

POZNAŃSKIE CENTRUM SUPERKOMPUTEROWO SIECIOWE



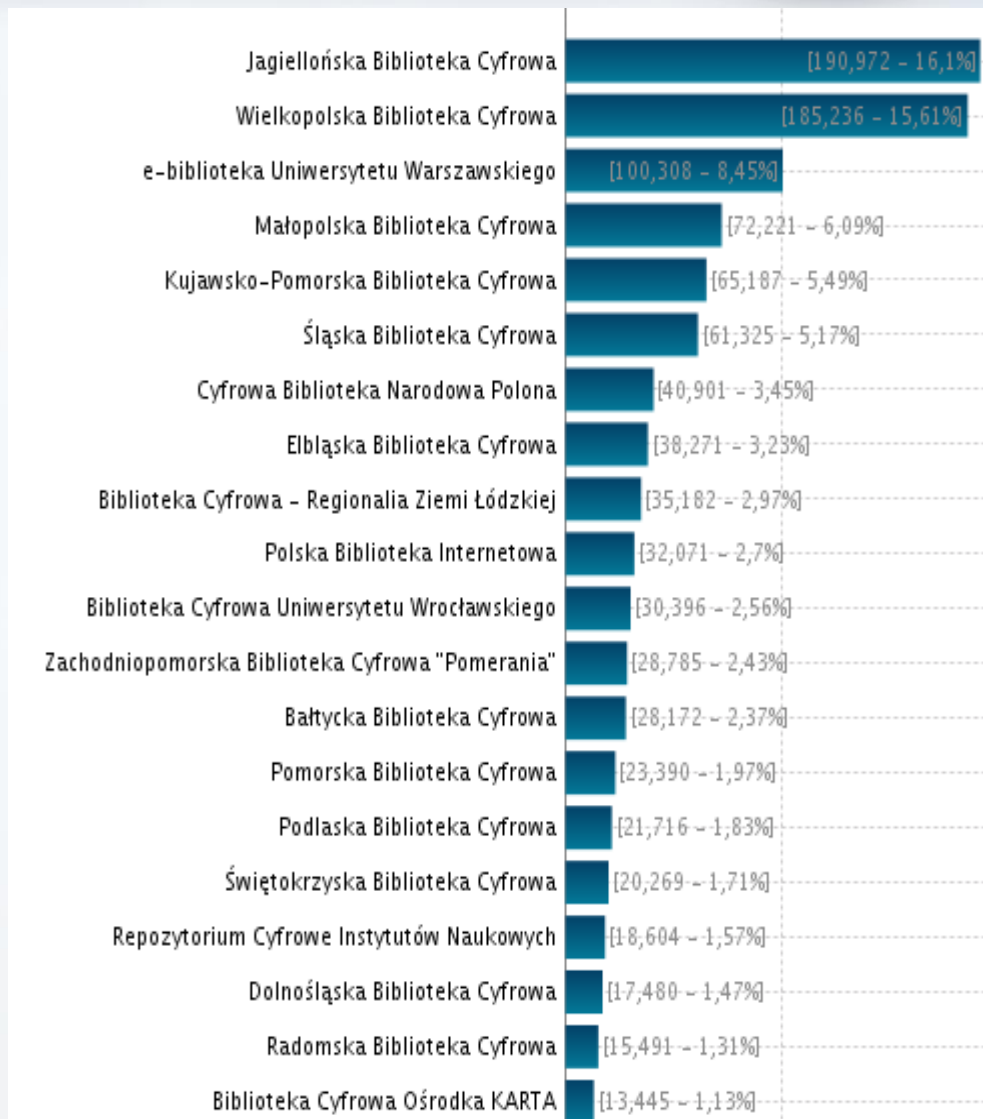
Możliwości rozwoju **dLibry**
dla potrzeb **dużych repozytoriów**

Marcin Werla
Zespół Bibliotek Cyfrowych PCSS

Co to właściwie jest duże repozytorium?

- Federacja Bibliotek Cyfrowych to obecnie około 1,2 mln obiektów z 82 źródeł
- Średnia rozmiaru biblioteki cyfrowej w FBC to 14 468 obiektów
- Mediana rozmiaru biblioteki cyfrowej w FBC to 2 301 obiektów

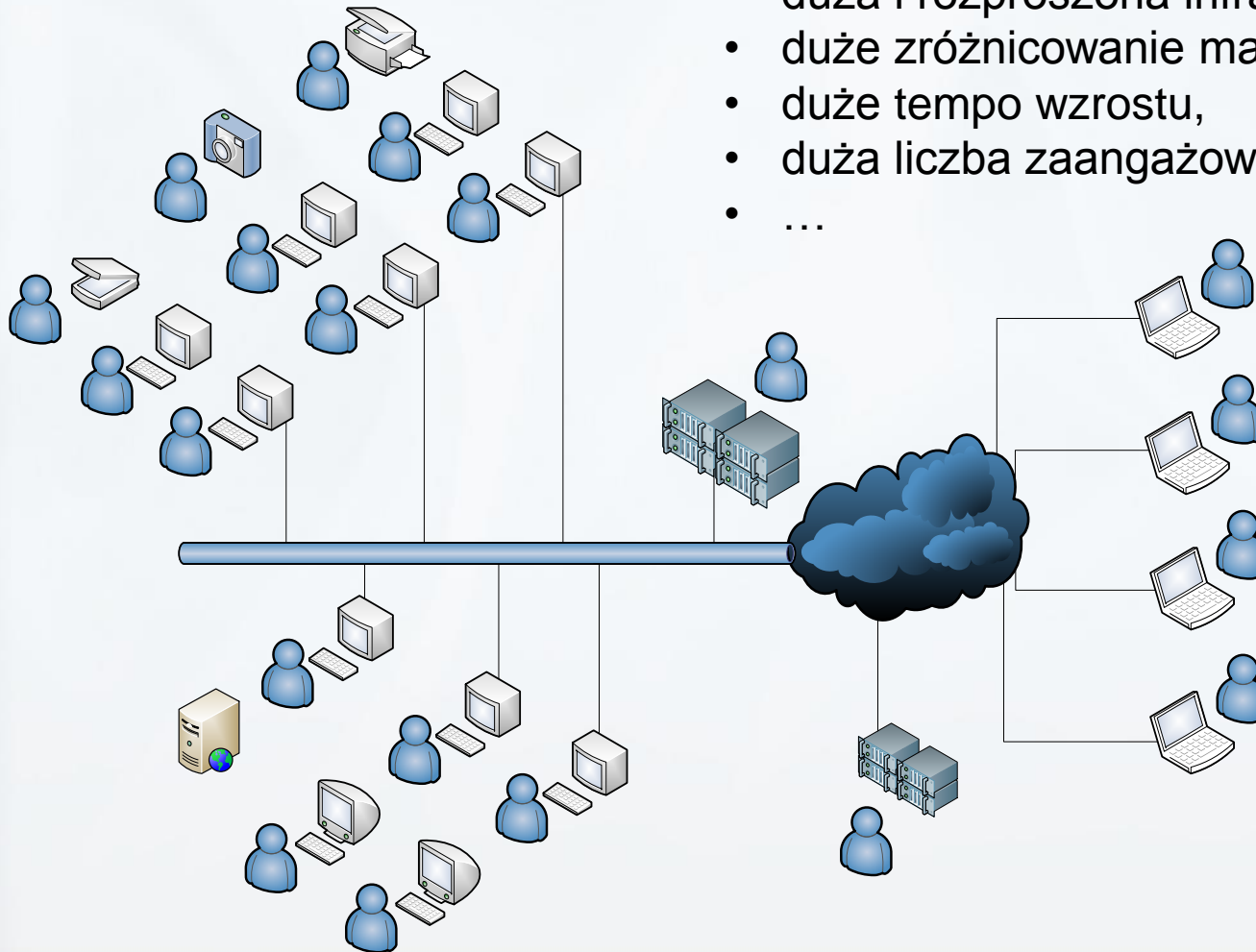
Dane z dnia 24.01.2013 wg <http://fbc.pionier.net.pl/owoc/list-libs>



Co to właściwie jest duże repozytorium?

