

Henryk Hollender

Uniwersytet Warszawski

Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego

Biblioteki cyfrowe w Polsce: warunki i zagrożenia

Świadomość utraty życia powoduje próby zatrzymywania jak największej ilości drobin i ocalania ich na wszelkie możliwe sposoby. To idea, która mnie prześladowa. Mówiłam kiedyś moim studentom i wszyscy się dość przerazili mojej koncepcji, więc staram się ją trzymać w tajemnicy, no ale chyba teraz pani powiem. Uważam mianowicie, że potrzebne jest archiwum egzystencji. Chodzi mi o archiwum egzystencji wszystkich. Żeby gdzieś wszystko było zapisane.

Maria Janion w wywiadzie z Barbarą N. Łopieńską, *Wysokie Obcasy* 16 X 2004, nr 41, s. 30.

Widmo bibliotek wirtualnych czy cyfrowych krąży po Europie. Można nawet stwierdzić, że one po prostu istnieją i mają się coraz lepiej. Nie znalazły tylko, jak dotychczas, fortunnej nazwy: oba terminy, nieznaczące przecież tego samego, dalekie są od precyzji. Logiczniej i ściślej historycznie byłoby "wirtualnymi" nazywać tylko takie, w których scalenie materiału następuje jedynie w procesie wyszukiwania i nie ma mowy o żadnym materiale rezydującym w pamięci jednego konkretnego aparatu, zaś "cyfrowym" przydać dopowiedzenie "elektroniczne"^[1]. Na wszelki też wypadek warto przypominać, że chodzi o biblioteki - prawdziwe składnice pełnych tekstów, a nie o katalogi. W każdym razie kierunek jest oczywisty: z książek, czasopism, wtórników typu *facsimile* czy mikrofilm, materiałów audiowizualnych itp. dokumentów, z których dawniej korzystaliśmy za pośrednictwem bibliotek, teraz będziemy korzystać za pomocą Internetu. W świecie anglosaskim dorasta właśnie pierwsze pokolenie inteligencji, które mogło sensownie się uczyć, nie odchodząc od komputera. I na nic argumenty, czego to w tym komputerze nie ma. Kiedy my dorastaliśmy czterdzieści czy dwadzieścia lat temu, w bibliotekach i księgarniach nie było ani ćwierci tego, co dziś można wyciągnąć z Google, jeśli katalog ten skrzyżujemy z lokalnymi zasobami edukacyjnymi przeciętnego college'u. Dzisiejsi polscy "ludzie książki" wchodzili w życie jako informacyjne pólsieroty.

Intuicyjnie wyczuwamy, że chodzi po prostu o przemianę cywilizacyjną, polegającą na zastąpieniu w komunikacji społecznej jednego nośnika innym. Taka zmiana czasem następuje, stając się faktem historycznym, jak przejście od rękopisu do druku, czasem zaś, uporczywie zapowiadana, wykonuje po drodze przedziwną wolte; w końcu dochodzi do niej, ale w stopniu mniejszym niż pierwotnie oczekiwano. Przygoda taka spotkała telewizję. Pomiedzy jednak istnieniem rozwiniętym a istnieniem zaczątkowym rozpościera się ogromna połać, którą wypełniają zupełnie zróżnicowane byty. Biblioteki cyfrowe bywają już duże, w przygotowaniu są wielkie, ale biblioteką cyfrową zwie się też przecież niejeden ekranik z łączami do kilku czy kilkunastu tekstów. Zawsze przecież można podłączyć więcej. Co więcej, niektóre z wielkich instalacji o strukturze bibliotek cyfrowych udostępniają nie wtórniki typowych dokumentów bibliotecznych czy archiwalnych, ale dane innego rodzaju. Czy wspaniały rozkład jazdy PKP to biblioteka cyfrowa? Zbiornice wzorów chemicznych? Katalogi handlowe wraz z zawartością sklepów internetowych? Urządzenia do pobierania muzyki? Biblioteki cyfrowe, które udostępniają książki naukowe czy beletrystykę, płatne czy bezpłatne, to dziś jedno z wielu baz danych. Nawet jeśli włączyć tu serwisy komercyjne i postraszyć trochę kosztami prenumeraty takiego np. *ScienceDirect*, stoi za nimi stosunkowo niewielki kapitał i bardzo mała część masowej wyobraźni; niełatwo je znaleźć w codziennych poczynaniach milionów internautów i w książkach poświęconych Internetowi.

Kluczem do zrozumienia bibliotek wirtualnych jest ich odrębność kulturowa od świata bibliotek "tradycyjnych". Gdyby narzędzia komunikacji rozwijały się pod dyktando tych ostatnich, a nie rynku i otwartej ludzkiej inwencji, epoka pełnych tekstów w przestrzeni cyfrowej pewno by i nastąpiła, ale jakże inne byłyby jej cechy. Wyobraźmy sobie rzeczywistość, w której pełny tekst można podłączyć tylko do opisu katalogowego. Wszystko wygląda tak, jak teraz w dojrzałych, rozwiniętych katalogach bibliotecznych, dla których normą jest prowadzenie użytkownika do takich dokumentów, jakich tylko może potrzebować. Poprzez pole 856 zapis bibliograficzny łączy z pełnym tekstem, opublikowanym w tym czy innym elektronicznym standardzie. Ale dzieje publikacji internetowej nie chcą ograniczyć się do dziejów katalogowania. Biblioteki wirtualne należą do świata wydawania tekstów w większym stopniu niż do świata gromadzenia, udostępniania i informacji. World Wide Web nie oglądał się na biblioteki. Nim pełne teksty dokumentów zaczęły otwierać się z poziomu katalogów, otwierały się z pobieranych z odpowiedniego

