

Dodatek C. Stałe odnośniki do obiektów w Aplikacji Czytelnika systemu dLibra

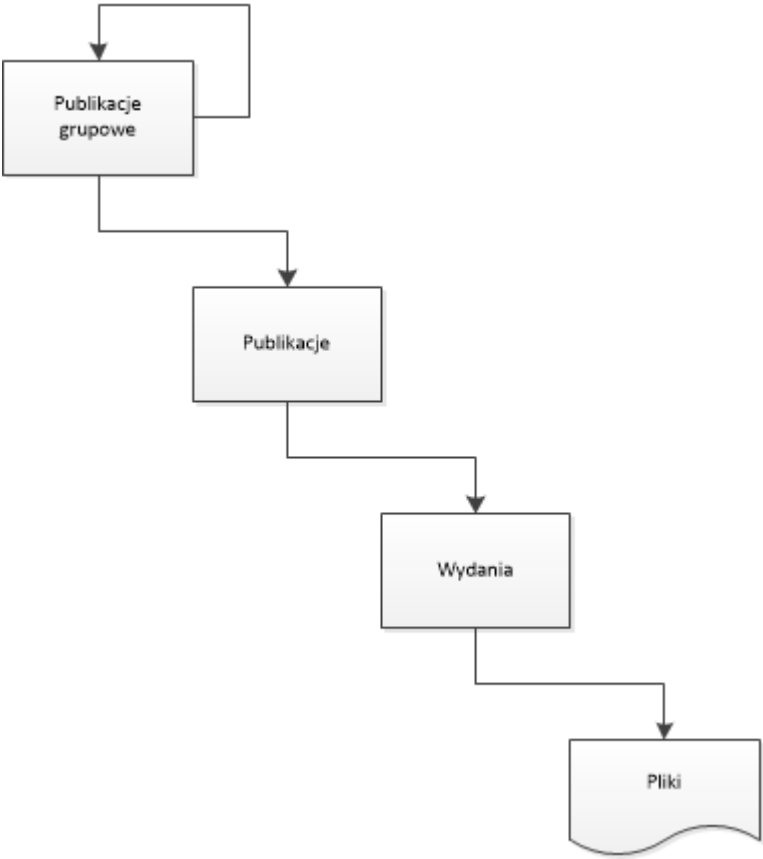
Wstęp

Bardzo istotnym aspektem udostępniania obiektów cyfrowych on-line jest możliwość tworzenia trwałych odnośników do tych obiektów. Odnośniki takie mogą być potrzebne np. w pracy naukowej jako element bibliografii, ale także w przypadku odnoszenia się do obiektów cyfrowych w zewnętrznych serwisach, takich jak katalogi biblioteczne, bazy bibliograficzne czy nawet Wikipedia. Poniżej opisano w jaki sposób tworzone są odnośniki do obiektów cyfrowych w systemie dLibra i z których z nich należy korzystać w celu zapewnienia maksymalnej trwałości referencji.

Model danych systemu dLibra

Poniższy rysunek prezentuje podstawowy model danych w systemie dLibra. W skład tego modelu wchodzi:

- Publikacje grupowe - elementy grupujące inne publikacje grupowe oraz zwyczajne publikacje. Publikacje grupowe mogą reprezentować np. całe czasopismo, lub jego konkretny rocznik czy miesiąc.
- Publikacje - elementy reprezentujące pojedynczy obiekt cyfrowy w systemie, który może posiadać wiele kolejnych wersji (wydań). Przykładem takiej publikacji może być zarówno książka, jak i pojedynczy numer czasopisma.
- Wydania - Konkretnie wersje publikacji. Mogą odpowiadać np. faktycznym kolejnym wydaniom książki, ale mogą być też tworzone np. w przypadku udostępniania innej postaci cyfrowej tego samego obiektu fizycznego. W systemie może istnieć równocześnie wiele wydań tej samej publikacji, jednak praktyka pokazuje że użytkownikom końcowym udostępniane jest najczęściej tylko ostatnie, najnowsze wydanie.
- Pliki - pliki wchodzące w skład konkretnych wydań publikacji. Zawierają dane/informacje, które pobiera czytelnik korzystający z biblioteki cyfrowej.



