

Open Archival Information System OAIS - standard w zakresie archiwizacji publikacji elektronicznych.

Streszczenie: Podstawą projektowanych i implementowanych systemów depozytowych dla publikacji elektronicznych jest obecnie referencyjny model organizacji i funkcjonowania archiwów elektronicznych OAIS, zaakceptowany przez *International Organization for Standardization* (ISO) jako norma postępowania w zakresie długoterminowej archiwizacji dokumentów elektronicznych. OAIS jest utożsamiany z organizacją, złożoną z ludzi i stosownej infrastruktury, której działania koncentrują się na długoterminowym przechowywaniu, zabezpieczaniu i udostępnianiu publikacji elektronicznych. Prezentacja struktury informacji OAIS, modelu otoczenia archiwum oraz schematu funkcjonowania archiwum dostarczą nazewnictwa i opisu kluczowych komponentów archiwum elektronicznego, a także zachodzących w nim procesów. Przez fakt uznania OAIS za standard ISO, zarówno model referencyjny, jak i zaproponowane w nim nazewnictwo zostały powszechnie przyjęte i są stosowane w międzynarodowej dyskusji fachowców, służącej wymianie doświadczeń w zakresie archiwizacji publikacji elektronicznych. Model referencyjny OAIS stał się również bazą dla opracowania systemu depozytowego publikacji elektronicznych *Deposit System for Electronic Publications* DSEP. System ten umożliwia realizację zadań związanych z długoterminową archiwizacją publikacji elektronicznych w ścisłej współpracy z funkcjonującym systemem bibliotecznym, archiwalnym lub muzealnym. Ma on stanowić jego integralną część i wykorzystywać jego infrastrukturę. Jednak OAIS i DSEP dostarczają tylko teoretycznego opisu wymagań stawianych organizacji i funkcjonowaniu bibliotecznych, archiwalnych oraz muzealnych systemów depozytowych dla dokumentów elektronicznych. Opracowania wymagają wciąż struktury dla implementacji tych modeli w praktyce.

Wprowadzenie

W piśmiennictwie dotyczącym problematyki gromadzenia, archiwizacji oraz udostępniania publikacji elektronicznych przytaczane są często nazwy: *Open Archives Initiative* OAI oraz *Open Archival Information System* OAIS. Ze względu na podobieństwo akronimów, zachodzi potrzeba wyjaśnienia istotnej różnicy między oznaczanymi przez nie przedsięwzięciami.

Open Archives Initiative OAI¹ jest projektem w ramach którego prowadzone są prace nad rozwojem i upowszechnianiem standardu wymiany danych pomiędzy archiwami preprintów elektronicznych. Celem OAI jest stworzenie zunifikowanej struktury organizacyjnej i technicznej, ułatwiającej dostęp do przechowywanych w archiwach preprintów poprzez ich metadane. Pod pojęciem „archiwum” należy w tym przypadku rozumieć serwer, na którym preprinty elektroniczne są gromadzone i przechowywane, a także na bieżąco udostępniane. Archiwa takie określane są również jako repozytoria preprintów.

Zawarte w nazwie inicjatywy OAI słowo *open* odnosi się do „otwartości” w sensie architektonicznym. Założeniem inicjatorów OAI jest zdefiniowanie i rozpowszechnienie mechanizmów, umożliwiających dostęp do treści preprintów elektronicznych, przechowywanych w repozytoriach instytucji naukowych na całym świecie. W ten sposób ma zostać zapewniony dostęp do niepublikowanej informacji naukowej oraz jej sprawna wymiana w skali globalnej. Słowa *open* nie należy kojarzyć zatem z wolnym - w sensie nieograniczonym prawami autorskimi - dostępem do treści dokumentów².

Podsumowując - w ramach inicjatyw OAI organizowane są archiwa elektroniczne, nazywane repozytoriami preprintów, służące zapewnieniu bieżącego, licencjonowanego dostępu do dokumentów elektronicznych, które z różnych powodów nie zostały jeszcze opublikowane, a które zawierają w swej treści informacje istotne dla postępu nauki.

Natomiast w przypadku *Open Archival Information System* OAIS rzecz dotyczy referencyjnego modelu organizacji i przebiegu procesu długoterminowej archiwizacji obiektów cyfrowych. W kluczowym dokumencie opisującym model referencyjny OAIS jest on definiowany jako organizacja, składająca się z ludzi i stosownej infrastruktury, której starania skoncentrowane są na długoterminowym przechowywaniu, zabezpieczaniu i udostępnianiu obiektów elektronicznych wyznaczonej grupie użytkowników (Designated Community). Słowo „open” oznacza, że proponowany standard organizacji archiwów elektronicznych został opracowany na forum otwartym, przy współudziale wielu

¹ <http://www.openarchives.org/> (dostęp: 06.04.2005r.)

² Szczegółowych informacji na temat inicjatywy *Open Archives Initiative* OAI oraz serwisów e-printowych dostarczają następujące publikacje: C. Lagoze, H. Van de Sompel, *The Open Archives Initiative: Building a low-barrier interoperability framework*. Dokument dostępny online pod adresem: <http://www.openarchives.org/documents/jcdl2001-OAI.pdf> (dostęp: 06.04.2005r.); A. Kamiński, *E-printowa rewolucja. Elektroniczny Biuletyn Informacyjny Bibliotekarzy EBIB* 4/2002 <http://ebib.oss.wroc.pl/2002/33/kaminski.php> (dostęp: 06.04.2005r.); M. Nahotko, *Licencjonowane udostępnianie wartości intelektualnych w Internecie. Elektroniczny Biuletyn Informacyjny Bibliotekarzy EBIB* 6/2003 <http://ebib.oss.wroc.pl/2003/46/nahotko.php> (dostęp: 06.04.2005r.)

specjalistów; nie implikuje ono otwartości, w sensie nieograniczonego dostępu do archiwum oraz do treści przechowywanych w nim dokumentów elektronicznych³.

Model referencyjny OAIS został stworzony przez *Consultative Committee for Space Data Systems* (CCSDS)⁴ na potrzeby archiwizacji i wymiany danych elektronicznych, zawierających informacje z badań przestrzeni kosmicznej. W maju 1999 roku zaprezentowana została pierwsza wersja modelu OAIS, a w lutym 2003 roku, po licznych poprawkach model OAIS został zaakceptowany przez *International Organization for Standardization* ISO jako norma postępowania w zakresie długoterminowej archiwizacji danych cyfrowych (ISO: nr 14721:2003). Pomimo, że model OAIS został stworzony głównie z myślą o archiwizacji jednego typu danych elektronicznych, jest on uznawany za uniwersalny model organizowania i funkcjonowania archiwów elektronicznych i stosowany do gromadzenia, przechowywania i udostępniania różnych typów publikacji elektronicznych⁵. OAIS jest wykorzystywany w wielu światowych bibliotekach, archiwach i muzeach, w których realizowane są projekty długoterminowej archiwizacji zbiorów elektronicznych.

Przedmiotem dalszych rozważań o modelu referencyjnym OAIS są struktura informacji deponowanych w OAIS, otoczenie archiwum OAIS oraz schemat funkcjonowania archiwum OAIS.

Struktura informacji OAIS

W archiwach zgodnych z modelem OAIS istotne jest rozróżnienie pomiędzy danymi cyfrowymi (*Data Object*) a obiektami informacyjnymi (*Information Object*). Na obiekt informacyjny oprócz danych cyfrowych składają się także narzędzia konieczne do przetworzenia surowych danych do postaci zrozumiałej przez użytkownika. Narzędzia te muszą pozwalać na przedstawienie treści w postaci dostosowanej do poziomu wiedzy użytkownika. W modelu OAIS wiedza ta określona jest terminem *Knowledge Base*, a narzędzia używane do przetwarzania danych nazywane są *Representation Information*⁶.

³ Consultative Committee for Space Data System, *Reference Model for an Open Archival Information System* OAIS, January 2002. Dokument dostępny w wersji online pod adresem:

<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf> ;
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html (dostęp: 06.04.2005r.)

