04. Wyszukiwanie i przeglądanie zasobów w bibliotece cyfrowej

System dLibra oferuje swoim użytkownikom wiele sposobów dotarcia do interesującego obiektu cyfrowego. Podstawowym sposobem jest wyszukiwanie. Ponadto możliwe jest przeglądanie zawartości biblioteki cyfrowej poprzez kolekcje czy indeksy słów. Dodatkowo każdy obiekt cyfrowy w systemie dLibra posiada swój własny, unikalny link dostępu oraz identyfikator OAI.

Wyszukiwanie lokalne

Wyszukiwanie lokalne obiektów cyfrowych w systemie dLibra jest podstawowym mechanizmem wykorzystywanym przez użytkowników stron internetowych. Jest to najprostszy i najłatwiejszy w wykorzystaniu mechanizm odnajdywania interesujących pozycji. Mechanizm ten wyszukuje w zgromadzonych lokalnie zasobach obiekty i prezentuje listę wyników wyszukiwania poczynając od elementów pasujących najbardziej do zapytania. Informacjami przeszukiwanymi są metadane oraz ewentualny tekst publikacji jeśli jest dostępny. Mechanizm wyszukiwania w systemie dLibra oparty jest na silniku wyszukiwawczym Solr. Wszelkie właściwości mechanizmu wyszukiwania (np. konstrukcja zapytań wyszukiwawczych, czynniki wpływające na kolejność wyników wyszukiwania) bazują na silniku Solr i w jego dokumentacji można odnaleźć więcej informacji na temat zasad jego działania.

Wykorzystanie słownika synonimów

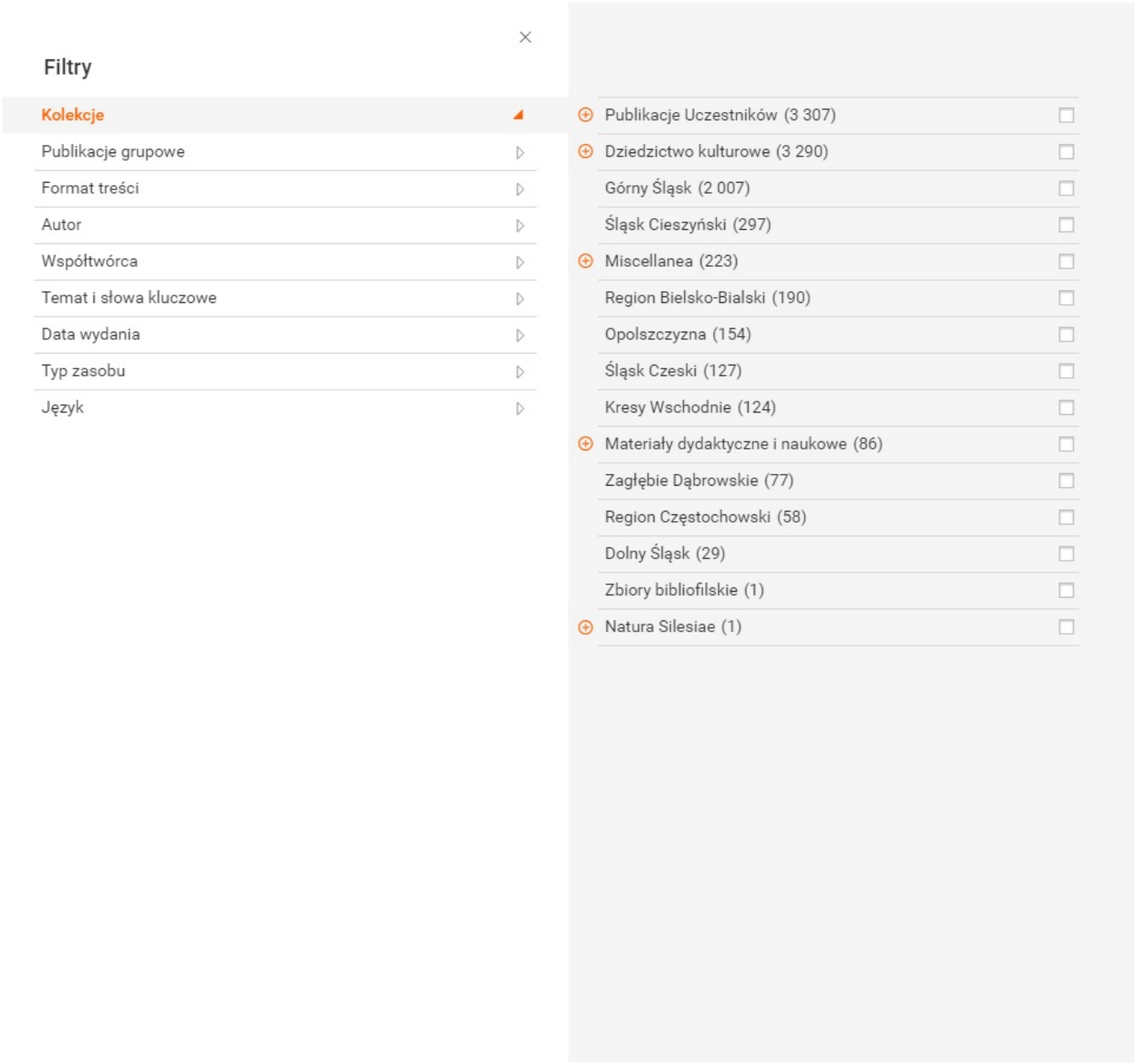
Wyszukiwanie lokalne oferuje również mechanizm rozszerzania zapytań o słowa/frazy. Budowa tego słownika zostały opisane tutaj.

