



Integracja systemu dLibra i Manuscriptorium

Marcin Werla, PCSS
mwerla@man.poznan.pl

ENRICH PROJECT



Motywacja

1. Konieczność manualnego przesyłania rekordów opisujących manuskrypty z BC UWr do Manuscriptorium
2. Odrębny interfejs prezentacyjny Manuscriptorium oparty na plikach JPG
3. Odrębny format opisu publikacji - MASTER

Motywacja

1. Konieczność manualnego przesyłania rekordów opisujących manuskrypty z BC UWr do Manuscriptorium
 - Manuscriptorium pobiera dane dostępne przez protokół OAI-PMH lub przesłane manualnie
 - Protokół OAI-PMH pozwala na wybór pobieranych rekordów jedynie na podstawie ich daty modyfikacji bądź przynależności do kolekcji (zbiorów)
 - Kryterium daty modyfikacji przydaje się przy regularnej aktualizacji informacji o bibliotece cyfrowej
 - Kryterium przynależności do kolekcji nie zawsze jest wystarczające aby wydzielić interesujący zbiór obiektów

Biblioteka	
▼	BCUWr
▼	Dziedzictwo kulturowe
>	Czasopisma
>	Dokumenty Życia Społecznego
>	Książki wydane w latach 1800 do 1945
>	Rękopisy
>	Stare Druk
>	Zbiory Bibliologiczne
▶	Zbiory Graficzne
▶	Zbiory Kartograficzne
>	Zbiory Muzyczne
▶	e-Książk
>	Inne
▶	Katalogi BUWr
▶	Kolekcje specjalistyczne
▶	Materiały edukacyjne
▶	Publikacje współczesne
▶	Regionalia