⁴ Komitet CCSDS został powołany w 1982 roku. Jest organizacją składającą się z przedstawicieli wielu światowych agencji badań przestrzeni kosmicznej i podlega bezpośrednio agencji NASA; <http://www.ccsds.org/> (dostęp: 06.04.2005r.)

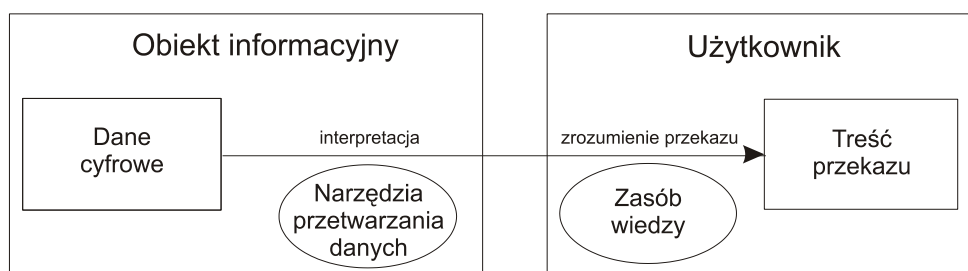
⁵ *Trusted Digital Repositories: Attributes and Responsibilities*. An RLG-OCLC Report. May 2002. Dokument dostępny w wersji online pod adresem: <http://www.rlg.org/longterm/repositories.pdf> (dostęp: 12.04.2005r.)

⁶ *Reference Model...*, op.cit.

Jako przykład, autorzy modelu opisują najprostszą sytuację, w której korzystanie z tekstów angielskojęzycznych uzależnione jest od znajomości języka angielskiego. Jeśli odbiorca tekstu nie posiada w swoich „zasobach” wiedzy (*Knowledge Base*) posługiwania się językiem angielskim, to oryginalny tekst angielski wymaga tłumaczenia przy pomocy słownika i podręcznika gramatyki, które w tym przypadku stanowią narzędzia używane do przetwarzania danych (*Representation Information*). Na tej samej zasadzie odbywa się korzystanie z publikacji elektronicznych. Dane zapisane w postaci kodu zerojedynkowego (*Data Object*), aby były użyteczne, muszą zostać przedstawione w formie czytelnej dla użytkownika, za pomocą odpowiednich narzędzi w postaci sprzętu oraz oprogramowania (*Representation Information*). Otrzymana w ten sposób treść przekazu powinna być zrozumiała dla użytkownika przy założeniu określonego zasobu jego umiejętności (*Knowledge Base*).

Przyjęcie w nazwie *Information Object* członu *Information* wynika zapewne z faktu, że – jak już wspomniano – model ten został stworzony dla celów archiwizacji i wymiany danych cyfrowych, zawierających w swej treści określone informacje (w pierwotnym zastosowaniu pochodzące z badań przestrzeni kosmicznej). Jednak z tej racji, że OAIS został zaakceptowany jako standard archiwizacji wszelkich dokumentów cyfrowych – bez względu na ich zawartość informacyjną – można przyjąć, że każdy dokument elektroniczny, należący do zbiorów bibliotecznych, archiwalnych, a także muzealnych, wraz z narzędziami umożliwiającymi jego odczyt będzie określany mianem *Information Object*.

Proces przetwarzania danych cyfrowych do postaci treści zrozumiałej dla użytkownika przedstawia schemat 1.



Schemat 1: Proces interpretacji i prezentacji treści danych cyfrowych. Na podstawie modelu referencyjnego OAIS⁷.

W referencyjnym modelu OAIS przyjęto zasadę, że dla prawidłowego przebiegu procesu archiwizacji istotne jest dokładne zidentyfikowanie obiektu cyfrowego oraz

⁷ Ibidem.

powiązanych z nim narzędzi. Bez odpowiedniego sprzętu i oprogramowania interpretacja przechowywanych danych cyfrowych nie będzie w przyszłości możliwa. Dlatego też w archiwach działających na podstawie modelu OAIS, oprócz danych cyfrowych (*Data Object*) przedmiotem archiwizacji są także narzędzia (*Representation Information*), umożliwiające ich odczytanie i udostępnienie użytkownikowi w postaci dla niego zrozumiałej.

Kolejnym kluczowym pojęciem w modelu referencyjnym OAIS jest pakiet informacyjny (*Information Package*)⁸. Składa się on z dwóch komponentów, tj: przechowywanej informacji (*Content Information*) oraz opisu przechowywania (*Preservation Description Information - PDI*). *Content Information* czyli przechowywana informacja zawiera dane cyfrowe wraz z narzędziami ich odczytu i prezentacji, natomiast *PDI* to w myśl modelu OAIS wszelkie informacje konieczne w procesie przechowywania. Zalicza się tu cztery typy informacji, określane jako: proveniencja (*Provenance*), kontekst (*Context*), identyfikatory (*Reference*) oraz mechanizmy ochrony autentyczności danych (*Fixity*)⁹.

- *Provenance*, czyli informacja o *pochodzeniu*, określa źródło obiektu informacyjnego, wskazuje na podmiot odpowiedzialny za opiekę nad obiektem od momentu jego powstania oraz dostarcza wiedzę na temat historii obiektu
- *Context*, opisuje związek obiektu informacyjnego z innymi obiektami, które nie należą do danego pakietu informacyjnego
- *Reference*, to informacje zapewniające jednoznaczną identyfikację obiektu informacyjnego. Najogólniej rzecz ujmując ich zadaniem jest dostarczenie identyfikatora publikacji elektronicznej, który odróżnia określoną publikację od innych. W archiwach elektronicznych identyfikatory występują pod nazwą *Digital Object Identifier* (DOI) czy też *Persistent Identifier* (PI)¹⁰. Identyfikatory obiektów elektronicznych są odpowiednikami numerów ISBN oraz ISSN
- *Fixity* to element wprowadzający mechanizmy ochronne, mające na celu zabezpieczenie autentyczności i integralności obiektów informacyjnych przed jakimikolwiek nieudokumentowanymi zmianami.

PDI jest więc zarówno pewnego rodzaju informatorem o pochodzeniu i historii obiektu informacyjnego, jego przynależności oraz powiązaniach z innymi obiektami w archiwum, jak i mechanizmem chroniącym jego integralność i autentyczność.

⁸ Ibidem.

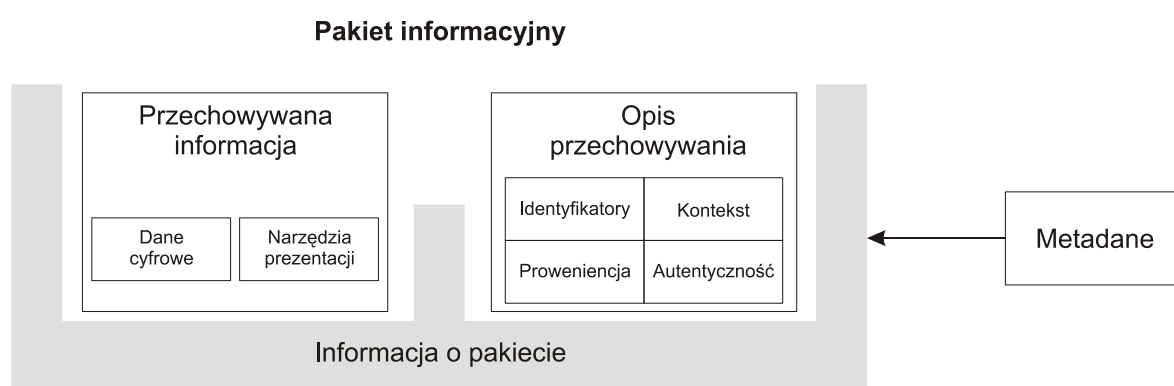
⁹ Ibidem.

¹⁰ *Persistent Identifier. Eindeutige Bezeichner für digitale Inhalte* [online]; Persistent Identifier: Eindeutige Bezeichner für digitale Inhalte. Der Service der Deutschen Nationalbibliothek. Frankfurt am Main 2008. [Dostęp: 16 lutego 2010]. Dostępny w World Wide Web: < <http://www.persistent-identifier.de/> >.

