

# Biblioteka cyfrowa jako otwarte, internetowe repozytorium publikacji



Marcin Heliński, Cezary Mazurek,  
Tomasz Parkoła, Marcin Werla

Poznańskie Centrum  
Superkomputerowo-Sieciowe



# Biblioteka cyfrowa



- Podstawowe wymagania funkcjonalne:
  - Umieszczanie i przechowywanie obiektów cyfrowych wraz z metadanymi
  - Konserwacja obiektów cyfrowych oraz metadanych oraz zarządzanie nimi
  - Udostępnianie i prezentacja przechowywanych obiektów cyfrowych i metadanych



# Biblioteka cyfrowa

- Wymagania pozafunkcjonalne – system musi być:
  - Bezpieczny
  - Wydajny i skalowalny
  - Ergonomiczny i efektywny

# Biblioteka cyfrowa



- Typowe rozwiązania – wariant 1 („Strony WWW” ):
  - Przechowywanie:
    - Treść: na serwerze WWW
    - Metadane: na serwerze WWW, w postaci stron HTML
  - Konserwacja
    - Ręczna – trzeba samodzielnie przetwarzać wszystkie publikacje i metadane
    - Brak kompleksowych mechanizmów zarządzania
  - Udostępnianie i prezentacja
    - Przez strony WWW
    - Brak kompleksowych mechanizmów kontroli dostępu
    - Brak zaawansowanych mechanizmów wyszukiwawczych



# Biblioteka cyfrowa



- Typowe rozwiązania – wariant 2 („Katalog obiektów cyfrowych”):
  - Przechowywanie:
    - Treść: na serwerze WWW
    - Metadane: w dedykowanym oprogramowaniu katalogowym
  - Konserwacja
    - Ręczna treści, automatyczna metadanych
    - Istnieją kompleksowe mechanizmy zarządzania metadanymi
  - Udostępnianie i prezentacja
    - Przez strony WWW
    - Brak kompleksowych mechanizmów kontroli dostępu do treści, dostępne dla metadanych
    - Brak zaawansowanych mechanizmów wyszukiwawczych dla treści, dostępne dla metadanych