Automatyczne rozszerzanie zapytań

Rozszerzanie zapytań odbywa się automatycznie w przypadku gdy zapytanie użytkownika zostanie dopasowane do grupy synonimów. W takim przypadku dLibra uzupełnia wpisane zapytanie o wszystkie słowa/frazy zawarte w wybranej grupie synonimów. Rozważmy następujący przykład. Użytkownik wpisuje słowo *Kopernik* w słowniku synonimów znajduje się grupa *Copernicus* zawierająca dwie wartości *Kopernik* i *Copernicus*. Zapytanie które trafi do indeksu będzie miało następującą postać: *Kopernik OR Copernicus*.

Filtrowanie zapytań

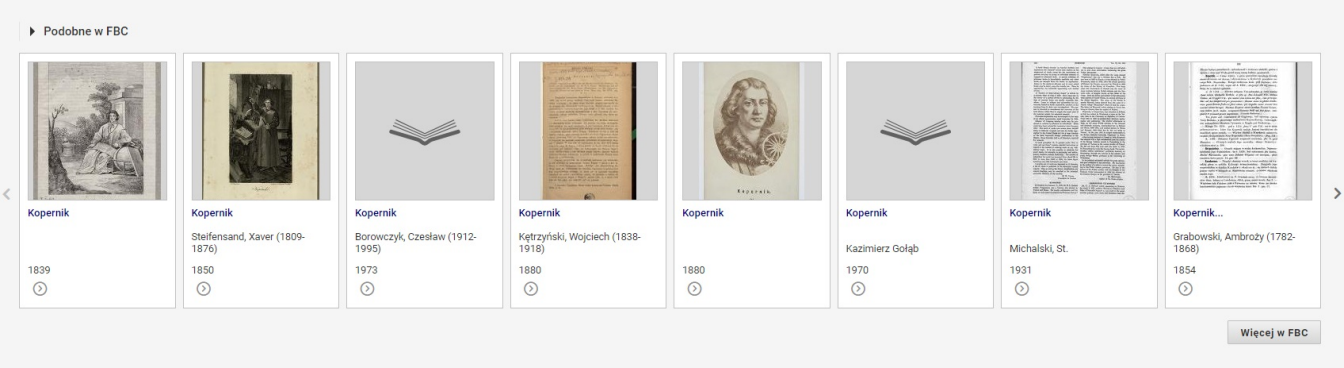
Wszystkie wyniki wyszukiwania mogą zostać zawężone dodatkowymi filtrami. Dostępne filtry wyświetlają się po każdym zapytaniu wyszukiwawczym na stronie wyników wyszukiwania. Wyniki możemy zawęzić po wartościach atrybutów, przynależności do kolekcji, przynależności do publikacji grupowej a także po formacie treści.



