

ANETA JANUSZKO-SZAKIEL

Krakowska Akademia im. Andrzeja Frycza Modrzewskiego

e-mail: ajanuszko-szakiel@afm.edu.pl

ANALIZA STANU ZBIORÓW ELEKTRONICZNYCH I WARUNKÓW ICH ARCHIWIZOWANIA W POLSKICH INSTYTUCJACH BIBLIOTECZNYCH I WYDAWNICZYCH



Aneta Januszko-Szakiel jest adiunktem w Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego. Jej zainteresowania naukowe skupiają się wokół organizacji i funkcjonowania repozytoriów cyfrowych, w szczególności dotyczą długoterminowej archiwizacji zasobów cyfrowych w instytucjach pamięci. Ważniejsze publikacje: *Archiwizacja publikacji elektronicznych jako wyzwanie dla bibliotek – zarys problematyki* (2003), *Open Archival Information System – standard w zakresie archiwizacji publikacji elektronicznych* (2005), *Wiarygodność archiwów cyfrowych* (2009).

SŁOWA KLUCZOWE: Publikacje elektroniczne. Długoterminowa archiwizacja publikacji elektronicznych. Polskie biblioteki. Biblioteka Narodowa. Biblioteka Jagiellońska. Poznańskie Centrum Superkomputerowo-Sieciowe. Ministerstwo Kultury i Dziedzictwa Narodowego. Polskie wydawnictwa publikacji elektronicznych. Sekcja Wydawnictw Elektronicznych.

ABSTRAKT: W artykule przedstawiono wyniki badań przeprowadzonych w 2009 r., dotyczące zasobów elektronicznych polskich instytucji bibliotecznych. Celem tych badań było ustalenie, w jakiej kondycji są materiały elektroniczne zgromadzone w bibliotekach tradycyjnych i cyfrowych oraz, jaki jest poziom świadomości i nastawienie przedstawicieli polskich instytucji tworzących, gromadzących i zarządzających polskimi bibliotecznymi zasobami elektronicznymi, odnośnie do kwestii zabezpieczenia ich długoterminowej użyteczności. Omówiono również wnioski z rozmów na temat zapewnienia długoterminowej archiwizacji elektronicznych zasobów bibliotecznych z pracownikami Poznańskiego Centrum Superkomputerowo-Sieciowego.

WPROWADZENIE

Jedne z najwcześniejszych, udokumentowanych badań dotyczących archiwizacji bibliotecznych materiałów cyfrowych zostały przeprowadzone

w 1996 r. przez amerykańską Research Library Group (RLG)¹. Polegały one na przeglądzie dokumentów elektronicznych, zgromadzonych głównie w bibliotekach, ale także w archiwach, muzeach i innych instytucjach sektora nauki i kultury. Ich celem było ustalenie liczby oraz jakości zgromadzonych publikacji (tj. zastosowanych nośników, formatów, otoczenia sprzętowo-programowego odczytu i prezentacji ich treści w formie zrozumiałej dla użytkownika), a także identyfikacja zasobów, do treści których dostęp został utracony. Z raportu RLG wynika, że po udokumentowaniu ilościowego i jakościowego stanu zasobów cyfrowych, instytucje biblioteczne powinny przystąpić do definiowania metod, narzędzi i technologii, umożliwiających długoterminowe utrzymanie użyteczności „czytelnych” dokumentów elektronicznych oraz ewentualne odzyskanie treści dokumentów „nieczytelnych” (Waters & Garrett, 1996).

Podobne badanie przeprowadzono w 2009 r., aby ustalić, w jakiej kondycji są zasoby elektroniczne zgromadzone w polskich bibliotekach, jaka jest świadomość i gotowość do podjęcia działań archiwizacyjnych wśród przedstawicieli polskich instytucji pamięci i od jakich zadań należałoby prace te rozpocząć.

METODA I ORGANIZACJA BADANIA

Badaniami postanowiono objąć przede wszystkim instytucje biblioteczne, których zbiory służą pracom naukowo-badawczym oraz edukacji i stanowią w całości lub części narodowy zasób biblioteczny, podlegający szczególnej, niekiedy wieczystej archiwizacji.

Ponadto, wobec popularyzacji projektów organizujących wolny dostęp do informacji naukowej i materiałów edukacyjnych o ponadczasowej wartości, a także wobec upowszechniania się projektów zdalnego nauczania, w ramach których tworzy się i przechowuje wartościowe zasoby cyfrowe, do udziału w badaniach zaproszono polskie inicjatywy Open Access, głównie biblioteki cyfrowe oraz polskie ośrodki akademickie oferujące kształcenie w trybie e-learningu.

Przy ustalaniu listy bibliotek, które należy objąć badaniem za punkt wyjścia przyjęto bazę BN Biblios – z uwagi na jej status centralnego informatora o wszelkich bibliotekach i ośrodkach informacji w Polsce. Założono wstępnie, że badania powinny obejmować wszystkie biblioteki polskie spełniające warunek sformułowany powyżej, stąd też wykaz utworzony na podstawie bazy BN Biblios uzupełniono na podstawie wykazów bibliotek naukowych, o którym mowa w rozdziale 6 Ustawy z 27 czerwca 1997 r. o bibliotekach (Ustawa..., 1997), w Rozporządzeniu Ministra Kultury i Sztuki z 19 marca 1998 r. w sprawie zasad i trybu

¹ W 1994 r. Commission on Preservation & Access (CPA)/RLG Research Library Group Task Force on Archiving of Digital Information rozpoczęła prace badawcze nad dokumentami cyfrowymi, a w 1996 r. opracowany został raport na temat obowiązku podjęcia działań na rzecz zagwarantowania długoterminowej dostępności cyfrowego dziedzictwa i budowy cyfrowego repozytorium. Więcej informacji na ten temat w: *Preserving Digital Information*, 1996.

zaliczania bibliotek do bibliotek naukowych oraz ustalania ich wykazu (MKiS, 1998a) oraz w Rozporządzeniu Ministra Kultury i Sztuki z 24 listopada 1998 r. w sprawie ustalenia wykazu bibliotek, których zbiory tworzą narodowy zasób biblioteczny, określenia organizacji tego zasobu oraz zasad i zakresu jego szczególnej ochrony (MKiS, 1998b). Ostatecznie zbiór bibliotek, które włączono do badania zawierał 138 instytucji bibliotecznych.

Kwalifikując do próby badawczej polskie biblioteki cyfrowe, wykorzystano wykaz bibliotek cyfrowych dostępny w serwisie WWW Federacji Bibliotek Cyfrowych (zob. FBC, 2008) oraz wykaz bibliotek cyfrowych dostępny w serwisie WWW Cyfrowej Platformy Stowarzyszenia Bibliotekarzy Polskich EBIB (zob. Wybrane kolekcje cyfrowe, 2008). W celu zapewnienia kompletności tej listy porównano ją z wykazami dostępnymi na stronach internetowych bibliotek akademickich: Biblioteki Głównej UMCS, Biblioteki Uniwersyteckiej w Warszawie, Biblioteki Uniwersytetu Śląskiego. W serwisach bibliotek UMCS oraz UW zamieszczone są odesłania do wykazów bibliotek cyfrowych udostępnianych w serwisie EBIB oraz FBC. Biblioteka UŚ natomiast zamieszcza obszerny wykaz bibliotek cyfrowych wraz z ich krótką charakterystyką oraz odesłaniami do określonych kolekcji. Nie ma w wykazie tym jednak bibliotek, które nie byłyby wymienione w serwisach EBIB i FBC. Kompletując zbiór badawczy polskich bibliotek cyfrowych, odniesiono się także do wykazu polskich bibliotek cyfrowych zamieszczonego w *The Directory of Open Access Repositories – OpenDOAR (The Directory..., 2010)*² oraz wykazu zamieszczonego w Wolnej Encyklopedii Wikipedia pod hasłem *Biblioteka cyfrowa* (Biblioteka cyfrowa, 2008)³. Na podstawie wymienionych źródeł, do zbioru badawczego włączono 53 biblioteki cyfrowe.

W przypadku inicjatyw zdalnego nauczania źródła doboru do próby badawczej stanowiły: aneks zamieszczony w opracowaniu Marka Nahotki (Nahotko, 2006), zawierający wykaz ośrodków akademickich e-learningu w Polsce działających w latach 1996-2005, katalog „Studia online” dostępny na stronach portalu internetowego Onet.pl (Studia online, b.d.), wykaz zamieszczony w Google Directory w kategorii *Nauka i edukacja*, w podkategorii *Nauczanie zdalne* (zob. Katalog WWW Google, 2008) oraz wykazy otrzymywane w wyniku wyszukiwania w zasobach Internetu⁴. Łącznie do zbioru badanych instytucji włączono 27 polskich inicjatyw, w ramach których organizuje się kształcenie w trybie zdalnym na pozio-

² W wykazie tym wymienia się 19 polskich inicjatyw (stan na 2 marca 2010 r.). Tendencja jest rosnąca, albowiem w 2008 r., kiedy rozpoczęto prace nad tworzeniem zbioru badawczego, w wykazie tym odnotowana była tylko jedna polska inicjatywa – Biblioteka Cyfrowa Politechniki Łódzkiej.

³ W porównaniu z wymienionymi już wykazami bibliotek cyfrowych, na stronie Wolnej Encyklopedii Wikipedii znajdują się odesłania tylko do 13 polskich bibliotek cyfrowych, głównie instytucjonalnych i regionalnych. Wymieniona została również Cyfrowa Biblioteka Narodowa Polona. <http://pl.wikipedia.org/wiki/Kategoria:Polskie_biblioteki_cyfrowe> [dostęp: 2.03.2008].

⁴ Podstawowe narzędzie wyszukiwania stanowiła wyszukiwarka Google z prostym i zaawansowanym trybem wyszukiwania, a zastosowane hasła wyszukiwawcze to różnorodne kombinacje następujących słów: centrum, edukacja, e-kształcenie, e-learning, internetowy, kształcenie, nauczanie, odległość, online, ośrodek, studia, szkolenia, wirtualny, zdalny etc., z każdorazowym uwzględnieniem, iż mają to być inicjatywy polskie.

mie szkolnictwa wyższego. Podczas ustalania inicjatyw e-learningowych w Polsce, których cyfrowe kolekcje powinny zostać objęte badaniem, poważną trudność stanowiło zidentyfikowanie kompletnego, na bieżąco aktualizowanego rejestru instytucji, które takie kształcenie oferują⁵. Zbiór badawczy starano się zatem utworzyć na podstawie niewielu istniejących i przypadkowo dobranych źródeł, dostępnych głównie w Internecie. Źródła te zidentyfikowano na podstawie przeszukiwania zasobów internetowych z zastosowaniem wyszukiwarki Google.

Obok popularnych w Polsce projektów związanych z tworzeniem bibliotek cyfrowych, do zbioru badawczego włączono mniej popularne inicjatywy, których celem jest organizowanie wolnego dostępu do informacji naukowej i materiałów edukacyjnych o ponadczasowej wartości, np. repozytoriów preprintów. Inicjatywy tego typu zyskują dopiero na popularności w naszym kraju, mają raczej charakter pionierski, niekiedy efemeryczny. Brakuje na bieżąco aktualizowanych źródeł je identyfikujących i ewidencjonujących. Włączone do zbioru badanych instytucji rodzime przedsięwzięcia tej kategorii zidentyfikowano głównie na podstawie wyników wyszukiwania w zasobach Internetu⁶. Uwzględniono także wykaz zasobów sieciowych przygotowany i udostępniony na stronie WWW Instytutu Informacji Naukowej i Bibliotekoznawstwa Akademii Pedagogicznej w Krakowie *Zasoby sieciowe (Zasoby sieciowe, b.d.)*. Do próby badawczej włączono ostatecznie 11 inicjatyw Open Access.

