Distributed Digital Libraries Platform In the PIONIER Network

Cezary Mazurek, Tomasz Parkoła, Marcin Werla

Poznan Supercomputing and Networking Center
Noskowskiego 12/14, 61-704 Poznań, POLAND
PIONIER NREN: http://www.pionier.gov.pl/, e-mail: office@pionier.gov.pl
dLibra Framework: http://dlibra.psnc.pl/, e-mail: dlibra@man.poznan.pl

1 Content-based Services

The national programme "PIONIER - Polish Optical Internet -Advanced Applications, Services and Technologies for the Information Society" was started in Poland in 2001. One of its main objectives was to enrich the content-based services in Polish NREN. In order to reach this goal a platform composed of several digital library installations has been started up. All of them are based on dLibra – digital library framework developed by Poznan Supercomputing and Networking Center.

Periodic metadata synchronization based on the OAI-PMH protocol is the core functionality applied to digital libraries platform in the Polish Optical Internet PIONIER network. This process is based on selective (incremental) metadata harvesting and requires information about deleted records for consistency. Custom collections defined in digital libraries are transformed to OAI-PMH sets and custom metadata schemas are transformed to DCMES attributes. As a result all PIONIER digital libraries create one distributed platform where each digital library has full information about all metadata in the platform. Each digital library uses this metadata to provide its users with means to access any resource from the digital libraries platform. It also allows to create new advanced content and metadata-based services such as distributed metadata search and virtual dynamic collections.

2 Distributed Metadata Search

Each digital library indexes metadata harvested from remote libraries and allows users to search through the indexed metadata. The entire process is performed transparently to the end user and, as a result, the user receives one list of search results with both local and remote resources which matched the user query. The distributed searching functionality is easily accessed by external network services and information portals with the RSS protocol. Such integration is very useful especially in Web 2.0 services dynamically aggregating content from various distributed sources.
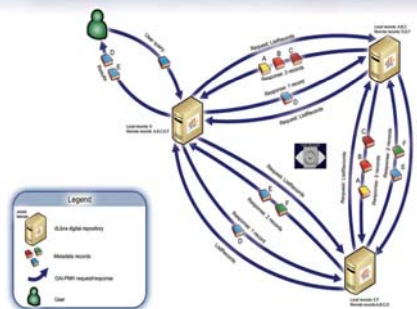

3 Virtual Dynamic Collections

Virtual dynamic collections group elements from different digital libraries across the platform. Those collections are defined by users as conditions that should be met by digital objects metadata. When a new digital object is published, it propagates through all digital libraries and all defined virtual collections. Such mechanism does not require any additional work of either the digital library administrator or the editor, in contradiction to static collections where all resources must by explicitly assigned. The virtual dynamic collections mechanism is based on RSS feeds. It is accessed from various user-selected external applications, portals and services. Such approach significantly increases the visibility of digital objects in the digital libraries platform.
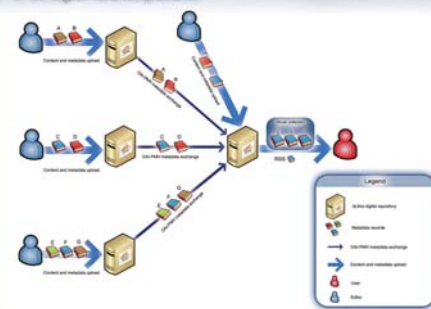
4 OAI-identifier

It is a platform based on multiple digital libraries that enables a system of unique and persistent digital object identifiers for all digital objects available in this platform. It is necessary for the platform consistency and may be a basis for new platform services like identifier resolvers. It was decided, that the unique identifiers system should be based on the OAI-identifier syntax. Therefore each identifier is composed of the "oai" scheme, namespace identifier which is the domain name for a given digital library and a local identifier which is unique in the scope of the specific digital library. Such identifiers are used to identify digital objects in communication between digital libraries, and are also used by readers as pointers to specific digital objects

Unique identifiers of objects in the Distributed Digital Libraries Platform
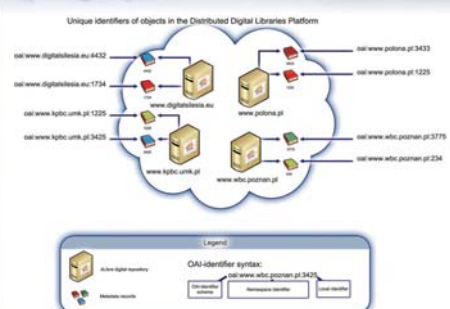
5 dLibra installations

Currently in the PIONIER network there are 9 regional and 7 institutional dLibra-based digital libraries and at least five more are planned to be available before the end of 2007. All these libraries together create a platform of distributed digital libraries with almost 65 000 digital objects (writing relicts, cultural heritage, regional and educational materials) and an average number of over 1 000 concurrent users.

Regional digital libraries:

Digital Library of Wielkopolska, Poznań
http://www.wbc.poznan.pl/

Dolnoslaska Digital Library, Wrocław
http://www.dbc.wroc.pl/

Jeleniogórska Digital Library, Jelenia Góra
http://jbc.jelenia-gora.pl/

Kujawsko-Pomorska Digital Library, Toruń
http://kpbc.umk.pl/

Małopolska Digital Library, Kraków
http://mbc.malopolska.pl/

Podlaska Digital Library, Białystok
http://pbc.biaman.pl/

Silesian Digital Library, Katowice
http://www.digitalsilesia.eu/

Zielonogórska Digital Library, Zielona Góra
http://zbc.uz.zgora.pl/

Wejherowska Digital Library, Wejherowo
http://biblioteka.wejherowo.pl/dlibra/

Institutional digital libraries:

Digital Library of The Federation
of the Church Libraries FIDES, Warszawa
http://digital.fides.org.pl/

Digital Library of The National In-Service Teacher
Training Center, Warszawa
http://bc.codn.edu.pl/

Lodz University of Technology Digital Library
http://ebipol.p.lodz.pl/

National Digital Library "Polona", Warszawa
http://www.polona.pl/

Pedagogical Digital Library, Krakow
http://www.ap.krakow.pl/dlibra/

Warsaw University of Technology Digital Library
http://bcpw.bg.pw.edu.pl/

Wroclaw University Digital Library
http://www.bibliotekacyfrowa.pl/