

Masowe pobieranie metadanych przez interfejs OAI-PMH

Często m.in. na potrzeby rozliczania projektów digitalizacyjnych zachodzi potrzeba masowego wyeksportowania metadanych obiektów z systemu dLibra. Najprościej w tym celu wykorzystać mechanizm kolekcji i protokół OAI-PMH.

1. W pierwszej kolejności należy utworzyć kolekcję zawierającą dokładnie te obiekty, których metadane chcemy wyeksportować. Wszystkie te obiekty powinny być opublikowane (choć dostęp do treści nie musi być publiczny).
2. Następnie należy użyć przygotowanego przez nas narzędzia do pobierania metadanych z repozytorium OAI-PMH. Narzędzie to można pobrać tutaj: [EuropeanaXMLBuilder-1.1-SNAPSHOT-bin.zip](#), po pobraniu należy je rozpakować.
3. Do uruchomienia narzędzia potrzebna jest maszyna wirtualna Java - można ją pobrać bezpłatnie pod adresem: <http://java.com/>
4. W celu uruchomienia programu należy wywołać wiersz poleceń, przejść do katalogu w którym znajduje się rozpakowane pobrane wcześniej narzędzie i wywołać następujące polecenie:

```
java -jar EuropeanaXMLBuilder-1.1-SNAPSHOT.jar -u "http://<adres biblioteki cyfrowej>/dlibra/oai-pmh-repository.xml?verb=ListRecords&metadataPrefix=<schemat metadanych>&set=<kolekcja>" -o <nazwa pliku wynikowego> -g
```

gdzie:

- <adres biblioteki cyfrowej> to np.: www.wbc.poznan.pl
- <schemat metadanych> to identyfikator schematu:
 - oai_dc - dla Dublin Core
 - mets - dla METS
 - dlibra_avs - dla schematu danej biblioteki cyfrowej
- <kolekcja> to identyfikator OAI-PMH kolekcji - jest on dostępny a Aplikacji Administratora po wybraniu danej kolekcji
- <nazwa pliku wynikowego> wskazuje na plik XML, w którym mają zostać zapisane pobrane dane

Ponadto można podać jeszcze dodatkowy parametr --csv, który spowoduje dodatkowo zapisanie metadanych w pliku CSV.

Przykładowe wywołanie dla kolekcji "Muzykalnia" z Wielkopolskiej Biblioteki Cyfrowej i metadanych w schemacie DublinCore wygląda następująco:

```
java -jar EuropeanaXMLBuilder-1.1-SNAPSHOT.jar -u "http://www.wbc.poznan.pl/dlibra/oai-pmh-repository.xml?verb=ListRecords&metadataPrefix=oai_dc&set=rootCollection:wbc:DigitalLibraryOfWielkopolska:MusicNotes" --csv -o Muzykalnia-WBC.xml -g
```