



# Wprowadzenie do tematyki bibliotek cyfrowych

Tomasz Parkoła

Poznańskie Centrum Superkomputerowo-Sieciowe

---

IV Warsztaty „Biblioteki cyfrowe”

Poznań, 2007

# Plan prezentacji

- Wstęp
- Podstawowe funkcje systemu dLibra
- Organizacja pracy
- Metadane w systemie dLibra
- Wielojęzyczność
- Kontrola dostępu do treści



# Wstęp

---

# Wstęp

- Zadania biblioteki to m.in.:
  - Gromadzenie materiałów
  - Przechowywanie materiałów
  - Udostępnianie materiałów
- Obszary działania w Internecie
  - Strona internetowa biblioteki
  - Katalog „on-line” (OPAC)
  - Biblioteka cyfrowa

# Biblioteka cyfrowa

„Biblioteka cyfrowa to organizacja, która dostarcza środki, włączając w to specjalizowane kadry, do wyboru, organizacji, klasyfikacji, rozpowszechniania, bezpiecznego i długoterminowego przechowywania kolekcji obiektów cyfrowych, tak by były łatwo dostępne dla określonej społeczności lub grupy społeczności”

Digital Library Federation [1998]

# Biblioteka cyfrowa – treść

„Treść biblioteki cyfrowej zawiera dane, metadane opisujące liczne aspekty tych danych (np. twórca, właściciel, prawa do wykorzystania) oraz metadane opisujące zależności w stosunku do innych wewnętrznych lub zewnętrznych danych lub metadanych.”

Ch. L. Borgman

# Zalety biblioteki cyfrowej

- Zdalny dostęp do biblioteki
- Ciągły dostęp do przechowywanych zasobów
- Łatwe wyszukiwanie i przeglądanie zasobów
- Przechowywanie i konserwacja informacji
  - Współdzielenie informacji
  - Bezpieczeństwo oryginalnego obiektu
- Łatwa aktualizacja informacji
- Możliwe nowe formy informacji
- Komunikacja z innymi bibliotekami



POZNAŃSKIE CENTRUM SUPERKOMPUTEROWO-SIECIOWE  
dLibra - PLATFORMA DO BUDOWY BIBLIOTEK CYFROWYCH



**dLibra**  
DIGITAL LIBRARY FRAMEWORK

=

oprogramowanie dla  
bibliotek cyfrowych



# dLibra – podstawowe funkcje

- Przechowywanie obiektów cyfrowych
  - Dowolny format treści (np.: PDF, HTML, DjVu, obraz, dźwięk, wideo, sceny 3D)
  - Metadane domyślnie w schemacie Dublin Core 1.1 z możliwością modyfikacji/dostosowania
    - Metadane mogą być wprowadzane w różnych językach
  - Hierarchiczna struktura przechowywania obiektów
  - Grupowanie powiązanych ze sobą obiektów

# dLibra – podstawowe funkcje

- Przechowywanie obiektów cyfrowych
  - Obsługa dużych ilości obiektów cyfrowych, metadanych i użytkowników
    - Możliwy podział systemu na usługi działające na różnych komputerach
    - Zarządzane na poziomie całego systemu słowniki metadanych
  - Rozbudowany mechanizm praw do poszczególnych funkcji systemu
  - Aktualizacja obiektu cyfrowego (wersjonowanie)
  - Ciągła kontrola spójności treści obiektu cyfrowego

# dLibra – podstawowe funkcje

- Udostępnianie obiektów cyfrowych
  - Poprzez strony internetowe
    - Wygląd może być odpowiednio dostosowany
    - Interfejs dostępny w różnych językach
  - Wyszukiwanie i przeglądanie opisów i treści obiektów cyfrowych
  - Prezentacja obiektów prostych i złożonych
    - Możliwe zabezpieczenie przed kopiowaniem niektórych typów obiektów (HTML, DjVu, PDF)
    - Metadane mogą być prezentowane w różnych językach niezależnie od języka interfejsu