W celu powiązania obu komponentów pakietu informacyjnego, model referencyjny OAIS przewiduje także element w postaci informacji o pakiecie (*Packaging Information*). Jego zadaniem jest identyfikacja poszczególnych składników pakietu informacyjnego.

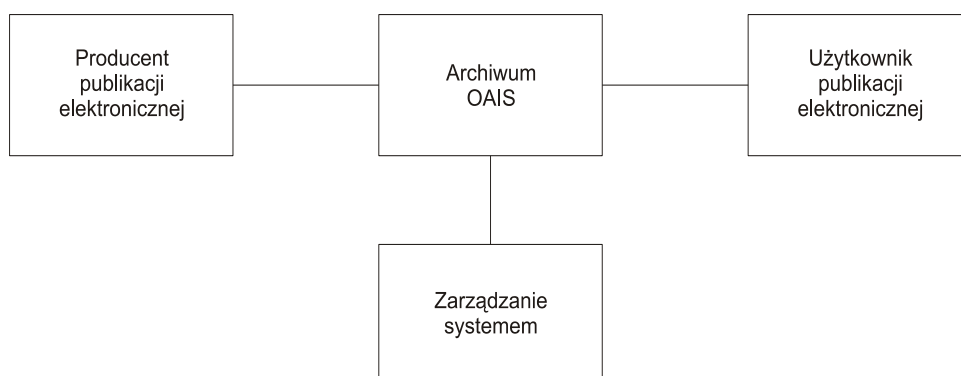
Elementem niezbędnym w archiwum elektronicznym są wreszcie metadane przechowywanych obiektów (*Information Packages*). W modelu referencyjnym OAIS określane są one terminem *Descriptive Information*. Metadane dostarczają informacje o zawartości pakietu informacyjnego oraz umożliwiają jego odnalezienie w archiwum.

Pakiet informacyjny wraz ze wszystkimi jego elementami składowymi należy traktować jako obiekt archiwizacji w archiwum elektronicznym OAIS (schemat 2).



Schemat 2: Pakiet Informacyjny (Information Package) jako obiekt archiwizacji w archiwum elektronicznym OAIS. Na podstawie modelu referencyjnego OAIS¹¹.

Otoczenie archiwum w modelu referencyjnym OAIS



Schemat 3: Otoczenie archiwum OAIS. Na podstawie modelu referencyjnego OAIS¹².

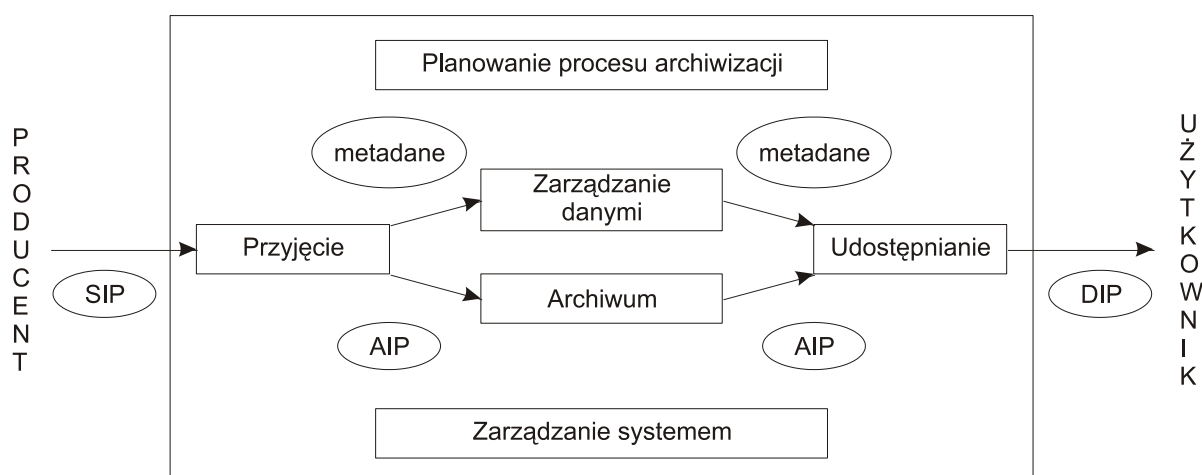
¹¹ *Reference Model...*, op.cit.

¹² Ibidem.

Widniejący na schemacie 3 model referencyjny OAIS uwzględnia elementy otoczenia archiwum. W punkcie centralnym modelu znajduje się archiwum ze swoimi zadaniami gromadzenia obiektów archiwizacji, ich długoterminowego zabezpieczania oraz udostępniania zainteresowanym kręgom użytkowników. Z archiwum ściśle powiązane są elementy określone jako „Producent”, czyli producenci publikacji elektronicznych (autorzy oraz wszelkie instytucje, w których one powstają), „Użytkownik”, czyli użytkownicy, z myślą o których organizuje się procesy archiwizacji publikacji elektronicznych, i wreszcie „Zarządzanie”, a więc jednostka zarządzająca, której działania koncentrują się wokół organizacji archiwum OAIS, organizacji procesów archiwizacji oraz czuwania nad ich prawidłowym przebiegiem.

Schemat funkcjonowania archiwum OAIS

Autorzy modelu OAIS dokonują rozróżnienia pomiędzy pakietami informacyjnymi, które są przechowywane w archiwum OAIS, następnie pakietami, które są dostarczane do archiwum przez producenta (twórcę) oraz pakietami wysyłanymi z archiwum w celu ich udostępnienia użytkownikowi. Pakiet informacyjny przesyłany od producenta do archiwum nazwany został w modelu OAIS zgłoszeniowym pakietem informacyjnym (*Submission Information Package-SIP*). Jego forma oraz zawartość są wcześniej dokładnie ustalone pomiędzy OAIS oraz producentem. Pakiet, który jest przechowywany w archiwum oznaczono jako archiwizowany pakiet informacyjny (*Archive Information Package - AIP*). Natomiast dla pakietu udostępnianego użytkownikowi przyjęto określenie udostępnianego pakietu informacyjnego (*Dissemination Information Package - DIP*). Na zamówienie OAIS przekazuje użytkownikowi całość lub odpowiednią część pakietu informacyjnego AIP w formie pakietu DIP.



Schemat 4: Funkcjonowanie archiwum OAIS. Na podstawie modelu referencyjnego OAIS¹³.

Zaprezentowany schemat 4 obrazuje istotę funkcjonowania archiwum publikacji elektronicznych zorganizowanego na podstawie modelu OAIS. Obejmuje on sześć funkcjonalnych jednostek oraz drogę obiektu archiwizacji od jego producenta do użytkownika. I tak, jednostka „Przyjęcie” odpowiedzialna jest za przyjęcie zgłoszeniowego pakietu informacyjnego SIP od producenta oraz za przygotowanie go do umieszczenia oraz administrowania nim w archiwum. W zakresie jej zadań znajduje się m.in. kontrola kompletności oraz autentyczności zgłoszeniowego pakietu informacyjnego, przekształcenie pakietu SIP w pakiet gotowy do archiwizacji oraz stworzenie do niego metadanych. Następnie archiwizowany pakiet informacyjny AIP przekazywany jest do jednostki zajmującej się archiwizacją tj. do „Archiwum”, a metadane odsyłane do jednostki „Zarządzanie danymi”, odpowiedzialnej za zarządzanie archiwalnymi zasobami.

Kolejną i bardzo istotną jednostką funkcjonalną systemu OAIS jest „Archiwum”, odpowiedzialne za zapis, właściwe przechowywanie pakietów informacyjnych (AIP) oraz możliwość ich odczytu. „Archiwum” odpowiada za długoterminowe przechowywanie i zapewnienie nienaruszalności pakietów AIP, okresowe przenoszenie danych na media nowszej generacji, migrację do aktualnie stosowanych formatów lub systemów, a w przypadku awarii systemu za ich rekonstrukcję. Na żądanie „Archiwum” przekazuje określony pakiet AIP do jednostki „Udostępnianie”.

