

# POZNAŃSKIE CENTRUM SUPERKOMPUTEROWO-SIECIOWE



**AFILIOWANE PRZY INSTYTUCIE CHEMII BIOORGANICZNEJ PAN**

---

**h t t p : / / w w w . m a n . p o z n a n . p l**



## Atomowe usługi w środowisku rozproszonych bibliotek cyfrowych

Cezary Mazurek,  
Tomasz Parkoła, Marcin Werla

---

# Biblioteki cyfrowe

„ ...łączą gromadzenie i organizowanie informacji, które od dawna prowadzone jest przez biblioteki i archiwa z cyfrową prezentacją tych informacji możliwą dzięki komputerom”

- Michael Lesk

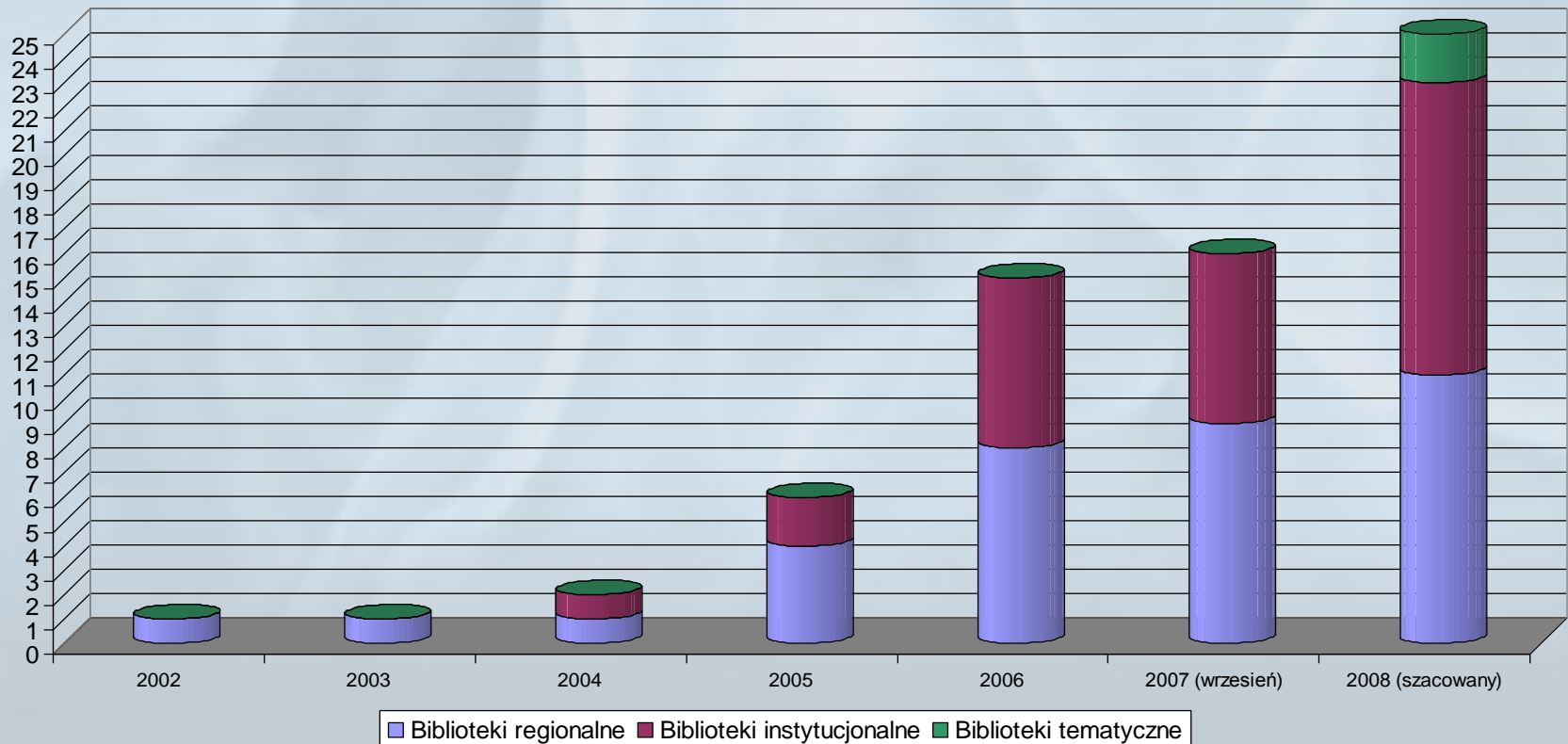
*„Practical digital libraries:  
Books, bytes and bucks”*









