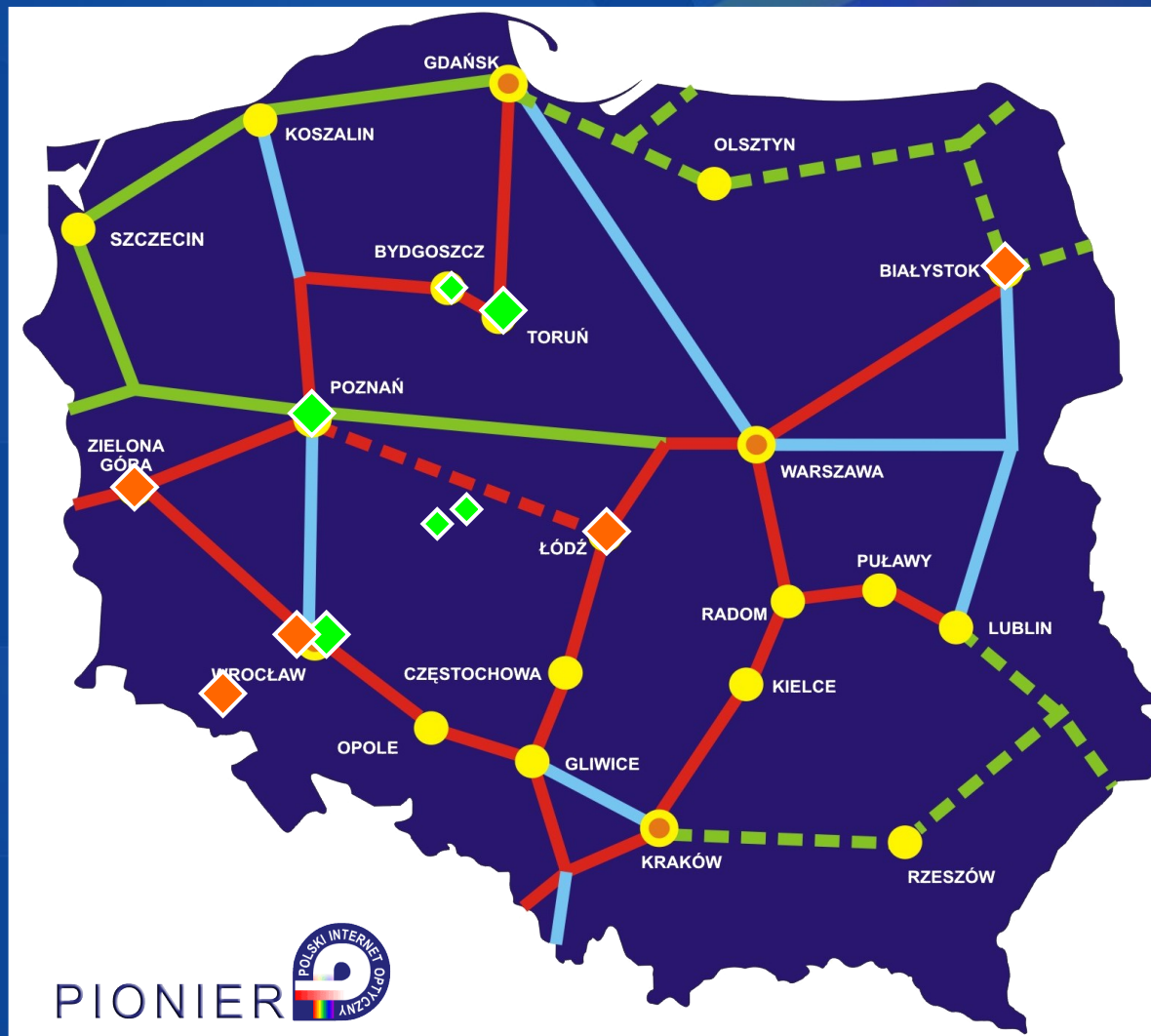


Wdrażanie regionalnych bibliotek cyfrowych w sieci PIONIER w oparciu o środowisko dLibra

Cezary Mazurek, Maciej Stroiński, Marcin Werla
{mazurek,stroins,mwerla}@man.poznan.pl
IChB PAN PCSS

dLibra w sieci PIONIER



Instalacje dLibry w sieci PIONIER:

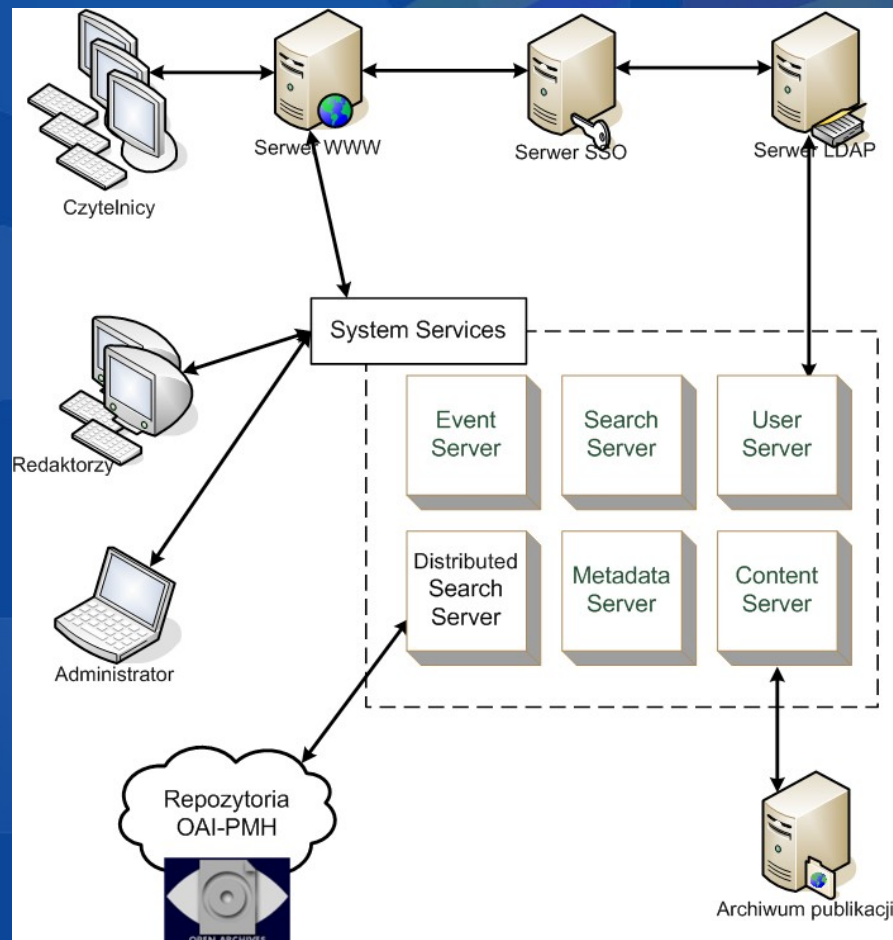
- ◆ Uruchomione
- ◆ W fazie konfiguracji

26 września 2005

Architektura systemu dLibra

Zalety:

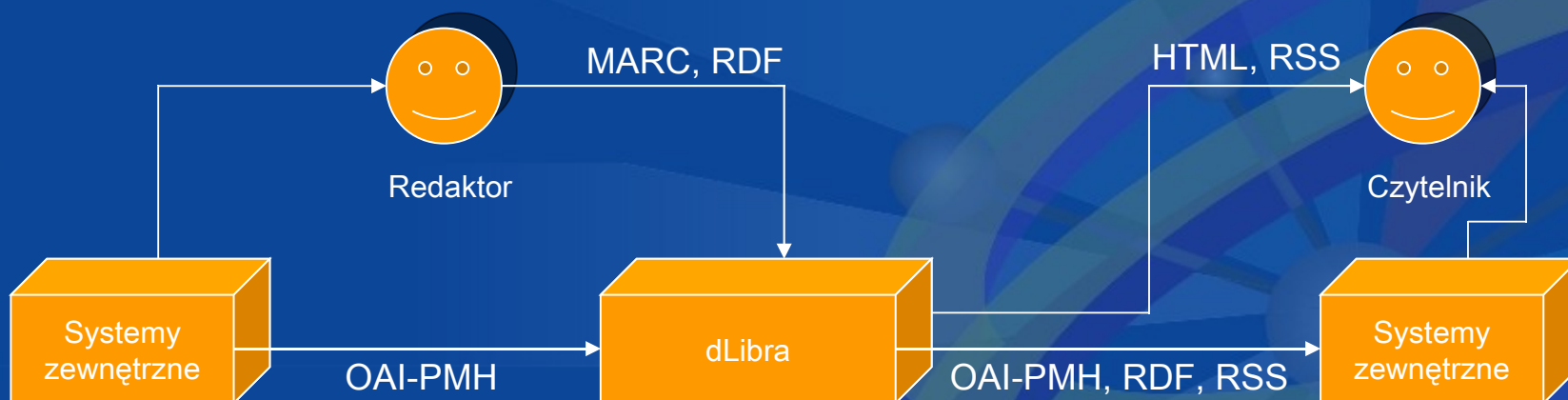
- Niezależność
 - od systemu operacyjnego – Java™
 - od rodzaju bazy danych - Hibernate
- Skalowanie i niezawodność
 - System rozproszonych usług
 - Zaawansowane mechanizmy dynamicznej konfiguracji i kontroli (m.in. Java™ JMX)
- Łatwość integracji
 - XML Schema, RSS, RDF, OAI-PMH
- Łatwość rozbudowy
 - Zestaw interfejsów opisujących usługi