W archiwum elektronicznym niezbędna jest jednostka „Zarządzanie danymi”. Jej zadaniem jest utrzymywanie i udostępnianie szerokiego wachlarza informacji. Przykładami mogą być katalogi i inwentarze, na podstawie których można uzyskać określone zasoby

¹³ Ibidem.

z archiwum, a także statystyki dotyczące udostępniania zbiorów. Do zadań tej jednostki należy także zaliczyć kontrolę bezpieczeństwa danych oraz inne procedury narzucane przez OAIS.

Poprawne funkcjonowanie całego archiwum jest uzależnione od prac jednostki administrującej procesami w nim zachodzącymi. „Zarządzanie systemem” zajmuje się negocjowaniem warunków z producentami, na których podstawie publikacje są transferowane do archiwum, czuwa nad kontrolą zgodności dostarczonych obiektów ze standardami archiwum oraz przejmuje odpowiedzialność za utrzymywanie sprawności sprzętu i oprogramowania w archiwum. Czyni także starania na rzecz rozwoju oraz nadzoru nad standardami, niezbędnymi dla funkcjonowania archiwum.

Archiwa organizowane oraz działające na podstawie modelu OAIS starają się zapewnić na przyszłość stabilny dostęp do przechowywanych w nich różnorodnych publikacji elektronicznych. W tym celu w ramach OAIS wyodrębniono jednostkę „Planowanie procesu archiwizacji”, która zajmuje się m.in.: obserwowaniem rozwoju rynku sprzętu i oprogramowania, testowaniem nowych rozwiązań, kontrolą, czy archiwizowane obiekty dadzą się uruchomić i odczytać. Jednostka ta jest odpowiedzialna za wszelkie decyzje dotyczące strategii postępowania, m.in. częstotliwości odświeżania danych, działań mających na celu dostosowanie rozwiązań do zmieniających się warunków sprzętowo-programowych (emulacja lub migracja) i udostępnienia treści publikacji w zmienionych warunkach.

Proces udostępniania dokumentów elektronicznych w archiwach OAIS określony został terminem *Access*. W ramach udostępniania zasobów użytkownikom umożliwia się przeglądanie zawartości archiwum poprzez katalogi online, określenie lokalizacji i dostępności określonych zbiorów. Na zamówienie użytkownika system tworzy i wysyła pakiety informacyjne typu DIP.

Na podstawie przytoczonych powyżej schematów i opisów możliwe staje się prześledzenie drogi cyfrowego obiektu przez archiwum budowane zgodnie z założeniami OAIS. Otóż producent, który chce odesłać do archiwum dokument cyfrowy w celu jego długoterminowego przechowania, powinien nadać dokumentowi właściwą – ustaloną wcześniej z archiwum – formę oraz dołączyć wszelkie dodatkowe informacje o dokumencie wraz z metadanymi. Jednak realizowane w zakresie archiwizacji projekty pokazują, że producenci nie zawsze podejmują się generowania metadanych; wówczas jest to dodatkowe zadanie pracowników archiwum. Dokument wraz z metadanymi przesyłany jest w postaci zgłoszeniowego pakietu informacyjnego SIP do działu przyjęcia, gdzie zostaje „rozpakowany” oraz sprawdzony pod względem kompletności i poprawności wszelkich informacji niezbędnych dla przyjęcia i długoterminowego zarchiwizowania dokumentu. W systemie archiwalnym

każdy dokument przyjmuje następnie postać AIP, jest zapisywany na serwerze archiwum i przechowywany w sposób umożliwiający jego długotrwałą, stabilną użyteczność. Wygenerowane metadane archiwizowanych dokumentów są odsyłane do działu, zajmującego się ich zarządzaniem. Archiwizowany obiekt AIP może być przekształcony w formę DIP, udostępnianą na żądanie użytkownikowi, w postaci umożliwiającej jego zrozumienie.

Model referencyjny OAIS wyjaśnia podstawowe terminy oraz identyfikuje kluczowe procesy z zakresu archiwizacji publikacji elektronicznych. Przyczynia się także do wzrostu teoretycznej wiedzy na temat projektowania, budowania i funkcjonowania archiwów elektronicznych, nastawionych na długoterminową ochronę przechowywanych obiektów cyfrowych i ich udostępniania w przyszłości. Model wskazuje również na niebagatelną rolę sprzętu i oprogramowania w utrzymaniu użyteczności cyfrowych obiektów informacyjnych. Uwzględnia potrzebę współpracy z instytucjami niepowiązanymi bezpośrednio z archiwum, jednak współodpowiedzialnymi za efektywność procesów archiwizacji. Dokładnie określa zakres odpowiedzialności za ochronę pakietów informacyjnych, spoczywającej na archiwum. Przede wszystkim model OAIS – przez fakt uznania go za standard ISO – dostarcza zunifikowanego i powszechnie stosowanego nazewnictwa, ułatwiającego międzynarodową dyskusję, wymianę pomysłów oraz doświadczeń w środowiskach zainteresowanych problematyką długoterminowej archiwizacji publikacji elektronicznych.

Autorzy modelu OAIS zapewniają, że może on być stosowany przy organizacji wszelkich elektronicznych archiwów, ze specjalnym przeznaczeniem dla archiwów odpowiedzialnych za długoterminowe przechowywanie dokumentów elektronicznych¹⁴. Przykładami potwierdzającymi tę tezę są realizowane z powodzeniem światowe projekty archiwizacji publikacji elektronicznych, wykorzystujące model referencyjny OAIS. Jedną z największych inicjatyw w tym obszarze przedsięwzięła narodowa biblioteka Holandii Koninklijke Bibliotheek (KB). W kooperacji z firmą IBM, Koninklijke Bibliotheek opracowała i zaimplementowała prototyp archiwum elektronicznego DIAS (Digital Informations and Archiving System), które w założeniu powinno zapewnić dostępność publikacji elektronicznych w przyszłości¹⁵. System DIAS stanowi jądro systemu

¹⁴ Ibidem.

¹⁵ T. van der Werf-Davelaar: Long-term Preservation of Electronic Publications. The Nedlib Project [online]; *D-Lib Magazine*, September 1999, Vol. 5, Iss. 9. [Dostęp: 16 lutego 2010]. Dostępny w World Wide Web: < <http://www.dlib.org/dlib/september99/vanderwerf/09vanderwerf.html> >; A. K. Amse: Zabezpieczanie historycznych zasobów dla przyszłości - archiwum cyfrowe holenderskiej biblioteki narodowej [online]; *Elektroniczny Biuletyn Informacyjny Bibliotekarzy EBIB*, 2003, nr 2(42). [Dostęp: 16 lutego 2010]. Dostępny w World Wide Web: < <http://ebib.oss.wroc.pl/2003/42/amse.php> >.

depozytowego dla cyfrowych zasobów zgromadzonych w Koninklijke Bibliotheek – „eDepot”¹⁶.

DSEP - Deposit System for Electronic Publications

W latach 1998 - 2000 Narodowa Biblioteka Holandii *Koninklijke Bibliotheek* kierowała projektem „Networked European Deposit Library - NEDLIB”, którego uczestnicy - tj. przedstawiciele ośmiu bibliotek narodowych, jednego archiwum, dwóch firm informatycznych oraz trzech wydawnictw naukowych - postawili sobie za cel szukanie rozwiązań, będących w stanie zapewnić długoterminowy dostęp do elektronicznych zbiorów bibliotecznych, archiwalnych oraz muzealnych¹⁷. W rezultacie projektu powstały między innymi założenia systemu depozytowego dla publikacji elektronicznych DSEP *Deposit System for Electronic Publications*, opracowane w oparciu o model referencyjny OAIS (Open Archival Information System)¹⁸. Z uwagi na fakt, że pierwotna wersja modelu OAIS z 1999 roku nie zakładała długoterminowej archiwizacji danych cyfrowych, a jedynie ich przechowywanie w celu bieżącego użytku¹⁹, w czasie realizacji projektu NEDLIB konieczne stało się wprowadzenie jednostki funkcjonalnej, odpowiedzialnej za długoterminowe utrzymanie dostępności i użyteczności zarchiwizowanego materiału cyfrowego. W konsekwencji powstała jednostka Planowanie Procesu Archiwizacji (*Preservation Planning*), wbudowana następnie do wersji modelu OAIS, zaakceptowanej i obowiązującej obecnie jako standard ISO w obszarze archiwizacji publikacji elektronicznych²⁰.