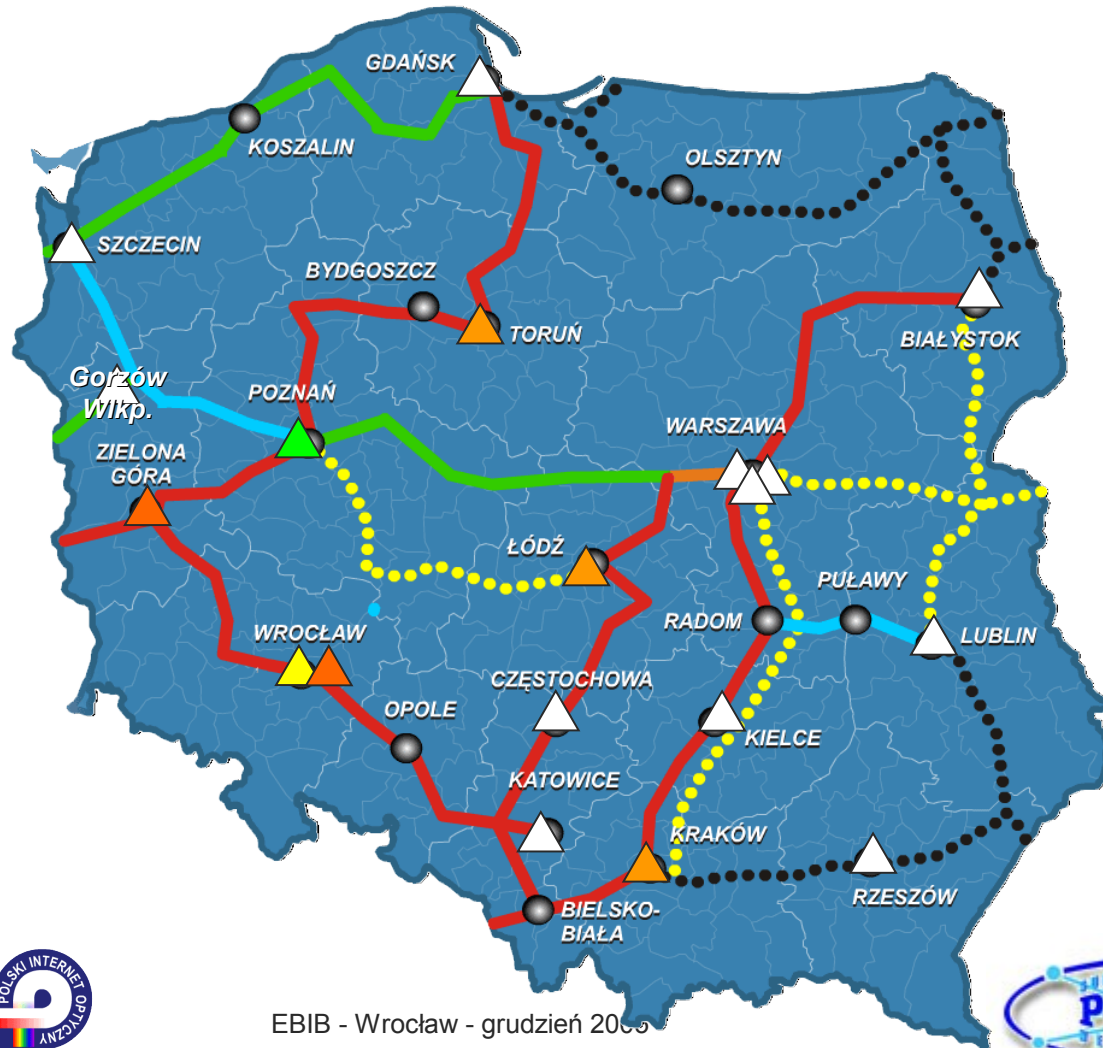
# Biblioteka cyfrowa

- Typowe rozwiązania – wariant 3 („Biblioteka cyfrowa”):
  - Przechowywanie:
    - Treść i metadane: w dedykowanym systemie biblioteki cyfrowej
  - Konserwacja
    - W pełni automatyczna
    - Dostępne kompleksowych mechanizmy zarządzania treścią i metadanymi
  - Udostępnianie i prezentacja
    - Przez strony WWW
    - Dostępne kompleksowe mechanizmów kontroli dostępu
    - Dostępne zaawansowane mechanizmy wyszukiwawcze dla treści i metadanych

# Instalacje systemu dLibra

dLibra

- ▲ 2002 (WBC)
- ▲ 2004
- ▲ 2005
- ▲ 2006?





# Wyszukiwanie zasobów rozproszonych w systemie dLibra

- Umożliwia przeszukiwanie opisów zasobów z wszystkich połączonych bibliotek cyfrowych równocześnie
- Wykorzystuje podejście typu P2P
- Do wymiany metadanych wykorzystywany jest protokół OAI-PMH





