

# WBC i dLibra



Marcin Werla

Poznańskie Centrum Superkomputerowo-Sieciowe

[dlibra@man.poznan.pl](mailto:dlibra@man.poznan.pl)



# [ Plan prezentacji



- Wielkopolska Biblioteka Cyfrowa
- Biblioteka cyfrowa – wymagania
- Biblioteka cyfrowa oparta o system dLibra
- Biblioteki cyfrowe w Polsce



# Wielkopolska Biblioteka Cyfrowa



- Wspólna inicjatywa poznańskiego środowiska akademickiego
- Nadzorowana przez PCSS oraz Poznańską Fundację Bibliotek Naukowych
- Uruchomiona 1 października 2002 r.
- Oparta o oprogramowanie dLibra



# Wielkopolska Biblioteka Cyfrowa



- Obecnie składa się z 4. kolekcji
  - Dziedzictwo kulturowe
  - Materiały dydaktyczne
    - Podręczniki z Wydziału Chemii UAM
  - Materiały regionalne
  - Muzykalia
- Cały katalog zaindeksowany również przez zewnętrzne wyszukiwarki (Google itp.)



# Wielkopolska Biblioteka Cyfrowa



- Ponad 1300 publikacji z 22 źródeł
- Średnio dziennie:
  - Ponad 230 wyszukiwań
  - Ponad 700 odwiedzających
  - Ponad 175 wyświetlonych publikacji
- Linki ze zintegrowanych systemów bibliotecznych i zewnętrznych serwisów WWW



# [ Źródła publikacji w WBC



- Biblioteka Kórnicka PAN
- Biblioteka Uniwersytecka w Poznaniu
- Biblioteka Raczyńskich
- Poznańskie Towarzystwo Przyjaciół Nauk
- BG AWF w Poznaniu
- Wydział Chemii UAM w Poznaniu
- Poznańska Fundacja Bibliotek Naukowych