# Biblioteki cyfrowe w Polsce

- Badania prowadzone od 1996 roku
- Od 1999 roku budowany jest system dLibra
- Obecnie w sieci PIONIER funkcjonuje 16 publicznie dostępnych bibliotek cyfrowych – regionalnych i instytucjonalnych

# Liczba instalacji bibliotek cyfrowych dLibra



# Biblioteki cyfrowe dLibra

- dLibra
-  2002
  -  2004
  -  2005
  -  2006
  -  do 05.2007
  -  od 06.2007

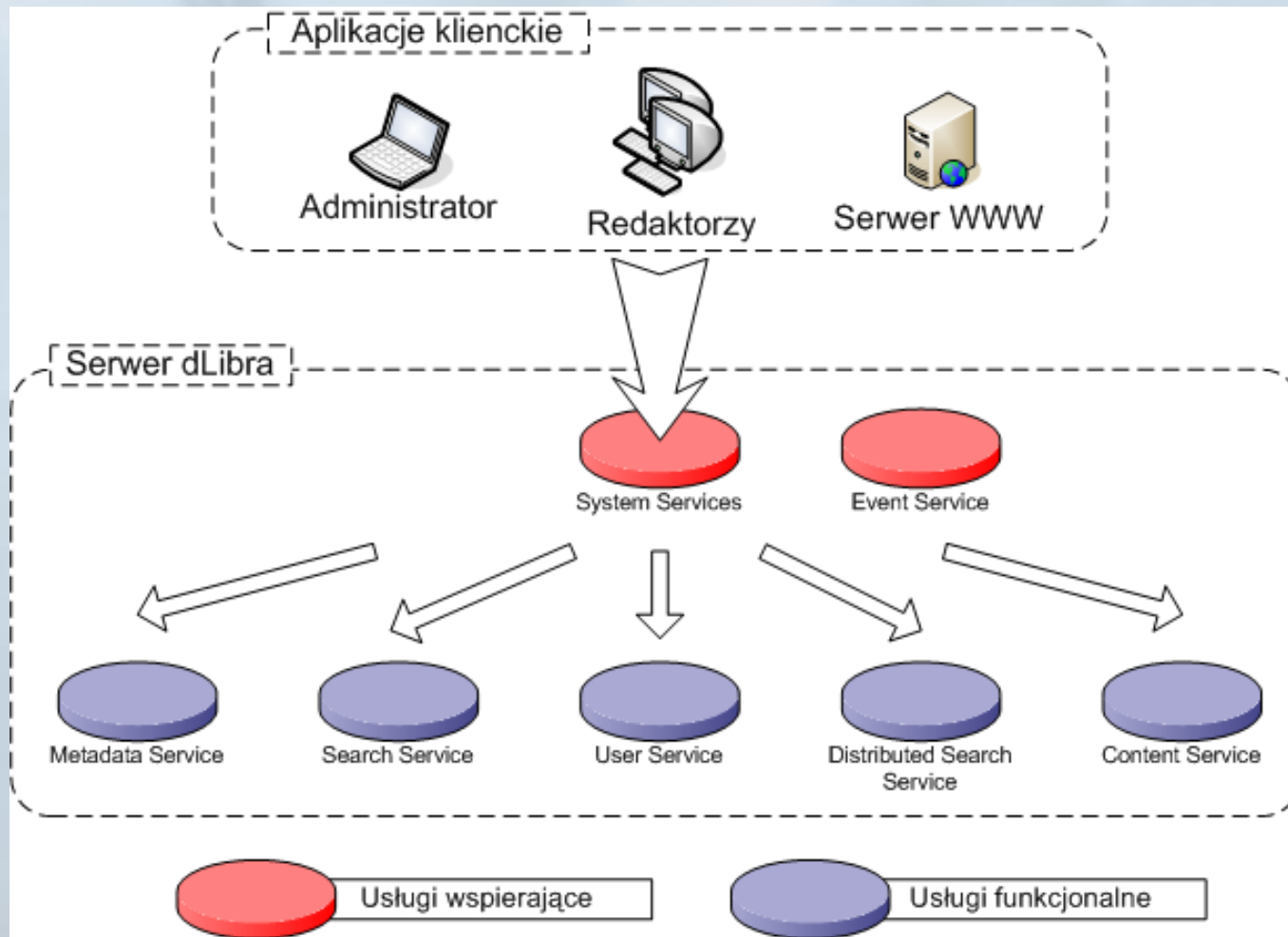




# Architektura systemu dLibra

- Wielowarstwowa
- Zorientowana na usługi
- Dwa typy usług:
  - funkcjonalne
  - wspierające
- Implementacja w języku Java