serwera dokumentów w którejś z odmian języka SGML, usianych hipertekstowymi leksjami, prowadzącymi do innych hipertekstowych leksji. Te leksje jakoś organizowano wizualnie. Narzędziem informacyjno-wyszukiwawczym stał się sam (hiper)tekst lub jakakolwiek wmontowana weń lista - byle dostarczyć miejsca do uruchomienia łącza. Później zaczęto zastanawiać się, jak zarządzać tymi listami, jak je porządkować, przeszukiwać, scalać. Biblioteki wirtualne rodziły się jako pochodna stron internetowych, pozostawiając katalogi biblioteczne z boku. Gdyby jakimś cudem niemożliwe było grupowanie "linków" na stronie WWW w jakimkolwiek logicznym układzie, katalogi biblioteczne, i tylko one, organizowałyby dostęp do dokumentów cyfrowych. Na pewno sprawdziłyby się w tej roli, i w końcu się sprawdzają wszędzie tam, gdzie są tak wykorzystywane. Ale łatwa w lekturze i układzie lista jest pokusą dla twórcy i dla odbiorcy informacji, zwłaszcza kiedy obsługuje niewielki zbiór. I zamiast powiedzieć, że skomputeryzowana biblioteka ma w swoich zasobach materiały w różnych postaciach, na różnych nośnikach, które znakomicie obsługuje katalog on-line, stworzyliśmy biblioteki cyfrowe jako oddzielne zbiornice dokumentów, często - odrębne instytucje, z odrębnymi, coraz bardziej skomplikowanymi narzędziami informacyjno-wyszukiwawczymi.

Choć procesu tego nie sposób odwrócić, a wobec katalogowania dokumentów elektronicznych biblioteki są z różnych powodów odporne, trudno nie rozczarować się takim obrotem sprawy. Mieliśmy bibliografie i było oczywiste, że muszą one być wyposażone we własne indeksy. Mieliśmy bazy danych bibliograficzne i abstraktowe i było oczywiste, że nijak się one mają do naszego katalogu, te bowiem odsyłał do innego rodzaju materiałów, mianowicie pełnych tekstów. A teraz traktujemy nasze coraz bogatsze zbiory cyfrowe jak kolejną bazę i zamiast powiedzieć, że włączenie dokumentu pełnotekstowego do zbiorów za pomocą skatalogowania go jest chwilą, na którą biblioteki czekały tysiące lat, wolimy uciekać od tego pięknego paradoksu i przeznaczamy wielki wysiłek i fundusze na projektowanie oprogramowania, które by obsługiwało tylko dokumenty cyfrowe, i na dobór metadanych, czytelnych dla tego oprogramowania. Jasne, że programy do obsługi bibliotek cyfrowych umożliwiają także inne rzeczy, których nie robią katalogi on-line, zwłaszcza zaś samą produkcję zasobu pełnotekstowego oraz zarządzanie dostępem do niego odbiorcom o różnych uprawnieniach. Wszystko to jednak jest tylko pogłębianiem podziałów, izolujących użytkownika katalogów bibliotecznych od świata pełnych tekstów w Internecie oraz izolujących internautę od zawartości bibliotek, w których pełne teksty na nośnikach tradycyjnych też przecież potrafią się otwierać!

Bibliotekarz, który liczył, że epoka rozległej informatyzacji przyniesie wielkie scalenia zasobów informacyjnych na bazie dotychczasowych, coraz świetniejszych katalogów, jest rozczarowany, widząc odrębne biblioteki cyfrowe obsługiwane przez wyspecjalizowane programy bazodanowe, nawet jeśli rozumie, że zdecydowana obrona integralności narzędzi informacyjno-wyszukiwawczych oznaczałaby bezsensowne powstrzymywanie naturalnych procesów. Mamy powody, by obawiać się, że katalog on-line będzie eliminowany z użytku w taki sam sposób, w jaki niedawno czytelnicy uczyli się nie dostrzegać zachowanych w holach księżnic nieskonwertowanych katalogów kartkowych. Przyjmujemy wszakże do wiadomości rozmaite powody techniczne i organizacyjne, które spowodowały, że bazy danych wsparte przez WWW przyniosły faktycznie powstanie nowych serwisów, a nie *upgrade* starych. Jesteśmy świadomi, że podstawowa zawartość bibliotek wirtualnych, to utwory samoistne piśmienniczo, lecz niesamoistne wydawniczo, z artykułami z czasopism naukowych na czele, których katalogi biblioteczne z reguły nie umiały i nie chciały uwzględnić. Większość zaś czasopism naukowych należy do świata STM ("science, technology, medicine"). Oznacza to istnienie milionów artykułów (dokładne dane - w materiałach reklamowych każdego dostawcy!), które muszą nadawać się do łatwego wyszukania i szybkiej lektury. Inaczej niż humanista, adept STM nie może napisać pracy, która pomijałaby istotne opracowania. Z tym coś trzeba było zrobić, no to zrobiono.