To nie tylko duża liczba obiektów, ale też:

- duża i rozproszona infrastruktura
- duże zróżnicowanie materiałów,
- duże tempo wzrostu,
- duża liczba zaangażowanych osób
- ...



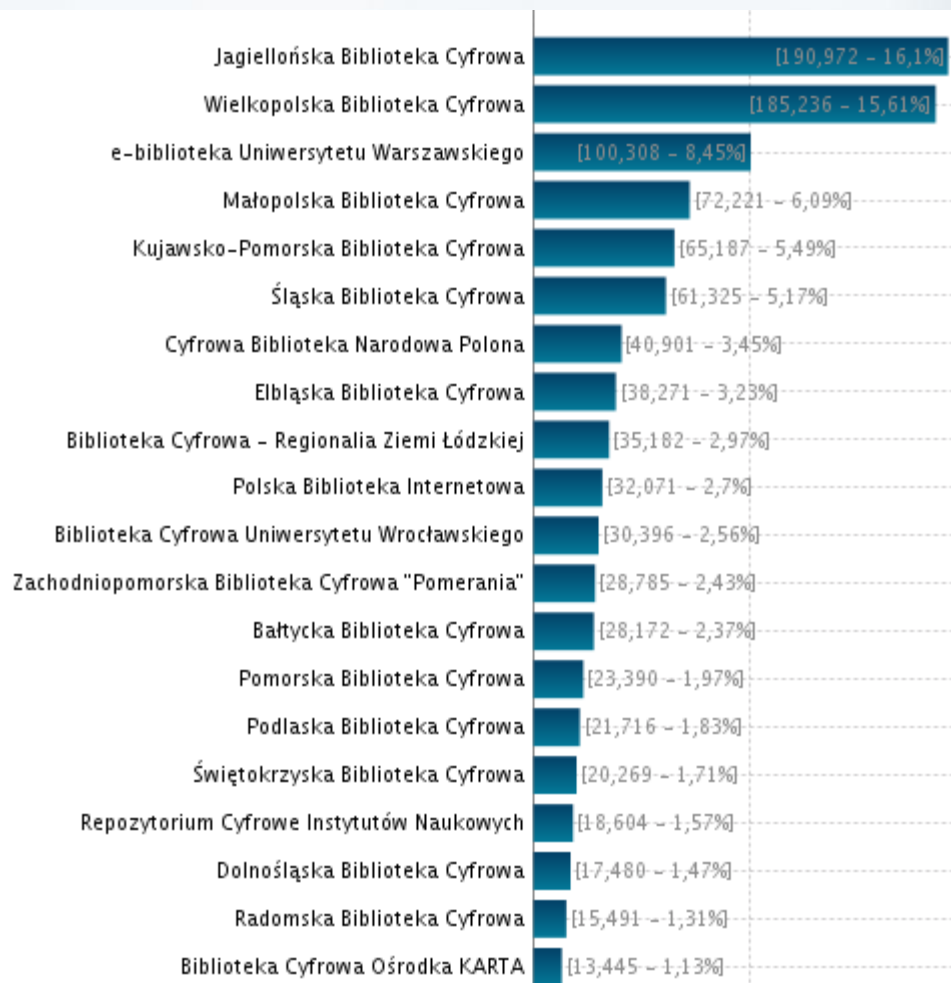
Co to jest dLibra?

- Oprogramowanie do budowy bibliotek cyfrowych rozwijane w PCSS od 1999 roku, wdrożone publicznie po raz pierwszy w 2002 roku w Wielkopolskiej Bibliotece Cyfrowej
- Rozwijana w ścisłej współpracy z użytkownikami, nastawione na innowacje oraz wspieranie współpracy instytucji kultury i nauki
- Obecnie podstawa funkcjonowania kilkudziesięciu bibliotek cyfrowych w państwowych instytucjach kultury, nauki, administracji, a także w organizacjach pozarządowych
- Więcej informacji: <http://dlibra.psnc.pl/>



dLibra i największe polskie biblioteki cyfrowe

- Obecnie 19 z 20 największych bibliotek cyfrowych w FBC jest opartych na oprogramowaniu dLibra
- Wyjątkiem (z wielu powodów) jest Polska Biblioteka Internetowa



Dane z dnia 24.01.2013 wg
<http://fbc.pionier.net.pl/owoc/list-libs>

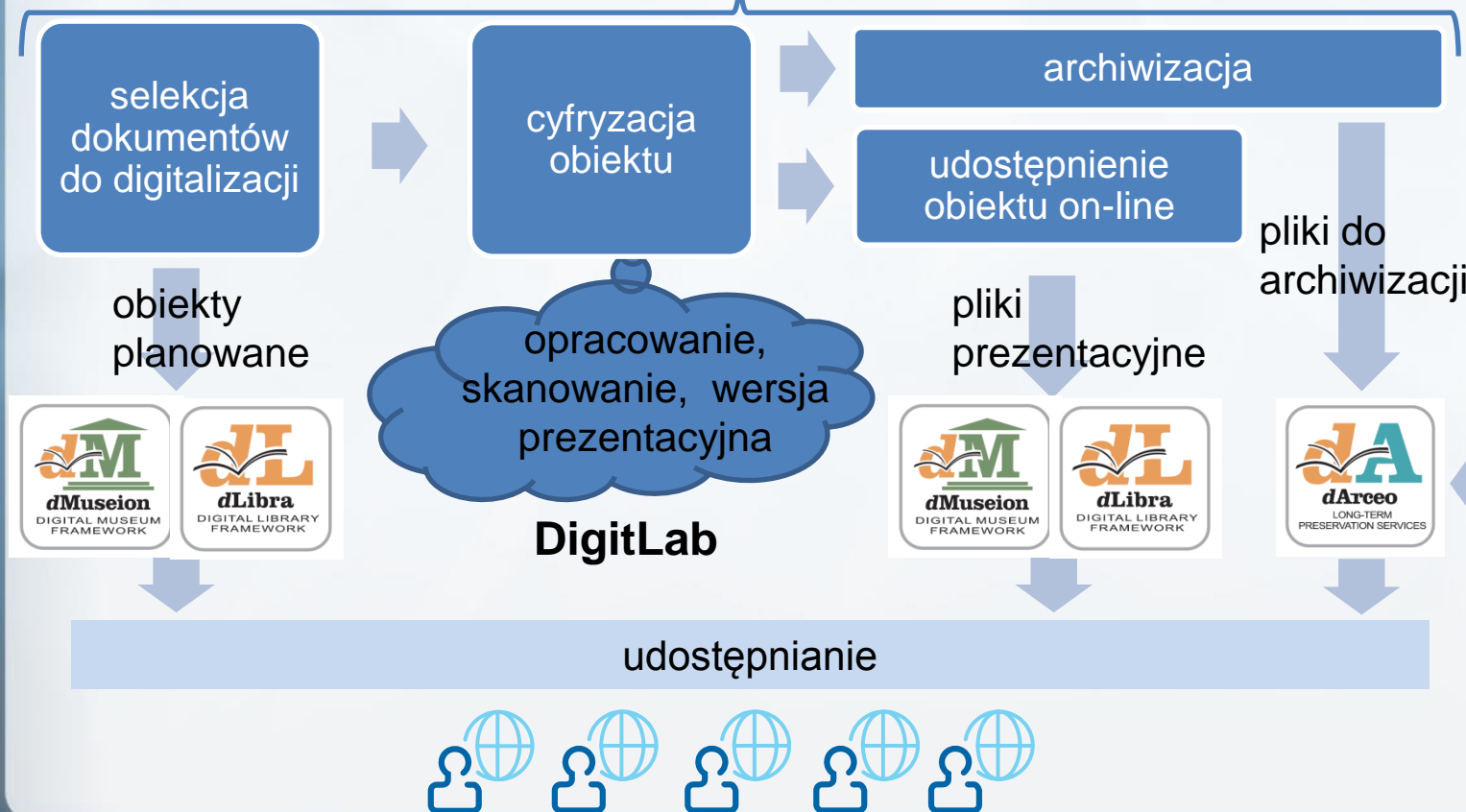
dLibra a wybrane aspekty dużych repozytoriów