W wyniku kompletowania próby badawczej na podstawie przedstawionych źródeł powstał wykaz 228 instytucji. Zdecydowaną większość w tym zbiorze stanowią instytucje biblioteczne, głównie biblioteki akademickie, naukowe oraz biblioteki, których zbiory stanowią narodowy zasób biblioteczny. Rozbudowaną grupę stanowią również biblioteki cyfrowe oraz inne inicjatywy, w wyniku których powstają polskie kolekcje dokumentów elektronicznych. Znacznie mniejsze liczebnie są grupy inicjatyw e-learningowych oraz repozytoriów Open Access.

Do instytucji, które utworzyły zbiór badawczy w marcu 2009 r. skierowano kwestionariusz ankiety wraz z pismem przewodnim, wyjaśniającym cel badań oraz prośbę o udzielenie odpowiedzi.

Spośród 228 instytucji, do których wysłano ankietę, odpowiedziało 10 instytucji. Wypełnioną ankietę zwrotną nadesłało tylko pięć instytucji. Spośród instytucji, które wypełniły ankietę, 2 przesłały dodatkową informację. Odpowiedzi 5 instytucji nie zawierały wypełnionej ankiety, a jedynie komentarze dotyczące samych badań.

Na ankietę nie odpowiedziało 218 instytucji. W stosunku do łącznej liczby rozesłanych kwestionariuszy wypełniono 2,2% ankiet. Z racji niskiej zwrotności – prawdopodobnie będącej odzwierciedleniem poziomu

⁵ We wrześniu 2006 r. utworzone zostało Stowarzyszenie E-learningu Akademickiego (SEA), realizujące m.in. projekt, w ramach którego budowana jest baza wiedzy na temat form aktywności polskich uczelni w obszarze e-learningu. Więcej informacji na ten temat na stronie Stowarzyszenia: <<http://www.sea.edu.pl/>>.

⁶ W wyszukiwaniu posłużono się wyszukiwarką Google oraz hasłami wyszukiwawczymi konstruowanymi z następujących słów: access, edukacja, lektury, materiały, open, otwarty, podręczniki, wiedza, wolny, zasób etc.

świadomości w polskich bibliotekach odnośnie do istoty i potrzeby działań archiwizacyjnych – badania zostały rozszerzone o wywiady zogniskowane na problematyce długoterminowej archiwizacji materiałów cyfrowych. Z prośbą o udzielenie wywiadu zwrócono się do przedstawicieli dwóch bibliotek o statusie księżnic narodowych – Biblioteki Narodowej oraz Biblioteki Jagiellońskiej. W wywiadzie posłużono się kwestionariuszem ankiety, rozsyłanym do wszystkich instytucji wytypowanych do badań.

Kwestionariusz ankiety adresowany do instytucji pamięci został podzielony na dwie części. Pierwsza służyła wysondowaniu, jaki jest orientacyjny stan zasobów cyfrowych przechowywanych w polskich instytucjach, tj. z którego roku pochodzą najstarsze publikacje, w jakich formatach i na jakich nośnikach są zapisane, jaka jest ich łączna objętość oraz czy obecnie możliwy jest ich odczyt. Jedno z pytań dotyczyło istnienia w badanych instytucjach strategii postępowania z zasobami cyfrowymi. Druga część kwestionariusza ankiety służyła poznaniu opinii przedstawicieli polskich instytucji pamięci na temat długoterminowej archiwizacji publikacji cyfrowych. Odpowiedzi miały umożliwić rozpoznanie, jaki jest stopień świadomości pracowników bibliotek, poziom ich kompetencji i gotowości do podjęcia działań w tym zakresie. Założono również, że badanie pozwoli ustalić preferencje dotyczące modelu polskiego archiwum cyfrowego, w tym wskazania instytucji, która powinna pełnić rolę krajowego koordynatora działań w Polsce. Celem jednego z pytań było też wysondowanie, czy w bibliotekach polskich została podjęta dyskusja w tej sprawie oraz jakie jest nastawienie do tej kwestii bibliotekarzy w badanych bibliotekach. Na podstawie badania starano się także poznać stanowisko pracowników instytucji pamięci wobec kwestii udziału i roli wydawców w procesach długoterminowej archiwizacji publikacji cyfrowych.

Omówienie wyników badań oparte jest na informacjach uzyskanych łącznie z siedmiu wypowiedzi przedstawicieli polskich instytucji gromadzących publikacje elektroniczne (z pięciu wypowiedzi ankietowych oraz dwóch wywiadów).

STAN ELEKTRONICZNYCH ZASOBÓW BIBLIOTECZNYCH

Tylko w jednej wypowiedzi otrzymanej w ankiecie, deklaruje się istnienie strategii długoterminowego zarządzania zgromadzonymi dokumentami cyfrowymi. Wypowiedź dotyczy zbioru publikacji cyfrowych udostępnianych w Sieci w ramach projektu Open Access, a strategia ta – zgodnie z odpowiedzią – polega na „systematycznym opracowywaniu dokumentów w ramach logicznie ułożonego systemu biblioteki, ze zdefiniowanymi działaniami i kategoriami, a także opisie każdego dokumentu w bazie danych”. W cytowanej odpowiedzi nie zawarto elementów typowych dla strategii zarządzania zbiorem elektronicznych dokumentów, w celu utrzymania ich długoterminowej użyteczności, tj. takich jak: systematyczny okreso-

wy odczyt dokumentów, sporządzanie ich kopii zapasowej, odświeżanie nośnika, zmiana generacji nośnika, migracja, emulacja etc. Jedną odpowiedź nie jest oczywiście podstawą do uogólnień, ale można traktować ją jako odzwierciedlenie wyobrażenia na ten temat osób zajmujących się gromadzeniem i zarządzaniem publikacjami elektronicznymi.

W pozostałych badanych instytucjach nie wskazano istnienia jakiegokolwiek formy planowego postępowania z publikacjami elektronicznymi. Instytucje deklarują natomiast, że na podstawie danych z katalogów oraz sporządzanych inwentarzy, są w stanie oszacować liczbę zgromadzonych publikacji elektronicznych, wskazać formaty ich zapisu, objętość oraz rok opublikowania poszczególnych dokumentów. Nie zgłoszono również problemów związanych z odczytem tychże publikacji, przy czym wypowiedzi te odnoszą się do zbiorów bibliotek, w których najstarsze publikacje datowane są na lata 1999-2000.

Zdecydowanie mniej korzystnie wygląda sytuacja w przypadku zbiorów bibliotek przechowujących starsze publikacje elektroniczne. Z informacji uzyskanych na podstawie wywiadów wynika, że w instytucjach o statusie bibliotek narodowych najstarsze publikacje elektroniczne pochodzą z lat 1994-1996 i ich odczyt jest poważnie utrudniony, w większości przypadków niemożliwy, bowiem instytucje te nie są w posiadaniu otoczenia sprzętowo-programowego, umożliwiającego odczyt i prezentację ich treści. Jest to ewidentny dowód zaniedbania zbiorów elektronicznych, wynikającego prawdopodobnie z braku stosownej świadomości, a także wiedzy, w jaki sposób postępować z tego typu materiałami bibliotecznymi. Polskie instytucje nie są jednak wyjątkiem; podobne przypadki utraty dostępu do treści najstarszych dokumentów elektronicznych odnotowano w wielu instytucjach bibliotecznych i archiwalnych na świecie i zwykle miały one pozytywne następstwa: dawały impuls do badań i działań w zakresie długoterminowej archiwizacji. W niektórych krajach miało to jednak miejsce przed piętnastu laty.

Z wywiadów wynika, że wśród publikacji elektronicznych, zgromadzonych w badanych bibliotekach istnieje wyraźny podział na dwie grupy. Pierwszą grupę stanowią dokumenty najstarsze, które należałoby określić jako „nieużyteczne” bądź „nieczytelne”, opublikowane w latach 1994-1999, na nośnikach i w formatach, które obecnie nie są w powszechnym użytku i których odczyt oraz prezentacja są zwykle niemożliwe z powodu braku starszych wersji sprzętu i oprogramowania. Druga grupa natomiast to dokumenty „użyteczne”, „czytelne”, opublikowane po 2000 r., w zdecydowanej większości w formatach i na nośnikach uznanych za standardowe, obecnie możliwe do odczytania i prezentacji przy użyciu powszechnie dostępnej platformy sprzętowo-programowej. W drugiej grupie bardzo poważną część publikacji stanowią zasoby bibliotek cyfrowych, przechowywane na serwerach, nad którymi fachową opiekę sprawują techniczni pracownicy bibliotek, często informatycy, przygotowani do prac nad zarządzaniem i archiwizacją kolekcji cyfrowych dokumentów. Archiwizacja ta polega zazwyczaj na zautomatyzowanej

kontroli sporządzanych sum kontrolnych oraz tworzeniu kopii zapasowych. Ze studiów rodzimej literatury przedmiotu oraz badań wynika, że kolekcje te mają bardzo duże szanse przetrwania w długim czasie i długoterminowego zachowania swej użyteczności. Potwierdzeniem tej opinii mogą być wypowiedzi z wywiadów: „W Bibliotece prace nad kwestiami bezpiecznej i długotrwałej archiwizacji trwają od 2006 r. Ze względu na obowiązek statutowy Biblioteki polegający na wieczystej archiwizacji jednego egzemplarza polskich materiałów bibliotecznych od początku budowy zasobów cyfrowych kwestia ich archiwizacji była traktowana z całą powagą i w oparciu o najlepsze wzorce opracowane w innych europejskich bibliotekach narodowych”. Podobnie dzieje się w drugiej badanej księżnicy narodowej, tzn. zbiory zdigitalizowane podlegają ochronie w sensie poddawania ich zabiegom typowym dla przechowywania danych w systemie informatycznym, np. sporządzanie kopii zapasowych.

O dużej dbałości i profesjonalnym zarządzaniu zasobami polskich bibliotek cyfrowych świadczą prace Poznańskiego Centrum Superkomputerowo-Sieciowego. Z literatury przedmiotu oraz na podstawie informacji uzyskanych podczas wywiadu z pracownikami PCSS, bezpośrednio związanymi z Federacją Bibliotek Cyfrowych wynika, że prace te realizowane są z uwzględnieniem przyjętych w świecie standardów. Pomimo znacznie skromniejszych możliwości finansowych, polskie biblioteki cyfrowe są organizowane i funkcjonują na poziomie podobnych kolekcji w świecie. Wprawdzie ich ochrona zapewniająca długoterminową użyteczność jest tematem podejmowanym dopiero od 2008 r. i wymaga dużego nakładu pracy, ale istnieją bardzo duże szanse, że przy przemyślanym, starannym i konsekwentnym działaniu uda się również w tej kwestii osiągnąć pożądaną efekt. Najważniejsza jest świadomość potrzeby działań archiwizacyjnych, a ta w środowisku PCSS jest na wysokim poziomie, co zostało podkreślone w wypowiedzi podsumowującej pierwszą polską konferencję oraz piąte warsztaty poświęcone polskim bibliotekom cyfrowym zorganizowane w Poznaniu w 2008 r.⁷.