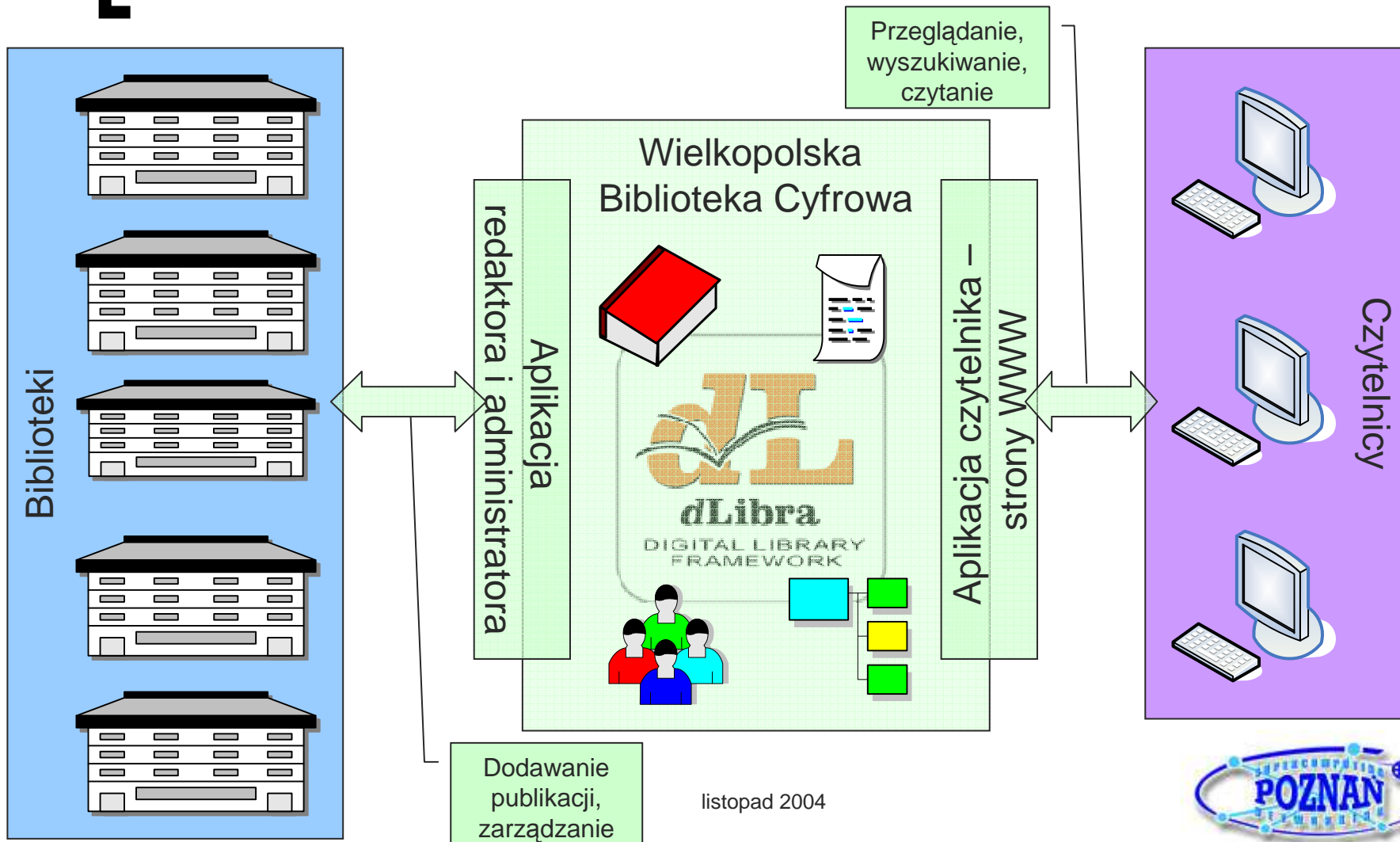
# Źródła publikacji w WBC



- Biblioteka Politechniki Poznańskiej
- Biblioteka Główna Akademii Muzycznej im. I. J. Paderewskiego w Poznaniu
- Wojewódzka i Miejska Biblioteka Publiczna im. dra Witolda Bełzy w Bydgoszczy
- Wojewódzki Specjalistyczny Szpital Gruźlicy i Chorób Płuc im. St. Staszica w Ludwikowie
- Osoby prywatne



# WBC – Struktura





# [ Początkowa motywacja



- Wady „klasycznych” bibliotek
  - Ograniczony dostęp do zgromadzonych zasobów
    - lokalizacja i czas otwarcia biblioteki
    - liczba egzemplarzy danego tytułu
    - brak możliwości wypożyczenia
  - Problemy z informacją o zawartości biblioteki



# [ Plan prezentacji



- Wielkopolska Biblioteka Cyfrowa
- Biblioteka cyfrowa – wymagania
- Biblioteka cyfrowa oparta o system dLibra
- Biblioteki cyfrowe w Polsce



# Biblioteka cyfrowa – wymagania



- Cztery obszary działania
  - Przechowywanie
  - Zarządzanie
  - Wyszukiwanie
  - Udostępnianie i prezentacja



# Biblioteka cyfrowa – wymagania



- Przechowywanie
  - Trwałość
  - Obsługa wielu formatów danych
  - Przechowywanie metadanych
  - Wsparcie dla całego okresu życia przechowywanego zasobu



# Biblioteka cyfrowa – wymagania



- Zarządzanie
  - Zarządzanie poszczególnymi komponentami systemu biblioteki (katalogi, kolekcje, ...)
  - Zarządzanie użytkownikami
  - Zarządzanie zgromadzonymi zasobami



# Biblioteka cyfrowa – wymagania



- Wyszukiwanie
  - Wyszukiwanie w treści i w metadanych
  - Wyszukiwanie przez ludzi i przez systemy zautomatyzowane
  - Wykorzystanie zaawansowanych rozwiązań technicznych w celu poprawy jakości uzyskiwanych wyników



# Biblioteka cyfrowa – wymagania



- Udostępnianie i prezentacja
  - Ergonomia
  - Autoryzacja dostępu
  - Trwałość „punktu dostępu” do zasobu
  - Wiarygodność treści i metadanych