- Kompleksowe wsparcie dla procesu cyfryzacji
- Obiekty born-digital (rdzennie cyfrowe?) i self-archiving (samoarchiwizacja?)
- Długoterminowa archiwizacja obiektów cyfrowych
- Obiekty zdalne
- Analiza wykorzystania bibliotek cyfrowych

Kompleksowe wsparcie dla procesu cyfryzacji



System opracowany na potrzeby projektu



dLab

Informacje podstawowe

- Usprawnienie prac związanych z cyfryzacją
 - Pomoc pracownikom w wykonywaniu przydzielonych zadań
 - Ułatwienie komunikacji między pracownikami
 - Automatyzacja czynności składających się na proces cyfryzacji
- Zarządzanie cyfryzacją
 - Przydział osób odpowiedzialnych za określone czynności w ramach procesu cyfryzacji
 - Realizacja procesu cyfryzacji zgodnie ze zdefiniowanym przebiegiem prac
- Monitorowanie prac
 - Raportowanie na podstawie zbieranych statystyk
 - Nadzór nad wykonaniem określonych czynności



dLab

Informacje podstawowe

- **Zadanie cyfryzacji obiektu**
 - Podstawowy element w systemie, dotyczący cyfryzacji konkretnego obiektu (książki, mapy, zabytku)
- **Czynność**
 - Składowa zadania, identyfikująca konkretne działanie do wykonania w ramach zadania
 - Wykonywana z uwzględnieniem zdefiniowanych ograniczeń kolejnościowych
- **Grupa zadań**
 - Element grupujący zadania, umożliwiający wykonanie określonych czynności „na raz” dla danej grupy zadań