# Architektura systemu dLibra





# Architektura systemu dLibra

- Usługi funkcjonalne:
  - Metadata Service – przechowuje metadane obiektów cyfrowych, informacje o strukturze biblioteki cyfrowej, kolekcjach itp.
  - Content Service – przechowuje treść obiektów cyfrowych
  - Search Service – indeksuje i przeszukuje treść (tekst) i metadane obiektów cyfrowych

# Architektura systemu dLibra

- Usługi funkcjonalne:
  - Distributed Search Service – pozwala na indeksowanie i przeszukiwanie obiektów cyfrowych ze zdalnych repozytoriów
  - User Service – przechowuje informacje o użytkownikach systemu i ich uprawnieniach

# Architektura systemu dLibra

- Usługi wspierające:
  - System Service – pośredniczy w dostępie do wszystkich pozostałych usług
  - Event Service – odpowiada za asynchroniczne przetwarzanie zdarzeń w systemie

# Atomowe usługi bibliotek cyfrowych

- Motywacja:
  - Wykorzystanie funkcjonalności bibliotek cyfrowych w innych systemach
    - Telemedycyna
    - E-learning
    - Wirtualne laboratoria
    - ...
  - Wzrost liczby bibliotek cyfrowych w sieci PIONIER
  - Rosnąca rola repozytoriów cyfrowych w e-Science



# Atomowe usługi bibliotek cyfrowych

- Efektem analizy funkcjonalnej wybranych modeli i systemów bibliotek cyfrowych było wyróżnienie 4 usług:
  - Usługa przechowywania obiektów cyfrowych
  - Usługa przechowywania metadanych
  - Usługa kompozycji obiektów cyfrowych i tworzenia powiązań między nimi
  - Usługa wyszukiwania obiektów cyfrowych

# Usługa przechowywania obiektów cyfrowych

- Umożliwia przechowywanie i wersjonowanie obiektów cyfrowych dowolnego typu
- Typ obiektu istotny dopiero w kontekście konkretnej aplikacji
- Możliwość implementacji np. w oparciu o:
  - serwer FTP/gridFTP
  - serwer HTTP (WebDAV)
  - bazę danych
  - serwer wersjonowania (np. CVS, SVN)

# Usługa przechowywania metadanych

- Przechowywanie metadanych powiązanych z obiektami cyfrowymi
- Dla każdego obiektu
  - dowolna liczba zestawów metadanych
  - dowolne zróżnicowanie (formatu) poszczególnych zestawów metadanych
- Implementacja np. w oparciu o XML i powiązane z nim standardy (XQuery, XPath)

# Usługa kompozycji obiektów cyfrowych

- Tworzenie powiązań pomiędzy obiektami cyfrowymi
- Tworzenie grup obiektów cyfrowych
- Może służyć do
  - budowania złożonych obiektów cyfrowych
  - budowania kolekcji obiektów cyfrowych
  - budowania ontologii w oparciu o obiekty cyfrowe



# Usługa wyszukiwania obiektów cyfrowych

- Umożliwia przeszukiwanie (wcześniej zdefiniowanego czy zaindeksowanego) zbioru obiektów cyfrowych
- Na bazie przesłanego zapytania generuje listę referencji do obiektów spełniających to zapytanie
- Usługa o największym możliwym zróżnicowaniu spośród zdefiniowanych wcześniej
  - Różne języki zapytań
  - Różne rodzaje przeszukiwanych formatów i informacji
- Stworzenie uniwersalnej implementacji nie jest obecnie możliwe

# Wymagania pozafunkcjonalne

- Dobór instancji usług w kontekście konkretnej aplikacji może mieć dodatkowo charakter pozafunkcjonalny związany np. z:
  - maksymalnym rozmiarem składowanych obiektów cyfrowych
  - wydajnością przetwarzania
  - bezpieczeństwem komunikacji i przetwarzania danych
  - ...

# Usługi atomowe w środowisku dLibra

- Mapowanie usług atomowych na usługi w systemie dLibra:
  - Przechowywanie obiektów cyfrowych – Content Service
  - Przechowywanie metadanych – Metadata Service
  - Kompozycja obiektów cyfrowych – Metadata Service
  - Wyszukiwanie – Search Service

# Usługi atomowe w środowisku dLibra

- Inne zmiany
  - Zmiana roli System Service z pośrednika do brokera informacji o dostępnych usługach
    - Typ usługi
    - Adres usługi
    - Dostępna (tak/nie)
  - Aplikacje klienckie w systemie dLibra stają się koordynatorami usług w rozumieniu SOA