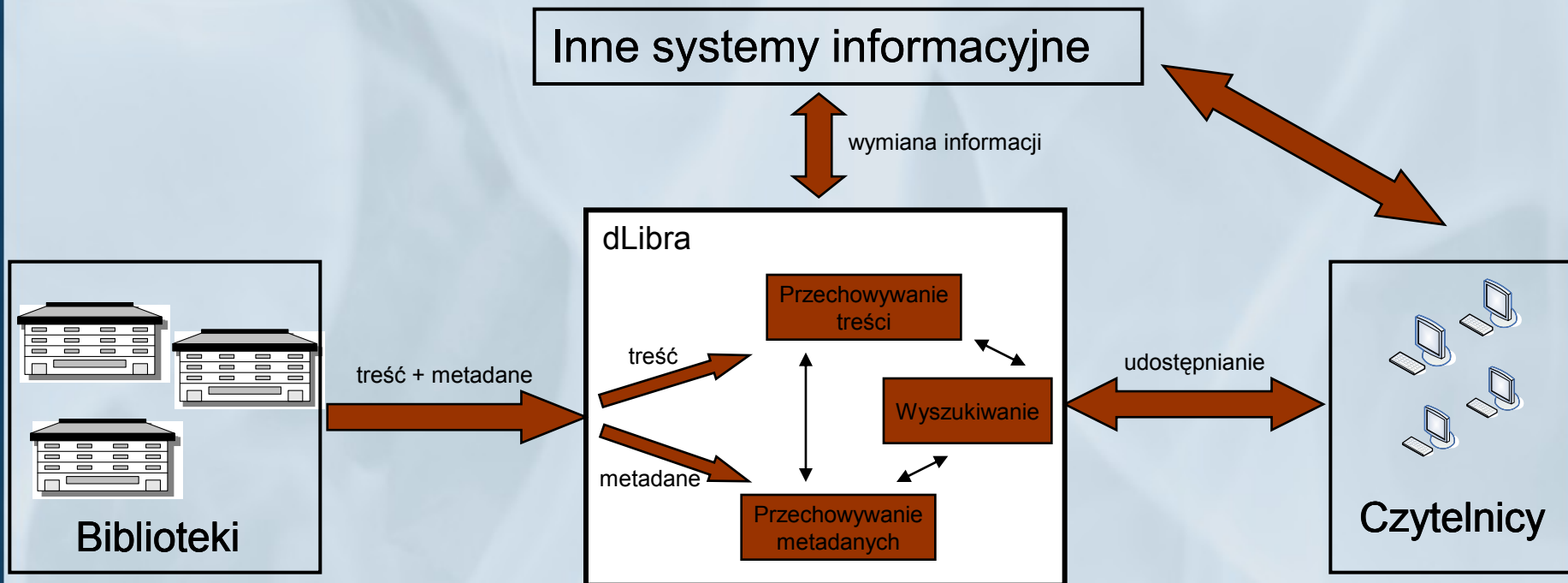
# dLibra – podstawowe funkcje

- Udostępnianie obiektów cyfrowych
  - Użytkownicy biblioteki cyfrowej
    - Konta użytkowników z sieciowym profilem
    - Biuletyny z nowościami, różnorodne kanały RSS, statystyki
    - Wkład w zawartość biblioteki przez tagowanie i oceny obiektów cyfrowych
  - Współpraca z innymi systemami w Internecie
    - Inne biblioteki cyfrowe
    - Wyszukiwarki internetowe
  - Rozbudowany mechanizm praw do poszczególnych obiektów w systemie

# Użytkownicy systemu dLibra

- Redaktor
  - Wprowadza publikacje do biblioteki cyfrowej
- Administrator
  - Zarządza biblioteką cyfrową
- Czytelnik
  - Korzysta z biblioteki cyfrowej

# Działanie z systemu dLibra



## Aplikacja Redaktora i Administratora:

Tworzenie struktury biblioteki,  
umieszczanie publikacji, zarządzanie  
biblioteką

## Aplikacja Czytelnika – WWW:

Przeglądanie zawartości,  
wyszukiwanie, powiadamianie,  
RSS, RDF, itp.





# Organizacja systemu dLibra

---

# Struktura biblioteki

## Katalogi – widok redaktora

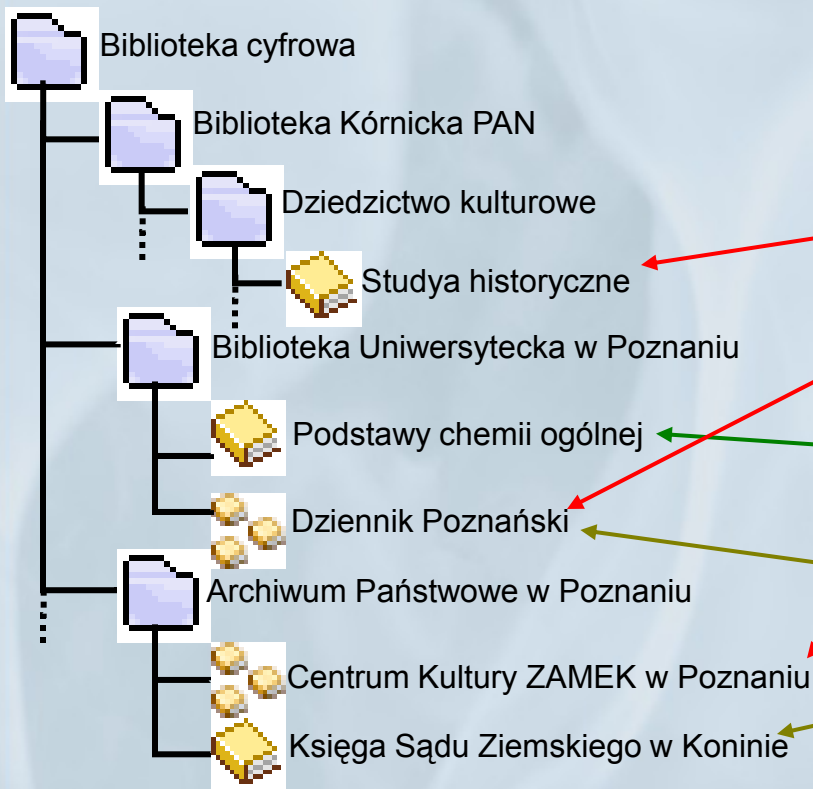
## Kolekcje – widok czytelnika

|  |   |
|--|---|
| Symbol katalogu:  | Symbol kolekcji:                     |
| Struktura drzewiasta (katalogi i publikacje)   | Struktura drzewiasta (tylko kolekcje)   |
| Miejsce przechowywania obiektów cyfrowych  | Lista obiektów  |
| Mechanizm niezależny od kolekcji   | Mechanizm niezależny od katalogów   |
| Widoczne tylko w aplikacji redaktora i administratora – organizacja pracy redaktora                | Widoczne na stronach internetowych (prezentacja) oraz w aplikacji redaktora i administratora (przypisywanie publikacji) |
| <i>W WBC każda biblioteka ma własny, zabezpieczony katalog</i>                                     | <i>W WBC kolekcje to listy obiektów o podobnej tematyce</i>   |

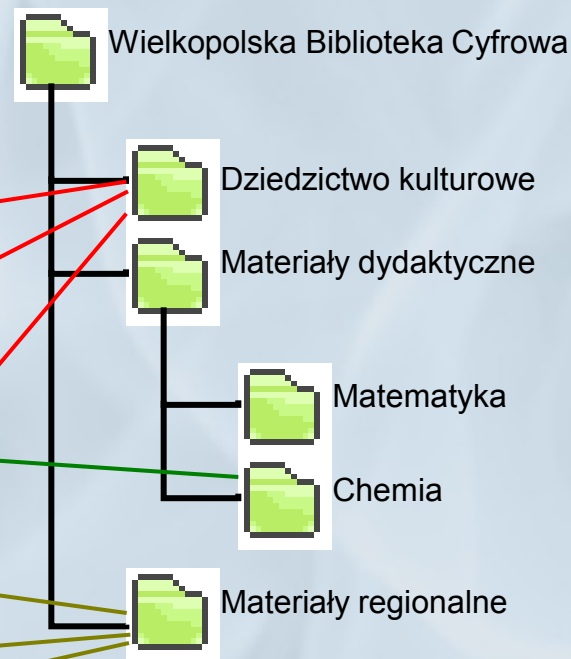


# Struktura biblioteki

## Katalogi – widok redaktora



## Kolekcje – widok czytelnika



# Katalog

- Nazwa
- Uwagi
- Opis bibliograficzny
  - Domyślny opis bibliograficzny elementów tworzonych w katalogu

# Kolekcja

- Nazwa
- Opis kolekcji
- Uwagi
  - Widoczne tylko w aplikacji redaktora i administratora