Wiemy też skądinąd, że biblioteki "tradycyjne" z różnych powodów, także, choć nie tylko, z powodu swojej nieudolności, braku poczucia misji oraz braku adekwatnego finansowania, opornie rozpoczynają katalogowanie czegokolwiek, co nie jest książką i czasopismem drukowanym; dobra kultury, które powinny być szerzej udostępniane, nie wchodzi do obiegu, bo nie mogą wejść do katalogów. Zjawisko to w Polsce ma charakter tak rozległy, że trudno o nim mówić bez zrobienia komuś przykrości. Poprzestańmy na stwierdzeniu, że skoro na razie brak rządowego raportu lub naukowych badań nad merytoryczną zawartością katalogów on-line, to powstrzymujemy się od szerszego komentarza. Jest jednak oczywiste, że za chwilę w Polsce będzie, jeśli już nie jest, więcej rękopisów dostępnych za pośrednictwem Wielkopolskiej Biblioteki Cyfrowej niż wszystkich katalogów bibliotecznych razem wziętych. Można wskazać na inne jeszcze postaci dokumentów, wobec których biblioteki elektroniczne okażą się bardziej chłonne niż tradycyjne instrumentaria wyszukiwawcze bibliotek, stosując uproszczony format opisu i prymitywne klucze wyszukiwawcze.

Jeszcze ważniejsze jest to, że zdominowane przez informatyków biblioteki są wolne od obyczajowych barier, zgodnie z którymi niektóre materiały rękopiśmienne i archiwalne oraz tradycyjne "skarby" bibliotek nie powinny być przedmiotem równego dostępu dla wszystkich. Albowiem otwartość biblioteki to pojęcie niezbyt głęboko przyswojone. Wszyscy jesteśmy otwarci, "służymy użytkownikowi" i prowadzimy "informację naukową", ale niekiedy okazuje się, że niektórzy badacze mają monopol na zajmowanie się pewnymi zespołami źródłowymi albo informacją katalogowa jest niewskazana, bo zachęci złodzieja, albo dostęp do specjalistycznych kartotek w zbiorach specjalnych nie jest publiczny, albo ogłaszanie spuścizn (drukami, elektronicznie, na mikrofilmie itd.) prowadziłoby do szargania świętości. Publikacja zapisu bibliograficznego w dostępnej sieciowo bazie danych nie zajmuje wysokiego miejsca na liście osiągnięć

zawodowych polskiego bibliotekarza; nadal słyszymy głosy - nie wiadomo, jak powszechne - że w międzynarodowych katalogach centralnych polskie materiały znajdują się wówczas, gdy zostaną od nas zakupione.

Dla zbiorów specjalnych digitalizacja i włączanie wtórników cyfrowych do bibliotek wirtualnych może więc być pewnego rodzaju drogą na skróty, pomijającą tradycyjne procedury biblioteczne, które okazały się czy wydały się być zbyt uciążliwe. Być może powinno to stanowić dodatkowy bodziec dla zakładania bibliotek wirtualnych, choć doprawdy lepiej byłoby znaleźć argumenty i środki na rzecz dokończenia komputeryzacji zwykłych (powtórzmy - zwykłych, ale bogatych w pełne teksty oryginalne i elektroniczne, spotykające się w jednolitych katalogach). Niczego zresztą nie można zacząć w sposób zorganizowany bez odpowiedniego przygotowania, a takim dla bibliotek w krajach zamożniejszych była długa tradycja kupowania wtórników źródeł historycznych na mikroformach i na CD-ROM-ach. Wtórники pełnotekstowe nie są wynalazkiem epoki WWW i digitalizacji! Rynek na takie materiały był od dawna, na naszym biurku pojawiły się właśnie nowe prospekty reklamowe, zaiste, niestrudzonego wydawnictwa Adam Matthew Publications: *East India Company factory records, Arthurian legends and the influence of French prose romance i Renaissance man - the books and manuscripts of John De* (wszystkie trzy z British Library), oraz ich katalog na bieżący rok, liczący ponad 30 stron^[2]. Można też, z pewną przesadą, powiedzieć, że polscy historycy będą w nieskończoność ubolewać nad ograniczeniami w korzystaniu z rosyjskich archiwów, nawet kiedy ich zawartość będzie już od dawna dostępna komercyjnie. Oczywiście możemy się tylko domyślać, że głównym odbiorcą takich produktów jest amerykański przemysł edukacyjny, zwłaszcza studia doktoranckie w zakresie humanistyki, prowadzone w tysiącach uczelni. Ale w Polsce też już mamy studentów-doktorantów i zapewne czas przełamać tradycję, że dostęp do źródeł oznacza konieczność znalezienia w porozumieniu z promotorem odpowiedniej "fizycznej" biblioteki lub archiwum z odpowiednio sprzyjającym personelem.

Tworząc biblioteki elektroniczno-cyfrowe, warto też uwzględnić wszystkie osobliwości polskiej nauki. W dużych bibliotekach tego typu, w Ameryce i Europie, jednym z zasadniczych składników są wytworzone na miejscu wypowiedzi naukowe. Szeroko znana California Digital Library, wśród swoich wielu kolekcji, oferuje *eScholarship Editions, eScholarship Publications* oraz *eScholarship Repository* - z wyjaśnienia możemy wywnioskować, że dla uczonych, którzy i tak przecież nie oczekują honorarium, biblioteka cyfrowa stwarza dodatkową okazję, aby zaistnieć w roli autorów^[3]; podobne serwisy (dla publiczności z jednej, dla autorów niepublikowanych jeszcze badań z drugiej strony) oferują i mniejsze uczelnie, jak np. Uniwersytet Erazma w Rotterdamie^[4].