Założeniem projektu NEDLIB było opracowanie takiego systemu, który umożliwiałby realizację zadań związanych z długoterminową archiwizacją publikacji elektronicznych nie w izolacji, lecz w ścisłej współpracy z funkcjonującym już elektronicznym systemem bibliotecznym, archiwalnym lub muzealnym. System depozytowy miałby stanowić dopełnienie procesów gromadzenia, opracowania, przechowywania, utrzymywania w należytym stanie oraz udostępniania zbiorów cyfrowych²¹. Bez względu na to, w jakiej

¹⁶ *The e-Depot system (DIAS)* [online]; Koninklijke Bibliotheek, National Library of the Netherlands, The Hague [b.d.] [Dostęp: 16 lutego 2010]. Dostępny w World Wide Web: < <http://www.kb.nl/dnp/e-depot/operational/background/index-en.html> >.

¹⁷ <http://www.kb.nl/coop/nedlib/index.html> (dostęp: 13.04.2005r.); Titia van Werf-Daveelaar, *Long-term Preservation...*

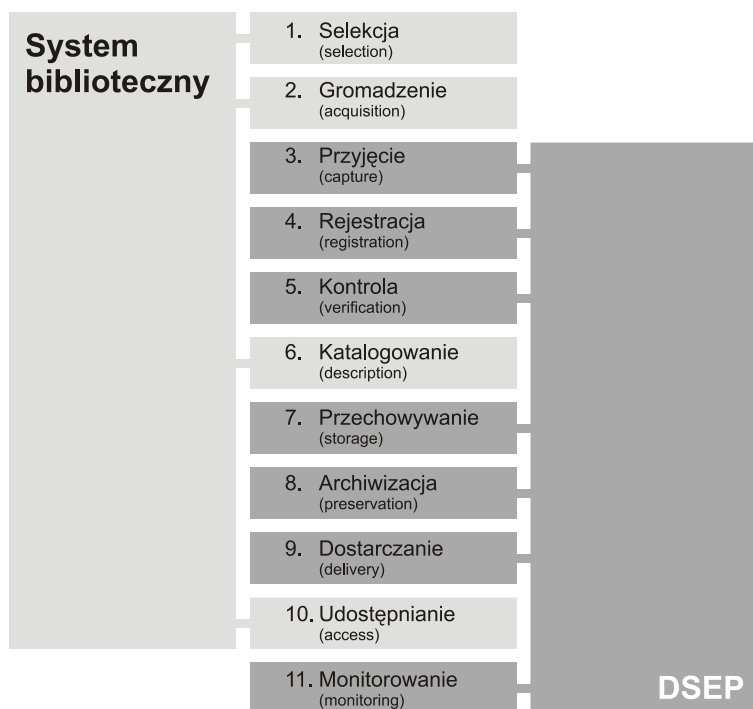
¹⁸ Consultative Committee for Space Data System: *Reference Model ...*

¹⁹ <http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/p2/CCSDS-650.0-R-1.pdf> (dostęp: 10.04.2005r.)

²⁰ Consultative Committee for Space Data System: *Reference Model ...*

²¹ Uwe. M. Borghoff et al., *Langzeitarchivierung. Methoden zur Erhaltung digitaler Dokumente*, Heidelberg: dpunkt.verlag, 2003, s. 32

instytucji archiwizującej miałby zostać zaimplementowany system DSEP, ma on wykorzystywać infrastrukturę funkcjonującego w niej już systemu i stanowić jego integralną część.



Schemat 5. Model systemu depozytowego dla publikacji elektronicznych DSEP²².

Pierwszy, wymieniony w modelu moduł **Selekcja** (*Selection*) reprezentuje wszelkie czynności, związane z wyborem spośród wszystkich publikacji elektronicznych odprowadzanych do biblioteki tych, które powinny znaleźć się w systemie depozytowym, w celu ich długoterminowej archiwizacji. W procesie selekcji decydującą rolę pełnią ludzie opracowujący strategię włączania lub wykluczania publikacji elektronicznej z kolekcji depozytowej. Strategia doboru publikacji jest kwestią indywidualną biblioteki i zwykle bazuje na narodowej polityce gromadzenia. Czynnikiem decydującymi mogą tu być również umowy z wydawcami, dostarczającymi publikacje oraz wytyczne informatyków, sprawujących opiekę techniczną nad archiwum elektronicznym²³. Chodzi tu o określenie parametrów, którymi powinny charakteryzować się publikacje elektroniczne, aby mogły zostać wprowadzone do depozytu i podlegać procesom archiwizacji. Jednym z nich jest likwidacja przez wydawców mechanizmów zabezpieczających publikacje przed jej nielegalnym kopiowaniem, oraz tzw. time-lock'ów, czyli blokad czasowych, regulujących okres

²² <http://www.kb.nl/coop/nedlib/index.html> (dostęp: 13.04.2005r.)

²³ Titia van der Werf, *The Deposit System for Electronic Publications. A Process Model*. Nedlib Report Series 6. Amsterdam: Koninklijke Bibliotheek, 2000, s. 7-8

dostępności publikacji²⁴. Twórcy systemu depozytowego utrzymują, że umożliwia on przechowanie dokumentów elektronicznych bez względu na format, w jakim zostały zapisane oraz nośnik, na którym występują. Ich zdaniem system depozytowy powinien, wg założeń, zachować też zdolność przyjmowania w przyszłości do depozytu publikacji elektronicznych w aktualnie nieznanych formatach²⁵. Świadczyłoby to o pewnej otwartości i elastyczności architektury systemu depozytowego, co w przypadku heterogeniczności publikacji elektronicznych ma niebywale istotne znaczenie.

Ważnym czynnikiem wpływającym na procesy selekcji jest wreszcie kondycja finansowa instytucji archiwizującej²⁶, bowiem z jej budżetu muszą zostać wyasygnowane środki na zaprojektowanie i zbudowanie systemu depozytowego o odpowiedniej pojemności, w dalszej kolejności na jego implementację, testowanie, stałe utrzymywanie oraz w razie potrzeby jego rozbudowę. Za idealną uznać by można sytuację, w której biblioteki narodowe mogłyby pozwolić sobie na wdrożenie i utrzymanie systemu depozytowego, mieszczącego wszystkie, opublikowane w kraju publikacje elektroniczne, tak by chociaż ta jedna instytucja w kraju zachowała pełny dorobek jego nauki i kultury. Pomimo że ani w literaturze przedmiotu, ani w innych istniejących źródłach nie wspomina się o nakładach finansowych, koniecznych dla uruchomienia takiego projektu, na jaki pozwoliła sobie holenderska *Koninklijke Bibliotheek*, pewnym jest, że instytucje biblioteczne, pretendujące do utworzenia i utrzymywania systemu depozytowego dla publikacji elektronicznych muszą liczyć się z poważnymi obciążeniami finansowymi²⁷.

Kolejny moduł systemu to **Gromadzenie** (*Acquisition*). Nazwą tą określa się w praktyce bibliotecznej wszelkie czynności związane z nabywaniem bądź otrzymywaniem zbiorów. W piśmiennictwie przedmiotu zaznacza się, że proces ten w przypadku publikacji elektronicznych nie różni się niczym szczególnym od procesu gromadzenia dokumentów tradycyjnych. Nieco bardziej skomplikowane może okazać się jedynie postępowanie

²⁴ A. K. Amse, *Zabezpieczanie historycznych zasobów dla przyszłości...*

²⁵ <http://www.kb.nl/coop/nedlib/index.html> (dostęp: 13.04.2005r.)