dLab

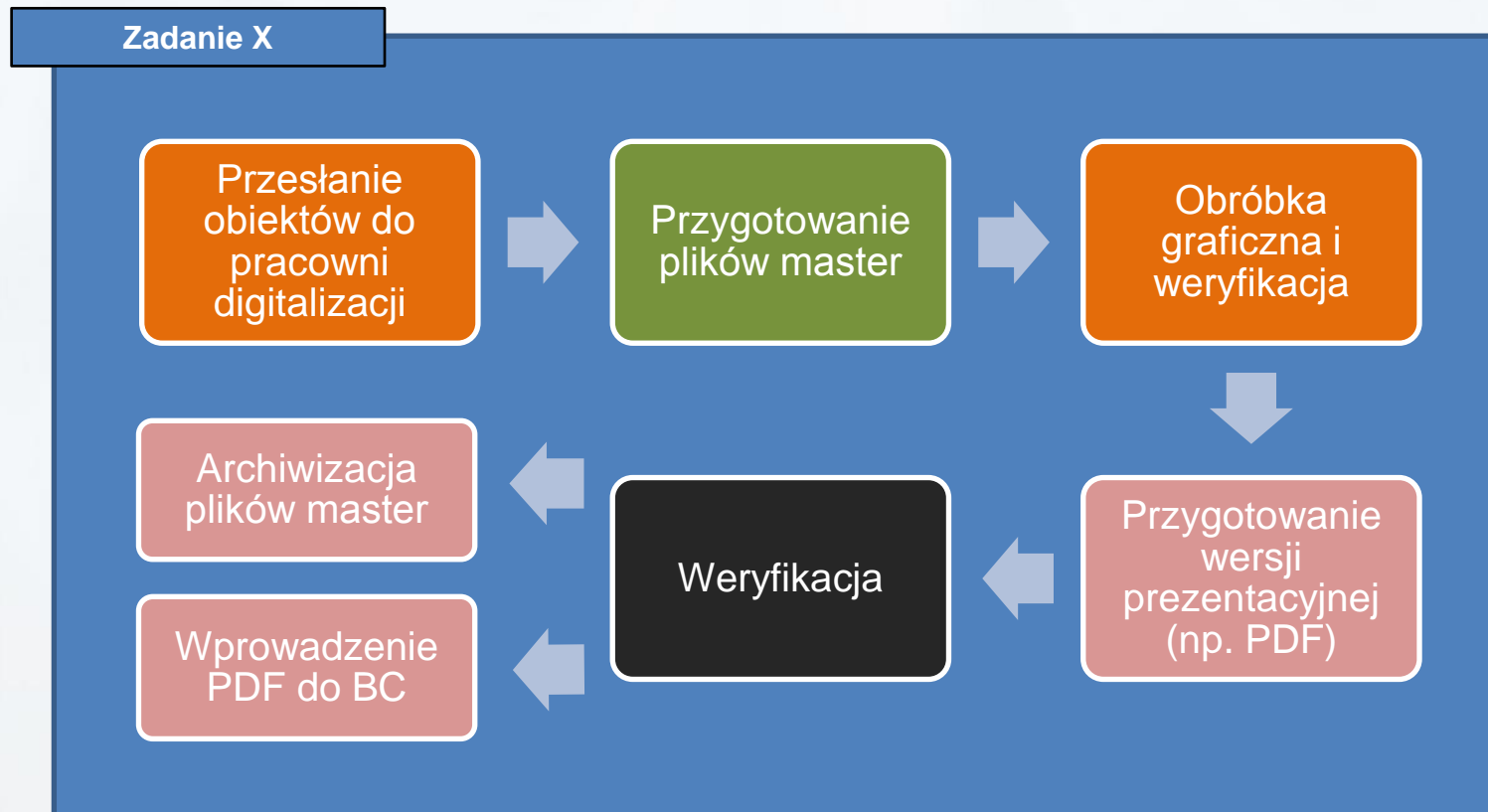
Przykład – zadanie i użytkownicy

Redaktor

Skanujący

Automat

Nadzorujący



Obiekty born-digital i self archiving


- Oprogramowanie dLibra daje możliwość deponowania obiektów cyfrowych samodzielnie np. przez pracowników naukowych uczelni
 - Deponowanie odbywa się przez stronę WWW i jest poprzedzone uwierzytelnieniem i autoryzacją (opcjonalna integracja z zewnętrznym systemem uwierzytelniania)
 - Formularz deponowania może być dostosowany do potrzeb konkretnej instytucji
 - Osoba deponująca przesyła pliki oraz określony zestaw metadanych
 - Obiekt trafia do moderacji do redaktora biblioteki cyfrowej (wsparcie dla dzielenia moderacji na wielu redaktorów)
 - Obiekt może być cofnięty do deponującego z podaniem powodu odrzucenia lub zaakceptowany i udostępniony w bibliotece cyfrowej
- Więcej informacji: <https://dl.psnc.pl/community/x/xwDQ>

Tworzenie nowej publikacji


Krok 1 z 3 - wybierz treść publikacji

 Wybierz pojedynczy plik lub archiwum ZIP zawierające zestaw plików do opublikowania.

Profil użytkownika


 Podstawowe

informacje

 Ulubione publikacje

 Tagi

 Twoje publikacje

 Dodaj publikację

Biblioteka

► Biblioteka Cyfrowa

dLibra

► Przykładowa_kolekcja

> Rocznicza

► Wystawy

Edytor Publikacji

Tworzenie nowej publikacji

Krok 2 z 2

Wybierz wartości dla atrybutów



* Tytuł

+ dodaj wartość

* Autor

+ dodaj wartość

Temat i słowa kluczowe

+ dodaj wartość

* Data wydania

Dzień: Miesiąc: Rok: 2000

+ dodaj wartość

Licencja

Licencja typu Creative Commons

+ dodaj wartość

Uwagi

Anuluj

Zapisz



obecnie czytających: 1

[STRONA GŁÓWNA](#)

[KOLEKCJE](#)

[KONTO](#)

[WYLOGUJ](#)

[KONTAKT](#)

Profil użytkownika

[Podstawowe informacje](#)

[informacje](#)

[Ulubione publikacje](#)

[Tagi](#)

[Twoje publikacje](#)

[Dodaj publikację](#)

Biblioteka

[Biblioteka Cyfrowa](#)

dLibra

[Przykładowa_kolekcja](#)

[Rocznica](#)

[Wystawy](#)

Twoje publikacje

Strona 1 z 2

[◀](#) [⏪](#) 1 | 2 [⏩](#) [▶](#)

Stan	Nazwa publikacji	Komentarz administratora	Wiadomości
	Agregator na każdą okazję	Proszę uzupełnić opis publikacji!	
	xxx4		

Komentarz administratora: Agregator na każdą okazję







Proszę uzupełnić opis publikacji!

Wiadomości

W tym panelu możesz odczytać adresowane do Ciebie wiadomości.

Strona 1 z 2

◀ ◀ 1 | 2 ▶ ▶

Typ	Temat	Nadawca	Data
	PROŚBA O KOREKTĘ : Kolejna publikacja	admin	01.12.2009
	Publikacja do zatwierdzenia : Kolejna inna...	mariusz	01.12.2009
	Publikacja do zatwierdzenia : Inna publika...	mariusz	01.12.2009
	Publikacja do zatwierdzenia : Kolejna publ...	radzik	01.12.2009
	Publikacja do zatwierdzenia : Publikacja p...	radzik	01.12.2009
	Publikacja : Publikacja publikowana - post	radzik	01.12.2009

Nadawca : **admin** Data : **01.12.2009**

 [Odpowiedz nadawcy](#)

Temat : **PROŚBA O KOREKTĘ : Kolejna publikacja**

Dotyczy publikacji : [Kolejna publikacja](#)


Treść :





Szanowny Panie/Pani!

Wprowadzony przez Pana/Panią obiekt cyfrowy, nazwany "Kolejna publikacja", został skierowany do korekty przez moderatora z następującym komentarzem:
Publikacja wymaga korekty ...

Aby skontaktować się z moderatorem należy odpowiedzieć na tę wiadomość.

Ta wiadomość została automatycznie wygenerowana po przekazaniu obiektu cyfrowego do korekty.