Jest prawdopodobne, że nie prowadząc badań nad rozwojem dokumentów pełnotekstowych w polskim Internecie, pomijamy jakieś istotne fakty. Wydaje się jednak, że tego typu materiałom bardzo jeszcze daleko w naszym kraju do bibliotek cyfrowych. Inna jest też kultura organizacyjna świata akademickiego: publikuje się książkę lub artykuł. Artykuł może być w czasopiśmie, może być w pracy zbiorowej, nie może być w bibliotece cyfrowej, bo to nie jest ani prestiż, ani honorarium. W ten sposób idzie m.in. na marne wysiłek autorów prac magisterskich, którzy produkują teksty nieczytane na ogół dosłownie przez nikogo. Byliśmy świadkami, jak na największej polskiej uczelni cała dyskusja na temat archiwizacji prac dyplomowych ograniczyła się wyłącznie do kwestii zabezpieczenia przed plagiatami. Pomysły, aby prace naukowe "luzem" publikowała Biblioteka Wirtualna Nauki (prowadzona przez jednostkę tej uczelni, jaką jest Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego), nigdy nie dotarły do ciał urzędowo zatroskanych powszechną praktyką kradzieży własności intelektualnej. Nie ma w tym niczyjej winy, zadziałały bezosobowe siły społeczne, które zawsze traktowały pracę magisterską jako coś, czego się nie udostępnia, bo się to pewnie skończy plagiatem (no i co za koszmarnie zawracanie głowy). Idea, aby pracować na rzecz uczciwości, z zasady publikując prace, z których wiele odpowiada objętością i ujęciem tradycyjnie publikowanym niemieckim doktoratom, okazała się bardzo odległa od sposobu, w jaki traktujemy w Polsce tok i sens studiowania. A przecież zmuszając studentów do pisania i egzekwując od nich tę umiejętność, można wyłączyć z rygorozów - tak jak w Stanach - pracę dyplomową, skoro staje się ona jednorazowym popisem i z założenia nie służy nauce, czyli nie może być legalnie i umiejętnie cytowana^[5].

Tak więc biblioteki cyfrowe w Polsce raczej nie będą wspierały nauki w tak jednoznaczny sposób, jak biblioteki w krajach bardziej rozwiniętych. Nic na to na razie nie wskazuje. Zwłaszcza nie staną się one publicznymi archiwami nauki. Nie podejmą też programu publikowania źródeł historycznych, co wymagałoby przemyślanej struktury, założonej kompletności i rozległego finansowania. Bliższa im będzie formuła konserwacji i zabezpieczania zbiorów, co w niektórych inicjatywach sprowadza się do wybiórczego, zaledwie promocyjnego (czyli nawet nie popularyzatorskiego), publikowania materiałów efektownych formalnie lub treściowo: tu znane czasopismo młodopolskie, tam kilka gazet z powstania styczniowego, piękne, choć niekoniecznie czytane stare druki, jeszcze gdzie indziej pocztówki czy fotografie. Takie wystawiennictwo, nawet na poziomie, ma z bibliotekami cyfrowymi niewiele wspólnego. Idea połączenia zawartości tysięcy importowanych tytułów czasopism naukowych, udostępnianych na zasadzie konsorcjalnej, z zasobami powstałymi w wyniku akcji pozyskiwania bieżących tytułów polskich i z zasobami powstałymi w wyniku rozległej akcji digitalizacji, z jaką wystartował ICM UW i jakiej źródłem była właśnie CDL, nie została na razie zrealizowana, a Biblioteka Wirtualna Nauki ma prawo kojarzyć się tylko z Elsevierem, Kluwerem, Springerem, kilkoma wielkimi bazami danych i zaczątkową kolekcją polskiej klasyki matematyczno-fizycznej.

W dodatku nie widać tutaj pytań o skalę i koordynację całego tego ruchu. Biblioteka Wirtualna Nauki, której początki sięgają 1998 r., miała prawo oczekiwać, że osiągnie coś w rodzaju wyłączności^[6]. Taka centralizacja, niestety, kojarzy się w Polsce z SINTO, tak jakby SINTO cokolwiek scentralizowało, nie kojarzy się natomiast z racjonalnie pomyślanymi krajowymi bibliotekami cyfrowymi dla nauki, odpowiednimi dla krajów o średnim potencjale i skromnym finansowaniu, takimi jak np.: duńska^[7], fińska^[8], rosyjska^[9] czy białoruska^[10]. Serwisom tym, prowadzonym na ogół przez centralne ksiąznice państwa lub przez grupy bibliotek (w przypadku białoruskiej - z udziałem uczelni zagranicznej), brakuje na ogół materiałów pochodzących z digitalizacji. Te oczywiście pojawiają się gdzie indziej. W Polsce, jak się zdaje, dominuje wątek zabezpieczenia poprzez cyfryzację publikacji dydaktycznej, ale jest on realizowany przez bardzo wiele instytucji i każda robi go inaczej; brakuje też jasności, czy chodzi o pokazywanie *facsimile* oryginałów, czy po prostu o reedycje tekstów. O ile można sobie wyobrazić sens istnienia regionalnych bibliotek cyfrowych (w których dominować będą materiały historyczne) i bibliotek uczelnianych (w których będą dominować materiały dydaktyczne), i o ile nie sposób zabronić jakiegokolwiek ksiąznicy publikowania w Internecie, a nawet należałoby do tego zachęcać, rozproszenie całego tego ruchu nie może nie budzić niepokoju. Przeglądanie małych katalogów on-line jest łatwiejsze niż przeglądanie małych katalogów kartkowych czy katalogów drukowanych (dla zbiorów specjalnych), ale o ileż mniej wydajne niż przeglądanie dużych katalogów on-line. W dodatku katalogi biblioteczne scalamy za pomocą wyszukiwań rozproszonych, zaś o możliwości scalania katalogów wewnętrznych bibliotek cyfrowych dopiero słyszymy, przy czym chodzi o biblioteki korzystające z tej samej platformy (dLibra). Nowe instalacje cyfrowe zwiększają możliwości dostępu do zasobów kultury, ale wobec braku polskich *gateway*ów i katalogów dziedzinowych skazują nas na korzystanie z zaiste prymitywnej heurystyki.