²⁶ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 7-8

²⁷ Z nieoficjalnych informacji, uzyskanych podczas wywiadów z fachowcami projektującymi systemy depozytowe dla elektronicznych kolekcji narodowych wynika, że zbudowanie systemu depozytowego o pojemności 20 TB wiązało się w 2004 roku z kosztami ok. 4 mln Euro. Jeden z raportów Narodowej Biblioteki Holenderskiej podaje przykład systemu o pojemności 1 TB; jest on w stanie pomieścić: 200-2.000 czasopism elektronicznych, zawierających 340 artykułów po 10 stron przez okres 10-ciu lat (w formacie PDF) lub 50-500 książek po 100 stron (w formacie TIFF) lub 2.000-20.000 książek po 100 stron (w formacie PDF); inaczej rzecz ujmując: 7-70 mln stron formatu A4 (PDF) lub 150 tys - 1,5 mln stron A4 (TIFF).

związane z nabywaniem praw użytkowania niektórych form publikacji²⁸. Przykładem są czasopisma elektroniczne, których wydawcy godzą się wprawdzie - w ramach udzielonych licencji - na ich umieszczanie i przechowywanie w archiwum elektronicznym biblioteki, jednak czynią przy tym pewne obostrzenia, dotyczące np. udostępniania treści dokumentów lub podejmowania czynności typowych dla procesów archiwizacji, np. eksperymentów migracji, które czasem okazują się niezbędne dla utrzymania użyteczności publikacji elektronicznych, mogą jednak prowadzić do zmian w dokumentach, naruszając tym samym ich autentyczność.

W ramach procesu gromadzenia publikacji elektronicznych następuje też wymiana informacji bibliograficznych pomiędzy wydawnictwami a depozytem bibliotecznym, dostarczanie przez wydawców metadanych archiwizowanych dokumentów, tudzież wszelkich informacji, na podstawie których metadane mogą zostać wygenerowane²⁹.

Z tej racji, że biblioteczne systemy depozytowe powstają głównie przy bibliotekach narodowych, wszelkie archiwizowane tam publikacje elektroniczne powinny być przekazywane bezpłatnie, w ramach przysługujących im egzemplarzy obowiązkowych, ewentualnie na bazie licencji lub innych rodzajów umów zawieranych pomiędzy biblioteką a wydawcami³⁰. Z doświadczeń Niemieckiej Biblioteki Narodowej, która od kilku lat czyni intensywne przygotowania do podjęcia zadań na rzecz zachowania niemieckiego dziedzictwa kultury i nauki w postaci cyfrowej, wynika jednak, że tylko nieliczni wydawcy są świadomi potrzeby współpracy w tym zakresie. Problem polega na tym, że ustawa o Niemieckiej Bibliotece Narodowej - podobnie zresztą jak i ustawy biblioteczne innych krajów - nie regulują kwestii odprowadzania egzemplarza obowiązkowego wszelkich ukazujących się w tym kraju publikacji elektronicznych; mowa tu głównie o publikacjach sieciowych. Pomimo usilnych starań i ponawianych prób zawierania umów z wydawcami, tylko nieliczni z nich zobowiązują się odprowadzać opublikowane przez nich dokumenty elektroniczne do Biblioteki w celu ich długoterminowej archiwizacji³¹. Nasuwa się zatem wniosek, że niezbędne jest nowelizowanie istniejących ustaw o bibliotekach narodowych, gdyż tylko poprzez nowe regulacje prawne biblioteki narodowe zyskają szansę na realizację zadań, do

²⁸ H. Liegmann, *Langzeitverfügbarkeit digitaler Publikationen*. In: *Bibliotheken - Portale zum Globalen Wissen*. 91. Deutscher Bibliothekartag in Bielefeld. Hrsg. M. Rützel-Banz. Frankfurt am Main: Vittorio Klostermann, 2001, s. 106-109

²⁹ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 7-8

³⁰ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 7-8

³¹ Na podstawie rozmów przeprowadzonych z pracownikami Niemieckiej Biblioteki Narodowej we Frankfurcie nad Menem.

których zostały powołane tj. gromadzenia, przechowywania oraz udostępniania kompletnej narodowej kolekcji dorobku nauki i kultury.

Trzeci moduł systemu, nazwany **Capture** - spotykany też pod nazwą *delivery* lub *harvest* - odpowiedzialny jest za przyjęcie publikacji elektronicznej do depozytu. W praktyce wygląda to tak, że publikacje elektroniczne *offline* dostarczane są do instytucji archiwizującej za pomocą tradycyjnych form transportu, natomiast publikacje *online* są transferowane przez Internet. W szczególnych przypadkach biblioteki cyfrowe dopuszczają dostarczanie publikacji sieciowych na elektronicznych mediach przenośnych, by następnie umieścić je na serwerze i udostępniać w sieci. Sposób dostarczenia publikacji do archiwum jest kwestią indywidualnych umów pomiędzy wydawnictwami a instytucjami archiwizującymi.

W źródłach opisujących system DSEP³² zwraca się uwagę na trzy następujące po sobie kroki postępowania w ramach procesu wprowadzania publikacji elektronicznej do depozytu. Są to:

- *Authentication* - czyli ustalenie źródła pochodzenia przysłanej do biblioteki publikacji elektronicznej. Celem tego zabiegu jest sprawdzenie, czy publikacja pochodzi od wydawcy, umieszczonego w rejestrze wydawców współpracujących z biblioteką,
- *Quality scan* - czyli działania, które mają zapewnić, że wprowadzana do depozytu publikacja jest kompletna i wolna od wirusów,
- *Transfer to deposit* - czyli dostarczenie do depozytu publikacji elektronicznej w postaci pakietu informacyjnego typu *Submission Information Package* SIP.

Biblioteki zwykle nie mają wpływu na to, w jakich formatach zapisu danych publikowane są dokumenty elektroniczne. Przyjmują wszystkie możliwe publikacje, troszcząc się następnie o ich konwersję do formatów, odpowiadających standardom DSEP i zapewniających ich przechowanie na przyszłość³³. Zarówno konwertowanie treści publikacji elektronicznych (w trakcie włączania ich do kolekcji depozytowej) z nietypowych do standardowych formatów zapisu danych jak i - niejednokrotnie konieczne w procesie archiwizacji - eksperymenty migracji do nowszych formatów budzą pewne zastrzeżenia, bowiem wiążą się ze zmianami oryginalnej formy publikacji. To z kolei kłóci się z głównym założeniem archiwizacji, odnośnie do zachowania autentyczności publikacji elektronicznych. W myśl istotnych opracowań przedmiotu, celem procesu długoterminowej archiwizacji jest przechowanie dla przyszłych pokoleń użytkowników kompletnego zbioru autentycznych i

³² <http://www.kb.nl/coop/nedlib/index.html> (dostęp: 13.04.2005r.)

³³ Uwe. M. Borghoff et al., *Langzeitarchivierung. Methoden zur ...*, op.cit., s. 34

integralnych publikacji elektronicznych³⁴, co oznacza, że należy unikać wszelkich działań, w których wyniku pierwotna treść i forma publikacji ulegną zmianom. Dla specjalistów projektujących systemy depozytowe i odpowiedzialnych za utrzymanie ich zawartości to trudne zadanie, któremu jednak trzeba wyjść na przeciw. Niestety nie jest to zresztą jedyne skomplikowane zadanie, z jakim wiąże się archiwizowanie publikacji elektronicznych i dlatego właśnie zostało ono określone mianem wyzwania dla instytucji archiwizujących oraz wydawnictw³⁵.