# [ Plan prezentacji



- Wielkopolska Biblioteka Cyfrowa
- Biblioteka cyfrowa – wymagania
- Biblioteka cyfrowa oparta o system dLibra
- Biblioteki cyfrowe w Polsce





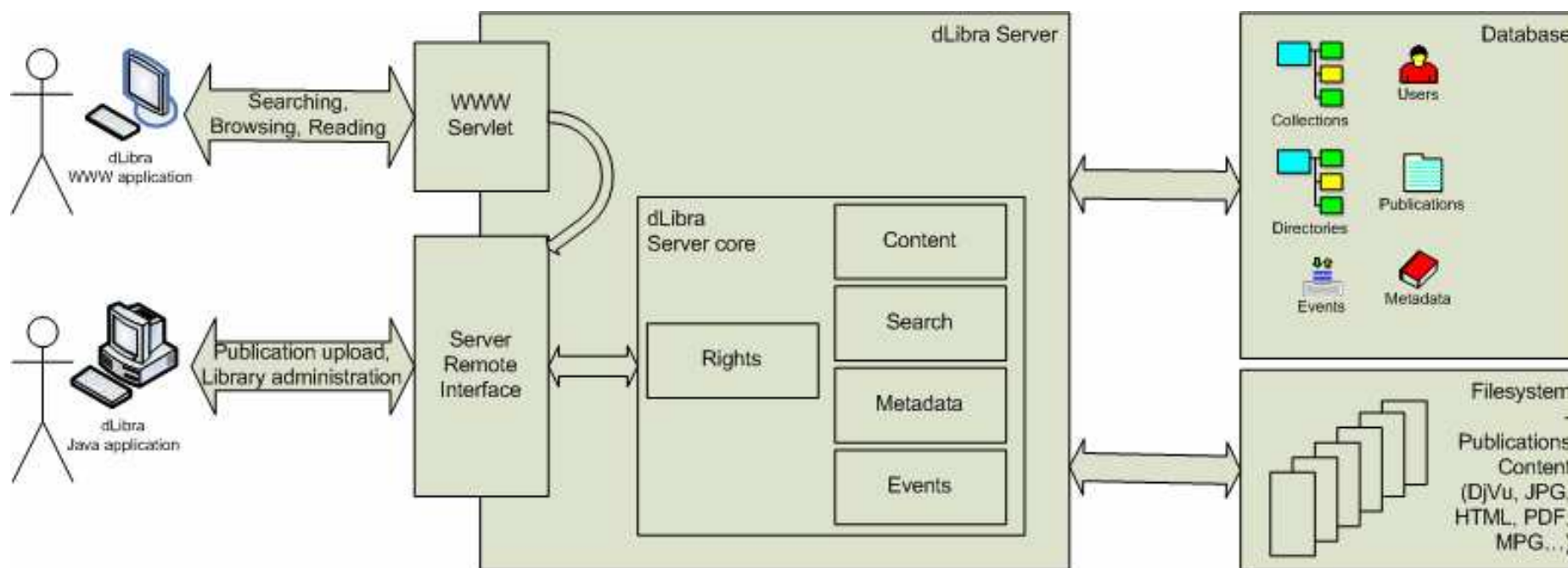
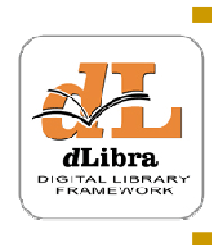
# [ Co to jest dLibra? ]



- Przenośny, rozproszony system do budowy bibliotek cyfrowych
- Rozwijany w PCSS od 1999 roku
- Używany w Wielkopolskiej Bibliotece Cyfrowej od 1.10.2002 r.