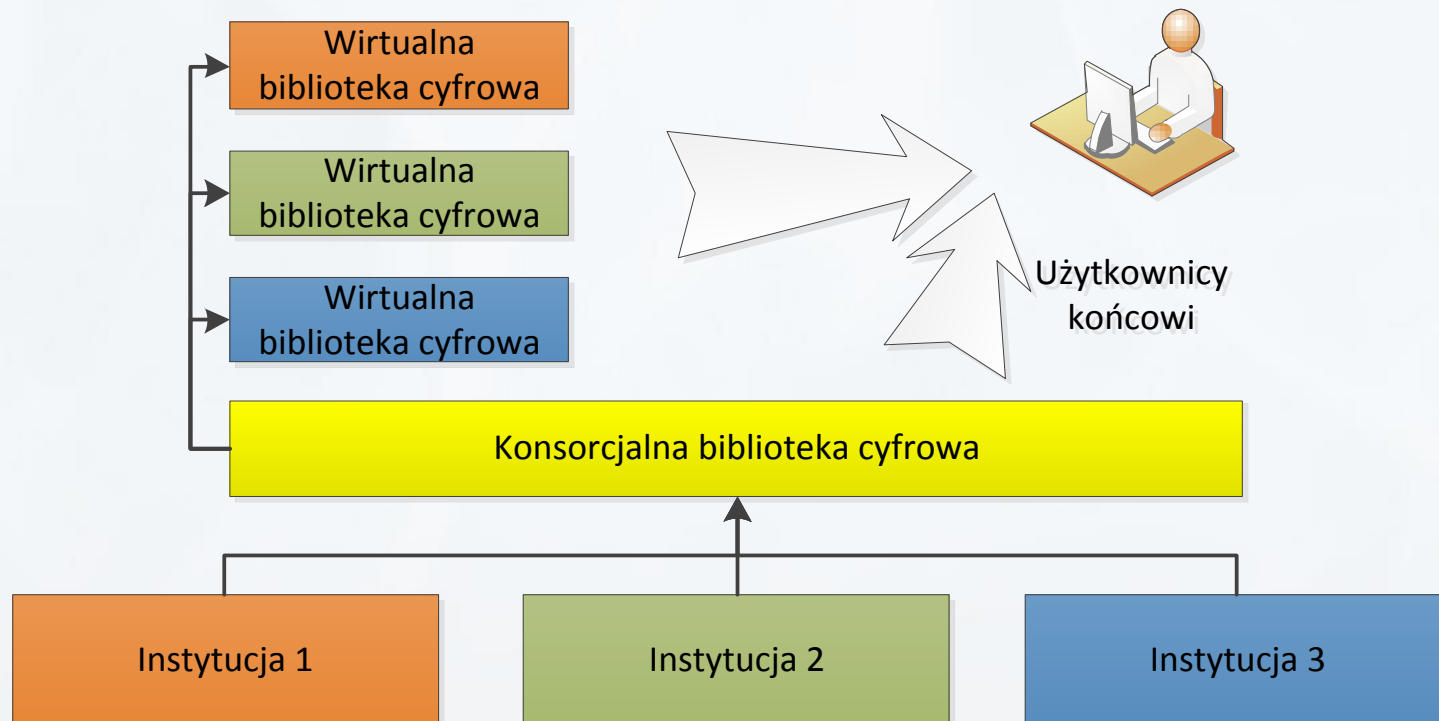
 **Wiadomości** _ □ ×

Typ wiadomości	Nadawca	Temat ▼	Data	..
	radzik	Publikacja do zatwierdzenia : Publikacja przykłado...	2009-12-01 1... ..	
	radzik	Publikacja do zatwierdzenia : Kolejna publikacja	2009-12-01 1... ..	
	mariusz	Publikacja do zatwierdzenia : Kolejna inna publikacja	2009-12-01 1... ..	
	mariusz	Publikacja do zatwierdzenia : Inna publikacja	2009-12-01 1... ..	

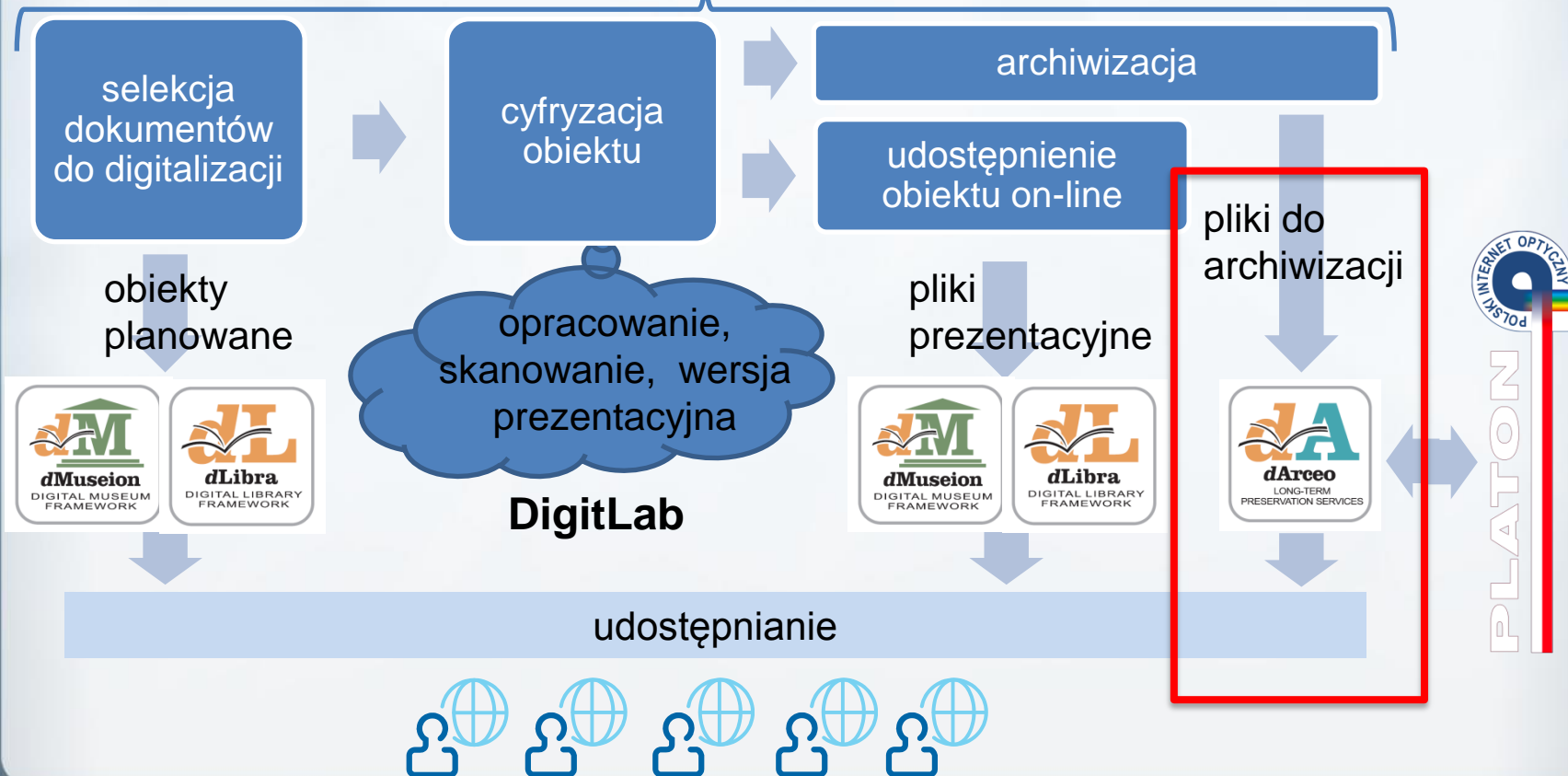
Temat: Publikacja do zatwierdzenia : Inna publikacja
Od: mariusz
Data: Tue Dec 01 10:32:50 CET 2009

Szanowny Panie/Pani,
Użytkownik "mariusz" utworzył nową publikację o nazwie "Inna publikacja".
Proszę sprawdzić czy wszystkie dane zostały wprowadzone poprawnie: czy publikacja zawiera prawidłowy zestaw atrybutów oraz właściwą treść.

Regionalne biblioteki cyfrowe a repozytoria instytucjonalne



Długoterminowa archiwizacja obiektów cyfrowych



dArceo

Informacje podstawowe



- Rozwijane przez Poznańskie Centrum Superkomputerowo-Sieciowe od 2011 roku
- Powstało na bazie wyników prac badawczo rozwojowych w ramach projektu SYNAT
- Dedykowane do przechowywania danych źródłowych w długim czasie, w szczególności
 - Tekstowych (np. PDF/A)
 - Graficznych (np. TIFF, JPEG2000)
 - Audiowizualnych (np. MPEG-4)

dArceo

Przechowywanie danych



- Wspiera różne systemy składowania danych
 - Lokalny system plików (np. macierz dyskowa)
 - (S)FTP
 - Usługa powszechnej archiwizacji PLATON U4
- Wykorzystuje uznane światowe standardy przechowywania informacji o zasobach
 - Formaty METS, PREMIS
 - Formaty TextMD, DocumentMD, VMD, MIX, AES57