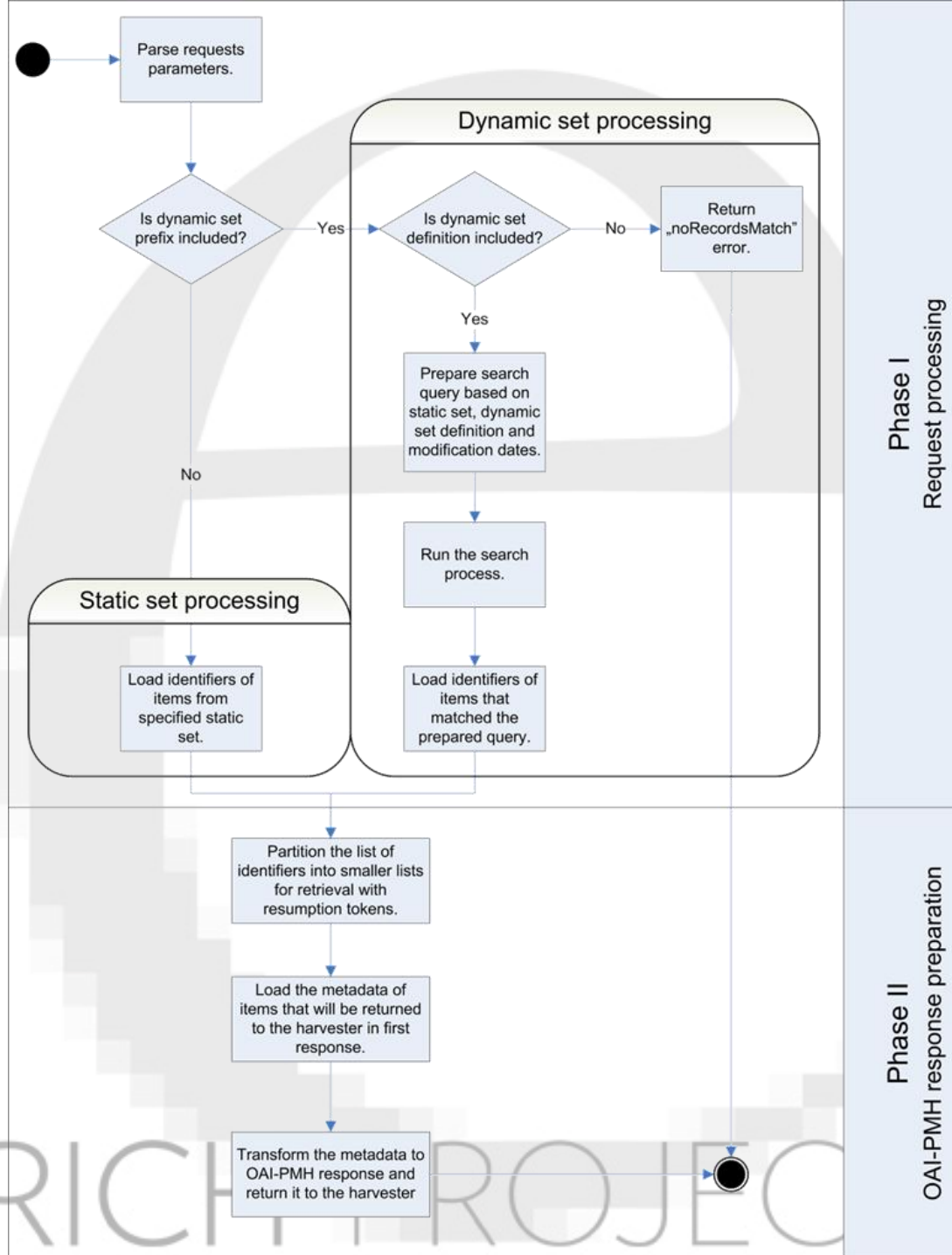
Integracja



- ZBC PCSS opracował rozszerzenie protokołu OAI-PMH pozwalające na elastyczne definiowanie kryteriów określających, które rekordy mają być pobrane
 - Rozszerzenie oparte jest na języku wyszukiwawczym CQL (Contextual Query Language) pochodzącym ze specyfikacji protokołu SRU/W (następca Z39.50)
 - Zmodyfikowany interfejs OAI-PMH jest nadal zgodny z podstawową specyfikacją protokołu
 - Nie ma konieczności modyfikowania oprogramowania po stronie pobierającej metadane

Jak to działa?

- Faza I – określenie listy identyfikatorów rekordów do pobrania
 - Na podstawie daty modyfikacji, przynależności do kolekcji i/lub dowolnego zapytania wyszukiwawczego
- Faza II – udostępnienie rekordów na podstawie listy identyfikatorów z fazy I
 - Opcjonalny podział wynikowej listy rekordów na „strony”
 - Wczytanie pełnych metadanych i wysłanie ich do klienta OAI-PMH





Integracja

- Kryteria BC UW r dla Manuscriptorium
 - dc.date >= 1000
 - dc.date <= 1850
 - not dc.type = czasop*
 - not dc.type = prog*
 - dc.format = (app* or pd* or dj* or im*)
- Wynik: **883** publikacje
 - Całkowita liczba publikacji w kolekcji Rękopisy: **161**
 - Całkowita liczba publikacji w kolekcji Stare druki: **388**
 - » Dane z 22.06.2009



Motywacja

2. Odrębny interfejs prezentacyjny Manuscriptorium

- Obiekt wyszukany w Manuscriptorium prezentowany jest na stronach Manuscriptorium
 - Treść pobierana jest bezpośrednio z serwera biblioteki cyfrowej współpracującej z Manuscriptorium
 - Elementy nawigacyjne pobierane są z serwera Manuscriptorium
 - Podstawą interfejsu nawigacji jest format JPG

Wydanie

- Opis
- Informacje
- Treść
- Treść (nowe okno)
- Pobierz
- Podobne wydania

O stronie...

Strona tworzona przez Bibliotekę Uniwersytecką we Wrocławiu

Ten serwis działa dzięki oprogramowaniu dLibra 4.0



Strona 4



Strona 5



Strona 6



Strona 7



LIZARDTECH

Cała strona

7/183

1490.

Am Montag nach Palmsonum ist geywehen
 König Matthias dem Gott gnädig sey.

1499.

Der selbige Nicolaus Blugwitz, jalle sinem Frister
 mit dem Namen Jovan Blugwitz, doliße
 Summa Goldt, mit doliße König von Neisschen
 zu saltz geywehen in sinem Laithen, das saltz er
 geywehen im Capittelhaus, usan die gemeins
 walt, London er sey ob usere sin halber theil,
 mit sinem Bruders, sonder in saltz von sinem
 geyweh, den das selbige gold salt mit dem Laithen zu
 dem selbigen Nicolaus Blugwitz geywehen, was
 das er saltz geywehen haben im Capittelhaus, den
 der selbige Frister was der Frister, Procurator
 zu der selbigen zeit, mit der Goldt ist geyweh
 geyweh 10000 Gulden, das der Kaysers Mann
 wann salt mozt, drago, das selbige gold nam den
 Blugwitz alle salt untrast als ein theil,