W ramach czwartego modułu tj. **Rejestracji** (*Registration*) następuje rejestracja pakietu informacyjnego wprowadzanego do depozytu. Rejestracja może obejmować też czynności związane z wysyłaniem potwierdzeń przyjęcia publikacji do depozytu, bądź jej odrzucenia³⁶. Wprowadzona do depozytu publikacja elektroniczna podlega kolejnym procesom kontrolnym, określanym w modelu jako **Kontrola** (*Verification*), których głównym celem jest ustalenie autentyczności oraz integralności publikacji elektronicznych. Weryfikacja przebiega w czterech następujących fazach³⁷:

- *Validate Package* - w której następuje kontrola integralności pakietu informacyjnego. System dokonuje też sprawdzenia, czy pakiet odpowiada ustalonym standardom,
- *Check logical Integrity* - to faza, w której następuje rozpakowanie pakietu informacyjnego, w celu ustalenia obecności wszystkich potrzebnych komponentów, tj. danych, reprezentujących treść publikacji, metadanych oraz narzędzi, niezbędnych do odczytania i interpretacji danych cyfrowych,
- *Establish authenticity* - jest fazą, w której następuje oznakowanie autentyczności kopii publikacji przyjętej do depozytu w celu archiwizacji,
- *Installation and testing* - to faza, w której publikacja wraz z oprogramowaniem wspomagającym jest ładowana do terminala w celu poddania jej testom na właściwe działanie.

Pakiety informacyjne, które pomyślnie przeszły proces weryfikacji są poddawane dalszym procesom w obrębie systemu depozytowego. Jednak specjaliści dopuszczają też sytuację, w której proces weryfikacji kończy się odesłaniem publikacji do wydawcy.

³⁴ *Attributes of a Trusted Digital Repository. Meeting the Needs of Research Resources*. RLG-OCLC Report. Mountain View, CA. August 2001. Dokument dostępny w wersji elektronicznej pod adresem:

<http://www.rlg.org/longterm/attributes01.pdf> (dostęp: 15.04.2005r.).

³⁵ *Die unendliche Bibliothek. Digitale Information in Wissenschaft, Verlag und Bibliothek*. Hrsg. Borsenverein des deutschen Buchhandels. Wiesbaden 1996.

³⁶ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 7-8

³⁷ <http://www.kb.nl/coop/nedlib/index.html> (dostęp: 15.04.2005r.)

Sporządzają wówczas stosowny raport o wykrytych błędach, uzasadniający odrzucenie publikacji przez system, a o zaistniałym fakcie informują dział gromadzenia *Acquisition*³⁸.

Weryfikacja należy - zdaniem fachowców - do jednego z trudniejszych i bardziej skomplikowanych procesów. Wynika to z faktu, że obiekty cyfrowe mogą zostać podrobione lub zmienione w dużo łatwiejszy sposób aniżeli analogowe. Problemów dostarcza też kontrola integralności publikacji sieciowych. Postępująca naprzód tendencja łączenia za pomocą hiperłączy treści różnych publikacji znacznie utrudnia określenie „granic” poszczególnych publikacji. Wciąż bez odpowiedzi pozostaje pytanie, jak w takiej sytuacji zautomatyzować proces kontroli integralności publikacji³⁹. W tym celu niewątpliwie konieczne są mechanizmy, radzące sobie z odróżnieniem zawartości publikacji, tzn. rozpoznających linki wewnętrzne, należące do podlegającego archiwizacji obiektu oraz zewnętrzne, stosowane na zasadzie odsyłaczy do treści innych publikacji sieciowych. Dostarczenie pomysłu i stworzenie takich mechanizmów to kolejna, otwarta wciąż kwestia w temacie archiwizacji publikacji elektronicznych.

Moduł **Katalogowanie** (*Description*) reprezentuje etap prac bibliotecznych, związanych z katalogowaniem zbiorów. Celem procesu katalogowania jest wprowadzenie opisów bibliograficznych publikacji elektronicznych wraz z metadanymi do systemu katalogowego zbiorów bibliotecznych (OPAC) tak, by mogły one być przez ten system wyszukiwane. Katalogowanie odbywa się zwykle w oparciu o narodowe normy katalogowania. Zakłada się, że opisy bibliograficzne publikacji elektronicznych archiwizowanych długoterminowo w depozytach, będą ujmowane w bibliografii narodowej⁴⁰. W procesie katalogowania generowane są też „techniczne” metadane, które stanowią nową kategorię danych o publikacji i są niezbędne dla właściwego przebiegu prac w zakresie długoterminowego utrzymania użyteczności publikacji elektronicznych⁴¹.

Po skatalogowaniu publikacji elektronicznych, następuje ich przyjęcie do depozytowego systemu przechowywania danych, reprezentowanego w modelu przez moduł **Przechowywanie** (*Storage*). W założeniu, depozytowy system przechowywania publikacji elektronicznych ma zapewniać regularną kontrolę możliwości ich odczytu i interpretacji; w przypadku zagrożenia utraty danych, powinny zostać podjęte wszelkie działania na rzecz ich utrzymania, natomiast w przypadku utraty danych, poczynione próby ich odzyskania.

³⁸ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 7-8

³⁹ H. Liegmann, *Langzeitverfügbarkeit digitaler Publikationen...* op.cit., s. 106-109

⁴⁰ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 7-8

⁴¹ H. Liegmann, *Langzeitverfügbarkeit digitaler Publikationen...* op.cit., s. 106-109

Głównym zadaniem **Storage** jest zapewnienie odczytu kodu zerojedynkowego w jego oryginalnej formie, co odbywa się za pomocą:

- *Periodic Refreshment* - czyli regularnego odświeżania nośnika zapisu danych cyfrowych,

oraz

- *Continuous backup* - tj. okresowego sporządzania kopii zapasowych danych⁴².

Jednocześnie zwraca się uwagę, że dla utrzymania kodu zerojedynkowego istotne jest zagwarantowanie bezpieczeństwa miejsca, w którym przechowywane są nośniki z jego zapisem. Konieczne są ustalenia, regulujące kwestie dostępu do archiwum, przeprowadzania prac na dokumentach elektronicznych, także dotyczące zabezpieczenia archiwum przed ewentualnymi katastrofami⁴³. Z rozmów ze specjalistami opracowującymi strategię postępowania z bibliotecznymi dokumentami elektronicznymi wynika, że jednym z elementów tej strategii ma być tworzenie systemów depozytowych o pojemności, która umożliwiałaby przechowanie, oprócz własnej, narodowej kolekcji elektronicznej wraz z jej kopią zapasową, także przechowanie kopii zapasowej kolekcji instytucji partnerskiej. Jest to sposób na zabezpieczenie kopii zapasowej archiwum elektronicznego w miejscu terytorialnie oddalonym od archiwum głównego⁴⁴. Cel ten jednak może zostać osiągnięty tylko i wyłącznie przy ścisłej współpracy bibliotek na rzecz wypracowania zunifikowanych - w skali międzynarodowej - metod archiwizacji publikacji elektronicznych, w szczególności przyjęcia i stosowania jednolitych parametrów dla budowanych systemów depozytowych.

Kolejny moduł - **Archiwizacja (Preservation)** - odzwierciedla wszelkie czynności, wymagane dla długoterminowej ochrony kolekcji depozytowej. H. Liegmann zauważa⁴⁵, że modułu tego nie można utożsamiać z jasno zdefiniowanym obszarem określonych zadań. Oznacza to, że nie ma sprecyzowanego planu działania, tudzież konkretnej recepty, stosowanej w celu zarchiwizowania publikacji elektronicznych. Ich długoterminowa ochrona - jego zdaniem - wymaga użycia różnych - w zależności od konkretnej sytuacji - metod archiwizacji. Autor zwraca jednocześnie uwagę, że warunkiem koniecznym dla przedsięwzięcia jakiegokolwiek strategii archiwizacji jest obserwowanie zmian zachodzących w technologii, co zostało określone terminem „Technology Watch”. W konsekwencji

⁴² <http://www.kb.nl/coop/nedlib/index.html> (dostęp: 15.04.2005r.)

⁴³ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 7-8

⁴⁴ Na podstawie rozmów przeprowadzonych z pracownikami Niemieckiej Biblioteki Narodowej we Frankfurcie nad Menem.