Być może byłoby dobrze, by wiele bibliotek korzystało przynajmniej z tego samego oprogramowania, np. dojrzałej już dLibry lub JeromeDL (Politechnika Gdańska), nawet jeśli co najmniej równie naturalne jest, by lokalny cyfrowy serwis pełnotekstowy obsługiwało raczej to oprogramowanie, które służy do automatyzacji ksiąznicy. Wspólny program prowadzi do wymiany doświadczeń i ujednolicenia praktyki, zwłaszcza jeśli chodzi o wyszukiwanie i prezentację materiału. Nie brakuje tu zamętu. Jasne ukazanie relacji między dokumentem pierwotnym i wtórnym (a jak np. pokazać wtórnik niekompletny?) i między metadanymi dla oryginału i wtórника sprawia na ogół wiele trudności projektantom bibliotek cyfrowych; stosowna terminologia jest niestandardowa i odchodzi często od przyjętych w bibliotekarstwie ustaleń, co to "przedmiot", "dokument", "hasło", "temat" itp. Gdybyśmy chcieli uwzględnić wszystkie aspekty bibliotek cyfrowych, rozważane np. w ramach programu Delos - Network of Excellence on Digital Libraries^[11], nigdy nie ośmielilibyśmy się rozpocząć pracy nad własną.

Wydaje się jednak, że wiele można osiągnąć poprzez zadbanie o współpracę informatyków z bibliotekarzami oraz potraktowanie publikacji w Internecie jako zadania wymagającego wysiłku redakcyjnego i korektorskiego. Jesteśmy pełni uznania dla serwisów pełnotekstowych Akademii Górniczo-Hutniczej, Politechniki Wrocławskiej czy dla Cieszyńskiej Biblioteki Wirtualnej; zapewne można by podać więcej budujących przykładów. Na ogół jednak nigdzie nie rozwiązuje się problemu nawigacji po tekście dokumentu, skazując użytkownika na mozolne przeglądanie stron. Jeśli dokument jest pozbawiony spisu treści albo jeśli numeracja stron w spisie treści nie odpowiada numeracji stron w dokumencie elektronicznym, wszelkie wyszukiwanie traci sens, obraża sprawność czytelnika użytkownika serwisu i stanowi antypropagandę bibliotek elektronicznych. Z łatwością znajdziemy jednak znane instalacje zagraniczne, takie jak przeroklamowany Project Gutenberg^[12] czy Austrian Literature Online^[13], które takich narzędzi nawigacji są również pozbawione. Jako przykład rozwiązania problemu można jednak podać np. Cyfrową Kolekcję Czasopism Polskich, realizowaną przez Bibliotekę Uniwersytecką w Warszawie we współpracy z Biblioteką Narodową^[14].

Nie mamy tu na razie na myśli OCR, techniki kosztownej, skrajnie trudnej w odniesieniu do dzieł dawnego druku, i wymagającej wysiłku redakcyjnego, choć bez wątpienia otwierającej nowe perspektywy w zakresie analizy tekstu. Mamy też powody, by przypuszczać, że kilka innych bibliotek cyfrowych, zwłaszcza posługujących się do prezentacji programem DjVu, umożliwi wkrótce przeszukiwanie dokumentów według artykułów, rozdziałów, podrozdziałów itp.

"Narodową" polską biblioteką cyfrową mogła stać się, i wciąż jeszcze może, Polska Biblioteka Internetowa^[15]. Jej zadania miały być jednak raczej oświatowe, i to według zachowawczego modelu historyczno-literackiego, toteż było oczywiste, że nie rozwiąże ona problemu dostępu do piśmiennictwa naukowego, zwłaszcza z zakresu nauk matematyczno-przyrodniczych i praktycznych. W jej zawartości przeważać miała klasika literatury polskiej, podręczniki akademickie, dokumenty archiwalne, publikacje dla niewidomych, rękopisy muzyczne oraz pozycje kartograficzne, dzieła malarstwa, grafiki i fotografii, a także zawartość czasopism naukowych. Niektóre teksty PBI przejął z Wirtualnej Biblioteki Literatury Polskiej - pożytecznego serwisu Instytutu Filologii Polskiej Uniwersytetu Gdańskiego, zawierającego ponad 120 dzieł lub zbiorów, opublikowanych w postaci transkrypcji html-owych, z pewną liczbą niekonsekwencji redakcyjnych i o zupełnie amatorskiej szacie graficznej^[16]. Ze względu na postępujące napełnienie (ponad 10,5 tys. pozycji), PBI staje się istotną i szanowaną instytucją polskiego bibliotekarstwa. Poprawiła się też nieco szwankująca technika organizacji i prezentacji materiału, i przynajmniej drukarz Gutenberg przestał występować jako autor... Biblii Gutenberga, co stanowiło najzabawniejsze potknięcie PBI jeszcze rok temu. Ale PBI pozostaje tworem niedopracowanym, który psuje swoje własne wartości edukacyjne. Obok Biblii