# Kolekcja wystaw

- Specjalna kolekcja
- Nie jest umieszczona wśród innych kolekcji na stronie internetowej
- Lista publikacji przeznaczonych do szczególnego wyeksponowania
- Wystawą jest każda publikacja przypisana do kolekcji wystaw

# Publikacja

- Obiekt cyfrowy umieszczony w systemie dLibra
- Reprezentuje pojedynczą pozycję (książka, numer gazety, mapa, film, itp.)
- Składa się z wydań – kolejne „wersje” publikacji



# Wydanie

- Nazwa
- Pliki z treścią
- Opis bibliograficzny – tytuł, autor, itp.
- Publikowanie
  - Stałe (jest opublikowane, bądź nie – do momentu ręcznej zmiany stanu)
  - Czasowe (do określonego dnia w przyszłości)
    - Automatyczna zmiana stanu na nieopublikowane w wyznaczonym terminie
    - Ułatwia przestrzeganie umów z właścicielami praw
- Uwagi administracyjne
  - Widoczne tylko w aplikacji redaktora i administratora

# Wydanie – informacje WWW

- Opis
  - Krótki, jednozdaniowy opis
  - Widoczny na WWW w wynikach wyszukiwania przy konkretnej pozycji
  - Zależny od języka interfejsu
- Komentarz
  - Dłuższa charakterystyka publikacji, ciekawostki, itp.
  - Widoczny na stronie WWW dotyczącej wydania
  - Zależny od języka interfejsu
- Miniaturka wydania
  - Zdjęcie, rysunek, skan obrazujący treść wydania

# Publikacja



- Nazwa
- Uwagi administracyjne
  - Widoczne tylko w aplikacji redaktora i administratora
- Lista kolekcji do których jest przypisana
  - Redaktor przypisuje publikację do kolekcji
  - Pośrednie przypisanie do kolekcji
- Prawa użytkowników do publikacji
  - Redaktor przyznaje prawa do publikacji użytkownikom
  - Możliwe prawa
    - Przeglądanie – widoczne tylko **opublikowane** wydania
    - Odczyt – widoczne **wszystkie** wydania
    - Zarządzanie – pełny dostęp do publikacji i jej wydań



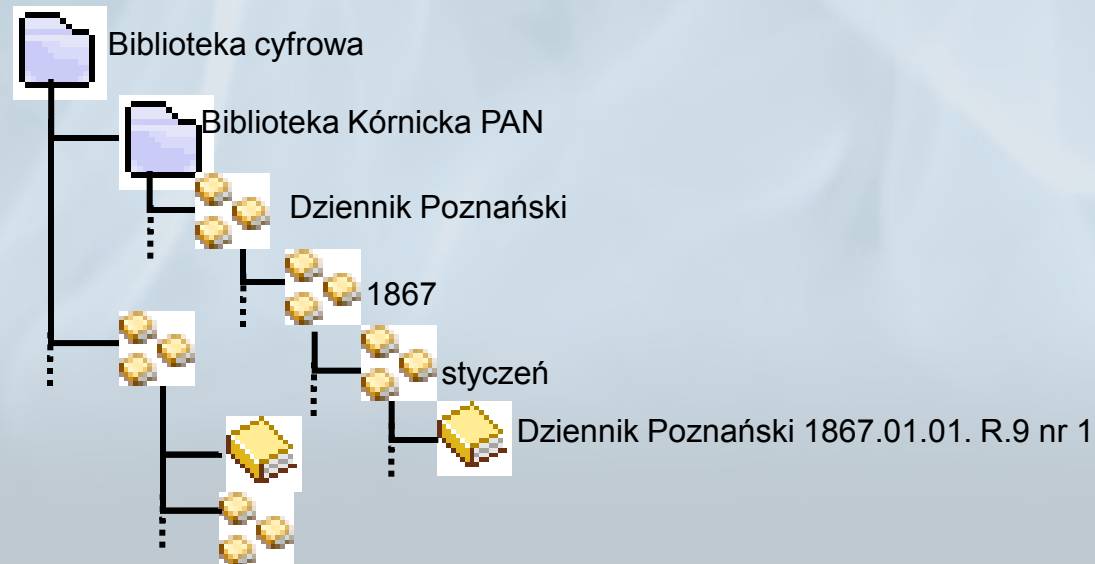
# Publikacja



- Zabezpieczanie publikacji
  - Uniemożliwia wydruk
  - Uniemożliwia kopiowanie treści
  - Ograniczona do treści w formacie
    - DjVu
    - HTML
    - PDF

# Publikacja grupowa

- Nie posiada treści
- Służy do łączenia (grupowania) powiązanych ze sobą publikacji, np. gazety, publikacje wielotomowe
- Hierarchiczna struktura
  - Może zawierać publikacje i publikacje grupowe
  - Widoczna na stronach internetowych biblioteki cyfrowej
  - Ułatwia nawigację użytkownikom