Na podstawie wywiadów oraz odpowiedzi ankietowych można wnioskować, że w środowisku polskich bibliotekarzy temat długoterminowej archiwizacji publikacji elektronicznych wywołuje skojarzenia związane z digitalizacją i budowaniem kolekcji cyfrowych; nie wspomina się o ochronie oryginalnych publikacji elektronicznych, zwłaszcza tych starszych, zapisanych na fizycznych nośnikach przenośnych, a to właśnie w ich przypadku potrzeba działań jest najpilniejsza. Należałoby podjąć starania na rzecz ustalenia ich liczby, objętości, rodzajów nośników oraz formatów zastosowanych do ich zapisu, a także zdefiniować sprzęt i oprogramowanie, potrzebne do ich odczytu i prezentacji, ewentualnie podjąć decyzję o zastosowaniu metody migracji bądź emulacji w celu odtworzenia treści tych dokumentów. Przede wszystkim jednak, należałoby zastanowić się,

⁷ Pierwsza konferencja „Polskie Biblioteki Cyfrowe” odbyła się w dniach 24-25 listopada 2008 r. w Poznaniu. Szczegółowe informacje na temat konferencji oraz materiały konferencyjne są dostępne w serwisie WWW, pod adresem: <<http://www.man.poznan.pl/PBC/2008/>> [dostęp: 2.03.2009].

czy treść tych publikacji jest warta ponoszonych kosztów wymienionych powyżej czynności oraz czy jest zgłaszana potrzeba ich użytkowania. Z wywiadu przeprowadzonego w BN wynika, że w instytucji tej były podejmowane próby odczytu treści plików metodą emulacji i prace te są kontynuowane.

Niepokojące są wyniki badań, dotyczące okresowego odczytu oraz tworzenia kopii zapasowej zbiorów elektronicznych opublikowanych na fizycznych nośnikach przenośnych. Dowodzą one, że w bibliotekach nie ma planowego odczytu publikacji elektronicznych. Odczyt następuje w procesie katalogowania dokumentu, a potem sporadycznie – w procesach udostępniania. Jest to podejście mało skuteczne i mogące prowadzić do sytuacji, w której zasoby obecnie użyteczne dołączą do grupy zasobów nieużytecznych. Istnieje zatem pilna potrzeba podjęcia stosownych działań w tym zakresie. Jedynie w Bibliotece Narodowej deklaruje się istnienie planów dotyczących okresowego odczytu publikacji elektronicznych, a także rozpoczyna się prace związane z procesami odświeżania nośników publikacji elektronicznych. Z pozostałych odpowiedzi wynika, że procesy odświeżania nośników nie były dotychczas przeprowadzane, a kopie bezpieczeństwa mają tylko kolekcje dokumentów zdigitalizowanych, przechowywanych na serwerach bądź macierzach dyskowych. Treści dokumentów z takich nośników jak dyskietki i CD-ROM-y nie były przenoszone ani na nośniki nowszych generacji, ani na serwery. Tego typu pomysły były wysuwane przez pracowników Biblioteki Jagiellońskiej, jednak przeszkodą okazał się brak narzędzi emulujących.

Odpowiedzi respondentów wskazują, iż w badanych instytucjach nie dokonuje się oceny i selekcji publikacji, w celu identyfikacji tych, które mają ponadczasową wartość, istotne znaczenie dla nauki i kultury oraz powinny podlegać szczególnej ochronie. Oznacza to, że w planach dotyczących odczytu publikacji i procesów odświeżania, bądź zmiany generacji nośnika, ewentualnie oddzielenia treści publikacji od nośników oryginalnych i umieszczenie ich na serwerze archiwalnym, instytucje powinny uwzględnić potrzebę oceny wartości treści publikacji. Prawdopodobne jest, że ze względów finansowych, na serwerach archiwalnych, nawet tych w bibliotekach narodowych, przechowywane będą tylko publikacje o szczególnym znaczeniu i wartości ponadczasowej. W jednej z ankiet zwrotnych zamieszczono komentarz, powołując się na zalecenie IBM oraz Digital Center, z którego wynika, że dużym ułatwieniem procesu długoterminowej archiwizacji publikacji elektronicznych jest „ucieczka” od fizycznych nośników przenośnych. W literaturze przedmiotu, głównie angielskojęzycznej, niejednokrotnie prezentowano podejście w archiwizacji polegające właśnie na oddzielaniu treści publikacji od oryginalnych nośników i ich archiwizowanie w systemach depozytowych.

ŚWIADOMOŚĆ POTRZEBY ZAPEWNIENIA DŁUGOTRWAŁEJ ARCHIWIZACJI PUBLIKACJI ELEKTRONICZNYCH I WIEDZA O TYM PROCESIE

Celem badania było również uzyskanie informacji na temat poziomu świadomości pracowników polskich instytucji pamięci w zakresie potrzeby i obowiązku długoterminowej archiwizacji publikacji elektronicznych, a także poziomu wiedzy w tym zakresie.

Generalnie zdecydowanie częściej poziom tej świadomości był przez uczestników badania oceniany jako niezadowolający, niż jako wystarczający. W jednej odpowiedzi sugeruje się nawet zupełny jej brak. W dwóch ankietach został on oceniony zadowolająco, jednak z komentarzem, iż odnosi się to do pracowników bibliotek gromadzących narodowy zasób biblioteczny.

Na podstawie literatury przedmiotu, komentarzy respondentów oraz rozmów prowadzonych w ramach badań, można wnioskować, że w Polsce poziom świadomości i wiedzy o archiwizacji publikacji elektronicznych wychodzi z fazy pierwszej, w której pojęcie długoterminowej archiwizacji utożsamiano z procesami digitalizacji. Digitalizacja sama w sobie miała zapewnić długoterminową użyteczność publikacji. Nie uświadamiano sobie wówczas, że samo zgromadzenie materiału w postaci cyfrowej nie jest jego długoterminowym zabezpieczeniem, gdyż materiał cyfrowy potrzebuje, być może staranniejszej ochrony, szczególniejszych warunków przechowywania i metod konserwacji, aniżeli dokumenty analogowe. W Polsce fazę pierwszą kształtowania się świadomości prawdopodobnie można datować na lata 1999-2004. Jej początek jest związany z pierwszymi projektami digitalizacji najbardziej zagrożonych zbiorów drukowanych i pierwszymi sieciowymi kolekcjami dokumentów cyfrowych, a koniec to okres pojawienia się w polskiej literaturze przedmiotu wypowiedzi dotyczących długoterminowej archiwizacji publikacji elektronicznych oraz projektów z tego zakresu realizowanych w krajach zachodnich. Krzywdząca byłaby jednak powyższa charakterystyka w odniesieniu do całego środowiska związanego zawodowo z instytucjami pamięci. Z całą pewnością istniały wówczas w Polsce jednostki, które wykazywały się dobrą znajomością tematu i świadomością potrzeby archiwizowania publikacji elektronicznych we właściwym znaczeniu tego terminu⁸. Przepuszczalnie, świadomość potrzeby archiwizacji publikacji elektronicznych w Polsce znajduje się

⁸ Do pierwszych polskich publikacji zawierających odniesienie do zagadnień długoterminowej archiwizacji zasobów cyfrowych należy opracowanie Juranda Czermińskiego (Czermiński, 2002). Autor przytacza przykłady zagrożeń istnienia dokumentów, podkreśla konieczność archiwizacji i słusznie zwraca uwagę, że „sam proces digitalizacji jest często zaliczany do czynności archiwizacyjnych i wielu osobom trudno jest pojąć, po co archiwizować archiwum”. Wówczas już J. Czermiński zwrócił uwagę na podejmowane w świecie próby rozszerzenia egzemplarza obowiązkowego o dokumenty publikowane w Internecie wraz z ich obowiązkową archiwizacją, przytaczając przykłady archiwizacji fińskich i szwedzkich witryn internetowych. Autor zauważył, że w Polsce ustalenia wymaga miejsce składowania i rotacji cyfrowych zapisów archiwalnych; utrzymywanie się koncepcji archiwizacji na poziomie lokalnym, bez terytorialnej wymiany kopii J. Czermiński – słusznie – uznał za wyjątkowo niefortunne. Zauważył też, że potrzebne są rządowe, samorządowe lub komercyjne instytucje, oferujące usługi przechowywania cyfrowych materiałów archiwalnych.

obecnie w fazie drugiej, w której pod pojęciem długoterminowej archiwizacji publikacji elektronicznych bądź długoterminowej archiwizacji zasobów cyfrowych rozumie się stosowną ochronę i zabezpieczenie na przyszłość głównie zasobów bibliotek cyfrowych, zgromadzonych i przechowywanych na serwerach. Archiwizację utożsamia się z zabiegami sporządzania kopii zapasowych, z okresowym sprawdzaniem sum kontrolnych, z terytorialnym oddaleniem jednej kopii od miejsca macierzystego archiwum, przy czym są to zadania cedowane na informatyków lub techników bibliotecznych. Rola bibliotekarzy w tym procesie kończy się na sporządzeniu metodą digitalizacji cyfrowego dokumentu, który ma być udostępniany użytkownikom. Z jednej wypowiedzi zamieszczonej w ankiecie wynika, że wśród pracowników instytucji kultury widoczny jest sposób myślenia i działania, zgodnie z którym „podstawowym celem działalności digitalizacyjnej jest udostępnienie zasobów cyfrowych. Powoduje to umieszczanie w bibliotekach cyfrowych plików o niskiej rozdzielczości, podczas gdy pliki macierzyste często są archiwizowane na nośnikach nietrwałych (dyski wymienne, płyty DVD)”. W fazie tej sporadycznie występuje świadomość na temat potrzeby objęcia ochroną publikacji zapisanych i przechowywanych na fizycznych nośnikach przenośnych. Bardzo rzadkie są opinie o pilnej potrzebie prób odczytu publikacji najstarszych.

W obecnej fazie rozwoju świadomości raczej nie mówi się o roli twórców, głównie wydawców publikacji elektronicznych w procesach archiwizacji i nie prowadzi się prac w zakresie opracowywania wytycznych dotyczących publikowania elektronicznego. Zwraca się uwagę na wytyczne i standardy, ale w odniesieniu do procesów digitalizacji i tworzenia bibliotek cyfrowych. Przedstawiciele bibliotek przekonują się jednak, że współpraca z twórcami zasobów cyfrowych jest nieunikniona. Z rozmowy przeprowadzonej w Bibliotece Narodowej wynika, że realizowany jest projekt tworzenia narodowego repozytorium BN oraz gromadzenia polskich witryn internetowych, czyli tzw. harwestowania polskiego Internetu. Szczegółowe dane na temat stanu zaawansowania tych prac i stosowanych w nich metod jednak nie zostały dotąd upublicznione. Doświadczenia instytucji bibliotecznych innych krajów pokazują, że tego typu przedsięwzięcia wymagają ścisłej współpracy, w różnych zakresach, twórców zasobów oraz instytucji, które je gromadzą, chronią i organizują do nich dostęp.

Przypuszczalnie, opisana druga faza rozwoju świadomości potrzeby archiwizacji długoterminowej materiału elektronicznego występuje w dość wąskich kręgach ludzi związanych zawodowo z instytucjami bibliotecznymi. Jej upowszechnienie wymaga zatem stosownych działań propagujących.