# Architektura systemu dLibra



# Cykl życia publikacji elektronicznej



- Powstanie postaci cyfrowej
- Konwersja do formatu odpowiedniego dla biblioteki cyfrowej
- Wprowadzenie publikacji do biblioteki cyfrowej
- Opis hasłowy
- Udostępnienie czytelnikowi
- Tworzenie kolejnych wersji publikacji



# Cykl życia publikacji elektronicznej



- Powstanie postaci cyfrowej
  - Skanowanie publikacji
  - lub
  - Tworzenie publikacji od razu w postaci cyfrowej



# Cykl życia publikacji elektronicznej



- Powstanie postaci cyfrowej
- Konwersja do formatu odpowiedniego dla biblioteki cyfrowej
- Wprowadzenie publikacji do biblioteki cyfrowej
- Opis hasłowy
- Udostępnienie czytelnikowi
- Tworzenie kolejnych wersji publikacji



# Cykl życia publikacji elektronicznej



- Konwersja do formatu odpowiedniego dla biblioteki cyfrowej
  - Format zależny od rodzaju udostępnianej treści
  - Format zależny od docelowego czytelnika
  - Każdy format ma swoje wady i zalety
  - Dla publikacji tekstowych: DjVu, PDF, HTML, XML



# Cykl życia publikacji elektronicznej



- Powstanie postaci cyfrowej
- Konwersja do formatu odpowiedniego dla biblioteki cyfrowej
- Wprowadzenie publikacji do biblioteki cyfrowej
- Opis hasłowy
- Udostępnienie czytelnikowi
- Tworzenie kolejnych wersji publikacji



# Wprowadzenie publikacji do biblioteki cyfrowej



- Struktura publikacji
  - Na dysku
    - Pliki
  - W bibliotece cyfrowej
    - Pliki
    - ???





# Wprowadzenie publikacji do biblioteki cyfrowej



- Wewnętrzna struktura biblioteki cyfrowej
  - Katalog
    - Główna jednostka organizacyjna
    - Może zawierać dowolną liczbę innych katalogów lub publikacji
    - W WBC dla celów porządkowych każda biblioteka ma własny, zabezpieczony katalog



# Wprowadzenie publikacji do biblioteki cyfrowej



- Katalog - przykład
  - Katalog główny
    - BK PAN
      - Regionalia
      - Testowe
    - PTPN
      - Regionalia
    - Akademia Muzyczna
      - Muzykalia



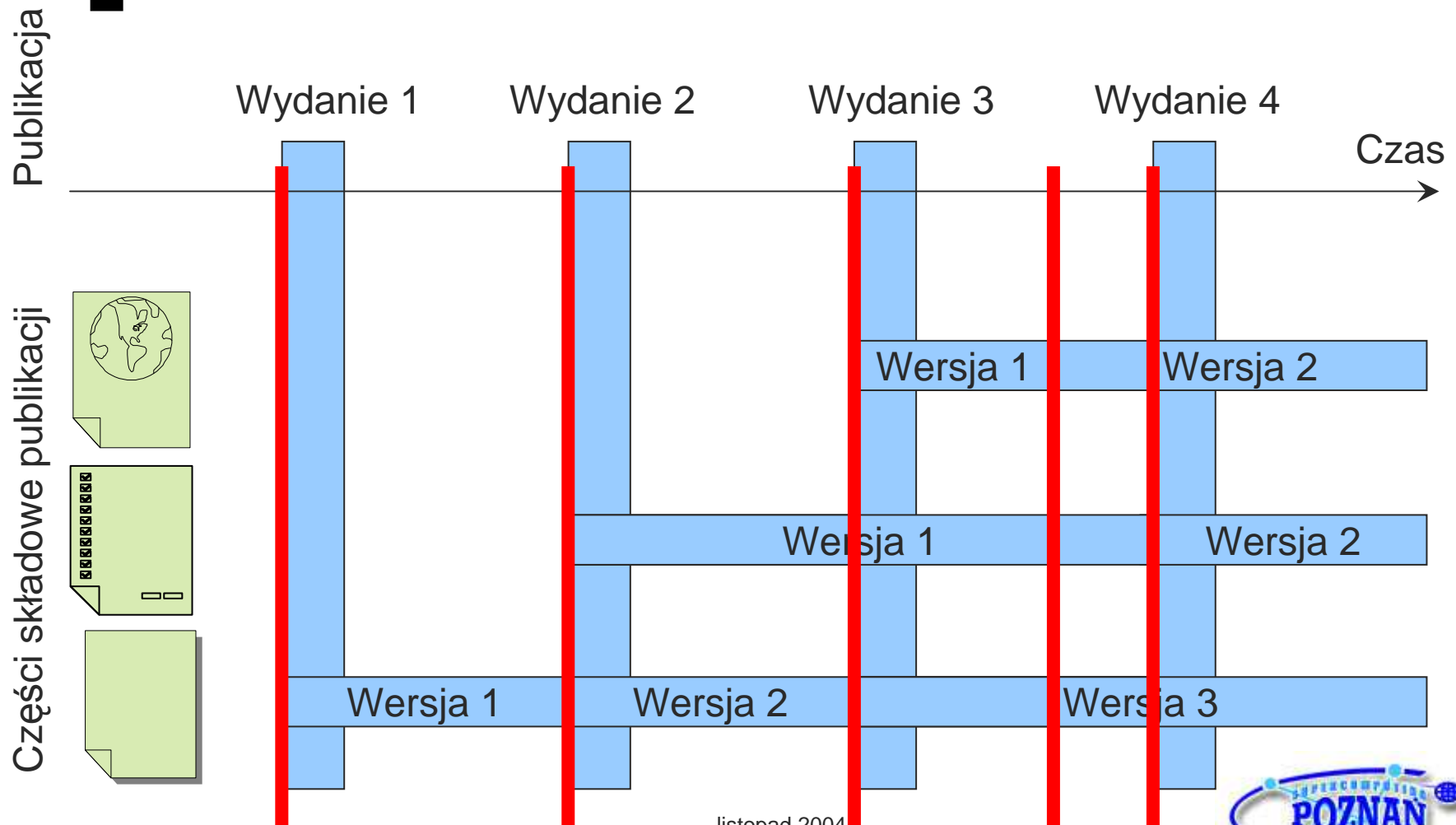
# Wprowadzenie publikacji do biblioteki cyfrowej