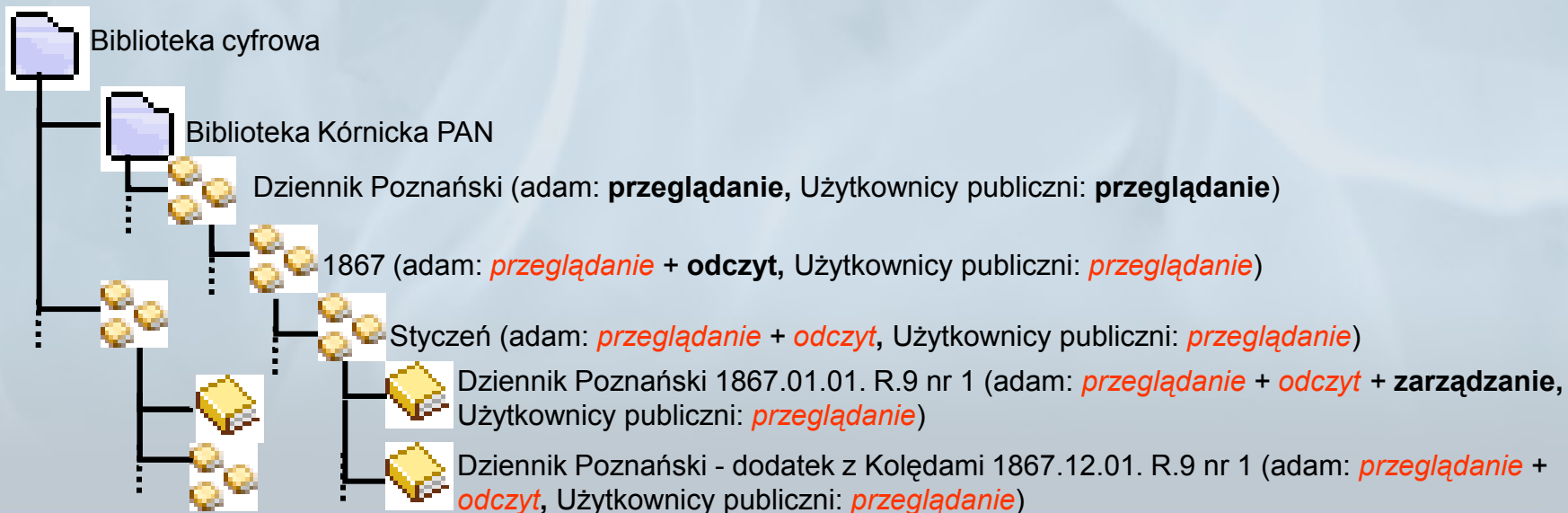
# Publikacja grupowa



- Nazwa
  - widoczna na stronie WWW w strukturze publikacji grupowej
- Uwagi
  - Widoczne tylko w aplikacji redaktora i administratora
- Informacje WWW
  - Opis
  - Komentarz
  - Miniaturka
- Lista publikacji
  - Możliwa ręczna zmiana pozycji na której znajduje się podrzędna publikacja

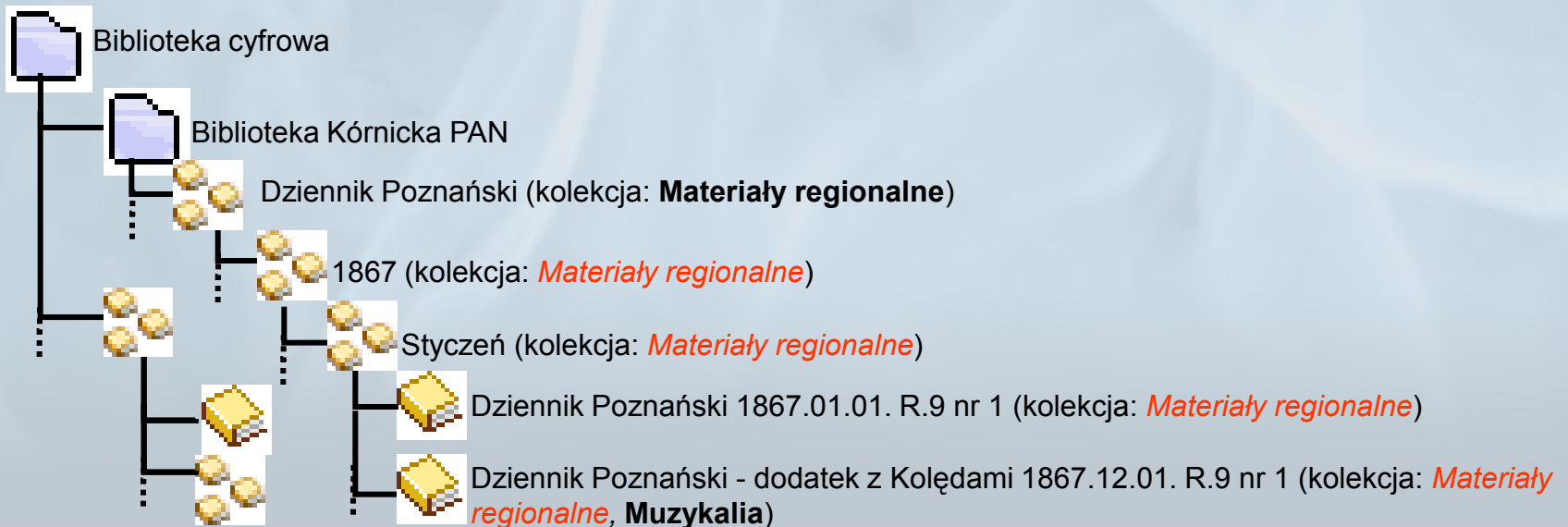
# Publikacja grupowa

- Dziedziczone właściwości
  - Prawa przyznane publikacji
  - Kolekcje do których przypisane są publikacje
  - Opis bibliograficzny publikacji



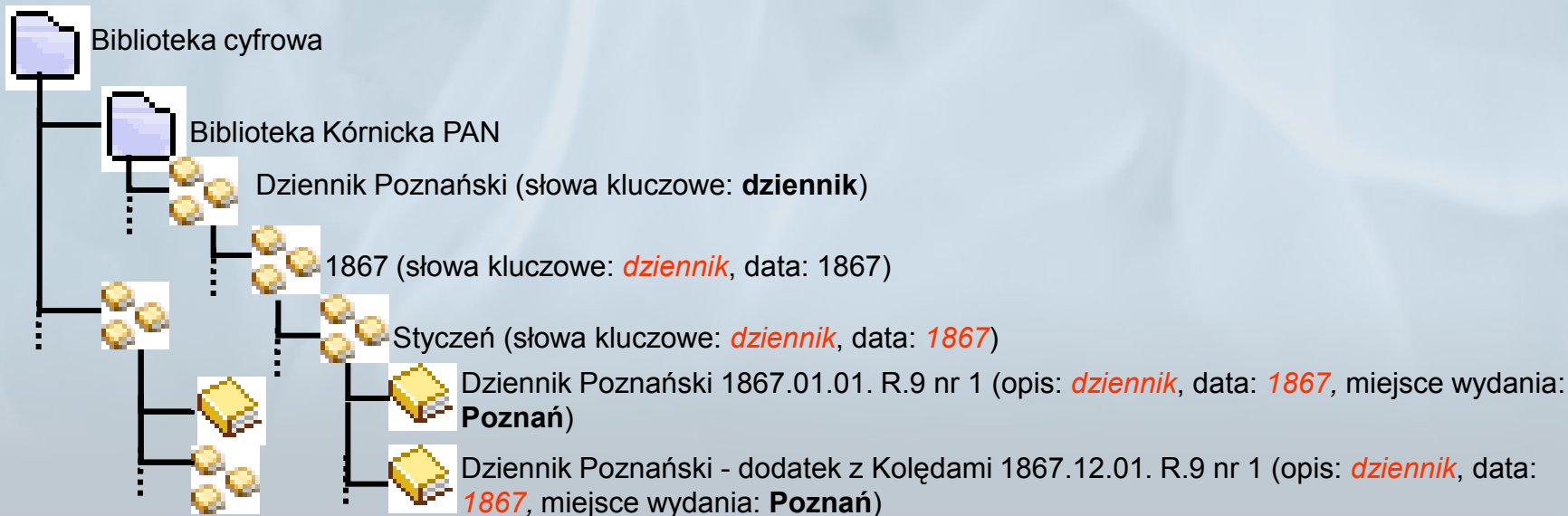
# Publikacja grupowa

- Dziedziczone właściwości
  - Prawa przyznane publikacji
  - **Kolekcje do których przypisane są publikacje**
  - Opis bibliograficzny publikacji