Ze wszystkich uzyskanych odpowiedzi jednoznacznie wynika, że w Polsce istnieje potrzeba propagowania zagadnień związanych z długoterminową archiwizacją zasobów cyfrowych. Mało znane są również oferty szkoleń, kursów, konferencji etc. z tego zakresu tematycznego. W dwóch ankietach wskazano na przykłady szkoleń i konferencji dotyczących tworzenia bibliotek cyfrowych, co jest potwierdzeniem, że nadal istnieje skłonność do

utożsamiania długoterminowej archiwizacji z digitalizacją i tworzeniem bibliotek cyfrowych. Respondenci nie wymienili przykładów oferty szkoleniowej z zakresu długoterminowej archiwizacji w ujęciu prezentowanym w badaniu. Osobiste rozmowy w ramach wywiadu dostarczyły dowodów, że w temacie archiwizacji istnieje terminologiczny nieład. W polskim środowisku bibliotecznym używane jest pojęcie elektronicznej archiwizacji i rozumie się je jako digitalizację, w sensie zabezpieczenia materiałów drukowanych. Prawdopodobnie przyjmuje się więc, że zdigitalizowanie dokumentu i składowanie w cyfrowej postaci zapewni jego przechowanie i użyteczność w długim czasie. Tymczasem zapewnienie długoterminowej dostępności i użyteczności cyfrowej postaci tego materiału to domena długoterminowej archiwizacji, czyli stosownych zabiegów głównie natury technicznej. Rozmówcom należało jednak uświadamić, że długoterminowa archiwizacja nie powinna oznaczać tylko ochrony materiałów zdigitalizowanych, składowanych na serwerach i na bieżąco udostępnianych, jak to ma miejsce w przypadku bibliotek cyfrowych. Co najmniej tak samo pilne jest objęcie procesami archiwizacji oryginalnych publikacji elektronicznych, zapisywanych na fizycznych nośnikach przenośnych, zwłaszcza tych najstarszych. Po takim wyjaśnieniu rozmówcy przyznawali, że temat długoterminowej archiwizacji publikacji elektronicznych podejmowano dotychczas w Polsce bardzo rzadko i że konieczne jest jego propagowanie. W jednej z odpowiedzi wskazano na wątki dotyczące długoterminowej archiwizacji pojawiające się na seminariach organizowanych przez Centrum Promocji Informatyki.

Świadomość potrzeby archiwizowania publikacji elektronicznych w Polsce w pełni zaistnieje wówczas, gdy procesami archiwizacji objęte zostaną zarówno publikacje zdigitalizowane, jak i pochodzące z procesów publikowania elektronicznego, zapisywane niegdyś na dyskietkach, obecnie najczęściej na dyskach typu CD oraz DVD. Nastąpi wówczas trzecia faza rozwoju świadomości, cechująca kraje i instytucje biblioteczne, w których od lat stosuje się świadomie zaplanowane i realizowane strategie archiwizacji długoterminowej. Regularnie odczytuje się tam treści publikacji elektronicznych i w zależności od przyjętej strategii, przenosi się je na serwery archiwalne bądź pozostawia w oryginalnym środowisku, jednak wykonuje się odświeżanie bądź zmianę generacji nośnika, planuje się, w razie potrzeby, operacje migrowania i emulowania. Bez względu na typ nośnika, sporządza się kopie zapasowe publikacji. Wszelkie operacje i zabiegi konserwatorskie przeprowadzane na całych kolekcjach cyfrowych, bądź jej elementach, są szczegółowo dokumentowane. Powołuje się centralną instytucję koordynującą, która kieruje działaniami archiwizacyjnymi w skali kraju, organizuje szkolenia, zarządza diagnozę stanu zasobów elektronicznych w instytucjach pamięci, prowadzi badania i prace w zakresie tworzenia narodowej strategii długoterminowej archiwizacji narodowego zasobu cyfrowego zgodnie z najnowszymi wytycznymi i standardami oraz uświadamia instytucjom rządowym i pozarządowym potrzebę finansowania działań archiwizacyjnych.

Z przeprowadzonych badań wynika, że do polskich instytucji bibliotecznych problematykę długoterminowej archiwizacji publikacji elektronicznych wprowadza się z około dziesięcioletnim opóźnieniem w stosunku do krajów zachodnich. Jest to oczywiście fakt niepokojący i prowokujący pytanie, czy polskim bibliotekarzom i archiwistom uda się zapewnić użyteczność najstarszych polskich publikacji elektronicznych, czy też polscy użytkownicy bezpowrotnie utracili już możliwość ich odczytu i interpretacji? Nie należy wykluczać, że nie ma podstaw do niepokoju; może okazać się bowiem, że treści najstarszych polskich publikacji elektronicznych nie przedstawiają szczególnej wartości bądź mają swoje odpowiedniki w formie drukowanej. Należałoby jednak przystąpić do stosownych prac diagnostycznych. Długoterminowa użyteczność polskich zasobów cyfrowych, głównie tych najstarszych, jest niewątpliwie zależna od świadomości przedstawicieli polskich instytucji pamięci dotyczącej pilnej potrzeby organizacji procesów archiwizacyjnych zgodnie z zasadami charakterystycznymi dla trzeciej fazy jej rozwoju.

Wyniki omawianego badania potwierdziły, że w polskich instytucjach bibliotecznych są zatrudnieni ludzie z wykształceniem i predyspozycjami, umożliwiającymi organizowanie prac w zakresie długoterminowej archiwizacji publikacji elektronicznych. W większości bibliotek zadeklarowano również posiadanie otoczenia sprzętowo-programowego, umożliwiającego odczyt i prezentację większości publikacji elektronicznych. Zarówno w Bibliotece Jagiellońskiej, jak i Bibliotece Narodowej przyznano, że poważny problem może stanowić odczyt najstarszych dokumentów. W obu instytucjach zgodzono się z opinią, że jest to wynik nieświadomości potrzeby stosownych działań we właściwym czasie. A zatem, w bibliotekach tych nie zadziałał, ponieważ nigdy nie został uruchomiony, tzw. mechanizm *technology watch*, którego zadaniem jest przypomnienie w stosownym czasie o potrzebie odświeżenia lub zmiany generacji nośnika, zastosowania emulacji bądź migracji. Stało się tak, pomimo że wszyscy badani respondenci zgadzają się z opinią, że materiał elektroniczny wymaga szczególnych warunków i metod przechowywania, w przeciwnym razie treści dokumentów zapisanych w cyfrowej postaci zostaną utracone. Dlatego tak istotne znaczenie ma dokładny opis publikacji elektronicznej, w którym m.in. zawiera się informację o typie nośnika, jego trwałości oraz planowej dacie testowego odczytu.

Jeden z respondentów przyznał w dodatkowym komentarzu, że w zasobach instytucji, w której pracuje, jest z pewnością sporo publikacji elektronicznych, zwłaszcza starszych, na dyskietkach 5,25 i 3,5 cala, z odczytem których z różnych powodów mogą być problemy i że instytucje biblioteczne powinny przystąpić do ich diagnozy, jednak dopóki nie robią tego biblioteki centralne i nie ma odgórnych wytycznych, działania prawdopodobnie nie zostaną podjęte. Biblioteki postawę taką tłumaczą brakiem środków. Autor wypowiedzi jest jednak pewien, że „diagnozę należałoby przeprowadzić nawet, jeśli nie ma na to środków”. W odniesieniu do systemu zarządzania publikacjami elektronicznymi, uzyskane odpo-

wiedzi ankietowe są dosyć zróżnicowane. Trzech respondentów zwróciło uwagę, że najważniejszą funkcją zarządzania jest zapewnienie odczytu zgromadzonych dokumentów elektronicznych i ich udostępnienie, najlepiej poprzez Sieć. Dodatkowo, w jednej odpowiedzi zwrócono uwagę na potrzebę stworzenia centralnego katalogu zasobów elektronicznych, w skali globalnej, i udostępnienie go online. Również w trzech ankietach istotne znaczenie przypisano dokładnemu, ujednoczonemu opisowi dokumentów elektronicznych, z uwzględnieniem metadanych bibliograficznych, technicznych i administracyjnych oraz sporządzaniu kopii bezpieczeństwa.

W trzech wypowiedziach na temat zarządzania zasobami archiwalnymi, pojawił się wątek potrzeby przechowywania oryginałów, nazywanych też plikami macierzystymi oraz ich kopii zapasowych w bezpiecznych repozytoriach cyfrowych. Jeden z respondentów zasugerował możliwość skorzystania z oferty zewnętrznych usługodawców, którzy przejęliby zadania archiwizacji zasobów elektronicznych.

Respondenci częściej preferują tworzenie lokalnych archiwów polskich zasobów cyfrowych, zorganizowanych i działających na podstawie ujednoczonych zasad. Zdaniem dwóch respondentów, w Polsce należałoby utworzyć jedno, centralne, narodowe archiwum, gromadzące kompletny zasób cyfrowy Polski. Jeden z respondentów argumentuje, że do zbudowania profesjonalnego systemu archiwalnego, zgodnego ze światowymi standardami potrzebna jest jednolitość procedur; w przypadku archiwów lokalnych, bardzo prawdopodobne są rozbieżności, które utrudnią funkcjonowanie jednorodnego, rozproszonego archiwum. Z badań wynika zatem, że nie ma zgodności co do modelu archiwum polskiego dziedzictwa cyfrowego. Niewątpliwie, potrzebne jest współdziałanie polskich instytucji pamięci, choćby właśnie po to, aby opracować i powołać powszechnie akceptowaną formę organizacji i funkcjonowania polskiego archiwum, bez względu na jego model. Prace takie, zgodnie z odpowiedziami większości respondentów, powinna inicjować i koordynować Biblioteka Narodowa. W pięciu wypowiedziach zadeklarowano chęć przyłączenia się do grupy opracowującej strategię archiwizacji cyfrowego dziedzictwa Polski.

Podczas wywiadu próbowano uzyskać opinię respondenta z Biblioteki Narodowej w kwestii ewentualnego podjęcia przez BN roli koordynatora działań archiwizacyjnych w Polsce. Z odpowiedzi wynika jednak, że za bardziej kompetentne przedstawiciel BN uznaje Ministerstwo Kultury i Dziedzictwa Narodowego. „Polskie dziedzictwo cyfrowe obejmuje zbiory archiwalne, biblioteczne i muzealne, archiwa audiowizualne oraz zbiory należące do fundacji, stowarzyszeń, osób prywatnych, instytucji kościelnych. Koordynacja archiwizacji tak różnorodnego materiału wymaga znajomości potrzeb różnych instytucji pamięci w Polsce, a jednocześnie powinna dawać możliwości zharmonizowanego finansowania prac”.

Nasuwa się zatem wniosek, że Biblioteka Narodowa z bardzo dużym prawdopodobieństwem byłaby właściwym w Polsce koordynatorem prac na rzecz długoterminowej archiwizacji elektronicznych zasobów bibliotecznych i takiej roli podjęłaby się. Natomiast w odniesieniu do komplet-

nego dziedzictwa cyfrowego Polski należałoby powołać, co sugeruje się również w jednej z odpowiedzi ankietowej, „...odrębną jednostkę, zrzeszającą pracowników instytucji gromadzących dokumenty elektroniczne (bibliotekarzy, archiwistów, muzealników) oraz informatyków”. Nie jest wykluczone, że takie zrzeszenie, działające przy właściwym ministerstwie i finansowane przez rząd mogłoby właściwie organizować prace archiwizacyjne i kierować nimi.