dArceo

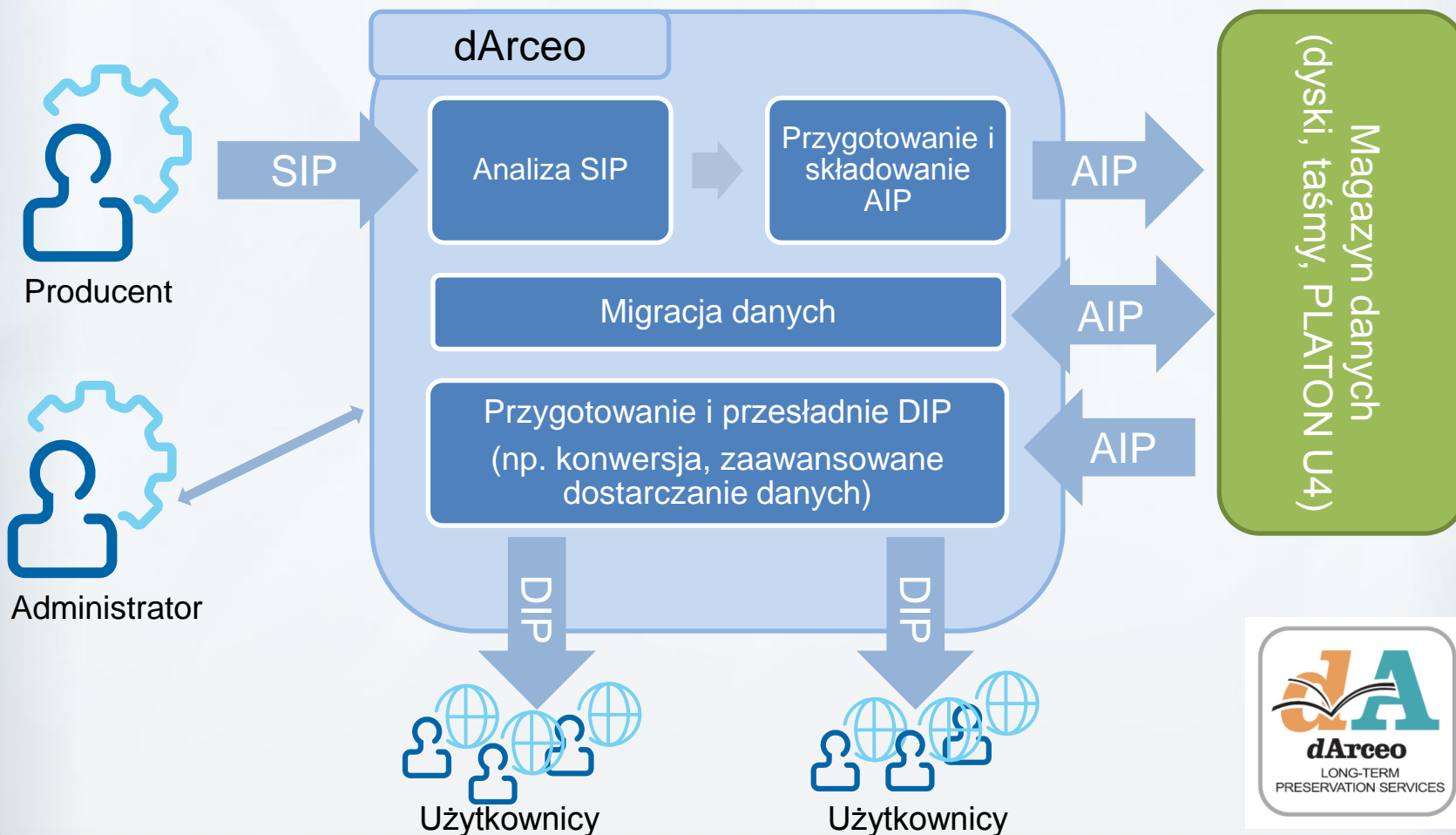
Migrowanie danych



- Monitorowanie danych pod kątem ich utraty
 - Weryfikacja sum kontrolnych SHA-512
 - Weryfikacja formatów z wykorzystaniem bazy UDFR (PRONOM+GDFR)
- Wykonywanie planów migracji
 - Migrowanie wielu obiektów z uwzględnieniem kryteriów administratora danych
 - Bieżące monitorowanie postępu migracji

dArceo

Zgodność z modelem OAIS



Obiekty zdalne

- Kolekcje w bibliotekach cyfrowych coraz częściej pełnią funkcję tematycznych katalogów obiektów cyfrowych, z których część jest umieszczona w danej bibliotece cyfrowej, a część w zewnętrznych serwisach
 - Serwisy komercyjne
 - Serwisy otwarte innych instytucji
 - Wydawnictwa uczelniane z systemem mikropłatności
 - Wartościowe zasoby Internetu społecznościowego – np. konta na Twitterze (np. <http://fbc.pionier.net.pl/id/oai:jbc.bj.uj.edu.pl:204509>)
- dLibra w wersji 6.0 poza typowymi obiektami cyfrowymi (treść + metadane) będzie wspierała też obiekty zdalne
 - Metadane + link
 - Metadane + treść (tylko do wyszukiwania) + link
- Powinno to umożliwić bezproblemowe osadzanie obiektów zdalnych w bibliotekach cyfrowych i budowanie internetowych bram tematycznych

Analiza wykorzystania bibliotek cyfrowych

- Temat wykorzystania zasobów i usług bibliotek cyfrowych jest bardzo interesujący zarówno dla twórców tych bibliotek jak i naukowców
 - Jest bardzo ważnym źródłem informacji na temat pożądaných kierunków udostępniania zbiorów i rozwoju funkcji biblioteki
- Podstawowe statystyki w systemie dLibra:
 - Na poziomie obiektu:
 - Łączna liczba wyświetleń treści obiektu od momentu jego udostępnienia
 - Na poziomie biblioteki:
 - Miesięczna liczba wyszukiwań w bibliotece
 - Miesięczna liczba wyświetlonych obiektów
 - Miesięczna liczba dodanych obiektów
 - Miesięczna liczba odwiedzających
- Popularne usługi takie jak Google Analytics są nastawione na analizę użycia stron WWW i nie biorą pod uwagę specyfiki bibliotek cyfrowych

Analiza wykorzystania bibliotek cyfrowych

- W ramach Federacji Bibliotek Cyfrowych uruchamiamy obecnie usługę FBC Analytics, której celem ma być monitorowanie polskich bibliotek cyfrowych
 - Oprogramowanie open-source PIWIK dostosowane do specyfiki bibliotek cyfrowych
- Obecnie w ramach pilotażu monitorowaniem objęte są:
 - Wielkopolska Biblioteka Cyfrowa
 - Kujawsko-Pomorska Biblioteka Cyfrowa
 - Cyfrowe Muzeum Narodowe w Warszawie
- Wszystkich zainteresowanych usługą proszę o kontakt

Tablica statystyk

Oglądający

Aktywność

Odsyłacze

Założenia

Strona WWW

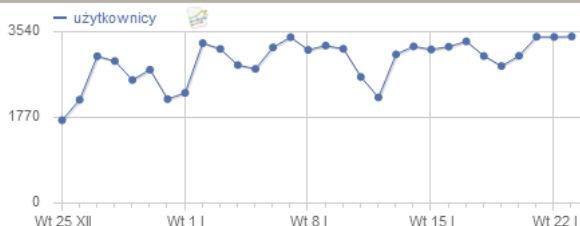
Wielkopolska Biblioteka Cyfrowa ▾

Zakres dat: 2013-01-23

Widżety i Tablica analiz ▾



Najnowsza aktualizacja: Piwik 1.10.1

Wykres z ostatnich odwiedzin







Odwiedzający w czasie rzeczywistym

Data	użytkownicy	Liczba odson
w czasie ostatnich 24 godzin	3240	17601
w czasie ostatnich 30 minut	134	915