# Publikacja grupowa


- Dziedziczone właściwości
  - Prawa przyznane publikacji
  - Kolekcje do których przypisane są publikacje
  - **Opis bibliograficzny publikacji**



# Inne rodzaje publikacji

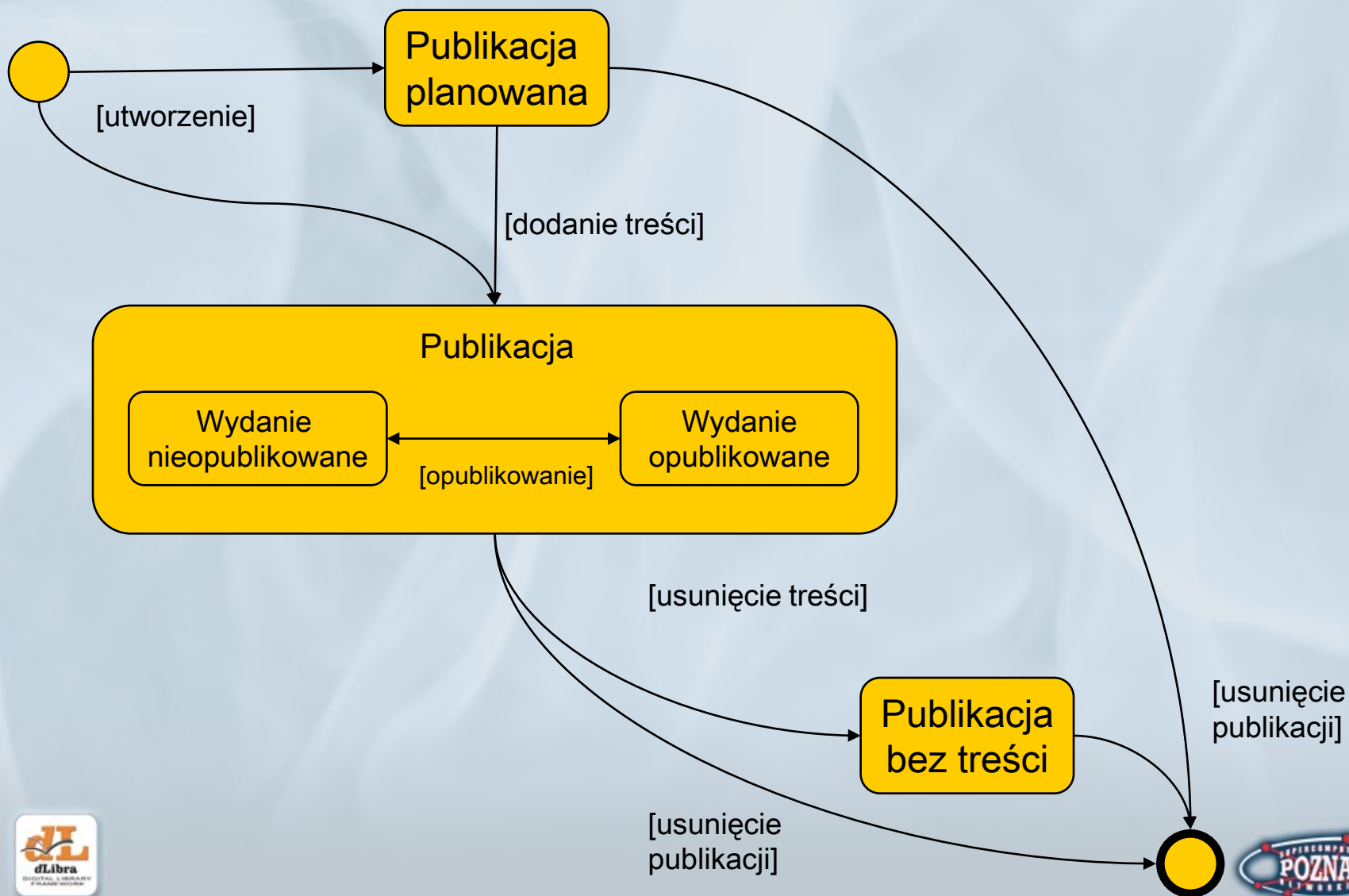
- Publikacje planowane 
  - Przedstawiają plany digitalizacji
    - Ułatwiają współpracę między bibliotekami
    - Informacja dla czytelników
  - Nie posiadają treści – nie posiadają wydań
  - Po dodaniu treści stają się zwykłymi publikacjami
    - Tworzone jest pierwsze wydanie publikacji
  - Właściwości
    - Przechodzące do publikacji
      - Nazwa
      - Przyznane prawa użytkowników
      - Przypisania do kolekcji
    - Przechodzące do pierwszego wydania
      - Opis bibliograficzny