Wyszukiwanie zasobów rozproszonych

Poza wyszukiwaniem lokalnym istnieje możliwość włączenia do biblioteki cyfrowej dLibra wyszukiwania zasobów rozproszonych. Wyszukiwanie rozproszonych zasobów bazuje na protokole OAI-PMH i możliwe jest przy wykorzystaniu usługi wyszukiwania zasobów rozproszonych portalu Federacji Bibliotek Cyfrowych.

Wraz z każdym zapytaniem wyszukiwawczym następuje wysłanie zapytania do serwisu FBC. Najwyżej punktowane wyniki wyświetlane są pod listą wyników z biblioteki lokalnej. Istnieje możliwość wyświetlenia wszystkich wyników w serwisie FBC. Wystarczy kliknąć przycisk "Więcej w FBC". Nastąpi przekierowanie do serwisu FBC, a następnie wyświetlona zostanie lista elementów znalezionych we wszystkich bibliotekach cyfrowych, które wchodzą w skład FBC (aktualna lista znajduje się tutaj). Serwis *Federacja Bibliotek Cyfrowych* (FBC) jest kolejnym etapem budowy sieci rozproszonych bibliotek cyfrowych i repozytoriów w Polsce. Nazwa serwisu FBC odzwierciedla jego charakter - serwis ten jest zbiorem zaawansowanych usług sieciowych opartych na zasobach cyfrowych dostępnych w polskich bibliotekach cyfrowych i repozytoriach uruchomionych w sieci PIONIER. Zasoby te współtworzone są przez wiele instytucji naukowych i publicznych, takich jak uczelnie wyższe, biblioteki, archiwa, muzea czy ośrodki badawcze. Serwis FBC utrzymywany jest przez Poznańskie Centrum Superkomputerowo-Sieciowe afiliowane przy Instytucie Chemii Bioorganicznej PAN.