Ciekawe i potwierdzające tezę o utożsamianiu w Polsce procesów długoterminowej archiwizacji, głównie z budową polskich bibliotek cyfrowych, są odpowiedzi na pytania ankiety, dotyczące znajomości polskich inicjatyw na rzecz opracowania projektu długoterminowej archiwizacji polskiego dziedzictwa cyfrowego. W kwestii tej wypowiada się pięciu respondentów, a jako przykłady projektów archiwizacji polskich zbiorów elektronicznych zgodnie wymieniają oni: Federację Bibliotek Cyfrowych i kilka nazw polskich bibliotek cyfrowych oraz Narodowe Archiwum Cyfrowe. W jednym z wywiadów zwrócono uwagę na działalność Ministerstwa Kultury i Dziedzictwa Narodowego oraz Zespołu ds. Digitalizacji.

Omawiane badanie miało też służyć poznaniu opinii pracowników polskich instytucji gromadzących i udostępniających zasoby cyfrowe, z jakich powodów długoterminowa archiwizacja została określona mianem wyzwania oraz na czym, ich zdaniem, polega trudność organizacji działań archiwizacyjnych w Polsce. W jednej ankiecie pytanie to pozostało bez odpowiedzi, w pozostałych respondenci wymieniali najczęściej: brak środków finansowych, brak wiedzy, jak reagować na zmiany technologiczne oraz brak planu działania, w sensie braku wzorców oraz zaniedbań organizacyjnych. Nieco rzadziej wymieniano szybkość zmian technologicznych, różnorodność publikacji elektronicznych, a także brak sprzętu i oprogramowania potrzebnych dla celów długoterminowej ochrony zbiorów cyfrowych.

Przy jednej z odpowiedzi zamieszczono dodatkowy komentarz, z którego wynika słuszne przekonanie, że realizacja wszelkich zadań związanych z archiwizacją zasobów elektronicznych wymaga zatrudnienia dodatkowej kadry, która mogłaby skupić się tylko na tych zadaniach. Spełnienie takiego postulatu jest oczywiście ściśle zależne od możliwości finansowych instytucji bibliotecznych, które zwykle są ograniczone. Stąd zapewne, na pierwszym miejscu w prawie wszystkich odpowiedziach wymieniono brak środków finansowych, jako czynnik utrudniający bądź uniemożliwiający podjęcie działań archiwizacyjnych w polskich instytucjach bibliotecznych.

Ostatnie pytanie zawarte w ankiecie dotyczyło ewentualnej roli wydawców publikacji elektronicznych w procesach długoterminowej archiwizacji. Respondenci bez wyjątku odpowiedzieli, że wydawcy mają lub mogą mieć wpływ na realizację procesów archiwizacji swoich produktów. W ich opinii wydawcy powinni publikować z uwzględnieniem standardów w zakresie stosowanych formatów, a także sprzętu i oprogramowania potrzebnego do odczytu publikacji. Jeden z respondentów proponuje, aby zobligować wydawców do dołączania oprogramowania niezbędnego do odczytu i prezentacji treści niestandardowych publikacji. Zwrócono również uwagę, że

nie wszystkie publikacje elektroniczne można odczytać po wykonaniu ich kopii zapasowej (dotyczy to zwłaszcza programów). Bibliotekarze wskazują również problemy z identyfikacją niektórych informacji potrzebnych do sporządzenia metadanych dla publikacji elektronicznych, stąd sugestia respondentów dotycząca większej staranności w tym zakresie, najlepiej ujednolicenia i konsekwentnego stosowania formy i miejsca umieszczania na publikacji podstawowych metadanych. Proponuje się także, aby wydawcy odsyłali publikacje do instytucji archiwalnych wraz z kompletem metadanych. Próby takie podjęto w niektórych bibliotekach na świecie, przy czym z inicjatywą wychodziły jednak instytucje biblioteczne, precyzując swoje oczekiwania wobec wydawców. Głównie dotyczy to elektronicznych publikacji sieciowych, które są zgłaszane przez wydawców do bibliotek na specjalnym formularzu, a zgłoszenie wymaga uzupełnienia przez wydawcę określonych metadanych. W przypadku publikacji na nośnikach fizycznych, pewnym rozwiązaniem i pomocą dla bibliotekarzy mogłyby być katalogi wydawnicze, w których wydawcy prezentowaliby nowo opublikowany dokument zgodnie z konwencją zaproponowaną przez instytucje archiwizujące.

Respondenci proponują również współpracę pomiędzy wydawcami oraz bibliotekarzami i archiwistami w zakresie przechowywania u wydawców tzw. matrycy („matki”) publikacji elektronicznych, co dawałoby dodatkowe zabezpieczenie dokumentów.

Wypowiadający się w ankiecie deklarują potrzebę współpracy na rzecz standaryzacji procesów wydawniczych, jednak z sugestią, że będzie to bardzo trudne w przypadku wydawnictw małych, działających efemerycznie i często nieprofesjonalnie.

BADANIA W INSTYTUCJACH WYDAWNICZYCH

W celu poznania stanowiska wydawców publikacji elektronicznych wobec ich ewentualnego udziału w procesach trwałego zachowania opublikowanego dziedzictwa nauki i kultury, odpowiedni formularz ankiety został rozesłany do instytucji wydawniczych, które publikują dokumenty cyfrowe o charakterze naukowym, które na mocy ustawy o egzemplarzu obowiązkowym powinny być przekazywane bibliotekom i przechowywane w nich jako narodowy zasób biblioteczny.

Za podstawowe źródło informacji o polskich wydawnictwach publikacji elektronicznych przyjęto bazę Biblioteki Narodowej *Bibliografia wydawnictw elektronicznych*, rejestrującą publikacje cyfrowe od 2001 r. Z obszernego indeksu wydawców (ok. 300 pozycji⁹), wybrano do badań te instytucje, które publikują cyfrowe dokumenty przydatne w procesach

⁹ Indeks wydawców bazy BN *Bibliografia Dokumentów Elektronicznych* liczył 457 pozycji (stan na dzień 25 stycznia 2009 r.), jednakże szereg wydawnictw występowało w bazie wielokrotnie, np. firma wydawnicza Bellona była odnotowana jako Bellona oraz jako Dom Wydawniczy Bellona. Należy zatem szacować, że w bazie faktycznie zarejestrowanych było ok. 300 instytucji wydawniczych.

edukacyjnych, pracy naukowej i badawczej. Jako kryteria selekcji przyjęto liczbę opublikowanych dokumentów (powyżej trzech publikacji) oraz profil przedmiotowy wydawcy, ustalany na podstawie informacji dostępnych na stronach internetowych poszczególnych wydawnictw oraz zamieszczanych w Portalu Rynku Wydawniczego WYDAWCA. Źródłami dodatkowymi informacji o wydawcach były serwisy informacyjne Polskiego Towarzystwa Wydawców Książek (PTWK, 2010) oraz Polskiej Izby Książki (PIK, 2010).

Ostatecznie do badań wybrano 39 instytucji wydawniczych ujętych w bazie BN. Zbiór ten został uzupełniony o 20 wydawnictw uczelni wyższych największych ośrodków akademickich w Polsce. Łącznie próba badawcza liczyła 58 instytucji wydawniczych.

Wydawnictwa objęte badaniem można przyporządkować do następujących kategorii wydawców: (1) uniwersalnych, (2) edukacyjnych, (3) akademickich i naukowych. W próbie badawczej znalazło się 12 firm wydawniczych z kategorii pierwszej, 14 – z kategorii drugiej oraz 32 – z trzeciej. Z opisu profilu wydawców wynika, że niektóre firmy publikują materiały o charakterze zarówno edukacyjnym, jak i uniwersalnym; można je zatem przyporządkować zarówno do pierwszej, jak i do drugiej kategorii.

Badanie w instytucjach wydawniczych przeprowadzono w dwóch etapach. Etap pierwszy polegał na wywiadzie telefonicznym z każdą wybraną do badań firmą wydawniczą, w celu potwierdzenia, czy firma rzeczywiście publikuje elektroniczne dokumenty przydatne w procesach edukacji i nauki, czy sama uznaje się za wydawcę publikacji elektronicznych i czy wyraziłaby chęć udziału w prowadzonych badaniach. Chęć udziału w badaniach wyraziły 23 firmy wydawnicze. W tej grupie 9 firm to wydawnictwa polskich uczelni wyższych, czyli wydawcy materiałów akademickich i naukowych, 10 to firmy publikujące materiały edukacyjne i 4 to wydawnictwa uniwersalne.

Dodatkowe informacje na temat stanowiska wydawców odnośnie do archiwizacji publikacji elektronicznych uzyskano od przedstawicieli Polskiego Towarzystwa Wydawców Książek oraz Polskiej Izby Książki. Z przeprowadzonych rozmów telefonicznych wynika, że temat długoterminowej archiwizacji, ochrony publikacji elektronicznych nie był w tych instytucjach dotychczas omawiany. Nie spotkano się również z głosami wydawców, które świadczyłyby o przekonaniu, że ochrona ich elektronicznych produktów powinna stać się tematem rozważań czy też przedmiotem dyskusji. Rozmówcy z PTWK oraz PIK wyrazili chęć zapoznania się z pytaniami ankietowymi. Ankiety zwrotnej jednak nie nadesłali.

Etap drugi badania stanowiło rozesłanie ankiet do tych instytucji wydawniczych, które wyraziły gotowość udziału w badaniu. Kwestionariusz ankiety rozesłano do 25 instytucji; jednak odpowiedzi otrzymano ostatecznie tylko od 2 wydawców. Z uwagi na tę bardzo niską zwrotność ankiet skierowano się z prośbą o konsultację do dyrekcji firmy wydawniczej Wolters Kluwer Polska, która brała udział w przeprowadzonym przez autorkę w 2007 r. podobnym badaniu sondażowym (por. Januszko-Szakiel, 2010).

Firma Wolters Kluwer Polska jest jednym z największych wydawców publikacji prawniczych i biznesowych na rynku polskim. Stanowi też jedną z największych firm wydawniczych w Polsce publikujących dokumenty elektroniczne, zarówno na nośnikach przenośnych, jak i w Sieci i ma długoletnie doświadczenie w tym zakresie¹⁰.

Omówiony poniżej materiał badawczy pochodzi, poza opinią z firmy Wolters Kluwer Polska, również z ankiet zwrotnych, nadesłanych z Wydawnictwa Instytutu Problemów Ochrony Zdrowia oraz Young Digital Planet. Przedstawiono też wnioski wynikające z telefonicznych rozmów przeprowadzonych z wydawcami w ramach pierwszego etapu badania.

Obie firmy wydawnicze, które wzięły udział w badaniu mają charakter wydawnictw specjalistycznych. Wydawnictwo Instytutu Problemów Ochrony Zdrowia z siedzibą w Warszawie skupia się w swych działaniach na zagadnieniach dotyczących ubezpieczeń zdrowotnych, prawa medycznego, zarządzania zakładami opieki zdrowotnej, zarządzania ryzykiem w opiece zdrowotnej, szkolenia kadry medycznej, pozyskiwania środków unijnych dla sektora medycznego. Firma jest właścicielem Portalu „Prawo i Medycyna.pl” oraz wydawcą kwartalnika „Prawo i Medycyna”¹¹. Firma wydawnicza Young Digital Planet SA działa głównie na rynku oprogramowania komputerowego. Jest dostawcą systemów szkoleniowych i treści edukacyjnych, opartych na technologiach informacyjnych, dla klientów indywidualnych, instytucji oświatowych oraz przedsiębiorstw. Celem firmy jest dostarczanie rozwiązań edukacyjnych, które charakteryzują się najwyższą jakością oraz sprawdzoną efektywnością. Firma zajmuje pozycję światowego lidera w produkcji interaktywnych programów do nauki języków obcych¹².