Czw 24 I - 14:08:45 (2 min 28s)   #6 -  - IP: 83.6.191.181

wejścia bezpośrednie

Strony: 

Czw 24 I - 14:08:43 (6 min 39s)   -  - IP: 195.16.46.130

z [Yandex](#) - "http:www.wbc.poznan.pl/dlibra/"

Strony: 

Czw 24 I - 14:08:43 (1 min 17s)   -  - IP: 80.53.80.150

wejścia bezpośrednie

Strony: 

Lista słów kluczowych






Słowo kluczowe	użytkownicy ▾
Słowo kluczowe nie zdefiniowany	380
wielkopolska biblioteka cyfrowa	5
kurier poznański	3
poradnik gospodarski wbc 1920	3
codex diplomaticus silesiae	2
ketrzyński przydomki szlachty pomorskiej	2
kurier poznański biblioteka cyfrowa	2
lakta grodzkie i zmiemie	2
nauka wiary i obyczajów	2
poemat krasickiego na sb	2
poklewski w bukareszcie	2
pomiar forii	2
"antoni bartłomiej raczyński" "opat"	1
"bolesław morawski"	1
'chciałeś panie naczelniku, abym nauczył lud, jak ...	1
"emergency worker" pracownik pogotowia	1
"eugeniusz konarzewski" "opat"	1
"generał zabołocki"	1
"hugo sgaller"	1
"konstanty mazurek"	1
"krwawa potyczka pod morzewem	1
"młyn walczak-rakowicz ostrów wielkopolski"	1

Visitor Location (mapa świata)



użytkownicy

Przeglądarki odwiedzających


Przeglądarka	Liczba niepowtarzających się użytkowników
 Firefox	1188
 Internet Explorer	747
 Chrome	607
 Opera	236
 Safari	112

1-5 z 14 Dalej >




Related report: Wersja przeglądarki

Najlepsze wyszukiwarki

Wyszukiwarka	użytkownicy ▾
 Google	1010

Tablica statystyk
Oglądający
Aktywność
Odsyłacze
Założenia
[Strony](#) | [Strony wejść](#) | [Strony wyjść](#) | **Tytuły stron** | [Site Search](#) | [Linki zewnętrzne](#) | [Pobrania](#)

 Zakres dat: **2013-01-23** 
Tytuły stron 

This report contains information about the titles of the pages that have been visited. The page title is the HTML <title> Tag that most browsers show in their window title.

Report generated 4 min 58s ago

Nazwa strony	Liczba odłon	Liczba odłon niepowtarzających się	Współczynnik rezygnacji	Średni czas przebywania na stronie	Współczynnik wyjść
+ Wielkopolska Biblioteka Cyfrowa	10975	6521	46%	1 min 51s	36%
Wielkopolska Biblioteka Cyfrowa	4273	1065	33%	4 min 44s	52%
+ Digital Library of Wielkopolska	636	400	48%	2 min 6s	46%
+ Großpolnische Digitale Bibliothek	480	256	41%	2 min 1s	46%
Digital Library of Wielkopolska	287	84	36%	5 min 30s	56%
Großpolnische Digitale Bibliothek	214	54	33%	4 min 1s	48%
- Bibliothèque numérique de la Grande Pologne	21	20	63%	39s	65%
Home Page	3	3	100%	1s	67%
Bajki	1	1	0%	11s	0%
Brevis descriptio historico-geographica ecclesiaru...	1	1	100%	0s	100%
Ilustracja Poznańska 1928.01.28 Nr4	1	1	100%	0s	100%
Las Polski. 1922 R.2 nr7-8	1	1	0%	8 min 10s	0%

Tablica statystyk
Oglądający
Aktywność
Odsyłacze
Założenia
[Strony](#) | [Strony wejść](#) | [Strony wyjść](#) | [Tytuły stron](#) | **Site Search** | [Linki zewnętrzne](#) | [Pobrania](#)

 Zakres dat: **2013-01-23** 


Site Search Keywords


Keyword	Searches ▼	Search Results pages	% Search Exits
Biblioteka Akademii Rolniczej w Poznaniu	9	1,1	89%
Przyjaciel ludu	4	1,3	25%
sternberg	4	2,5	25%
bluszcz	3	1,3	0%
browar	3	32,3	67%
Ramułt	3	14	0%
Tukatsch	3	3,3	33%
"Towarzystwo Pomocy Polonii Zagranicznej"	2	1	0%
1952	2	1	0%
adam z Życka	2	1	50%


 1-10 z 500 [Dalej >](#)
 


Pages Following a Site

Destination Page

 + **dlibra**
 /dlibra

 /dlibra?action=Change

 /dlibra?action=Change

 + **publication**


Related repo

Search Keywords with No Results

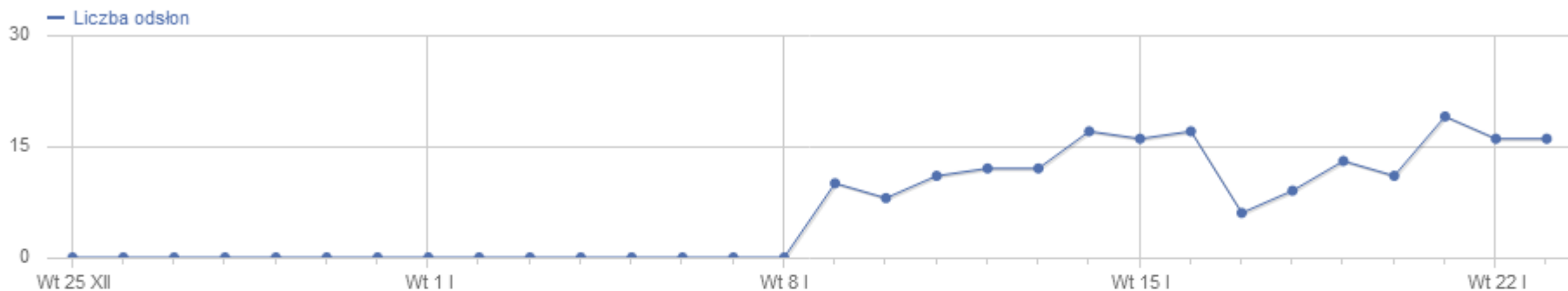
Tytuły stron

Nazwa strony	Liczba odślon	Liczba odślon niepowtarzających się	Współczynnik rezygnacji	Średni czas przebywania na stronie	Współczynnik wyjść
id	42742	41774	67%	24s	8%
lib www.wbc.poznan.pl	8923	8766	57%	24s	6%
lib jbc.bj.uj.edu.pl	5416	5391	71%	7s	2%
lib ebuw.uw.edu.pl	3092	3065	68%	10s	3%
lib www.sbc.org.pl	2913	2847	72%	34s	12%
lib kpbc.umk.pl	2847	2825	74%	20s	6%
lib mbc.malopolska.pl	2430	2407	78%	17s	7%
lib pbc.gda.pl	2363	2342	55%	13s	3%
lib dlibra.bibliotekaelblaska.pl	1168	1156	56%	18s	3%
lib www.bibliotekacyfrowa.pl	1160	1145	73%	27s	13%
lib bc.wimbp.lodz.pl	1111	1111	100%	2s	2%
lib bibliotekacyfrowa.eu	959	958	81%	10s	5%
lib zbc.ksiaznica.szczecin.pl	878	873	59%	9s	4%
lib rcin.org.pl	857	836	68%	45s	18%
lib www.dbc.wroc.pl	799	751	74%	50s	12%
lib www.polona.pl	705	623	58%	1 min 39s	39%
lib bc.mbpradom.pl	509	497	61%	23s	6%
lib bcpw.bg.pw.edu.pl	468	455	45%	29s	7%
lib www.dlibra.karta.org.pl	373	370	50%	13s	2%
lib www.pbc.rzeszow.pl	391	370	72%	1 min 14s	19%
lib cyfrowa.chbp.chelm.pl	363	361	63%	35s	5%
lib www.kbc.krosno.pl	339	336	100%	17s	2%