- Wewnętrzna struktura biblioteki cyfrowej
  - Publikacja
    - Reprezentuje pojedynczą publikację
    - Składa się z plików publikacji połączonych w wydania



# Wprowadzenie publikacji do biblioteki cyfrowej



# Wprowadzenie publikacji do biblioteki cyfrowej



- Wewnętrzna struktura biblioteki cyfrowej
  - Wydanie
    - Składa się z plików publikacji w określonych wersjach
    - Może być opublikowane bądź nie opublikowane, nie może zostać usunięte
    - W WBC większość publikacji ma tylko jedno wydanie (specyfika gromadzonych zasobów)

# Wprowadzenie publikacji do biblioteki cyfrowej



- Wewnętrzna struktura biblioteki cyfrowej
  - Plik
    - Reprezentuje najmniejszą część składową publikacji
  - Wersja pliku
    - Konkretna wersja – zawartość – danego pliku publikacji



# Wprowadzenie publikacji do biblioteki cyfrowej



- Wewnętrzna struktura biblioteki cyfrowej
  - Publikacja grupowa
    - Służy do łączenia powiązanych ze sobą publikacji
    - Ma strukturę drzewiastą, widoczną dla czytelnika



# Wprowadzenie publikacji do biblioteki cyfrowej



- Przykładowa publikacja grupowa
  - Dziennik Poznański (opis)
    - 1924 (opis)
      - Styczeń (opis)
        - Nr 1 (treść i opis)
        - Nr 2 (treść i opis)
      - Luty (treść i opis)
    - 1925 (opis)
    - 1926 (opis)





# Wprowadzenie publikacji do biblioteki cyfrowej



- Przykładowa publikacja grupowa
  - Biblia (opis)
    - Wydanie A (treść PDF i opis)
    - Wydanie B (opis)
      - Stary Testament (treść PDF i opis)
      - Nowy Testament (treść PDF i opis)
    - Wydanie C (treść HTML i opis)



# Cykl życia publikacji elektronicznej



- Powstanie postaci cyfrowej
- Konwersja do formatu odpowiedniego dla biblioteki cyfrowej
- Wprowadzenie publikacji do biblioteki cyfrowej
- Opis hasłowy
- Udostępnienie czytelnikowi
- Tworzenie kolejnych wersji publikacji