Metadane w systemie dLibra

- Dowolny zestaw atrybutów, którymi można opisywać gromadzone obiekty np. DublinCore czy MARC
- Hierarchia atrybutów – wsparcie np. dla kwalifikatorów w DublinCore
- Słowniki wartości atrybutów z obsługą synonimów
- Współpraca z zewnętrznymi systemami
- Obsługa wielu języków

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

Przechowywanie treści cyfrowych

- Za przechowywanie odpowiada usługa ContentServer
- Możliwość przechowywania dowolnego typu obiektów
- Możliwość pojedynczego lub wsadowego umieszczania publikacji
- Wersjonowanie umieszczanej treści
- Statystyki dostępu do poszczególnych obiektów cyfrowych

Przechowywanie treści cyfrowych

- Zabezpieczenia przed utratą danych
 - Automatyczne tworzenie kopii bezpieczeństwa publikacji w postaci archiwum ZIP w lokalnym systemie plików lub na zdalnym serwerze (FTP, planowany GridFTP)
 - Struktury używane do przechowywania treści dodatkowym zabezpieczeniem przed awarią bazy danych
- Opcjonalna możliwość odczytu treści z pominięciem usługi ContentServer – możliwa tylko dla usług, zwiększa wydajność

Kontrola dostępu do treści

- Rozbudowany system praw
 - 6 rodzajów praw administracyjnych
 - 7 rodzajów praw do katalogu
 - 3 rodzaje praw do pojedynczej publikacji
- Dziedziczenie uprawnień z obiektów nadrzędnych i grup
- Uprawnienia implikowane
- Zabezpieczenie treści przed kopiowaniem (HTML, planowane PDF i DjVu)

Kontrola dostępu do treści

- Sposoby autoryzacji
 - Nazwa użytkownika i hasło
 - Przechowywane w usłudze UserServer
 - W oparciu o serwer(y) LDAP
 - W oparciu o system Single Sign-On (CAS, planowany Shibboleth)
 - Nazwa użytkownika i adres (zakres adresów) IP - przyznanie lub odmowa dostępu

Kontrola dostępu do treści

- Klasy użytkowników
 - Czytelnik niezarejestrowany
 - Każdy kto przegląda strony biblioteki cyfrowej
 - Czytelnicy zarejestrowani
 - Czytelnik posiadający konto w bibliotece
 - Redaktorzy/Administratorzy
 - Dynamiczne grupy LDAP
 - Grupy składające się z użytkowników LDAP posiadających pewne właściwości (np. student 4 roku kierunku informatyka)

Dalsze kierunki rozwoju

- Rozproszone wyszukiwanie
 - W opisach bibliograficznych (wkrótce zostanie udostępnione)
 - W treści publikacji
- Lepsza integracja z wyszukiwarkami internetowymi
- Integracja z systemami gridowymi
- Mechanizmy wykrywania duplikatów m.in. w oparciu o protokół OAI-PMH

Dalsze kierunki rozwoju

- Rozbudowa mechanizmu słowników, import słowników z systemów bibliotecznych
- Rozbudowa funkcji związanych z przechowywaniem metadanych w formacie MARC
- Obsługa protokołu Z39.50 (obecnie możliwa przez mapowanie interfejsu OAI-PMH)
- Rozproszony system identyfikatorów obiektów cyfrowych

Dalsze kierunki rozwoju

- Analiza zachowań czytelników
- Zwiększenie ergononii interfejsów użytkownika
 - Funkcja „czy chodziło Ci o...”
 - Podpowiadanie wyrażeń ze słownika
- Spójny rozproszony system obsługi profili czytelników

Dziękuję za uwagę!

Kontakt:

dlibra@man.poznan.pl