# 01. Pobieranie metadanych poprzez rozszerzenie Z39.50

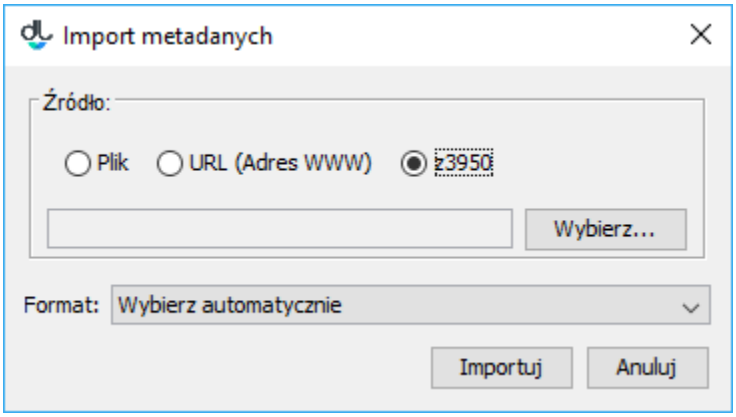
> **Notatka**
>
> W celu lepszego zrozumienia działania mechanizmu pobierania metadanych poprzez protokół Z39.50 czytelnik powinien zapoznać się z rozdziałem opisującym import metadanych w edytorze opisu bibliograficznego. Ponadto czytelnik powinien być zapoznany z protokołem Z39.50.

> **Uwaga**
>
> Biblioteka programistyczna dla tego rozszerzenia to `dlibra-app-extension-sl-z3950`.

Wiele systemów bibliotecznych udostępnia metadane poprzez protokół Z39.50. Opisywane rozszerzenie aplikacji redaktora umożliwia pobieranie metadanych z systemów, które udostępniają je poprzez protokół Z39.50. Protokół Z39.50 jest bardzo rozbudowanym standardem, który może przesyłać metadane w różnych formatach. Opisywane rozszerzenie pozwala na pobieranie metadanych tylko z formatu komunikacyjnego MARC 21 (konkretnie jest to format usmarc). Metadane pobrane przy pomocy tego rozszerzenia mogą być następnie zaimportowane przy użyciu rozszerzenia importującego pliki w formacie MARC.

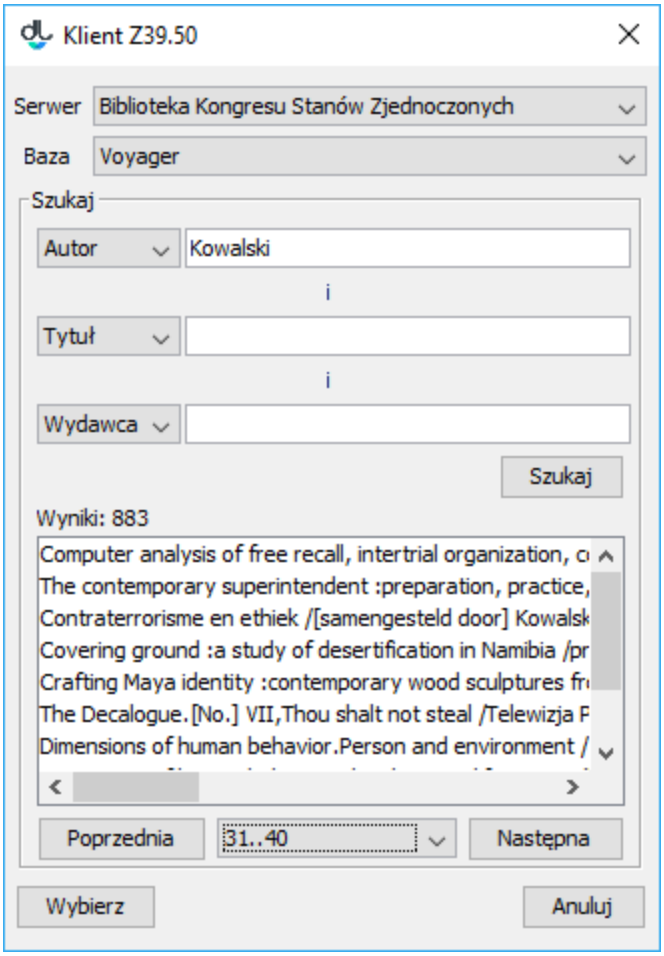
Aby pobrać metadane używając rozszerzenia Z39.50 należy:

1. W edytorze metadanych na panelu Import/Eksport metadanych wybrać przycisk Importuj..., który wyświetli okno importu metadanych (rysunek poniżej). Na oknie importu metadanych znajduje się opcja Z39.50, którą należy wybrać. Opcja ta determinuje wykorzystanie rozszerzenia Z39.50 w celu wyboru pliku metadanych do zaimportowania.




Okno importu metadanych

2. Aby wybrać metadane, które mają być zaimportowane należy wybrać przycisk Wybierz.... Pojawi się okno wyszukiwania metadanych (rysunek poniżej). W oknie tym użytkownik może wybrać serwer oraz bazę do której będą skierowane zapytania wyszukiwawcze. Zapytanie wyszukiwawcze budowane jest na podstawie trzech wartości/fraz, które wprowadzane są przez użytkownika do pól znajdujących się na panelu Szukaj. To jakie atrybuty będą przeszukiwane może być dostosowane przez użytkownika poprzez wybór odpowiedniego atrybutu z listy rozwijanej. Na przedstawionym rysunku ( Rysunek B.2.) wybrane do wyszukiwania atrybuty to autor, tytuł oraz wydawca. Aby rozpocząć wyszukiwanie należy wybrać przycisk Szukaj. Po zakończeniu procesu wyszukiwania możliwe jest przeglądanie wyników przy użyciu przycisków Poprzednia i Następna, które przechodzą odpowiednio do poprzedniej lub następnej strony wyników wyszukiwania. Redaktor może również wybrać bezpośrednio określoną stronę wyników wyszukiwania używając listy rozwijanej, która znajduje się pomiędzy przyciskami Poprzednia i Następna.




Okno wyszukiwania metadanych poprzez protokół Z39.50

3. Aby wybrać metadane należy zaznaczyć konkretną pozycję na liście wyników wyszukiwania a następnie wybrać przycisk Wybierz. Plik zostanie pobrany na lokalny dysk komputera, okno wyszukiwania Z39.50 zostanie zamknięte, a w oknie importu metadanych zostanie automatycznie wprowadzona ścieżka do pobranego pliku metadanych. Kolejne kroki należy wykonać zgodnie z informacjami zawartymi w sekcji import metadanych (w najprostszym przypadku wystarczy wybrać przycisk Importuj na oknie importu metadanych).

Rozszerzenie Z39.50 może być dostosowane do potrzeb konkretnej instalacji poprzez odpowiednią jego konfigurację. Konfiguracja rozszerzenia znajduje się w pliku konfiguracyjnym `z3950_servers.xml`. Plik ten definiuje m. in. serwery oraz bazy danych w których redaktor będzie mógł wyszukiwać metadane, atrybuty po jakich redaktor będzie mógł przeszukiwać konkretny serwer oraz operator logiczny jaki będzie użyty do połączenia zapytań z poszczególnych pól wyszukiwawczych. Poniżej przedstawiono domyślny plik konfiguracyjny rozszerzenia Z39.50.

```
<?xml version="1.0" encoding="UTF-8"?>
<servers>
        <server>
                <name>Library of Congress</name>
                <name lang="en">Library of Congress</name>
                <name lang="pl">Biblioteka Kongresu Stanów Zjednoczonych</name>
                <host>z3950.loc.gov</host>
                <port>7090</port>
                <databases>
                        <database>Voyager</database>
                </databases>
                <recordencoding>MARC-8</recordencoding>
                <queries>
                        <attrset>@attrset bib-1</attrset>
                        <operator>@and</operator>
                        <query>
                                <name>Author</name>
                                <name lang="pl">Autor</name>
                                <name lang="en">Author</name>
                                <searchquery>@attr 1=1003 $1</searchquery>
                        </query>
                        <query>
                                <name>Title</name>
                                <name lang="pl">Tytu</name>
                                <name lang="en">Title</name>
                                <searchquery>@attr 1=4 $1</searchquery>
                        </query>
                        <query>
                                <name>Publisher</name>
                                <name lang="pl">Wydawca</name>
                                <name lang="en">Publisher</name>
                                <searchquery>@attr 1=1018 $1</searchquery>
                        </query>
                        <query>
                                <name>Everywhere</name>
                                <name lang="pl">Wszdzie</name>
                                <name lang="en">Everywhere</name>
                                <searchquery>@attr 1=1035 $1</searchquery>
                        </query>
                </queries>
        </server>
</servers>
```