# Inne rodzaje publikacji

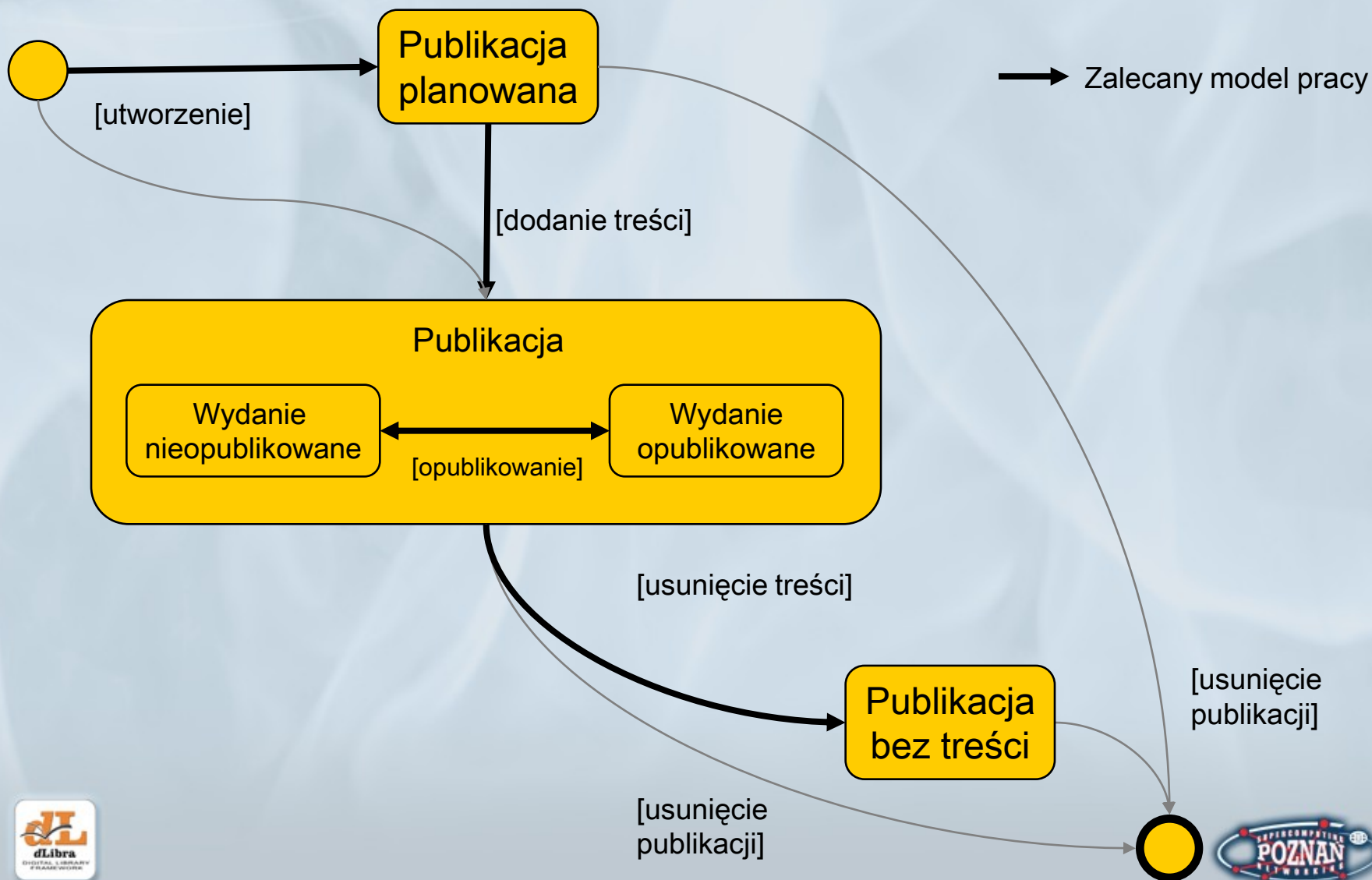
- Publikacje bez treści 
  - Powstają po usunięciu treści z publikacji
    - Aby usunąć treść należy podać powód
    - Nie można dodać treści do publikacji bez treści
  - Zawierają pełen opis, przypisane prawa itd.
  - Nie można wyszukać ani znaleźć na liście publikacji
  - Umożliwiają zachowanie trwałości odnośników
- Publikacje usunięte
  - Nie są widoczne w aplikacji redaktora i administratora
  - Można podać powód całkowitego usunięcia
  - Zachowują trwałość odnośnika
    - Pod linkiem do publikacji znajduje się powód i data całkowitego usunięcia publikacji



# Publikacje w systemie dLibra



# Publikacje w systemie dLibra



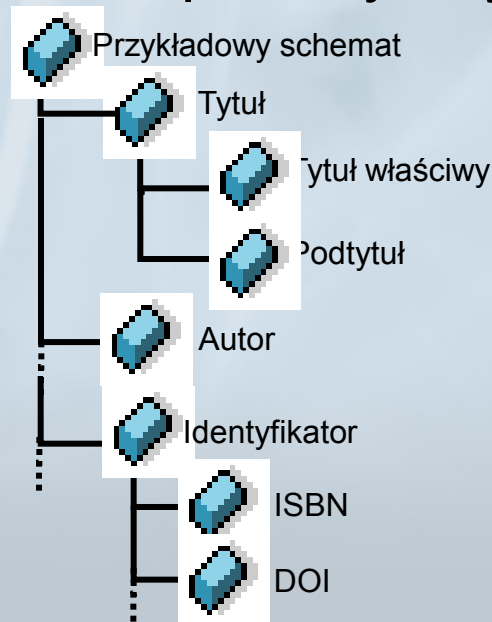


# Metadane

---

# Metadane – schemat atrybutów

- Administrator biblioteki zarządza używanym schematem atrybutów
  - Dodawanie atrybutu
  - Usuwanie atrybutu
  - Modyfikacja atrybutu
  - Każdy element, który ma metadane opisywany jest przy użyciu zdefiniowanego schematu
- Atrybuty mogą posiadać podatrybuty na wielu poziomach



# Metadane – schemat atrybutów

- Predefiniowany zestaw (schemat) atrybutów zgodny ze standardem DublinCore Metadata Element Set (DCMES) wersja 1.1
  - Wersja oryginalna
    - Ostatnio zmieniono 2006-12-18
    - <http://dublincore.org/documents/dces/>

# Metadane – schemat atrybutów

- DublinCore Metadata Element Set 1.1
  - Tłumaczenie standardu
    - Tłumaczono 2000-10-28 na podstawie oryginału z 1999-07-02
    - <http://ebib.oss.wroc.pl/standard/dc.html>
  - e-Poradnik redaktora zasobów cyfrowych
    - Interpretacja schematu Dublin Core wraz z materiałami pomocniczymi dla redaktorów zasobów cyfrowych Biblioteki Cyfrowej Uniwersytetu Wrocławskiego
    - <http://fbc.pionier.net.pl/id/oai:www.bibliotekacyfrowa.pl:17703>

# DCMES wersja 1.1

- **Tytuł**
- **Autor**
- **Temat i słowa kluczowe**
- **Opis (zawartości)**
- **Wydawca**
- **Współtwórca**
- **Data wydania**
- **Typ zasobu** (charakter, rodzaj treści)
- **Format**
- **Identyfikator zasobu**
- **Źródło**
- **Język**
- **Powiązania**
- **Zakres**
- **Prawa**