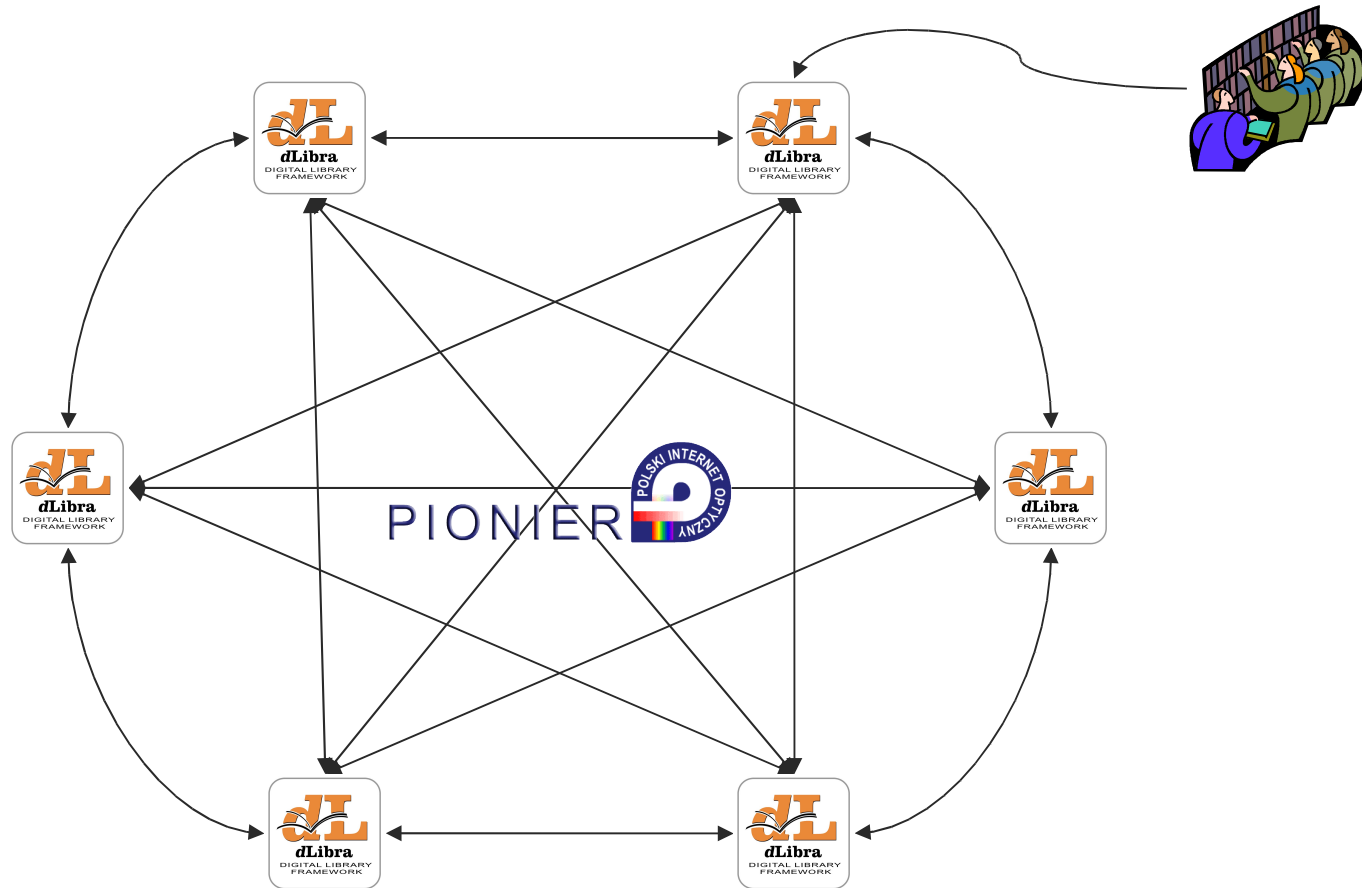
# Protokół OAI-PMH

- **Open Archives Initiative**  
– **Protocol for Metadata Harvesting**
- <http://www.openarchives.org>
- Otwarty, popularny na świecie, nie ogranicza do określonych technologii





# Wyszukiwanie zasobów rozproszonych w systemie dLibra





# Wyszukiwanie zasobów rozproszonych w systemie dLibra

- Każda biblioteka decyduje, jakie inne biblioteki indeksować
- Biblioteki są indeksowane okresowo, w sposób przyrostowy
  - Niezbędne są informacje o usuniętych publikacjach
- Proces przeszukiwania odbywa się lokalnie



# Wyszukiwanie zasobów rozproszonych w systemie dLibra

- Takie podejście:
  - Pozwala zrównoważyć obciążenie poszczególnych bibliotek (unikamy efektu KARO)
  - Jest możliwe przy indeksowaniu metadanych ze względu na ich niewielki rozmiar

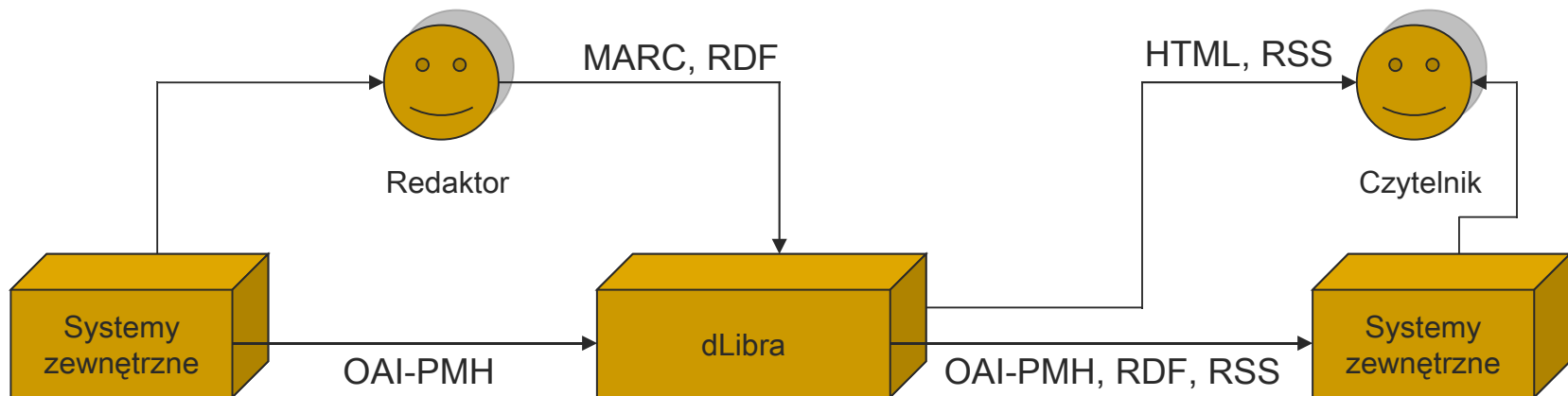
# Opis hasłowy

- Każda publikacja cyfrowa może być opisana przy użyciu dowolnej liczby atrybutów (tytuł, autor, etc.)
- Atrybuty mogą się dzielić na podatrybuty (np. identyfikator na ISBN, DOI, URI itd.)

# Opis hasłowy

- Każda biblioteka udostępnia opis przechowywanych obiektów w dwóch schematach:
  - DublinCore wersja 1.1 (wymiana danych z zewnętrznymi systemami)
  - Schemat własny biblioteki

# 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:

- Usługa MetadataServer
- 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

# Dodatkowe informacje

- Plakat i pokaz systemu **dLibra 2.2** w sesji posterowej
- Projekt dLibra
  - <http://dlibra.psnc.pl/>
  - [dlibra@man.poznan.pl](mailto:dlibra@man.poznan.pl)

Dziękuję za uwagę