Document Summary

Bibliographic Description

Browsing

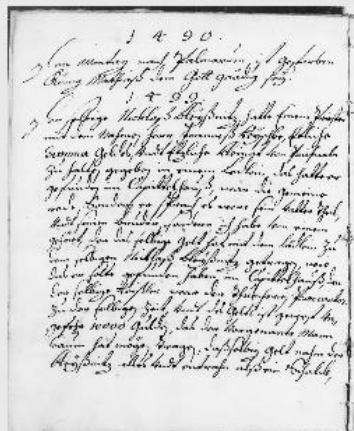
Bookmarks

Help



big

100%



Manuscriptorium.com
IV Q 143 - 7
(7/183)

1490.
Am Montag nach Palmsonum ist gegeben
König Matthias dem Gott gnädig sey.

1499.
Der selbige Nicolaus Blugwitz, alle seinem Erben
mit dem Namen seiner Schwagerin Königin, Katharina
Summa Goldes, und Katharina Königin von Saizfialn
zu saltz gegeben in einem Lehen, das saltz er
gepunden in Capittelstanz, was die gemeine
was, Sonder er sprach ob waser sein halbes Eitel,
und seinen Erben, sonder ist saltz von einem
gepunden, das das selbige gold salt mit dem Lehen zu
dem selbigen Nicolaus Blugwitz gegeben, was
das er saltz gepunden haben in Capittelstanz, dem
der selbige Richter was der Ertzherzog Procurator
zu der selbigen Zeit, und das Gold ist gegeben An
gepunden 10000 Gulden, das der Königinwants Mann
traum salt moze, drager, das selbige gold natm die



5



6



7



8





Integracja

- Automatyczna konwersja DjVu -> JPG
 - Kilka różnych jakości konwersji
- Automatyczne przesłanie informacji o strukturze publikacji
 - Podział na pliki i na „strony”
 - Informacje zarówno o plikach w formacie DjVu, jak i JPG
 - Dane zakodowane w formacie METS, przesyłane przez protokół OAI-PMH