⁴⁵ H. Liegmann, *Langzeitverfügbarkeit digitaler Publikationen...* op.cit., s. 106-109

obserwacji postępu technologicznego podejmowane są - zależnie od potrzeby - następujące działania na publikacjach elektronicznych: zmiana formatu dokumentu, emulacja jego systemowego otoczenia, odświeżenie medium, zmiana rodzaju medium, aktualizacja „technicznych” metadanych, a także kontrola jego integralności oraz autentyczności. Na uwagę zasługuje fakt, że niektóre z wymienionych powyżej działań mogą prowadzić do zmian w formie i treści publikacji elektronicznych, tym samym do utraty ich autentyczności. Zachodzi zatem konieczność dokładnego dokumentowania wszelkich prac podejmowanych w procesie archiwizacji oraz ich efektów, bowiem tylko poprzez dokładny opis każdej przeprowadzanej na obiekcie operacji, z uwzględnieniem porównania wersji oryginalnej z tą, którą uzyskujemy w rezultacie przeprowadzanych operacji, możliwe będzie w przyszłości odtworzenie jego historii, śledzenie wszelkich zmian, porównanie wersji oryginalnej z zachowaną.

W raportach NEDLIB⁴⁶ procesowi *Preservation* poświęca się szczególnie wiele uwagi. Zaznacza się w nich, że spełnia on kluczową rolę w systemie depozytowym i pieczołowicie opisuje jego części składowe. I tak na proces *Preservation* składają się tzw. sub-procesy, tj. *Preservation Planning* oraz *Preservation Activities*. *Preservation Planning* to proces odpowiedzialny za identyfikację wszelkich problemów i zadań, które wiążą się z długoterminową archiwizacją publikacji elektronicznych oraz za dostarczanie pomysłów i narzędzi do ich rozwiązywania. W jego ramach prowadzi się stosowne badania i opracowuje strategię ochrony danych cyfrowych w następujących obszarach:

- *Development Preservation Standards & Strategies* - czyli w zakresie rozwoju standardów i strategii umożliwiających archiwum aktywne uczestnictwo w procesach publikowania elektronicznego, stosowne reagowanie na rozwój technologiczny i dostosowywanie polityki archiwizacji do zmieniających się form publikacji elektronicznych,
- *Development Packaging Designs* - czyli w obszarze projektowania i dopasowywania do zmieniających się okoliczności, standardów i formatów pakietów informacyjnych typu SIP, AIP oraz DIP,
- *Reference Platforms Defining* - a więc w dziedzinie identyfikacji platformy programowo-sprzętowej, koniecznej do uruchomienia i udostępnienia publikacji elektronicznej. W procesie tym następuje przyporządkowanie

⁴⁶ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 52-58

konkretnego sprzętu i oprogramowania do każdego typu publikacji elektronicznych, przechowywanych w archiwum,

- *Technology Monitoring* - znanym także jako *Technology Watch*, zajmujący się obserwowaniem zmian i rozwoju branży IT, a w szczególności tych jej gałęzi, które są powiązane z przechowywaniem danych cyfrowych oraz formatami zapisu dokumentów cyfrowych.

Natomiast proces *Preservation Activities* odpowiedzialny jest za koordynację wszelkich czynności i zadań podejmowanych w ramach procesu archiwizacji publikacji elektronicznych oraz kontrolę ich zgodności ze standardami DSEP. W zakresie *Preservation Activities* dokonuje się następujących działań:

- *Create New Reference Platform* - czyli tworzenia i implementacji nowej platformy programowo-sprzętowej w systemie depozytowym DSEP. Działania te mają miejsce wówczas, gdy zachodzi potrzeba przeniesienia kolekcji depozytowej do nowego otoczenia sprzętowo-systemowego.
- *Archival Information Update* - a więc przekształcanie przechowywanych w archiwum pakietów typu AIP do nowszej postaci w przypadku zmiany projektu archiwalnego pakietu informacyjnego lub aktualizowanie ich zawartości w związku z przeprowadzaniem procesów emulacji lub migracji.

Kolejny, dziewiąty moduł **Dostarczanie** (*Delivery*) reprezentuje działania na rzecz przygotowania kopii archiwizowanej publikacji elektronicznej do jej udostępnienia zainteresowanym użytkownikom. Proces ten umożliwia wyszukiwanie przechowywanych w depozycie publikacji, w razie potrzeby sporządzenie kopii określonej publikacji oraz jej przesłanie do działu udostępniania. W opinii specjalistów publikacja ma zostać tak przygotowana, aby użytkownik mógł otrzymać w zależności od zapotrzebowania, jej całość lub odpowiedni fragment, w celach przeglądnięcia, wydruku lub zapisu na nośnik. System przewiduje też opcję zlecenia wydruku publikacji na żądanie (*print on demand*)⁴⁷. Liegmann dodaje, że dostarczenie publikacji użytkownikowi powinno odbywać się w jak najbardziej standardowej formie, tj. takiej, która stawia najmniej specyficzne wymagania odnośnie warunków, w jakich ma ona być użytkowana. Pakiet informacyjny *Dissemination Information Package* (DIP) powinien zawierać wszelkie dodatkowe elementy, konieczne do użytkowania publikacji.

⁴⁷ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 7-8

W modelu zawarty jest też charakterystyczny dla systemów bibliotecznych moduł **Udostępniania** (*Access*), reprezentujący te elementy infrastruktury systemu bibliotecznego, które odnoszą się do otoczenia użytkownika końcowego oraz gwarantują mu dostęp do bibliotecznych zbiorów. *Access* obejmuje dostępność narzędzi wyszukiwawczych, identyfikację użytkownika, zarządzanie prawami użytkowników, określanie profilu użytkownika etc.⁴⁸. Dzięki temu, że system depozytowy publikacji elektronicznych korzysta z katalogu systemu bibliotecznego, użytkownik nie jest zmuszony do przeszukiwania dwóch oddzielnych katalogów. Jako zaletę traktuje się możliwość zamówienia dokumentu depozytowego poprzez istniejący i znany już użytkownikom system biblioteczny⁴⁹.

Ostatni moduł w modelu systemu depozytowego publikacji elektronicznych to **Monitorowanie** (*Monitoring*), który należy utożsamiać z kontrolą jakości pracy zarówno całego systemu, jak i efektywności poszczególnych, zachodzących w nim procesów⁵⁰.

Zaprezentowany graficznie i opisany powyżej model, skomponowany z modułów systemu biblioteki cyfrowej i systemu depozytowego publikacji elektronicznych pokazuje, jakie zadania przejmują system depozytowy DSEP oraz w jaki sposób współtworzy on wraz z funkcjonującym już systemem biblioteki cyfrowej jedną całość. Wzajemnie uzupełniające się moduły obu systemów mają zapewnić nie tylko zgromadzenie, opracowanie i bieżące udostępnianie publikacji elektronicznych, lecz również ich przechowanie dla przyszłych pokoleń użytkowników.

W literaturze przedmiotu⁵¹ dostrzega się - niewątpliwie niebagatelne - znaczenie zarówno modelu referencyjnego OAIS, jak i DSEP, jednocześnie jednak podkreślając, że modele te dostarczają zaledwie teoretycznego opisu wymagań dla funkcjonowania bibliotecznych, tudzież archiwalnych systemów depozytowych dla publikacji elektronicznych. Pomimo dość dużego stopnia szczegółowości opisu obu systemów zaznaczyć należy, że są to opisy tylko modeli funkcjonalnych, natomiast nie dostarczają one gotowych struktur dla implementacji w praktyce. OAIS oraz DSEP są zatem jedynie bazą teoretycznej wiedzy o organizacji i funkcjonowaniu archiwów elektronicznych.

⁴⁸ Titia van der Werf, *The Deposit System for Electronic Publications...*, op.cit., s. 7-8

⁴⁹ H. Liegmann, *Langzeitverfügbarkeit digitaler Publikationen...* op.cit., s. 106-109

⁵⁰ <http://www.kb.nl/coop/nedlib/index.html> (dostęp: 15.04.2005r.)

⁵¹ Titia van der Werf, *The Deposit System for Electronic Publications...*; Uwe. M. Borghoff et al., *Langzeitarchivierung. Methoden zur ...* op.cit., s. 36