wydrukowanej przez Johanna Gutenberga, bez autora, jest jeszcze Biblia Tysiąclecia, która ma już autora, mianowicie "Zbiorowego"; "Zbiorowy" występuje także jako autor czasopism i można go sobie z dobrym skutkiem wprowadzić do pola dialogowego przy wyszukiwaniu tychże w katalogu. Który jest katalogiem tylko z nazwy: Gutenberga - drukarza trzeba się domyślić, bo "strona informacyjna" powiada tylko "Moguncja 1452"; nic też nie mówi o bibliotece, szczycącej się posiadaniem takiego dzieła. "Strona informacyjna" nie dosyć, że wyjątkowo uboga, jest jeszcze trudna do znalezienia, szczególnie dla użytkownika, który nic nie wie o katalogach, nie mówiąc już o metadanych. W ten sposób PBI wspomaga polski system oświatowy w masowej produkcji ludzi, którzy jak już coś znajdą, to nie potrafią zacytować, a konieczność dokumentowania własnych poczynąń autorskich wydaje im się niepotrzebnym zawracaniem głowy. Za to Moguncję można wywołać jako "wydanie" w "Szukaniu zaawansowanym", tak jakby nie było tam miejsca na wpisanie "miejsca wydania". Na próżno też szukać w tym katalogu rozwiązania pseudonimów (wprowadźcie np. "Hajota" albo "Marion"). Całe przedsięwzięcie jest zaopatrzone w wyjątkowo ubogą informację na temat zasad doboru tekstów oraz dalszych zamiarów wydawcy w zakresie harmonijnego budowania zawartości. Jest natomiast trochę ułatwień informatycznych, które stanowią o przewadze bibliotek cyfrowych nad fizycznymi. Cóż z tego, skoro towarzyszy im taka, na przykład, polszczyzna: "wybór ilości listowanych dzieł na jednej stronie". Z uznaniem należy powitać zniknięcie wyjątkowo niefortunnego "katalogu przedmiotowego". Widać jednak, że ze skądinąd sympatycznymi młodymi facetami w czerni, odpowiedzialnymi za PBI, nie usiadł na dwa dni żaden bibliotekarz, aby nauczyć ich swojej terminologii i rzemiosła.

W tej sytuacji za najsilniejszą polską bibliotekę wirtualną uważa się zwykle Wielkopolską Bibliotekę Cyfrową, tworzoną przez Poznańskie Centrum Superkomputerowo-Sieciowe wraz z Poznańską Fundacją Bibliotek Naukowych. Liczy ona sobie prawie 3500 dokumentów, reprezentujących: dziedzictwo kulturowe, materiały dydaktyczne, materiały regionalne i muzykalia, przy czym jeden dokument może należeć (słusznie) do więcej niż jednej kolekcji. Liczbowo przeważają, co zrozumiałe, materiały z zakresu dziedzictwa kultury i materiały regionalne. Nawigacja po zasobie jest wieloaspektowa, prezentacja - konsekwentnie *facsimilowa* (na ogół html lub DjVu), wyjaśnienia dotyczące całego zbioru i poszczególnych kolekcji - dobrze zredagowane. I tu jednak domagalibyśmy się kilku ingerencji świadomego bibliotekarza. Trudno np. zgodzić się z traktowaniem "biblii brzeskiej" jako "tematu" (patrz indeks), gdy po wprowadzeniu tego hasła faktycznie otwiera się *Biblia Vetus et Novum Testamentum: Biblia święta to jest Księgi Starego y Nowego Zakonu własnie z Żydowskiego, Greckiego y Łacińskiego nowo na Polski język... wyłożone*. Indeks tematów i słów kluczowych WBC to, niestety, prawdziwy quasiinternetowy śmietnik. Dalej, trudno zaakceptować fakt, że redakcja "spisem treści" nazywa zwykle spisy stron. Właściwie żadne jest też powiązanie wtórnika, opisanego zgodnie z formatem Dublin Core (co w przypadku starych druków jest frustrujące, gdyż nie dowiadujemy się imienia drukarza) z oryginałem i jego opisem bibliograficznym w katalogu prowadzonym według MARC-a^[17].

Jak wszystkie chyba biblioteki cyfrowe, WBC nie obsługuje żaden system klasyfikacyjny - poza podziałem na cztery kolekcje. Trudno więc uzyskać szybko wgląd w dziedzinowy charakter i strukturę całości. Zbiór o uniwersalnym polu tematycznym, to piękny pomysł, ale czy czytelnik, widząc historyczny i źródłowy charakter Biblioteki, faktycznie wie, że wśród 52 "publikacji" (dokumentów) w kolekcji "Materiały dydaktyczne" są i takie, jak *Oleje, paliwa i smary dla motoryzacji i przemysłu* (2001)? Jak na to trafi - wprowadzając ufnie słowa kluczowe "oleje smarowe" i "paliwa" właśnie do katalogu WBC? Czyżbyśmy znowu tworzyli serwis dla takich użytkowników, którzy wiedzą, że dane dzieło w nim jest, bo im powiedział "pan" od smarów, "pani" od romantyzmu i tak dalej?