Document Summary

Bibliographic Description

Browsing

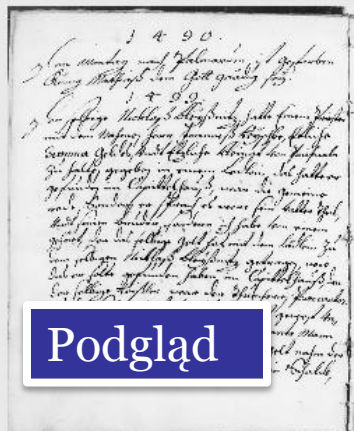
Bookmarks

Help

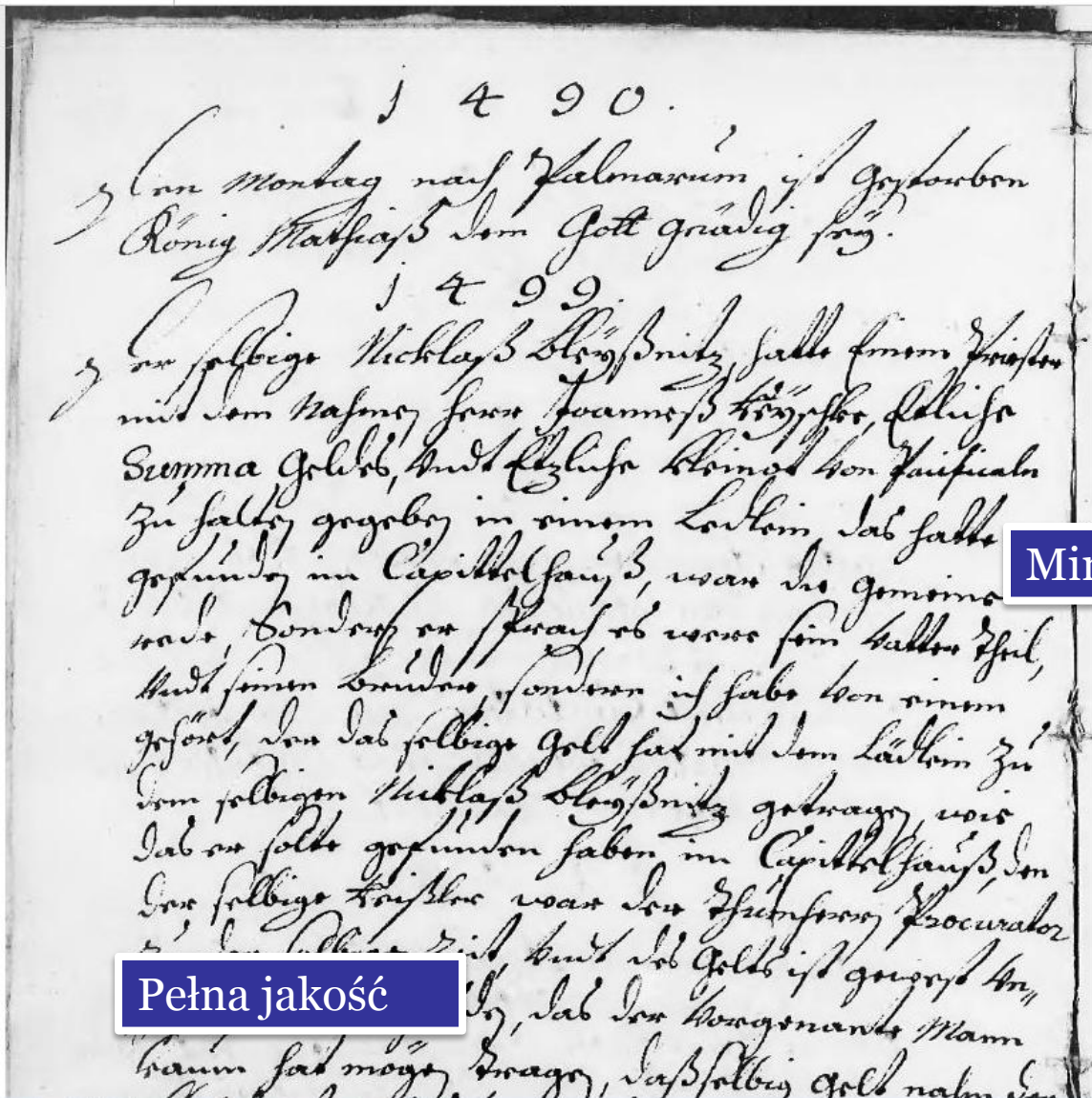


big

100%



Podgląd

 Manuscriptorium.com
 IV Q 143 - 7
 (7 / 183)


Pełna jakość

Miniaturki

Integracja

- Oryginał • http://www.bibliotekacyfrowa.pl/Content/11780/0037_0001.djvu
- Miniaturka • http://www.bibliotekacyfrowa.pl/Content/11780/d2j:thumbnail,36/0037_0001.djvu.jpg
- Podgląd • http://www.bibliotekacyfrowa.pl/Content/11780/d2j:preview,36/0037_0001.djvu.jpg
- Pełna jakość • http://www.bibliotekacyfrowa.pl/Content/11780/d2j:big,36/0037_0001.djvu.jpg

Integracja

- http://www.bibliotekacyfrowa.pl/Content/11780/d2j:preview,36/0037_0001.djvu.jpg
 - d2j = DjVu -> JPG
 - Mechanizm łatwo rozszerzalny na inne formaty
 - thumbnail, preview, big = predefiniowane ustawienia konwersji określone przez administratora BC
 - Powiększenie przy odczytywaniu DjVu (np. 25% czy 100%)
 - Stopień kompresji JPG (1 - 100)
 - 36 = strona w pliku DjVu
 - Ma znaczenie tylko w przypadku plików zawierających wiele stron