Rys. 1. Podstawowy model danych w systemie dLibra.

Poniższa tabela przedstawia przykłady dla poszczególnych elementów modelu danych wraz z odnośnikami.

Typ obiektu	Nazwa/tytuł obiektu	Dane niezbędne do jednoznacznej identyfikacji w systemie dLibra	Odnośnik wykorzystywany w interfejsie WWW	Stały odnośnik

Publikacja grupowa	Studies in Physical Culture and Tourism	Id publikacji: 4081	http://www.wbc.poznan.pl/dlibra/publication?id=4081	http://www.wbc.poznan.pl/publication/4081
Publikacja grupowa	Studies in Physical Culture and Tourism 2012	Id publikacji: 263176	http://www.wbc.poznan.pl/dlibra/publication?id=263176	http://www.wbc.poznan.pl/publication/263176
Publikacja grupowa	Studies in Physical Culture and Tourism 2012 Vol.19 No.3	Id publikacji: 290787	http://www.wbc.poznan.pl/dlibra/publication?id=290787	http://www.wbc.poznan.pl/publication/290787
Publikacja grupowa	Part II - Biology and exercise physiology	Id publikacji: 291297	http://www.wbc.poznan.pl/dlibra/publication?id=291297	http://www.wbc.poznan.pl/publication/291297
Publikacja	Insulin sensitivity and blood lipid profile in women recreationally practicing horseback riding	Id publikacji: 291316	http://www.wbc.poznan.pl/dlibra/publication?id=291316	Opis: http://www.wbc.poznan.pl/publication/291316 Treść (plik główny): http://www.wbc.poznan.pl/publication/291316/content
Wydanie	Insulin sensitivity and blood lipid profile in women recreationally practicing horseback riding	Id wydania: 239649	Opis: http://www.wbc.poznan.pl/dlibra/docmetadata?id=239649 Treść (plik główny): http://www.wbc.poznan.pl/Content/239649/	<i>Brak - zaleca się korzystanie ze stałego odnośnika do publikacji</i>
Plik	7_Studies_2012_3_135.pdf	Id wydania: 239649 Pełna nazwa pliku: 7_Studies_2012_3_135.pdf	http://www.wbc.poznan.pl/Content/239649/7_Studies_2012_3_135.pdf	<i>Brak - zaleca się korzystanie ze stałego odnośnika do publikacji</i>

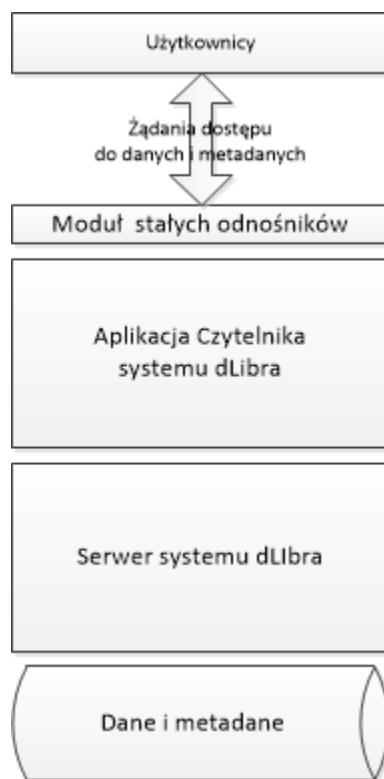
Tabela 1. Przykłady poszczególnych elementów modelu danych w systemie dLibra na przykładzie czasopisma dostępnego w Wielkopolskiej Bibliotece Cyfrowej.