# [ Opis hasłowy



- Każde wydanie lub element publikacji grupowej mogą zostać opisane przy użyciu dowolnej liczby atrybutów (tytuł, autor, etc.)
- Każdy atrybut może mieć wiele wartości
- Każda publikacja może być opisana w wielu językach



# [ Opis hasłowy



- Administrator biblioteki zarządza używanym zestawem atrybutów
- Predefiniowany zestaw atrybutów zgodny ze standardem DublinCore Metadata Element Set ver. 1.1
  - tytuł, autor, temat i słowa kluczowe, opis, wydawca, współtwórca, data wydania, typ zasobu, format, identyfikator zasobu, źródło, języki, powiązania, zakres, prawa



# [ Opis hasłowy



- Opis może być tworzony automatycznie
  - Tworzony na podstawie informacji w systemie (prawa, format zasobu)
  - Importowany z zewnętrznych systemów
    - RDF (również eksport do RDF)
    - MARC
  - Dziedziczony z elementów nadrzędnych



# [ Opis hasłowy



- Dziedziczenie elementów opisu w publikacjach grupowych
  - Dziennik Poznański (DP, Poznań)
    - 1924 (*DP, Poznań, 1924*)
      - Styczeń (*DP, Poznań, 1924, styczeń*)
        - Nr 1 (*DP, Poznań, 1924, styczeń, ...*)
        - Nr 2 (*DP, Poznań, 1924, styczeń, ...*)
      - Luty (*DP, Poznań, 1924, luty*)
    - 1925 (*DP, Poznań, 1925*)
    - 1926 (*DP, Poznań, 1926*)



# [ Opis hasłowy



- Słowniki wartości atrybutów
  - Każdy atrybut ma słownik występujących w systemie wartości
  - Słownik zawiera mechanizm obsługi wyrazów bliskoznacznych
  - Zaawansowane mechanizmy zarządzania słownikiem



# [ Opis hasłowy



- **Inne elementy opisu**
  - **Widoczne dla czytelnika**
    - Skrócony opis
    - Komentarz
    - Miniaturka
  - **Wewnętrzne**
    - Uwagi administracyjne
    - Kto ostatnio modyfikował





# [ Kolekcje



- Struktura drzewiasta
- Tematyczny podział zebranych zasobów
- Mechanizm kolekcji zupełnie niezależny od katalogów
- Kolekcje są widoczne dla czytelnika – katalogi nie



# [ Kolekcje



- Publikacja może być przyporządkowana do wielu kolekcji
- Przynależność do kolekcji jest dziedziczona w strukturze publikacji grupowej



# [ Opis hasłowy



- Dziedziczenie przynależności do kolekcji w publikacjach grupowych
  - Dziennik Poznański (Dzienniki)
    - 1924 (*Dzienniki*)
      - 22.12.1924 r. (*Dzienniki*)
        - DP (*Dzienniki*)
        - Dodatek z kolędami (*Dzienniki*, Muzykalia)
    - 1925 (*Dzienniki*)



# Cykl życia publikacji elektronicznej



- Powstanie postaci cyfrowej
- Konwersja do formatu odpowiedniego dla biblioteki cyfrowej
- Wprowadzenie publikacji do biblioteki cyfrowej
- Opis hasłowy
- Udostępnienie czytelnikowi
- Tworzenie kolejnych wersji publikacji



# Udostępnienie czytelnikowi



- Rozbudowany system uprawnień
  - Prawa administracyjne – zarządzanie elementami biblioteki
    - Użytkownicy
    - Atrybuty
    - ...
  - Prawa do poszczególnych elementów biblioteki – przeglądanie, odczyt, modyfikacja, zarządzanie



# Udostępnienie czytelnikowi



- Użytkowników można łączyć w grupy
- Aby użytkownik miał dostęp do treści zasobu musi mieć prawo odczytu określonego wydania
- Domyślny użytkownik reprezentujący wszystkich czytelników



# Udostępnienie czytelnikowi



- Dostęp do publikacji może wymagać podania hasła lub być ograniczony do określonej grupy komputerów
- Publikacje HTML mogą być zabezpieczone przed kopiowaniem treści i wydrukiem