Wydaje się też, że w obrębie uniwersalnych bibliotek cyfrowych o silnej zawartości humanistycznej sensowne byłoby wydzielenie kolekcji typu "księgozbiór podręczny", rekonstruującej elektronicznie dzieła pomocnicze, które miał pod ręką uczony czy student w dużej czytelni naukowej. Zbiory czytelni ulegają rozproszeniu w bibliotekach z wolnym dostępem - nie wiemy, jak to będzie zorganizowane w nowych gmachach bibliotek uniwersyteckich we Wrocławiu i Łodzi, ale w Warszawie wielu użytkownikom, przywiązanych do dawnego układu książek, bardzo zabrakło bliskiego sąsiedztwa kanonu wydawnictw informacyjnych i podręczników, zwłaszcza historycznych i filologicznych. Taki zbiór zaczęliśmy tworzyć w ramach Biblioteki Wirtualnej Nauki. Możliwe też są np. kolekcje narzędzi inżynierskich, zwłaszcza obliczeniowych i prezentacyjnych, jak choćby to, które jako "Productivity tools" prowadzi atrakcyjna komercyjna biblioteka elektroniczna Knovel, testowana obecnie m.in. w Bibliotece Głównej Politechniki Warszawskiej. Biblioteki cyfrowe nie muszą prowadzić tylko dokumentów - mogą także prowadzić oprogramowanie użytkowe i na pewno będą to robić w przyszłości.

Pierwsze dziesięć lat "zawieszania w Internecie" i digitalizacji nadało już dość wyrazisty kształt bibliotekom cyfrowym w Polsce - zaczynają one wchodzić w fazę napełnienia i mogą dość szybko pozbyć się swoich wad. Dalszy ich rozwój nie powinien odbywać się tylko poprzez realizowanie inicjatyw oddolnych. Mamy dość narzędzi i badaczy, aby spojrzeć całościowo i statystycznie na polskie zbiory, możliwości naukowe, potrzeby dydaktyczne. Nasuwa się tu wiele możliwych dróg postępowania. Na pewno potrzebna jest krajowa polityka digitalizacyjna - ustalenie priorytetów, a także listy obiektów, którymi "podziela" się biblioteki, archiwa, muzea i inne instytucje prowadzące dokumentację. Tu trzeba stanowiska wielu instytucji, z tymi, które odtrąbiły początek ery "społeczeństwa informacyjnego" na czele. Nie może upiec się ani Biblioteka Narodowa, ani Krajowej Radzie Bibliotecznej, ani ministerstwu i placówkom naukowym. Bez poważnego rozpoznania zasobów i dokonań wszelkie dywagacje, tak jak cały ten wywód, będą co najwyżej publicystyką wynikającą z osobistych doświadczeń i potocznych obserwacji.

Wśród najatrakcyjniejszych materiałów dla bibliotek cyfrowych widzimy niekoniecznie publikacje książkowe, wielokrotnie wznowiane i dostępne w wielu miejscach, może raczej i przede wszystkim zawartość dziewiętnastowiecznych czasopism, zwłaszcza humanistycznych oraz bogate zbiory drukowane, rękopiśmienne i archiwalne, także w mało znanym języku rosyjskim. Dalej - zbiory specjalne w języku polskim z Wilna, Mińska, Lwowa i Kijowa, głównie rękopiśmienne. Dalej - źródła do historii nauki i techniki, jako dyscypliny pozbawionej w Polsce reprezentacji akademickiej w sposób nieliczący z wyobrażeniem, iż mamy rozwinięte życie i instytucje naukowe. Digitalizacja może i powinna objąć też obiekty muzealne, materiały ikonograficzne i materiały należące do tzw. dokumentacji mechanicznej. A może wizja prawdziwie publicznej ekspozycji byłaby atrakcyjna dla niektórych spadkobierców, nieumiejących zdecydować się na ofiarowanie cennej spuścizny tej czy innej bibliotece? Może nawet w ten sposób przywrócony zostałby polskiej kulturze, schowany dotychczas głęboko, zbiór Jerzego Grotowskiego albo rękopisy Jerzego Stempowskiego przechowywane w Szwajcarii?

Z drugiej strony mamy jeszcze do rozwiązania problem bieżącej polskiej produkcji naukowej, słabo włączonej do obiegu naukowego ze względu na niskie nakłady i brak odpowiednich baz danych. Przejęcie przez biblioteki cyfrowe (zwłaszcza przez przedsięwzięcia realizowane dla dobra ogólnego) dużej części powstających w Polsce prac naukowych, to jeden z nie tak znowu wielu prezentów, jakie możemy zrobić Europie i jej sąsiadom. Jeśli zawartość bibliotek cyfrowych nie będzie bogata, humaniści nigdy nie zdołają zrobić z niej "archiwum egzystencji".

Przypisy

[1] Por. dyskusję J. Czermińskiego nad "cyfrowym dokumentem elektronicznym" w: *Cyfrowe środowisko współczesnej biblioteki*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego, 2002, s. 174 i wcześniejsze.

[2] Por. stronę internetową: *Adam Matthew Publications* [on-line]. Adam Matthew Publications, 2005 [dostęp 10 marca 2005]. Dostępny w World Wide Web: <http://www.ampltd.co.uk/>.