Krótkiego komentarza wymaga jeszcze widoczne w Tabeli 1 rozróżnienie pomiędzy plikiem, a plikiem głównym publikacji. Obiekty cyfrowe mogą składać się z jednego lub więcej plików i w tym drugim przypadku niezbędne jest określenie, od którego pliku czytelnik powinien zacząć przeglądanie obiektu. Dla stron HTML przyjęto się, że plik taki nazywany jest index.html, jednak ze względu na to że w systemie dLibra składować można różne rodzaje obiektów, plik główny (plik początkowy) wskazywany jest przez redaktora biblioteki cyfrowej na etapie umieszczania obiektu w bibliotece. Rozróżnienie to dobrze ilustruje następujący przykład:

- <http://www.wbc.poznan.pl/Content/8711/> - Link do treści wydania o identyfikatorze 8711 w Wielkopolskiej Bibliotece Cyfrowej, spowoduje przekierowanie do pliku głównego publikacji.
- <http://www.wbc.poznan.pl/Content/8711/index.html> - Link do pliku głównego treści wydania o identyfikatorze 8711 w Wielkopolskiej Bibliotece Cyfrowej, uzyskany w efekcie przekierowania z adresu <http://www.wbc.poznan.pl/Content/8711/>.
- <http://www.wbc.poznan.pl/Content/8711/tom1-014.jpg> - Link do jednego z plików składowych treści wydania o identyfikatorze 8711 w Wielkopolskiej Bibliotece Cyfrowej, uzyskany w efekcie przeglądania publikacji.

Rodzaje odnośników i sposób przetwarzania żądań czytelników

W Tabeli 1 przedstawiono dwa rodzaje odnośników do poszczególnych elementów modelu danych w systemie dLibra. Rozróżnienie to wynika z przyjętego sposobu przetwarzania żądań czytelników zilustrowanego na Rysunku 2. Kluczowe jest tutaj rozróżnienie pomiędzy Aplikacją Czytelnika, a modułem stałych odnośników. Moduł ten ma za zadanie uniezależnić część odnośników funkcjonujących w systemie dLibra od zmian technicznych zachodzących w Aplikacji Czytelnika, związanych z jej rozwojem, udostępnianiem kolejnych wersji i wprowadzanymi modernizacjami technologicznymi. Ostatnia kolumna w Tabeli 1 przedstawia przykłady takich stałych odnośników. Zadaniem modułu obsługującego te odnośniki jest zapewnienie, aby funkcjonowały one w taki sam sposób niezależnie od zmian wprowadzanych w kolejnych wersjach Aplikacji Czytelnika. Przedostatnia kolumna w Tabeli 1 przedstawia odnośniki funkcjonujące niezmiennie w systemie dLibra od wersji 0.4 wydanej w 2004 roku do chwili obecnej, jednak teoretycznie w przyszłości ze względów technicznych mogą one ulec zmianie, dlatego też przy tworzeniu referencji do obiektów zaleca się stosowanie stałych odnośników zaprezentowanych w ostatniej kolumnie.



Rys. 2. Uproszczony schemat architektury systemu dLibra.

Stałe odnośniki a sytuacje wyjątkowe

W pewnych specyficznych przypadkach zapewnienie trwałych odnośników nie jest możliwe wyłącznie w oparciu o system biblioteki cyfrowej. Patrząc na przykładowy stały odnośnik (np.: <http://www.wbc.poznan.pl/publication/290787>) można stwierdzić, że posiada on dwa elementy które mogą ulec zmianie. Są to