Na podstawie rozmów telefonicznych stwierdzono, że w Polsce przeważa publikowanie tradycyjne, tj. w formie druku na papierze. Z 58 prowadzonych rozmów telefonicznych z wydawcami, tylko 12 odpowiedziało jednoznacznie twierdząco, że uważają się za wydawców publikacji elektronicznych. Jednak nawet w największych firmach wydawniczych, które mają rozbudowane działy publikacji elektronicznych i publikują dokumenty na nośnikach przenośnych oraz w Sieci, publikowanie elektroniczne jest działalnością marginalną, aczkolwiek traktowaną poważnie. Rozmówcy pozostałych firm wydawniczych twierdzili, że procesy publikowania elektronicznego występują u nich sporadycznie i najczęściej nie są to publikacje samoistne wydawniczo, lecz materiały uzupełniające do publikacji drukowanych.

Z przeprowadzonych rozmów wynika też, że wydawnictwa uczelni w zasadzie nie publikują w wersji elektronicznej. W dwóch wydawnictwach odnotowano potrzebę dołączenia do publikacji drukowanej pliku na CD.

¹⁰ Na podstawie informacji zamieszczonych w serwisie WWW firmy (Wolters Kluwer Polska, 2010).

¹¹ Na podstawie informacji zamieszczonych w serwisie WWW Instytutu (Instytut Problemów Ochrony Zdrowia, b.d.).

¹² Na podstawie informacji zamieszczonych w serwisie WWW firmy (Young Digital Planet, 2010).

Były to rozprawy doktorskie z dużym objętościowo materiałem, który stanowił część aneksu. Przedstawiciele wydawnictw uczelnianych wyrażali przekonanie, że znacznie częściej to biblioteki uczelniane podejmują próby publikowania elektronicznego w związku z pracami digitalizacyjnymi (jako przykład wskazano Bibliotekę Główną AGH w Krakowie) oraz tworzeniem repozytoriów cyfrowych materiałów dydaktycznych i prac pracowników naukowo-dydaktycznych, udostępnianych na stronach WWW. Wydawnictwo Uniwersytetu Mikołaja Kopernika przyznało, że planuje publikację czasopism cyfrowych w Sieci.

Interesujące jest również podejście wydawców podręczników i materiałów elektronicznych, głównie w postaci systemów multimedialnych do nauczania języków obcych. Wydawcy w rozmowie tłumaczyli, że nie zaliczają siebie do wydawców publikacji elektronicznych, gdyż publikują podręczniki głównie w wersji drukowanej, a dodatki w postaci CD do podręczników bądź samoistne programy multimedialne do nauczania zamawiają u zewnętrznych producentów, pełniąc tylko rolę dystrybutora.

Bardzo podobnym podejściem charakteryzuje się wiele wydawnictw wyszczególnionych w bazie BN. Ich przedstawiciele twierdzą, że opublikowanie kilku dokumentów w postaci elektronicznej, głównie na CD, nie czyni ich wydawcami publikacji elektronicznych.

Na podstawie rozmów ustalono, że wiele firm z bazy BN zawiesiło swoją działalność bądź zaliczają się raczej do grupy firm dystrybucyjnych, aniżeli wydawniczych. Wydawcy nie są też świadomi istnienia bazy BN, w której są odnotowani.

Warto również zauważyć, że niektóre wydawnictwa przyznały się w rozmowie do znacznie większej liczby opublikowanych elektronicznych dokumentów w stosunku do liczby odnotowanej w bazie BN. Może to wynikać z zaniedbania obowiązku odsyłania do bibliotek egzemplarza obowiązkowego bądź z polityki BN dotyczącej prowadzenia bazy wydawnictw elektronicznych. Z rozmowy, którą przeprowadzono z pracownikami BN zajmującymi się tą bazą wynika, że jest ona uaktualniana na bieżąco, na podstawie wpływu egzemplarza obowiązkowego. Zatem opcja pierwsza, świadcząca o bagatelizowaniu przepisów prawnych przez wydawców, wydaje się być bardziej prawdopodobna.

Z ankiety zwrotnej nadesłanej z Wydawnictwa Instytutu Ochrony Zdrowia wynika, że dokumenty w formie elektronicznej zaczęto publikować tam w 2006 r., głównie na płytach CD i DVD. Nie zgłoszono dotychczas problemów z odczytem tych publikacji. Respondent zdecydowanie zgodził się z opinią, że zastosowanie przez wydawców odpowiedniego nośnika zapisu publikacji może mieć wpływ na efektywność procesu długoterminowej archiwizacji: „...wybór odpowiedniego nośnika jest niezwykle istotny w czasach bardzo dynamicznych zmian w dziedzinie technologii cyfrowych”. Decydując się na zapis dokumentu w odpowiednim formacie i na nośniku, wydawnictwo kieruje się głównie dwoma kryteriami, tj. ceną i powszechnością korzystania z konkretnego nośnika. Respondent ten przyznaje też, że nie spotkał się dotychczas z jakąkolwiek formą współpracy

instytucji wydawniczych z bibliotekami i archiwami w zakresie archiwizacji publikacji elektronicznych. Brak również jednoznacznej deklaracji chęci współpracy: „na pewno istotnym czynnikiem decydującym o podjęciu takiej współpracy byłyby ewentualne koszty dostosowania procesu wydawniczego do nowych standardów”. Istotne jest jednak to, że wydawca takiej współpracy nie wyklucza.

Znacznie wcześniej, bo w 1990 r., pierwsze dokumenty elektroniczne opublikowała firma Young Digital Planet. Publikacje te ukazały się na dyskietkach 5,25 cala i w badaniu przyznano, że firma nie ma obecnie możliwości odczytu treści tych dokumentów. I ten wydawca zgadza się również z opinią, że opublikowanie dokumentu w odpowiednim formacie i na odpowiednim nośniku może mieć wpływ na zachowanie jego użyteczności w dłuższym czasie, a decydując się na format i nośnik zapisu publikacji, podstawowym kryterium jest dostępność technologii.

W kwestii współpracy z instytucjami bibliotecznymi i archiwalnymi, firma YDP stwierdziła, że „przekazuje egzemplarze nowo wydanych tytułów do Biblioteki Narodowej i Biblioteki Jagiellońskiej zgodnie z ustawą o przekazywaniu egzemplarzy obowiązkowych”, bez odpowiedzi pozostawiając pytanie o możliwość dostosowania procesu wydawniczego do wytycznych bibliotekarzy i archiwistów, w celu nadania publikacjom parametrów umożliwiających ich długoterminową archiwizację. Przedstawiciel firmy zauważył natomiast, że wydawcy mogliby tworzyć wirtualne serwery dla publikacji elektronicznych. Pomysł niewątpliwie interesujący, jednak bliżej nieokreślony, zatem trudno jednoznacznie odczytać zamiar autora wypowiedzi.

Podobną koncepcję archiwizacji publikacji elektronicznych zaproponował przedstawiciel firmy Wolters Kluwer Polska: „...wspaniałym archiwum dla publikacji elektronicznych wydaje się obecnie przestrzeń Internetu. Oczywiście jest to rozwiązanie niedoskonałe pod wieloma względami, choćby z uwagi na brak kompletności, wiarygodności, systematyki. Podstawowymi zaletami Internetu jest ogólnodostępność, brak ograniczeń co do objętości, niski koszt umieszczania w Internecie publikacji i najważniejsza zaleta – publikacja przechowywana jest «wiecznie». Co dzisiaj trafi do sieci, jest przechowywane jako kopia, nawet po usunięciu pierwotnego pliku na wielu serwerach (choćby wyszukiwarka Google). Pozostaje oczywiście kwestia kompletności i autentyczności tak «przechowywanych» dokumentów. Do rozwiązania byłby więc problem, jak przy wykorzystaniu opisanej wyżej cechy Internetu, zbudować system dostępu do danych z jakimś systemem oceny wiarygodności źródła”. Pomimo wielokrotnie podejmowanych prób, nie udało się uzyskać od autora wypowiedzi, dookreślenia tego pomysłu.

Na podstawie przeprowadzonych badań można wnioskować, że w polskich firmach wydawniczych temat ochrony publikacji elektronicznych nie był dotychczas podejmowany, a jedyna forma współpracy z instytucjami bibliotecznymi polega na przekazywaniu egzemplarza obowiązkowego. Warto tu zwrócić uwagę na podjętą przez wydawców inicjatywę odnośnie

do obowiązkowego egzemplarza elektronicznego. Otóż 20 kwietnia 2009 r. w Biurze Polskiej Izby Książki w Warszawie odbyło się pierwsze posiedzenie powołanej tego samego dnia Sekcji Publikacji Elektronicznych Polskiej Izby Książki (SPE PIK). Sekcję utworzyły następujące firmy wydawnicze: Dom Wydawniczy MEDIUM, Young Digital Planet, SW Czytelnik, Key TExt, Biblioteka Analiz, LexisNexis Polska, Wydawnictwa Szkolne i Pedagogiczne, Wydawnictwo Naukowe PWN, Novae Res – Wydawnictwo Innowacyjne, ZamKor P. Sagnowski i Wspólnicy, Wolters Kluwer Polska, Porozumienie Wydawców, IPS International Publishing Service, NetPress Digital. Celem Sekcji ma być podjęcie działań dla ochrony interesów wydawców publikacji elektronicznych, w tym docelowo powołanie stowarzyszenia, które powinno posiadać prawa organizacji zbiorowego zarządu prawami autorskimi w odniesieniu do publikacji elektronicznych. Planowany jest również cykl szkoleń dla członków PIK w dziedzinach związanych z publikacjami elektronicznymi. Na posiedzeniu Sekcji Publikacji Elektronicznych PIK, 17 lutego 2010 r. jej członkowie podjęli dyskusję na temat projektu zastąpienia dotychczasowo ustawowego przepisu o egzemplarzu obowiązkowym (zgodnie z którym każdy wydawca powinien nieodpłatnie przekazywać 17 egzemplarzy każdej swojej publikacji bibliotekom wyznaczonym w ustawie), przepisem, w myśl którego liczba egzemplarzy przekazywanych nieodpłatnie do bibliotek zostałaby znacznie zmniejszona, a równocześnie wprowadzony zostałby obowiązek przekazywania Bibliotece Narodowej elektronicznego egzemplarza obowiązkowego każdej publikacji. Członkowie Sekcji odnieśli się krytycznie do tego projektu i poparli oświadczenie Sekcji opublikowane w formie Komunikatu PIK z 5 lutego 2010 r., którego podstawowe stwierdzenie brzmi: „Sekcja Publikacji Elektronicznych Polskiej Izby Książki stanowczo sprzeciwia się wprowadzeniu obowiązkowego egzemplarza elektronicznego w miejsce dotychczasowego egzemplarza papierowego, wysyłanego do bibliotek uprawnionych na mocy Ustawy z dnia 7 listopada 1996 r. o obowiązkowych egzemplarzach bibliotecznych (Dz.U. 1996, nr 152, poz. 722). Brak jakichkolwiek gwarancji dotyczących skutecznego zabezpieczenia plików przez ich dysponenta, Bibliotekę Narodową, rodzą ryzyko całkowitej utraty kontroli nad dostarczonymi treściami. Wydawcy mogą zgodzić się na przekazywanie jednego zabezpieczonego egzemplarza dla potrzeb archiwum narodowego, dopiero po przedstawieniu przez Bibliotekę Narodową systemu zabezpieczeń zbiorów, formatu zapisu pliku i sposobów archiwizacji, a także po przedstawieniu przez Bibliotekę Narodową umowy ubezpieczeniowej, która w przypadku złamania zabezpieczeń przechowywanych w jej zbiorach treści, gwarantowałaby wydawcom wypłatę odszkodowań z tytułu utraconych korzyści. Jednocześnie członkowie Sekcji Publikacji Elektronicznych Polskiej Izby Książki stanowczo sprzeciwiają się, by przekazywane przez wydawców pliki mogły być udostępniane lub przechowywane gdziekolwiek poza gmachem Biblioteki Narodowej”¹³.