Tytuły stron

Nazwa strony	Liczba odślon	Liczba odślon niepowtarzających się	Współczynnik rezygnacji	Średni czas przebywania na stronie	Współczynnik wyjść
id	42742	41774	67%	24s	8%
lib www.wbc.poznan.pl	8923	8766	57%	24s	6%
lib jbc.bj.uj.edu.pl	5416	5391	71%	7s	2%
oai:jbc.bj.uj.edu.pl:8023	2	2	0%	2 min 54s	0%
oai:jbc.bj.uj.edu.pl:39736	2	2	0%	21s	0%
oai:jbc.bj.uj.edu.pl:83598	2	2	0%	3s	50%
oai:jbc.bj.uj.edu.pl:94238	2	2	0%	4s	0%
oai:jbc.bj.uj.edu.pl:111304	2	2	0%	20 min 43s	0%
oai:jbc.bj.uj.edu.pl:117188	5	2	0%	3 min 20s	50%
oai:jbc.bj.uj.edu.pl:132165	2	2	100%	1s	50%
oai:jbc.bj.uj.edu.pl:144907	2	2	0%	10 min 47s	0%
oai:jbc.bj.uj.edu.pl:144911	2	2	0%	4s	0%
oai:jbc.bj.uj.edu.pl:145083	2	2	0%	7 min 15s	0%
oai:jbc.bj.uj.edu.pl:157985	2	2	0%	1 min 54s	0%
oai:jbc.bj.uj.edu.pl:179564	2	2	0%	7 min 42s	0%
oai:jbc.bj.uj.edu.pl:185766	3	2	0%	33s	0%
oai:jbc.bj.uj.edu.pl:10016	1	1	0%	1s	0%
oai:jbc.bj.uj.edu.pl:10026	1	1	0%	1s	0%
oai:jbc.bj.uj.edu.pl:10065	1	1	0%	0s	0%
oai:jbc.bj.uj.edu.pl:10092	1	1	0%	6s	0%
oai:jbc.bj.uj.edu.pl:10106	1	1	0%	6s	0%
oai:jbc.bj.uj.edu.pl:10116	1	1	0%	5s	0%

id - lib www.wbc.poznan.pl - oai:www.wbc.poznan.pl:65793



30 ▾

Metrics for Nazwa strony: id - lib www.wbc.poznan.pl - oai:www.wbc.poznan.pl:65793

Click the metrics to display them in the large evolution graph. Use shift-click to display multiple metrics at once.

- Liczba odsłon:** between 6 and 19
- Liczba odsłon niepowtarzających się:** between 3 and 16
- Współczynnik rezygnacji:** between 33% and 100%
- Średni czas przebywania na stronie:** between 0s and 286s
- Współczynnik wyjść:** between 42% and 100%

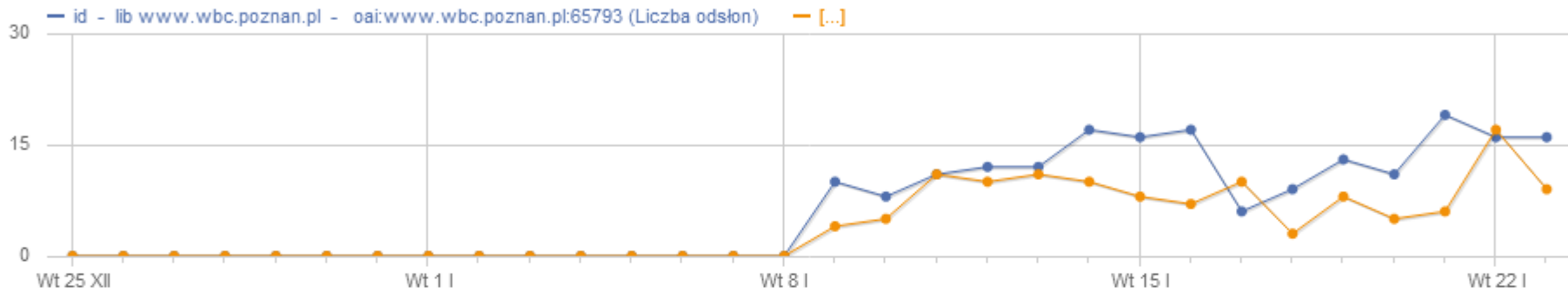
Compare records

Click the link below and open this popup for another row from the same table to compare multiple records. Use shift-click to mark the row for comparison without opening this popup.

» Pick a row to compare

167439	5	4	0%	4 min 11s	0%
--------	---	---	----	-----------	----

Evolution of multiple rows



30 ▾

Nazwa strony: Comparing 2 rows



id - lib www.wbc.poznan.pl - oai:www.wbc.poznan.pl:65793 (Liczba odsłon)
between 6 and 19



id - lib www.wbc.poznan.pl - oai:www.wbc.poznan.pl:51405 (Liczba odsłon)
between 3 and 17

» Pick another row to compare

Available metrics

Liczba odsłon ▾

71476	4	4	0%	16s	25%
107664	4	4	67%	27s	50%
130832	5	4	67%	7s	75%
167439	5	4	0%	4 min 11s	0%
7397	3	3	0%	9s	33%
19624	3	3	50%	1 min 15s	33%

Kompleksowe wsparcie dla dużych repozytoriów

- Narzędzia dINGO



- Infrastruktura i usługi



PLATFORMA OBSŁUGI NAUKI

PLATON



FEDERACJA

BIBLIOTEK CYFROWYCH

- Projekty badawczo-rozwojowe

- EuropeanaCloud
- LoCloud
- SUCCEED



Dziękuję za uwagę!

- Więcej informacji:
 - Strony informacyjne Federacji Bibliotek Cyfrowych:
 - <http://fbc.pionier.net.pl/pro/>
 - Bezpłatne kursy e-learningowe FBC na temat digitalizacji:
 - <http://fbc.pionier.net.pl/elearning/>
 - Nowy serwis FBC – Pytania i odpowiedzi
 - <http://fbc.pionier.net.pl/qa/>
 - Strona WWW Zespołu Bibliotek Cyfrowych PCSS:
 - <http://dl.psnc.pl/>
- Zapraszamy również na konferencję Polskie Biblioteki Cyfrowe 2013 i imprezy towarzyszące – 15-18 kwietnia 2013 r., Poznań
 - Więcej informacji od lutego na stronie WWW <http://www.man.poznan.pl/PBC/>