- Adres domenowy (w przykładzie: www.wbc.poznan.pl) - może ulec zmianie, na przykład gdy:
 - właściciel domeny zapomni przedłużyć jej ważności i po wygaśnięciu własności domeny wykupi ją ktoś inny i zacznie stosować do innych celów;
 - instytucja macierzysta zmieni swoją nazwę i domenę (np. z Akademii na Uniwersytet);
 - biblioteka cyfrowa zmieni swój charakter (np. z instytucjonalnej na regionalną) i domenę.
- Identyfikator systemowy publikacji w ramach biblioteki cyfrowej (w przykładzie: 290787) - może ulec zmianie na przykład w efekcie awarii sprzętu powodującej utratę danych. Gdy administrator nie posiadał kopii zapasowej biblioteki cyfrowej, ale posiada kopie poszczególnych obiektów i ich metadanych to jedynym rozwiązaniem jest zbudowanie biblioteki cyfrowej od nowa. W takiej sytuacji mimo, że adres domenowy biblioteki cyfrowej i oprogramowanie nie ulegną zmianie, to systemowe identyfikatory poszczególnych obiektów najprawdopodobniej zmienią się, gdyż są one nadawane zależnie od kolejności wprowadzania obiektów.

Aby zapobiec tego typu wyjątkowym problemom stosuje się systemy identyfikatorów zewnętrzne w stosunku do biblioteki cyfrowej. Najpopularniejszym tego typu komercyjnym systemem są **identyfikatory DOI**, a jedną z bezpłatnych alternatyw jest **system Handle**. W polskich bibliotekach cyfrowych przyjęty się powszechnie oparte o specyfikację **OAI Id** identyfikatory obsługiwane przez **Federację Bibliotek Cyfrowych**. Tego typu zewnętrzne identyfikatory wprowadzane są jako element metadanych opisowych obiektu, dzięki czemu stają się niezależne od identyfikatorów systemowych. Uniezależnienie od adresu domenowego uzyskuje się poprzez globalne usługi lokalizujące (tzw. resolvery).

Dla przykładu - do artykułu "The use of SEM morphoscopy in researching the litho-pedogenetic environments evolution of Late Pleistocene and Holocene" dostępnego w Repozytorium Cyfrowym Instytutów Naukowych PAN można w sposób trwały odwołać się na następujące sposoby:

- <http://rcin.org.pl/igipz/publication/13932> - Odnośnik na poziomie biblioteki cyfrowej
- <http://fbc.pionier.net.pl/id/oai:rcin.org.pl:2289> - Odnośnik wykorzystujący identyfikator OAI, realizowany w oparciu o usługę Federacji Bibliotek Cyfrowych
- <http://fbc.pionier.net.pl/id/10.7163/GPol.2011.S1.3> - Odnośnik wykorzystujący identyfikator DOI, realizowany w oparciu o usługę Federacji Bibliotek Cyfrowych
- <http://dx.doi.org/10.7163/GPol.2011.S1.3> - Odnośnik wykorzystujący identyfikator DOI, realizowany w oparciu o usługę dx.doi.org.

Podsumowanie

W ramach podsumowania poniżej przedstawiono tabelę zawierającą wady i zalety poszczególnych rodzajów odnośników do obiektów dostępnych w systemie dLibra.