# Metadane obiektów

- Wydanie
  - Opis treści widoczny na stronie internetowej
- Publikacja planowana
  - Opis pierwszego wydania, które zostanie utworzone po dodaniu treści
- Publikacja grupowa
  - Opis dziedziczony przez podrzędne publikacje i publikacje grupowe
- Katalog
  - Domyślny opis każdego elementu tworzonego w tym katalogu



# Metadane obiektów

- Wprowadzanie metadanych
  - Ręczne przy użyciu edytora metadanych
  - Automatycznie określone metadane
    - Format zasobu
  - Importowanie gotowych metadanych
    - MARC (również z serwerów Z39.50)
    - XML (np. RDF, MASTER)
    - BibTeX
  - Tagi użytkowników stron internetowych
    - Opcjonalny proces moderacji

# Atrybut

- Właściwości
  - Nazwa – widoczna na stronach internetowych
  - Opis – widoczny na stronach internetowych
  - Nazwa RDF – unikalny identyfikator atrybutu
  - Rola – przypisana rola ze schematu Dublin Core
- Lista podatrybutów
  - Kolejność podatrybutów możliwa do modyfikacji

# Atrybut

- Słownik wartości z opisów publikacji
  - Możliwe dzięki używaniu tego samego schematu do opisu każdego elementu
  - Odrębny słownik wartości dla każdego atrybutu w poszczególnych językach

Metadane elementów



wprowadzanie nowych wartości



wykorzystywanie istniejących wartości

Atrybut Opis



Słownik wartości

góry  
samoloty  
wojna

# Atrybut

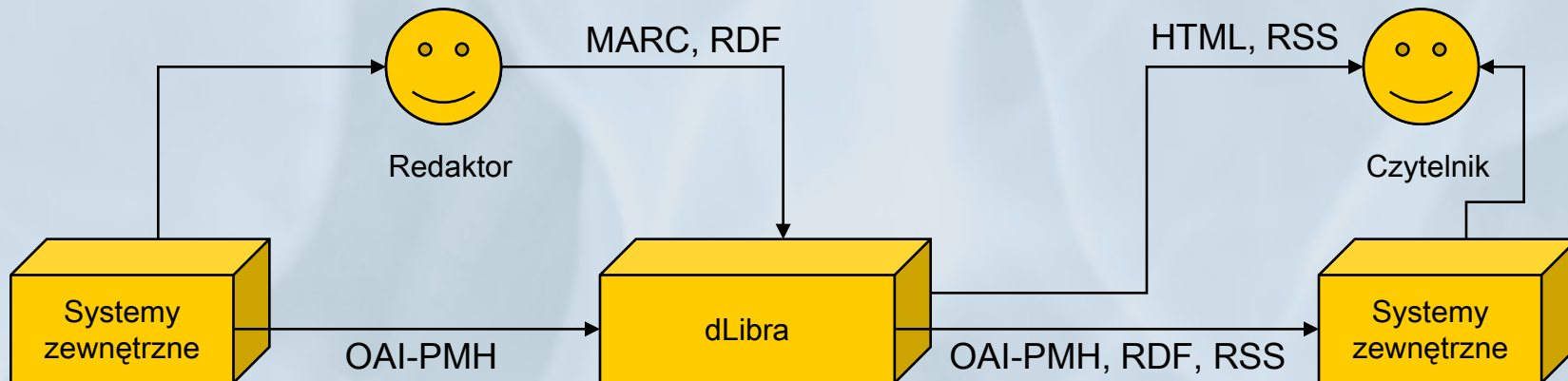
- Słownik wartości – synonimy
  - Każda wprowadzona wartość w słowniku może mieć synonimy
  - Synonimy dodaje redaktor posiadający odpowiednie to tego prawa
  - Synonimy polepszają wyniki wyszukiwania na stronach internetowych biblioteki cyfrowej

Atrybut Opis – słownik wartości PL

| Wartość bazowa | Grupy synonimów  |
|----------------|--|
| pałac          |  ← grupy synonimów  |
| zatrzask       |  ← grupy synonimów |

- Wartość bazowa może wystąpić w słowniku tylko raz
- Synonimy mogą się powtarzać w różnych grupach synonimów

# Metadane w systemie dLibra



## Wprowadzanie metadanych:

- Import MARC
- Import RDF (schemat DC lub dowolny inny)
- Indeksowanie repozytoriów OAI-PMH (na potrzeby wyszukiwania rozproszonego)

## Przechowywanie metadanych:

- Dowolny zdefiniowany hierarchiczny zestaw atrybutów
- Słowniki wartości atrybutów

## Prezentacja metadanych:

- Przeglądarki WWW
- Czytniki RSS
- Systemy OAI-PMH
- Wyszukiwarki
- Portale
- Schemat DC i schemat zdefiniowany w BC



# Wielojęzyczność

---

# Dwa poziomy wielojęzyczności

- „Interfejs”
  - Nazwa biblioteki, nazwy kolekcji, nazwy atrybutów, elementy interfejsu (opisy przycisków, formularzy, instrukcje, etc.)
  - Informacje zmieniane rzadko, zazwyczaj przed publicznym udostępnieniem w Internecie
  - Tłumaczenia mogą być opracowane przez osoby zewnętrzne (np. wynajętych tłumaczy)

# Dwa poziomy wielojęzyczności

- „Metadane”
  - Opisy, komentarze itp. publikacji i wydań
  - Informacje dotyczące wprowadzanych zasobów
  - Tłumaczenia powinny dokonywać osoby na bieżąco zajmujące się biblioteką cyfrową (najczęściej redaktorzy)