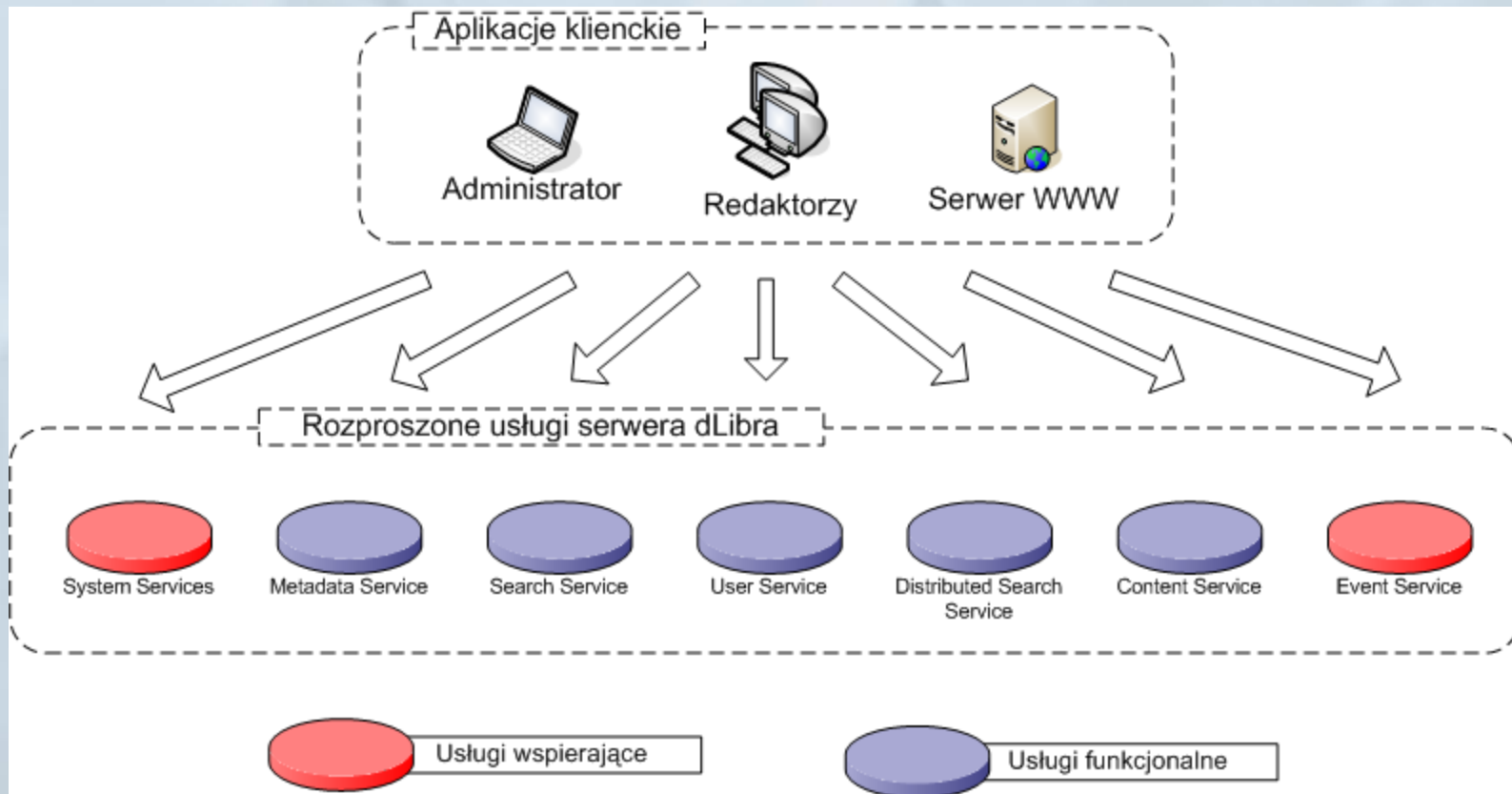
# Usługi atomowe w środowisku dLibra

- Inne zmiany
  - Uspójnienie podziału funkcjonalnego pomiędzy Search Service i Distributed Search Service
    - Pozwoliło na uruchomienie serwisu Federacja Bibliotek Cyfrowych opartego o dwie instancje nowej postaci usługi Distributed Search Service



<http://fbc.pionier.net.pl/>

# Usługi atomowe w środowisku dLibra



## Dalsze prace

- Wybór otwartych protokołów komunikacyjnych pozwalających na wykorzystanie poszczególnych usług
- Określenie podstawowych parametrów pozafunkcyjnych dla poszczególnych usług i zasad doboru tych usług na potrzeby konkretnych aplikacji



Atomowe usługi w środowisku  
rozproszonych bibliotek cyfrowych

Dziękuję za uwagę!

---