¹³ Zob. PIK, 2010. Informacja potwierdzona przez przewodniczącego Sekcji Publikacji Elektronicznych PIK, Łukasza Gołębińskiego.

Stanowisko Sekcji jednoznacznie wskazuje, jakie są oczekiwania wydawców od Biblioteki Narodowej w zakresie archiwizacji zasobów cyfrowych, opublikowanych przez polskich wydawców. Zatem kształt procesów ich gromadzenia i archiwizowania będzie zależeć głównie od decyzji i działań Biblioteki Narodowej.

POZNAŃSKIE CENTRUM SUPERKOMPUTEROWO-SIECIOWE WOBEC PROBLEMU DŁUGOTERMINOWEJ ARCHIWIZACJI CYFROWYCH MATERIAŁÓW BIBLIOTECZNYCH

Poznańskie Centrum Superkomputerowo-Sieciowe, afiliowane przy Instytucie Chemii Bioorganicznej PAN, działa od 1993 r. i świadczy m.in. usługi związane z Internetem dla środowiska naukowego. W 1996 r. PCSS rozpoczęło prace badawcze w zakresie bibliotek cyfrowych, które miały na celu stworzenie oprogramowania do budowy bibliotek cyfrowych. W rezultacie, w 2002 r. powstał system dLibra jako pierwszy zastosowany do budowy Wielkopolskiej Biblioteki Cyfrowej i z czasem do kolejnych polskich bibliotek cyfrowych, tworzących obecnie Federację Bibliotek Cyfrowych.

Oprócz projektu dotyczącego oprogramowania dLibra i budowy polskich bibliotek cyfrowych, PCSS jest zaangażowane w europejskie projekty związane z bibliotekami cyfrowymi: ENRICH oraz EuropeanaLocal¹⁴.

Podstawowym celem wywiadu, który został przeprowadzony w maju 2008 r. w Poznańskim Centrum Superkomputerowo-Sieciowym było ustalenie, czy PCSS przyłączyłoby się do opracowania strategii długoterminowej archiwizacji polskich zasobów cyfrowych i podjęło się technicznej opieki nad archiwum polskiego zasobu cyfrowego. Zamiarem było również poznanie strategii ochrony zasobów zgromadzonych w polskich bibliotekach cyfrowych i ewentualnego programu ich długoterminowej archiwizacji.

W rozmowie przedstawiciele PCSS wyrazili zainteresowanie udziałem w pracach nad ustaleniem strategii długoterminowej archiwizacji elektronicznych zasobów polskich bibliotek oraz zwrócili uwagę na projekt PCSS o nazwie Krajowy Magazyn Danych. Celem projektu KMD jest zbudowanie rozproszonego systemu przechowywania danych o zasięgu krajowym, który ma zapewnić wydajne metody dostępu do danych oraz operacji na danych, a także wysoki poziom wiarygodności, niezawodności i bezpieczeństwa danych. Głównym zastosowaniem systemu ma być usługa tworzenia kopii zapasowych i archiwizacji m.in. dla instytucji akademickich i edukacyjnych¹⁵. Zgodnie z opiniami prezentowanymi przez przedstawicieli PCSS, tworząc system archiwalny dla polskich zasobów cyfrowych należałoby założyć ich składowanie na serwerze archiwalnym, tj. profesjonalnym archiwizerze. Treści dokumentów opublikowanych na

¹⁴ Na podstawie: PCSS – badania i projekty, b.d.

¹⁵ Zob. KMD, 2009; A. Januszko-Szakiel (2010), s. 422.

nośnikach fizycznych, najsensowniej byłoby oddzielić od nośników i również umieścić na serwerze. Jeśli natomiast program ochrony zakładałby archiwizację na nośniku oryginalnym, należałoby przynajmniej co pięć lat dokonać zmiany nośnika. Do zapisu treści publikacji cyfrowych, informatycy z PCSS doradzali stosowanie formatu DjVu do celów prezentacji oraz TIFF jako formatu archiwalnego. Ich zdaniem, zawsze należy decydować się na standardy otwarte.

Śród problemów, z którymi najtrudniej uporać się w procesach długoterminowej archiwizacji respondenci wymienili tylko brak środków finansowych. Nawet najbardziej skomplikowane techniczne zadania można wykonać, pod warunkiem, że są fundusze na stosowne badania i projekty testowe.

W rozmowie pytano również o wskazania dotyczące modelu archiwum dla polskiego dziedzictwa cyfrowego. Z odpowiedzi wynika, że łatwiejsze i sprawniejsze organizacyjnie wydaje się utworzenie jednego centralnego archiwum. Należy jednak rozważyć, w jakim stopniu zaletą, a w jakim wadą jest złożenie funduszy i odpowiedzialności „w jedne ręce”. Prawdopodobnie lepszy byłby model jednorodnego archiwum rozproszonego, z rozproszoną odpowiedzialnością. Archiwum rozproszone wydaje się też lepiej spełniać wymagania bezpieczeństwa, np. w przypadku awarii, zagrożenie uszkodzenia bądź utraty zasobów dotyczy fragmentu kolekcji, a nie całości, jak w przypadku archiwum centralnego.

W kwestii metod ochrony zasobów polskich bibliotek cyfrowych stwierdzono, że odbywa się ona w sposób powszechnie przyjęty w systemach informatycznych, tzn. wykonywane są sumy kontrolne oraz kopie bezpieczeństwa obiektów cyfrowych. Zasoby bibliotek cyfrowych nie są objęte programem archiwizacji długoterminowej; nie ma w architekturze i funkcjach systemu dLibra komponentu odpowiedzialnego za długoterminową użyteczność zasobów cyfrowych. Pomysł jego utworzenia bardzo spodobał się pracownikom PCSS i wyrazili opinię, że warto uzupełnić architekturę dLibry. Interesująca wydała się również idea selekcjonowania zasobów i utworzenia w systemie modułu zajmującego się przechowywaniem treści ważnych.

Kompetencje i dotychczasowe doświadczenia PCSS wskazują, że byłoby ono właściwym partnerem w działaniach archiwizacyjnych polskich instytucji bibliotecznych. Od wielu lat pracownicy PCSS są skupieni na pracach związanych z tworzeniem i zarządzaniem polskimi bibliotekami cyfrowymi oraz z rozwojem projektu Federacja Bibliotek Cyfrowych, tym samym znają polskie środowisko biblioteczne oraz jego potrzeby. Są również świadomi potrzeby działań na rzecz długoterminowej archiwizacji zasobów cyfrowych, posiadają wiedzę i doświadczenie, potrzebne w tych procesach. Prawdopodobnie byłiby cennymi doradcami w zakresie technicznych zagadnień długoterminowej archiwizacji i tworzenia systemu depozytowego. Ponadto, badania przeprowadzone w bibliotekach wskazują, że PCSS jest firmą znaną i wiarygodną dla polskiego środowiska bibliotecznego. Bibliotekarze chętnie powierzyliby właśnie PCSS opiekę nad polskim zasobem cyfrowym.

PODSUMOWANIE

Wyniki omówionych powyżej badań upoważniają do wniosku, że w polskich instytucjach bibliotecznych, nawet w tych o statusie bibliotek narodowych, w których najstarsze publikacje elektroniczne pochodzą z lat 1994-1996, nie wypracowano dotąd strategii długoterminowej ochrony użyteczności materiałów elektronicznych. Odczyt i prezentacja treści najstarszych publikacji elektronicznych, zgromadzonych w polskich bibliotekach jest utrudniona, a nawet niemożliwa z racji braku potrzebnego sprzętu i oprogramowania. Mniej niepokojący jest stan publikacji wydanych po 2000 r. Odczyt i prezentacja ich treści są w zasadzie wykonalne, aczkolwiek konieczne i pilne jest podjęcie decyzji o sposobie długoterminowego utrzymania ich użyteczności. W dobrej kondycji są publikacje elektroniczne tworzące zasoby polskich bibliotek cyfrowych. Z uwagi na to, że kolekcje te są tworzone, składowane i zarządzane w profesjonalnych systemach informatycznych i posiadają kopie bezpieczeństwa, mają duże szanse przetrwania do czasu, kiedy zostaną umieszczone i zarchiwizowane w profesjonalnych systemach depozytowych.

Generalnie poziom świadomości pracowników polskich instytucji bibliotecznych dotyczący potrzeby długoterminowej archiwizacji polskich zasobów cyfrowych jest wciąż niezadowalający. Procesy archiwizacyjne nadal bywają utożsamiane z digitalizacją, czyli ochroną materiałów analogowych i tworzeniem bibliotek cyfrowych. Ewentualnie, długoterminową archiwizację publikacji elektronicznych, rozumianą jako sporządzanie kopii zapasowych, bibliotekarze odnoszą jedynie do zasobów bibliotek cyfrowych. Zupełnie pomijana jest natomiast kwestia archiwizacji materiałów bibliotecznych opublikowanych na nośnikach przenośnych.

Mimo tej ogólnej obserwacji stwierdzić należy, że w środowisku polskich bibliotekarzy, bibliotekoznawców i informatologów są osoby właściwie rozumiejące pojęcie długoterminowej archiwizacji publikacji elektronicznych i świadome tego, że materiały elektroniczne powinny być przechowywane w odpowiednich warunkach, okresowo odczytywane w celu zidentyfikowania ewentualnych utrudnień odczytu i prezentacji ich treści, poddawane właściwym metodom archiwizacyjnym. Przykładem są specjaliści z Biblioteki Narodowej, gdzie planuje się i podejmuje działania archiwizacyjne. Inicjatywą, która zasługuje na szczególną uwagę w kontekście niniejszych rozważań jest Narodowe Repozytorium Publikacji Elektronicznych BN, w którym ma być składowany i archiwizowany polski zasób cyfrowy. Drugą tego typu inicjatywą jest Krajowy Magazyn Danych, tworzony przez PCSS, głównie z myślą o ochronie zasobów polskich bibliotek cyfrowych.

W celu pogłębienia świadomości oraz propagowania zagadnień związanych z długoterminową archiwizacją polskich zasobów cyfrowych niezbędne jest stworzenie oferty wszelkich form kształcenia i dokształcania, w wyniku których pracownicy polskich bibliotek zyskają wiedzę i umiejętności potrzebne do organizacji i wykonywania czynności archiwizacyjnych. Brak wiedzy, jak organizować i realizować prace archiwizacyjne jest obok

braku środków finansowych najpoważniejszą trudnością w procesach długoterminowej archiwizacji.

Instytucje biblioteczne w Polsce stoją przed koniecznością ustalenia, czy polski biblioteczny zasób cyfrowy będzie archiwizowany w lokalnych archiwach cyfrowych, czy też w jednym centralnym archiwum. W tej kwestii potrzebne jest podjęcie współpracy i stworzenie powszechnie akceptowanego, narodowego programu działań archiwizacyjnych. Bez względu na to, czy będzie to model archiwum rozproszonego, czy centralnego, potrzebne jest powołanie instytucji bądź organizacji, która przyjmie rolę inicjatora i koordynatora działań archiwizacyjnych w Polsce. Jest bardzo prawdopodobne, że właściwym koordynatorem działań archiwizacyjnych w Polsce byłaby BN wraz MKiDN oraz PCSS.