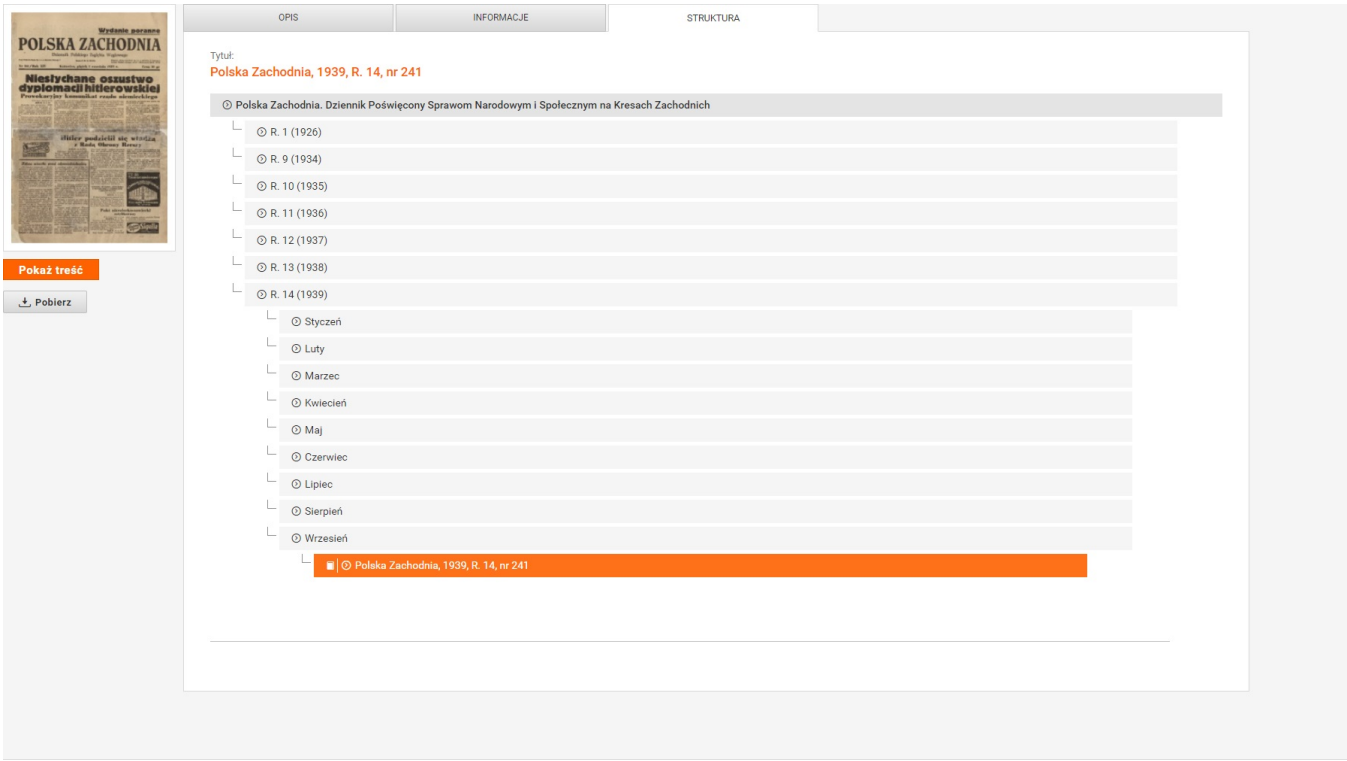


Metody przeglądania zasobów biblioteki cyfrowej

Poza wyszukiwaniem obiektów cyfrowych istnieje jeszcze kilka innych sposobów odnajdywania i przeglądania obiektów cyfrowych. Jest to przede wszystkim zestaw kolekcji, który w zorganizowany sposób dzieli obiekty cyfrowe na mniejsze zbiory. Naturalnym podziałem wydaje się być podział tematyczny i tak też wykorzystuje kolekcje większość bibliotek cyfrowych opartych na oprogramowaniu dLibra. Dzięki hierarchicznej organizacji kolekcji podział może rozpoczynać się od bardzo ogólnych zbiorów (np. *Materiały dydaktyczne*) , które podzielone są na bardziej zawężone podzbiory (np. *Matematyka*, *Fizyka*), itd.

Innym sposobem przeglądania obiektów w bibliotece cyfrowej są indeksy utworzone z wszystkich wartości wybranego atrybutu metadanych. To jakie atrybuty będą wykorzystywane do generowania indeksów dostosowywane jest do konkretnych potrzeb instytucji wykorzystującej bibliotekę cyfrową. Przykładowo mogą to być indeksy utworzone na bazie wartości z atrybutów *Autor*, *Tytuł* czy *Temat i słowa kluczowe*.

Obiekty złożone, będące wynikiem zastosowania publikacji grupowej, są przedstawiane w systemie dLibra jako pewna hierarchiczna struktura, po której użytkownik może się swobodnie poruszać. Dzięki temu użytkownik ma dostęp do dowolnej części całego obiektu cyfrowego. Przykładowo jeśli jest to struktura gazety to może ona mieć postać jak przedstawiono na rysunku poniżej.



Stały odnośnik i identyfikator obiektu cyfrowego

Każda publikacja w systemie biblioteki cyfrowej dLibra posiada unikalny odnośnik do strony internetowej pod którą zawsze znajdują się informacje na jej temat. Strona ta zbudowana jest z adresu strony internetowej biblioteki cyfrowej po którym następuje ciąg znaków `/publication/` oraz identyfikator publikacji. Przykładowo może to być następujący odnośnik: https://sbc.org.pl/publication/257008, który składa się z adresu strony internetowej https://sbc.org.pl, ciągu znaków `/publication/` oraz identyfikatora publikacji w tej konkretnej bibliotece cyfrowej 257008.

Ponadto każda publikacja w systemie dLibra posiada automatycznie wygenerowany globalny i unikalny identyfikator który jest oparty na otwartym standardzie OAI Identifier. Każdy taki identyfikator składa się z przedrostka `oai`, domeny repozytorium w którym się dany obiekt znajduje oraz unikalnego identyfikatora w ramach tego repozytorium. Wszystkie te elementy powinny być oddzielone znakiem dwukropka `:`. Zgodnie z tą zasadą publikacja z poprzedniego przykładu miałaby następujący identyfikator: oai:sbc.org.pl:243088. Szczegóły można znaleźć w samym standardzie OAI Identifier.