorientowana jest na usługi (SOA)³⁰. Takie podejście do projektowania bibliotek umożliwia ich elastyczną integrację z innymi systemami (Dudczak 2007), czyli sprawia, że są zdolne do współpracy z dowolnymi systemami, które powstaną w przyszłości.

Bibliografia

- Athanasopoulos, G., Ioannidis, Y., Fox, E. (2010): A functionality perspective on digital library interoperability. *Lecture Notes in Computer Science*, Vol. 6273, p. 405-408. Dostęp: <http://www.springerlink.com/content/52u4r1524v664187/> [2011.06.01].
- Ayris, P. (2009). The EU-funded Europeana Travel project. *Liber Quarterly: The Journal of European Research Libraries*, 19 (2), p. 63-102. Dostęp: <http://www.europeanatravel.eu/downloads/articleTravel.pdf> [2011.06.02].
- Bakker, T. (1999): Ongoing digital library projects in the Netherlands (1995-2001). *Betriebswirtschaftliche Forschung und Praxis*, Jg. 23, H. 3, s. 375-379. Dostęp: www.bibliothek-saur.de/1999_3/375-382.pdf [2011.06.01].
- Bush, V. (1945): As We May Think, *The Atlantic Monthly*, no 176, p. 101-108. Dostęp: <http://www.ps.uni-saarland.de/~duchier/pub/vbush/vbush-all.shtml> [2011.05.19].

³⁰ SOA (ang. 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 (Dudczak 2007).

- Candela, L., Castelli, D. (2006): *Digital Libraries Manifesto*, [S.l.]: DELOS. Dostęp: http://www.delos.info/index.php?option=com_content&task=view&id=345 [2011.05.21].
- Candela, L., Castelli, D. & Pagano P. (2007): A Reference Architecture for Digital Library Systems: Principles and Applications. *Lecture Notes in Computer Science*, Vol. 4877, p. 22-35, DOI: 10.1007/978-3-540-77088-6_3 [2011.05.21].
- Candela, L., Castelli, D. & Pagano P. (2011): History, Evolution and Impact of Digital Libraries. [in:] Iglezakis, I.; Synodinou, T. & Kapidakis, S. (ed.): *E-Publishing and Digital Libraries: Legal and Organizational Issues*, IGI Global, Hershey, PA: Information Science Reference, p. 1-30.
- Candela, L., Nardi, A. (ed.) (2011): *The Digital Library Reference Model* [ver. D 3.2b]. Dostęp: <http://www.dlorg.eu/index.php/outcomes/reference-model> [2011.05.21].
- Dudczak, A. Heliński, M., Mazurek, C., Parkoła, T., Werla M. (2007): Analiza funkcjonalności wybranych modeli i systemów zarządzania bibliotekami cyfrowymi. *Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej*. Nr 5, Seria: Technologie Informacyjne, s. 1-10.
- Englebart, D.C. (1963). Conceptual framework for the augmentation of man's intellect. [in:] Howerton, P. W. & Weeks, D.C. (Eds.), *Vistas in Information Handling*, Washington, D.C: Spartan Books, p. 1-20.
- Fox, E. A., ed. (1993): *Source Book on Digital Libraries*. Blacksburg: Virginia Tech. Dostęp: <http://fox.cs.vt.edu/DLSB.html> [2011.05.20].
- Fox, E. A. (1999): The Digital Libraries Initiative: Update and Discussion, *Bulletin of the American Society of Information Science*, Vol. 26, no 1, p. 7-11.
- Fox, E. A. (2001): Digital Libraries: Topical Outline. Blacksburg: Virginia Tech. Department of Computer Science. Dostęp: <http://ei.cs.vt.edu/~dlib/tut/FoxAll.pdf> [2011.05.21].
- Franke, J. (2005): Warianty rozwoju cyfrowych kolekcji, [in:] *Bibliologia, biblioteki, bibliotekarze*. Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, s. 73-89.
- Franke, J. (2007): Googletheca Universalis?, [in:] *Biblioteki cyfrowe : projekty, realizacje, technologie*. Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, s. 121-166. Toż: <http://bbc.uw.edu.pl/publication/4> [2011.05.21].
- Garrison, A. (2010): OAISTER: The Roots and the Resource. *PnLA Quarterly*, 74 (2), p. 46-51. Dostęp: <http://unllib.unl.edu/LPP/PnLA%20Quarterly/garrison74-2.htm> [2011.06.01].
- Greenstein, D., Thorin, S. E. (2002): *The digital library : a biography*. Washington: Digital Library Federation. Dostęp: www.clir.org/pubs/reports/pub109/pub109.pdf [2011.05.21].
- Griffin, S., Peters, C. & Thanos, C. (2005): Toward the new-generation digital libraries: recommendations of the NSF/EU-DELOS working groups. *International Journal on Digital Libraries*, 5 (4), p. 253-254. Dostęp: DOI: 10.1007/s00799-004-0093-9 [2011.06.02].
- Hauben, J. (2004): *Vannevar Bush and JCR Licklider: Libraries of the Future 1945-1965*, Conference Paper – Conference in the Institute for Library Science of Humboldt University in Berlin. Dostęp: <http://www.ais.org/~jrh/acn/acn15-2.articles/jhauben.pdf> [2011.05.21].
- Hunter, P. (2005): *OAI and OAI-PMH for absolute beginners: a non-technical introduction*. Geneva: CERN. Dostęp: <http://eprints.rclis.org/archive/00005512/> [2011.05.21].