Typ obiektu	Odnosnik wykorzystywany w interfejsie WWW	Efekt wywołania	Komentarz	Zalecane?
Publikacja grupowa	http://www.wbc.poznan.pl/publication/4081	Wyświetlenie struktury publikacji grupowej.	Odwołanie poprzez stały odnośnik niezależny od konkretnej wersji Aplikacji Czytelnika.	+
	http://www.wbc.poznan.pl/dlibra/publication?id=4081	Wyświetlenie struktury publikacji grupowej.	Odwołanie wprost do Aplikacji Czytelnika, teoretycznie struktura URL może w przyszłości ulec zmianie.	×
Publikacja	http://www.wbc.poznan.pl/dlibra/publication?id=291316	Wyświetlenie metadanych jednego dostępnego wydania danej publikacji lub listy odnośników do wydań tej publikacji w przypadku, gdy dostępne jest więcej niż jedno wydanie (w praktyce ta sytuacja występuje bardzo rzadko).	Odwołanie wprost do Aplikacji Czytelnika, teoretycznie struktura URL może w przyszłości ulec zmianie.	×
	http://www.wbc.poznan.pl/publication/291316	Wyświetlenie metadanych jednego dostępnego wydania danej publikacji lub listy odnośników do wydań tej publikacji w przypadku, gdy dostępne jest więcej niż jedno wydanie (w praktyce ta sytuacja występuje bardzo rzadko).	Odwołanie poprzez stały odnośnik niezależny od konkretnej wersji Aplikacji Czytelnika.	+
	http://www.wbc.poznan.pl/publication/291316/content	Wyświetlenie treści jednego dostępnego wydania danej publikacji lub listy odnośników do wydań tej publikacji w przypadku, gdy dostępne jest więcej niż jedno wydanie (w praktyce ta sytuacja występuje bardzo rzadko).	- Odwołanie poprzez stały odnośnik niezależny od konkretnej wersji Aplikacji Czytelnika. - Powoduje wyświetlenie treści w interfejsie biblioteki cyfrowej, pozwalając czytelnikowi na dostęp do metadanych prezentowanego obiektu, listy powiązanych obiektów itp.	+
Wydanie	http://www.wbc.poznan.pl/dlibra/docmetadata?id=239649	Wyświetlenie metadanych danego wydania publikacji.	- Odwołanie wprost do Aplikacji Czytelnika, teoretycznie struktura URL może w przyszłości ulec zmianie. - Odwołanie do konkretnego wydania publikacji, które może przestać być dostępne dla czytelników po utworzeniu i udostępnieniu kolejnego wydania tej samej publikacji.	×
Plik	http://www.wbc.poznan.pl/Content/239649/	Wyświetlenie pliku głównego danego wydania publikacji.	- Odwołanie do konkretnego wydania publikacji, które może przestać być dostępne dla czytelników po utworzeniu i udostępnieniu kolejnego wydania tej samej publikacji. - Odwołanie wprost do treści, powoduje wyświetlenie pliku bez interfejsu biblioteki cyfrowej, uniemożliwiając tym samym czytelnikowi zapoznanie się z metadanymi obiektu.	×
	http://www.wbc.poznan.pl/Content/239649/7_Studies_2012_3_135.pdf	Wyświetlenie określonego pliku PDF danego wydania publikacji.	- Odwołanie do konkretnego wydania publikacji, które może przestać być dostępne dla czytelników po utworzeniu i udostępnieniu kolejnego wydania tej samej publikacji. - Odwołanie wprost do treści, powoduje wyświetlenie pliku bez interfejsu biblioteki cyfrowej, uniemożliwiając tym samym czytelnikowi zapoznanie się z metadanymi obiektu. - Odwołanie wprost do nazwy pliku, która może ulec zmianie w przypadku zmiany plików wchodzących w skład danego obiektu cyfrowego (niezależnie od zmiany identyfikatora wydania), np. ze względu na wyjście z użycia formatu cyfrowego w którym pierwotnie umieszczono plik.	×
	http://www.wbc.poznan.pl/Content/239649/7_Studies_2012_3_135.pdf#page=3	Wyświetlenie konkretnego fragmentu (str. 3) określonego pliku PDF danego wydania publikacji	- Odwołanie do konkretnego wydania publikacji, które może przestać być dostępne dla czytelników po utworzeniu i udostępnieniu kolejnego wydania tej samej publikacji. - Odwołanie wprost do treści, powoduje wyświetlenie pliku bez interfejsu biblioteki cyfrowej, uniemożliwiając tym samym czytelnikowi zapoznanie się z metadanymi obiektu. - Odwołanie wprost do nazwy pliku, która może ulec zmianie w przypadku zmiany plików wchodzących w skład danego obiektu cyfrowego (niezależnie od zmiany identyfikatora wydania), np. ze względu na wyjście z użycia formatu cyfrowego w którym pierwotnie umieszczono plik. - Odwołanie do konkretnego fragmentu pliku - strony nr 3 - następuje przez dodatkowy parametr, który nie musi być obsługiwany przez przeglądarkę plików, której używa czytelnik.	×

Tabela 2. Wady i zalety poszczególnych rodzajów odnośników do obiektów w systemie dLibra