# Dwa poziomy wielojęzyczności

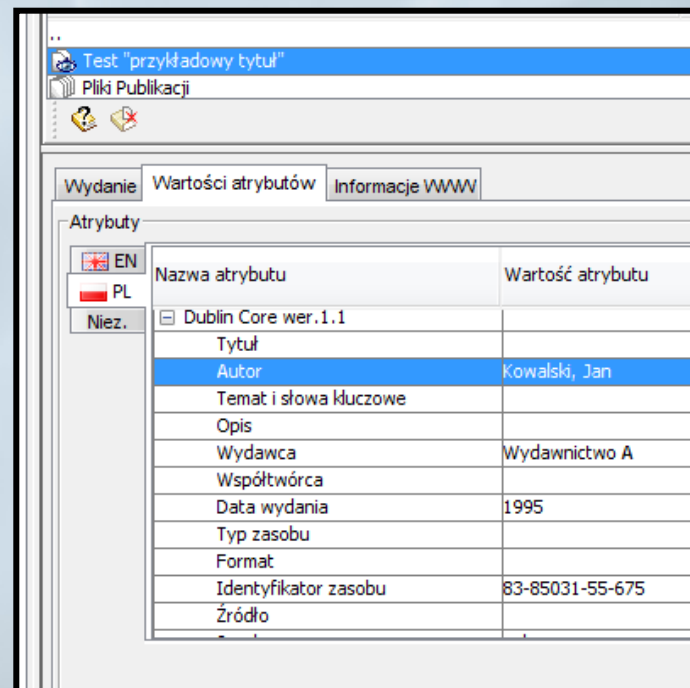
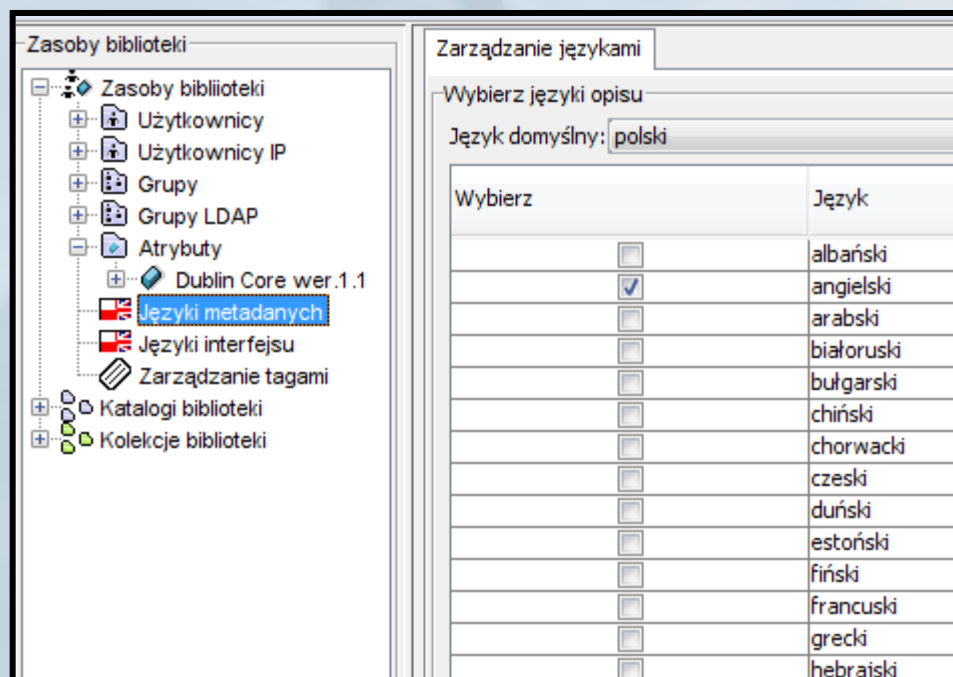
- Wielkopolska Biblioteka Cyfrowa:
  - „Interfejs”
    - 5 kolekcji
    - 23 atrybuty
  - „Metadane”
    - Około 500 publikacji grupowych
    - Około 1 500 podrzędnych publikacji grupowych
    - Około 46 000 wydań
    - Około 120 000 haseł w słownikach

## Wielojęzyczność

- Czytelnik na stronach WWW wybiera niezależnie język interfejsu i język metadanych
- Z reguły istnieje więcej języków interfejsu, niż języków metadanych
- Decyzję o językach metadanych najlepiej podjąć na początku istnienia biblioteki cyfrowej

# Wielojęzyczność

- Używane języki metadanych i interfejsu definiuje administrator systemu



# Języki domyślne

- Preferowane języki użytkownika pobierane są automatycznie z ustawień przeglądarki WWW
- Jeśli żaden z preferowanych języków użytkownika nie znajduje się na liście dostępnych języków to prezentowany jest interfejs w języku domyślnym
- Administrator określa
  - Język domyślny metadanych
  - Język domyślny interfejsu

# Wartości niezależne od języka

- Pozwalają na jednokrotne wprowadzenie elementów takich jak np. identyfikator (1234) czy format (text/html)
- Można wprowadzać, gdy jest zdefiniowany więcej niż jeden język interfejsu
- Wyświetlane są zawsze, niezależnie od wybranego języka metadanych

## Stałe elementy interfejsu

- Wersja polska i angielska opracowywana jest w PCSS
- Inne wersje opracowywane we współpracy z instytucjami wykorzystującymi system dLibra
  - francuski, niemiecki i czeski: Biblioteka Śląska



Udostępnianie zasobów  
—  
kontrola dostępu

---

# Kontrola dostępu do treści

## Dwa etapy

- Uwierzytelnianie – określenie tożsamości użytkownika
- Autoryzacja – określenie uprawnień do danego zasobu dla użytkownika o uwierzytelnionej tożsamości



# Kontrola dostępu do treści

- Uwierzytelnianie (określenie tożsamości)
  - Na podstawie nazwy użytkownika i hasła
  - Na podstawie zakresu adresów (adresu) IP
    - Dostęp bez hasła
    - Odmowa dostępu
  - Przy wykorzystaniu serwerów katalogowych LDAP
    - Dynamiczne grupy użytkowników
  - Przy wykorzystaniu usług Single Sign-On

# Kontrola dostępu do treści

- Autoryzacja (dostęp do obiektów)
  - Rozbudowany system uprawnień
    - 8 rodzajów praw administracyjnych
    - 7 rodzajów praw do katalogu
    - 3 rodzaje praw do pojedynczej publikacji
  - Dziedziczenie uprawnień z obiektów nadrzędnych i grup użytkowników
  - Uprawnienia implikowane

# Kontrola dostępu do treści

- Typowe warunki udostępnienia treści użytkownikom
  - Przyznane prawo **przeglądania** do publikacji
    - **Użytkownicy publiczni** – grupa reprezentująca wszystkich zarejestrowanych i anonimowych czytelników
  - Wydanie powinno być **opublikowane**
  - Publikacja powinna być przypisana do **kolekcji**

# Rodzaje kont w systemie dLibra

- Administrator/redaktor
  - Ma dostęp do aplikacji redaktora/administratora
  - Ograniczony przyznanymi mu prawami
- Czytelnik
  - Może modyfikować swój profil
  - Brak dostępu do aplikacji redaktora/administratora
- Czytelnik z ograniczeniem
  - Nie może modyfikować swojego profilu (brak dostępu)
  - Brak dostępu do aplikacji redaktora/administratora



Dziękuję za uwagę!

---