Plik konfiguracyjny zapisany jest w formacie XML. Głównym węzłem w pliku jest `<servers>`, w którym znajdują się definicje serwerów Z39.50 (liczba serwerów nie jest ograniczona). Każdy serwer zdefiniowany jest w ramach znacznika `<server>` w którym znajdować się powinny następujące informacje w poszczególnych znacznikach:

- `<name>` - jest to nazwa serwera wyświetlana redaktorowi w oknie wyszukiwania metadanych. Nazwę można wprowadzić dla konkretnego języka poprzez zastosowanie w znaczniku `<name>` atrybutu `lang`. Wartością atrybutu `lang` powinien być dwuliterowy symbol języka (zgodnie ze standardem ISO 639), w którym specyfikowana jest nazwa serwera. Domyślna nazwa serwera określana jest w znaczniku `<name>`, który nie posiada atrybutu `lang`. Proszę zauważyć, że poszczególne nazwy definiowane są w odrębnych znacznikach `<name>`.
- `<host>` - jest to adres internetowy serwera Z39.50 (adres IP lub adres domenowy).
- `<port>` - jest to port serwera Z39.50.
- `<user>` - nazwa użytkownika, jeśli serwer wymaga uwierzytelniania
- `<password>` - hasło użytkownika, jeśli serwer wymaga go do uwierzytelniania
- `<databases>` - w tym znaczniku znajdują się nazwy baz danych, które redaktor może przeszukiwać. Poszczególne nazwy powinny być wprowadzane w znacznikach `<database>`.
- `<recordencoding>` - jest to kodowanie rekordów metadanych, które udostępnia definiowany serwer Z39.50.
- `<queryencoding>` - jest to kodowanie, w jakim mają być wysyłane frazy w zapytaniach do serwera. Znacznik jest opcjonalny, domyślnie stosowane jest kodowanie US-ANSI.
- `<setname>` - jest to nazwa zbioru właściwości, który ma być zwrócony dla każdego znalezionego elementu. Jest to znacznik opcjonalny, domyślnie przyjmuje wartość F (od full - pełen zbiór). Często jest też używana wartość B (od brief - skrócone). Jeśli nazwa zbioru nie jest poprawnie ustawiona, pobieranie danych zakończy się błędem, a w konsoli Java pojawi się komunikat: `Z3950SearchTask - Non surrogate diagnostics [25]`.
- `<queries>` - jest to węzeł w którym definiuje się informacje związane z panelem wyszukiwawczym. Wszystkie wartości, które rozpoczynają się od znaku @ związane są z językiem zapytań stosowanym w bibliotece JZKit 2 (http://jzkit.org/). I tak:
    - `<attrset>` - określa zestaw atrybutów jaki jest używany do wyszukiwania.
    - `<operator>` - określa operator logiczny, który będzie używany do łączenia specyfikowanych wartości/fraz w polach wyszukiwawczych.

- `<query>` - specyfikuje nazwę atrybutu po którym możliwe będzie wyszukiwanie oraz zapytanie wyszukiwawcze związane z tym atrybutem. Nazwa atrybutu podobnie jak w przypadku nazwy serwera może być wprowadzona w wielu językach (użycie atrybutu `lang`). Domyślna nazwa atrybutu to wartość zawarta w znaczniku `<name>`, który nie posiada atrybutu lang. Zapytanie wyszukiwawcze specyfikuje się w znaczniku `<searchquery>`, gdzie w miejsce znacznika `$1` wstawiana jest wartość z odpowiedniego pola wyszukiwawczego.

Aby zmienić konfigurację rozszerzenia Z39.50 należy postępować zgodnie z informacjami zawartymi w sekcji konfiguracja aplikacji.