# Udostępnienie czytelnikowi



- **Możliwości czytelnika**
  - Przeglądanie zawartości biblioteki z podziałem na kolekcje
  - Przeszukiwanie ogólne
    - W treści
    - W opisie
    - W treści i opisie





# Udostępnienie czytelnikowi



- **Możliwości czytelnika**
  - Zaawansowane konstruowanie zapytań – rozszerzalny formularz z operatorami logicznymi i listą wartości poszczególnych atrybutów
  - Przeszukiwanie przy pomocy indeksów
    - Indeks autorów
    - Indeks tytułów
    - Indeks słów kluczowych



# Udostępnienie czytelnikowi



- Wykorzystanie słownika wartości atrybutów w celu polepszenia wyników wyszukiwania
  - Opcjonalne wykorzystanie synonimów do rozszerzania zapytań
  - Działa najlepiej przy wykorzystaniu *rozbudowanego* słownika wartości



# Wykorzystanie słownika...



# Wykorzystanie słownika...



Zapytanie  
użytkownika



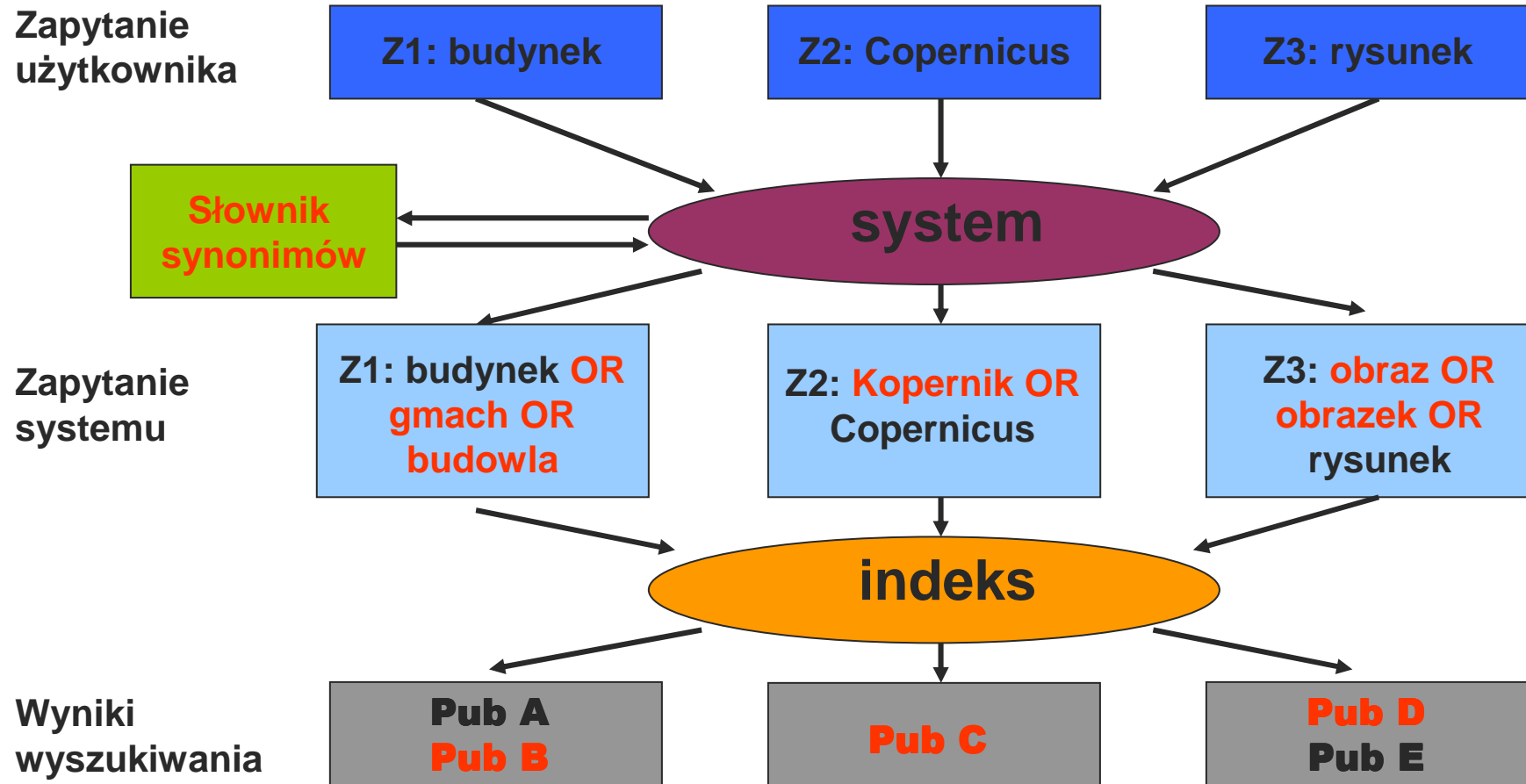
Zapytanie  
systemu



Wyniki  
wyszukiwania



# Wykorzystanie słownika...



# Wykorzystanie słownika...



- Niejednoznaczność wartości atrybutów

- W ramach tego samego atrybutu

Słownik synonimów

*klódka*

zamek

*zamek*

pałac

- W ramach różnych atrybutów

Słownik synonimów

*Kapusta*

Prokurator  
Zygmunt K.

*kapusta*

brukselka  
bigos



# Udostępnienie czytelnikowi



- **Możliwości czytelnika**
  - Przeszukiwanie przy pomocy zewnętrznych systemów
    - Czytelnik może nie wiedzieć o istnieniu biblioteki, ale poszukiwać materiałów w niej zgromadzonych
    - „Ukryte zasoby Internetu”
    - Wymaga specjalnego wsparcia ze strony biblioteki cyfrowej



# Udostępnienie czytelnikowi



## ■ Możliwości czytelnika

### ○ Przeszukiwanie przy pomocy zewnętrznych systemów

#### ■ Przykład – wg Google (z dnia 31.10.2004 r.):

- WBC – 1 340 publikacji – około 4 200 wyników (2 890 w języku polskim)