Z przeprowadzonych badań wynika, że w środowisku polskich wydawców dotychczas nie było podejmowane, na szeroką skalę, zagadnienie długoterminowej ochrony publikacji elektronicznych. Byliby oni jednak skłonni zapoznać się z koncepcją instytucji bibliotecznych w tej sprawie i przyłączyć się do działań na rzecz opracowania strategii archiwizacji publikacji elektronicznych.

BIBLIOGRAFIA

- About RLG Programs* (b.d.). [online]; RLG Mountain View, CA USA 2006 [dostęp: 12.02.2010]. Dostępny w World Wide Web: <<http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070511/viewer/file1939.html>>.
- Bibliografia Dokumentów Elektronicznych* (2008). *Indeks wydawców Bazy BN* [online]. Biblioteka Narodowa [dostęp: 25.01.2008]. Dostępny w World Wide Web: <<http://mak.bn.org.pl/cgi-bin/makwww.exe?BM=27&IZ=WYDAWCA>>.
- Biblioteka cyfrowa (2008). [online]. W: Wikipedia [dostęp: 25.01.2008]. Dostępny w World Wide Web: <http://pl.wikipedia.org/wiki/Biblioteka_cyfrowa>.
- Biblioteka Główna Uniwersytetu Marii Curie Skłodowskiej (b.d.). [online]. Biblioteka Główna UMCS, Lublin 2008. [dostęp: 02.03.2008]. Dostępny w World Wide Web: <<http://www.bg.umcs.lublin.pl/>>.
- Biblioteka Uniwersytecka w Warszawie [online]. Biblioteka Uniwersytecka w Warszawie, Warszawa 2008. [dostęp: 02.03.2008]. Dostępny w World Wide Web: <http://www.buw.uw.edu.pl>.
- Biblioteka Uniwersytetu Śląskiego [online]. Biblioteka Uniwersytetu Śląskiego, Katowice 2008. [dostęp: 02.03.2008]. Dostępny w World Wide Web: <http://www.bg.us.edu.pl/>.
- CRL (2010). Center for Research Libraries [online]; [dostęp: 11.02.2010]. Dostępny w World Wide Web: <<http://www.crl.edu/>>.
- Czermiński, Jurand B.(2002). *Cyfrowe środowisko współczesnej biblioteki*. Gdańsk: Wydaw. Uniwersytetu Gdańskiego.
- DPC(2010). Digital Preservation Coalition, Heslington, York [online]; [dostęp: 11.02.2010]. Dostępny w World Wide Web: <<http://www.dpconline.org/graphics/index.html>>.
- FBC (2008). Federacja Bibliotek Cyfrowych [online]. Poznańskie Centrum Superkomputerowo-Sieciowe [dostęp: 2.03.2008]. Dostępny w World Wide Web: <<http://fbc.pionier.net.pl/owoc>>.
- Hedstrom, Margaret; Montgomery, Sheon (1999). *Digital Preservation Needs and Requirements in RLG Member Institution*. December 1998. RLG. Research Libraries Group, Mountain View, CA [online]. OCLC Research [dostęp: 08.02.2010]. Dostępny w World Wide Web: <www.oclc.org/programs/ourwork/past/digpresneeds/digpres.pdf>.

- Instytut Problemów Ochrony Zdrowia (b.d.). [online]. Instytut Problemów Ochrony Zdrowia [dostęp: 03.03.2008]. Dostępny w World Wide Web: <<http://www.ipoz.pl/>>.
- It's about time: research challenges in digital archiving and long term preservation* (2003). *Final report: Workshop on Research Challenges in Digital Archiving and Long-term Preservation*, April 12-13, 2002. Washington, D. C.: Library of Congress, 2003.
- Januszko-Szakiel, Aneta (2010). Problemy organizacji działań w zakresie długoterminowej archiwizacji zasobów cyfrowych w Polsce. *Przegląd Biblioteczny*, z. 4, s. 405-428.
- Katalog WWW Google (2008). Nauka i edukacja – Nauczanie zdalne [online]; [dostęp: 02.03.2008]. Dostępny w World Wide Web: <http://www.google.com/Top/World/Polski/Nauka_i_edukacja/Nauczanie_zdalne/>.
- Katalogi, kartoteki i bazy danych (2009). Bazy danych Ewidencja oraz Biblios. Biblioteka Narodowa [online]; [dostęp: 10.01.2009]. Dostępny w World Wide Web: <<http://www.bn.org.pl/index.php/zbiory/zbiory-bibliologiczne/katalogi,-kartoteki-i-bazy-danych>>.
- KMD (2009). Krajowy Magazyn Danych [online]. Poznańskie Centrum Superkomputerowo-Sieciowe [dostęp: 11.02.2010]. Dostępny w World Wide Web: <<http://kmd.pcss.pl/index.html>>.
- Nahotko, Marek (2006). Nauczanie zdalne (e-learning) w polskich szkołach wyższych / e-ten euridice [online]. W: Kocójowa, Maria [red.]: *E-włączenie czy e-wyobcowanie?* Kraków: IINiB UJ 2006, s. 62-66. Seria: ePublikacje Instytutu IINiB UJ, nr 2. [dostęp: 02.03.2010]. Dostępny w World Wide Web: <http://www.inib.uj.edu.pl/wyd_inib/s3_z2/nahotko-n.pdf>.
- PCSS (2009). Poznańskie Centrum Superkomputerowo-Sieciowe (PCSS) [online]; [dostęp: 12.02.2010]. Dostępny w World Wide Web: <<http://www.man.poznan.pl>>.
- PCSS – badania i projekty (b.d.). [online]. Poznańskie Centrum Superkomputerowo-Sieciowe [dostęp: 12.06.2009]. Dostępny w World Wide Web: <<http://www.man.poznan.pl/badania/index.html>>.
- PIK (2010). Polska Izba Książki [online]; [dostęp: 20.02.2010]. Dostępny w World Wide Web: <<http://www.pik.org.pl/>>.
- PTWK (2010). Polskie Towarzystwo Wydawców Książek [online]. WYDAWCA. Portal Rynku Wydawniczego [dostęp: 20.02.2010]. Dostępny w World Wide Web: <<http://www.wydawca.com.pl/index.php?s=info&kat=8&dzial=23&poddzial=0>>.
- Preserving Digital Information* (1996). *Report of the Task Force on Archiving of Digital Information* [online]. RLG Mountain View, CA USA [dostęp: 19.02.2010]. Dostępny w World Wide Web: <<http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>>.
- Reports on Digital Archives and Repositories* (2010). [online]. Center for Research Libraries, Chicago, IL [dostęp: 14.02.2010]. Dostępny w World Wide Web: <<http://www.crl.edu/archiving-preservation/digital-archives/digital-archives-reports>>.
- RLG Partnership (2009). [online]. OCLC Dublin, Ohio USA [dostęp: 12.02.2010]. Dostępny w World Wide Web: <<http://www.oclc.org/programs/default.htm>>.
- MKiS (1998a). *Rozporządzenie Ministra Kultury i Sztuki z dnia 19 marca 1998 r. w sprawie zasad i trybu zaliczania bibliotek do bibliotek naukowych oraz ustalania ich wykazu*. 1998, nr 44, poz. 269.
- MKiS (1998b). *Rozporządzenie Ministra Kultury i Sztuki z dnia 24 listopada 1998 r. w sprawie ustalenia wykazu bibliotek, których zbiory tworzą narodowy zasób biblioteczny, określenia organizacji tego zasobu oraz zasad i zakresu jego szczególnej ochrony*. Dz. U. Nr 146, poz. 955.
- SEA (b.d.). [online]. Stowarzyszenie E-learningu Akademickiego [dostęp: 03.03.2010]. Dostępny w World Wide Web: <<http://www.sea.edu.pl/>>.
- Studia online* (b.d.). Onet.pl. Katalog [online]; [dostęp: 02.03.2008]. Dostępny w World Wide Web: <<http://katalog.onet.pl/6173,studiaonline,k.html>>.
- The Digital Preservation Coalition* (2010). [online]. Digital Preservation Coalition, Hellington, York [dostęp: 11.02.2010]. Dostępny w World Wide Web: <<http://www.dp-conline.org/>>.

- The Directory of Open Access Repositories – OpenDOAR* (2010). [online]. University of Nottingham [dostęp: 02.03.2010]. Dostępny w World Wide Web: <<http://www.opendoar.org/>>.
- Ustawa z dnia 27 czerwca 1997 r. o bibliotekach* (1997). Dz. U. Nr 85, poz. 539.
- Waters, Donald; Garrett, John (1996). *Preserving Digital Information. Report of the Task Force on Archiving of Digital Information*. Washington D. C.: The Commission on Preservation and Access, The Research and Library Group, 1996 [online]. OCLC, Online Computer Library Center [dostęp: 01.09.2009]. Dostępny w World Wide Web: <<http://www.oclc.org/programs/ourwork/past/digpresstudy/final-report.pdf>>.
- Wolters Kluwer Polska (2010). Informacje o wydawnictwie [online]; [dostęp: 02.03.2010]. Dostępny w World Wide Web: <http://www.wolterskluwer.pl/?cmd=info_o_wydawnictwie>.
- Wybrane kolekcje cyfrowe (2008). [online]. *EBIB. Elektroniczna Biblioteka – platforma cyfrowa Stowarzyszenia Bibliotekarzy Polskich* EBIB [dostęp: 02.03.2008]. Dostępny w World Wide Web: <<http://www.ebib.info/content/view/275/79/>>.
- Young Digital Planet* (2010). *O firmie* [online]. Young Digital Planet [dostęp: 03.03.2010]. Dostępny w World Wide Web: <<http://www2.ydp.com.pl/ydp/multimedia/1064,O-firmie.html>>.
- Zasoby sieciowe (2006). [online]. *Instytut Informacji Naukowej i Bibliotekoznawstwa Akademii Pedagogicznej im. Komisji Edukacji Narodowej w Krakowie* [dostęp: 02.03.2008]. Dostępny w World Wide Web: <<http://www.wsp.krakow.pl/kbin/zasoby.html>>.

ANETA JANUSZKO-SZAKIEL

Andrzej Frycz Modrzewski College in Cracow
e-mail: ajanuszko-szakiel@afm.edu.pl

CURRENT STATE OF ELECTRONIC RESOURCES
AND THEIR PRESERVATION IN POLISH LIBRARIES
AND PUBLISHING HOUSES – AN ANALYSIS

KEYWORDS: Electronic publications. Long-term preservation of electronic publications. Polish libraries. National Library of Poland. Jagiellonian Library. Poznań Supercomputing and Networking Center. Ministry of Culture and National Heritage. Polish publishing houses for electronic publications. Electronic Publications Division.

ABSTRACT: The author discusses the results of 2009 research on electronic resources of Polish libraries. The research was focused on defining the state of electronic resources collected in traditional and digital libraries and evaluating the awareness and attitudes of Polish electronic resources library managers towards the issue of long-term preservation of resources in question. Finally, the author analyzes conclusions reached as a result of conversations with Poznań Supercomputing and Networking Center employees on long-term preservation of electronic library resources.

Artykuł wpłynął do Redakcji 5 lipca 2010 r.