[3] *California Digital Library, Collections & Services* [on-line]. [dostęp 10 marca 2005]. Dostępny w World Wide Web: <http://www.cdlib.org/collect/index.html>.

[4] *dEAR: The digital Erasmus Academic Repository* [on-line]. [dostęp 11 marca 2005]. Dostępny w World Wide Web: <http://eps.eur.nl/>. We wstępie czytamy: dEAR aims at making all research results of the Erasmus University of Rotterdam available through the Internet, for everyone, at all times, and without charge.

[5] Warto jednak odnotować pojawiające się w Polsce w Internecie materiały dydaktyczne, np. w obrębie Akademickiej Biblioteki Internetowej Uniwersytetu im. Adama Mickiewicza. Choć ABI publikuje także teksty do czytania, przewagę mają w niej materiały dźwiękowe. *Akademicka Biblioteka Internetowa* [on-line]. Poznań: Uniwersytet Adama Mickiewicza w Poznaniu, 2003 [dostęp 11 marca 2005]. Dostępny w World Wide Web: <http://abi.amu.edu.pl/>.

[6] Por. np. H. Hollender: Biblioteki cyfrowe w Polsce i BWN: w poszukiwaniu drogi racjonalnego rozwoju. In *Konferencja na temat: Polskie biblioteki akademickie w Unii Europejskiej: Łódź, 23-25 czerwca 2004 r.: Materiały Konferencyjne: Suplement*. Łódź: Biblioteka Główna Politechniki Łódzkiej, 2004, s. 89-99.

[7] Por. stronę internetową: *deff.dk Danmarks Elektroniske Forskningsbibliotek* [on-line]. Danmarks Elektroniske Forskningsbibliotek, 2002-2005 [dostęp 10 marca 2005]. Dostępny w World Wide Web: <http://deff.dk/>.

[8] Por. stronę internetową: National Electronic Library z rozproszonymi informacjami nt. Finnish National Portal Nelli: Helsingin yliopiston kirjasto KANSALLISKIRJASTO [on-line]. [dostęp 10 marca 2005]. Tämä on FinELib. Dostępny w World Wide Web: <http://www.lib.helsinki.fi/finelib/>.

[9] Por. stronę internetową biblioteki elektronicznej: *Orel* [on-line]. [dostęp 10 marca 2005]. Dostępny w World Wide Web: <http://orel.rsl.ru/>; warto dodać, że Rosyjska Biblioteka Państwowa prowadzi odrębnie elektroniczną bibliotekę dysertacji: *Digital Library of dissertations* [on-line]. [dostęp 10 marca 2005]. Dostępny w World Wide Web: <http://diss.rsl.ru/>.

[10] Por. stronę internetową wirtualnej biblioteki *VirLib* [on-line]. [dostęp 10 marca 2005]. Dostępny w World Wide Web: <http://elib.bsuir.unibel.by/>.

[11] Por. np. zawartość Resersach Clusters w informatorze: *DELOS Network of Excellence on Digital Libraries* [on-line]. Last update: 2005-04-05 [dostęp 11 marca 2005]. Dostępny w World Wide Web: <http://www.delos.info/>.

[12] Por. stronę internetową: *Project Gutenberg* [on-line]. Project Gutenberg, 1971-2005 [dostęp 11 marca 2005]. Dostępny w World Wide Web: <http://www.gutenberg.org/>.

[13] Por. stronę internetową: *alo Austrian Literature Online* [on-line]. [dostęp 11 marca 2005]. Dostępny w World Wide Web: <http://www.literature.at/>.

[14] Por. stronę internetową: Cyfrowa kolekcja czasopism polskich. In *Biblioteka Uniwersytecka w Warszawie* [on-line]. BUW [dostęp 11 marca 2005]. Dostępny w World Wide Web: <http://www.buw.uw.edu.pl/zasoby/ckcp.htm>.

[15] Por. stronę internetową: *Polska Biblioteka Internetowa* [on-line]. PBI, 2003 [dostęp 11 marca 2005]. Dostępny w World Wide Web: <http://www.pbi.edu.pl/>; uwagi na podstawie oględzin w dniu 10 stycznia 2005.

[16] Por. stronę internetową: *Biblioteka Literatury Polskiej w Internecie Instytut Filologii Polskiej Uniwersytet Gdański* [on-line]. [dostęp 11 marca 2005]. Dostępny w World Wide Web: <http://monika.univ.gda.pl/~literat/>.

[17] Por. stronę internetową: *Wielkopolska Biblioteka Cyfrowa* [on-line]. Poznań: Poznan Supercomputing and Networking Center; Poznańska Fundacja Bibliotek Naukowych [dostęp 11 marca 2005]. Dostępny w World Wide Web: <http://www.wbc.poznan.pl/>; uwagi na podstawie oględzin w dniu 11 marca 2005.



Biblioteki cyfrowe w Polsce: warunki i zagrożenia / Henryk Hollender// W: Biuletyn EBIB [Dokument elektroniczny] / red. naczelny Bożena Bednarek-Michalska. - Nr 4/2005 (65) kwiecień. - Czasopismo elektroniczne. - [Warszawa] : Stowarzyszenie Bibliotekarzy Polskich KWE, 2005. - Tryb dostępu: <http://www.ebib.pl/2005/65/hollender.php>. - Tyt. z pierwszego ekranu. - ISSN 1507-7187