Motywacja

3. Odrębny format opisu publikacji – MASTER

– Konieczność dwukrotnego opisywania publikacji

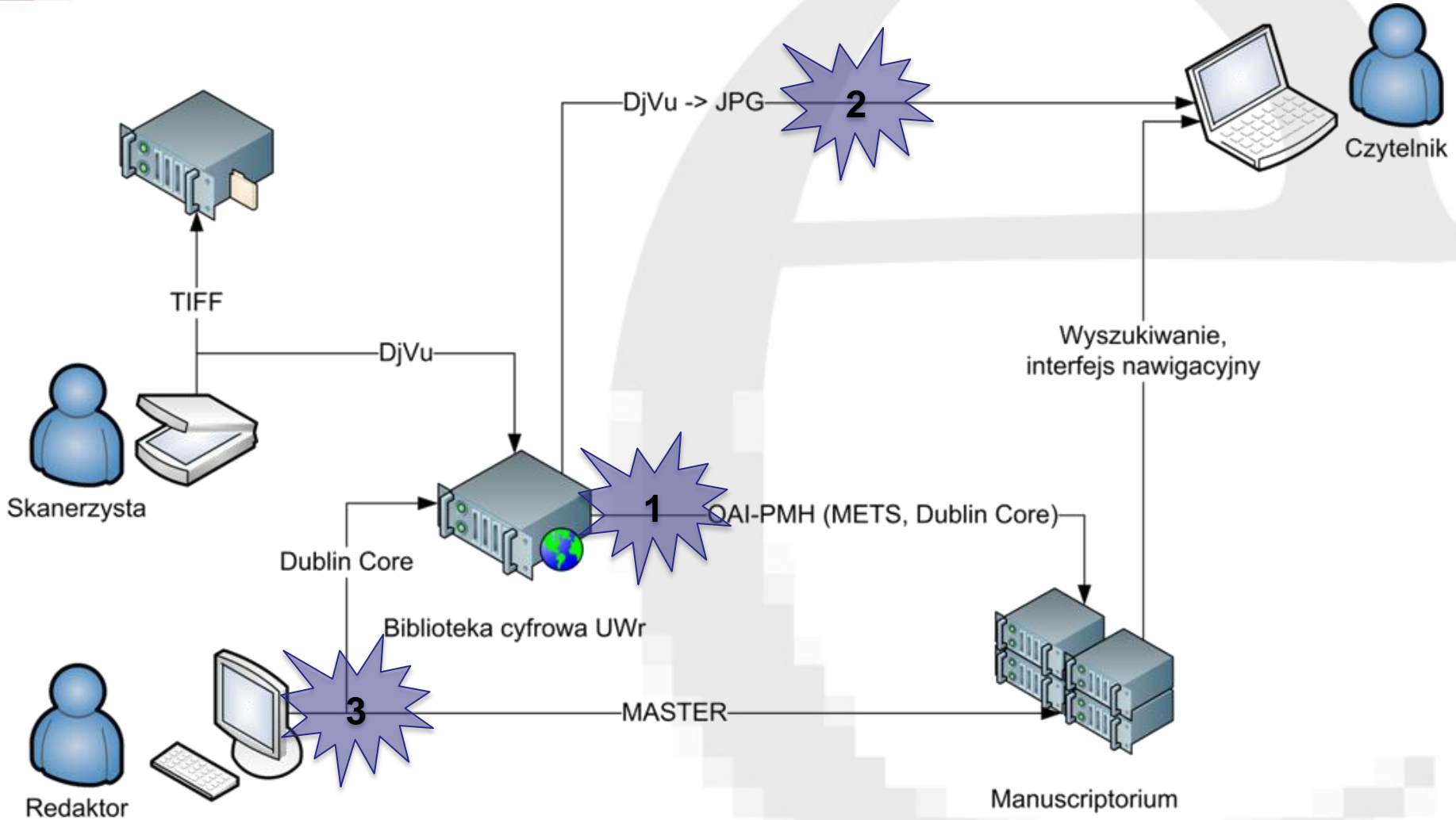
- Raz dla Manuscriptorium
- Raz dla biblioteki cyfrowej



Integracja

- Dodatek do dLibry pozwalający na import metadanych z dowolnego pliku XML
 - MASTER to format metadanych zapisywany w XML
 - Konfiguracja dodatku definiuje jakie elementy pliku XML mają zostać zaimportowane do poszczególnych pól opisu w BC
- Manuscriptorium przechodzi obecnie z formatu MASTER na format TEI P5
 - Podstawą jest nadal XML, więc dodatek importujący może spełniać swoje zadanie

Podsumowanie





Podsumowanie

- Zmiany w systemie dLibra pozwoliły na wysoce zautomatyzowaną współpracę pomiędzy BC UWr a Manuscriptorium
- Zmiany zostały wprowadzone na potrzeby BU UWr, ale są dostępne w standardowej dystrybucji dLibry
- Opisane mechanizmy są oparte na powszechnie używanych standardach
 - Dzięki temu mogą zostać wykorzystane nie tylko na potrzeby Manuscriptorium, ale również innych projektów
 - Projekt CACAO
 - The Networked Digital Library of Theses and Dissertations (NDLTD)



Dziękuję za uwagę!

Pytania?

Integracja systemu dLibra i Manuscriptorium

Marcin Werla, PCSS

mwerla@man.poznan.pl

ENRICH PROJECT