- PBI – 10 506 publikacji – 385 wyników (382 w języku polskim)





# Udostępnienie czytelnikowi



- **Możliwości czytelnika**
  - Trwały punkt dostępu do zasobów
    - **Treść wydania**
      - <http://www.wbc.poznan.pl/Content/1234/>
    - **Publikacja**
      - <http://www.wbc.poznan.pl/publication/123>



# Udostępnienie czytelnikowi



- **Możliwości czytelnika**
  - Powiadamianie o nowościach (wątki RSS)
  - Wielojęzyczny interfejs
  - Rankingi publikacji oraz bibliotek i statystyki dostępu do biblioteki



# Udostępnienie czytelnikowi



- **Możliwości czytelnika**
  - Lista planowanych publikacji
  - Pobieranie treści całej publikacji w postaci archiwum ZIP
  - Przeglądanie katalogu całej biblioteki bez konieczności dostępu do Internetu



# [ Plan prezentacji



- Wielkopolska Biblioteka Cyfrowa
- Biblioteka cyfrowa – wymagania
- Biblioteka cyfrowa oparta o system dLibra
- Biblioteki cyfrowe w Polsce



# Biblioteki cyfrowe w Polsce



- dLibra
- ▲ 2002 (WBC)
  - ▲ 2004
  - ▲ 2005, 2006?
  - ◆ Inne systemy BC



listopad 2004

# Biblioteki cyfrowe w Polsce



- Powstawanie nowych bibliotek cyfrowych
  - Elektroniczne filie „tradycyjnych” bibliotek
    - Najczęściej biblioteki akademickie
    - Potrzeba wsparcia bibliotek od strony technicznej
  - Specjalizowane biblioteki tematyczne (muzyczne, medyczne itp.)
  - Biblioteki specjalne (biblioteka dla niewidomych)



# [ Adresy



- Projekt dLibra
  - <http://dlibra.psnc.pl/>
  - [dlibra@man.poznan.pl](mailto:dlibra@man.poznan.pl)
- Wielkopolska Biblioteka Cyfrowa
  - <http://www.wbc.poznan.pl/>

Dziękuję za uwagę