- Ianella, R. (1996): Australian Digital Library Initiatives. *D-Lib Magazine*. Vol. 2 Iss. 12. Dostęp: <http://www.dlib.org/dlib/december96/12ianella.html> [2011.06.01].
- Kolasa, W. M. (2007): dLibra Digital Library Framework – platforma do budowy bibliotek cyfrowych, [in:] *Biblioteki cyfrowe : projekty, realizacje, technologie*. Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, s. 67-88. Dostęp: <http://bbc.uw.edu.pl/publication/4> [2011.05.21].
- Kolasa, W. M. (2009): Biblioteki cyfrowe : wstęp do problematyki, [w:] *Nowe technologie w bibliotekach publicznych*. Warszawa : SBP, s. 183-196.
- Lebert, M. (2008): Project Gutenberg (1971-2008). Dostęp: <http://www.gutenberg.org/cache/epub/27045/pg27045.html> [2011.05.20].
- Lewandowska, A., Werla, M. (2009) OAI-ORE : Interoperacyjność bibliotek cyfrowych. *PIONIER Magazine*, nr 3 (04), s. 22-23. Dostęp: <http://www.pionier.net.pl/magazine/pl/artykuly/1042/OAI-ORE.html> [2011.05.23].
- Leye, Y. & Ping, Z. (2009). Digital libraries in China: progress and prospects. *Electronic Library*, 27(2), 308-318. Dostęp: DOI: 10.1108/02640470910947656 [2011.06.01].
- Madalli, D. P. (2003): Digital Libraries and Digital Library Initiatives, [in:] *Digital Libraries: Theory and Practice*. Bangalore: Documentation Research & Training Centre, p. 1-8.
- McCray, A. T., Gallagher, M.E. (2001): Principles for digital library development, *Communications of the ACM*. Vol. 44, no 5, p. 49-53.
- O'Leary, M. (2009): Open Content Alliance Embodies Open Source Movement. *Information Today*, 26 (1), p. 37-43.
- Nahotko, M. (2007): *Naukowe czasopisma elektroniczne*. Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich.
- Nahotko, M. (2010): *Komunikacja naukowa w środowisku cyfrowym : globalna biblioteka cyfrowa w informatycznej infrastrukturze nauki*. Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich.
- Raitt, D. (2000): Digital initiatives across Europe. *Computers Libraries*, Vol. 20, no 11/12, p. 26-34. Dostęp: <http://www.infoday.com/cilmag/nov00/raitt.htm> [2011.06.01]
- Swanepoel, M. (2008): *Digitization initiatives: A reconnaissance of the global landscape*. [Lethbridge]: University of Lethbridge. Dostęp: <http://uleth.academia.edu/MarinusSwanepoel/Papers> [2011.05.20]
- Tedd, L. A. & Large, J. A. (2005): *Digital libraries: principles and practice in a global environment*. München: K.G. Saur.
- Toobins, J. (2007): Google's moon shot: The quest for the universal library. *The New Yorker*. (5 February).
- Urbaniec, M. (2006): Open Content Alliance – cyfrowa biblioteka dla wszystkich. *EBIB*. Nr 3 (73). Dostęp: <http://www.ebib.info/2006/73/urbaniec.php> [2011.05.21].
- Wells, H.G. (1938). *World Brain*. Garden City, New York: Doubleday, Doran & Co. Toż skr.: https://sherlock.ischool.berkeley.edu/wells/world_brain.html [2011.05.19].
- Zhen, X. (2010): Overview of Digital Library Development in China. *D-Lib Magazine*, Vol. 16 no 5/6. Dostęp. <http://www.dlib.org/dlib/may10/zhen/05zhen.html> [2011.05.23]