Repozytorium

Krótki opis usługi

Uzupełnieniem usług syndykacji i przetwarzania dużych zbiorów danych jest usługa udostępniania zbiorów danych (repozytorium Complex Networks), zarówno pobranych jak i tych, które są uważane jako referencyjne. Dane pobrane za pomocą usługi syndykacji, są automatycznie dodawane do repozytorium. Dzięki temu użytkownik w prosty i szybki sposób może uzyskać dostęp do cyfrowych danych, których pobieranie zlecił. Dodatkowo, umieszczenie ich w repozytorium, daje możliwość odwoływania się do nich w publikacjach i innych pracach naukowych. Dostęp do danych udostępnianych w ramach usługi, pozwala na szybką i łatwą weryfikację hipotez będących wynikiem prac badawczych oraz porównywanie między sobą kolejnych opracowywanych metod.

Repozytoria zbiorów danych cieszą się dużą popularnością, gdyż zebranie wielu zbiorów przyspiesza i ułatwia prowadzenie oraz weryfikację badań. Podział danych ze względu na kategorię oraz dodanie do nich metadanych umożliwia filtrowanie zbiorów ze względu na potrzeby i zainteresowania użytkownika.

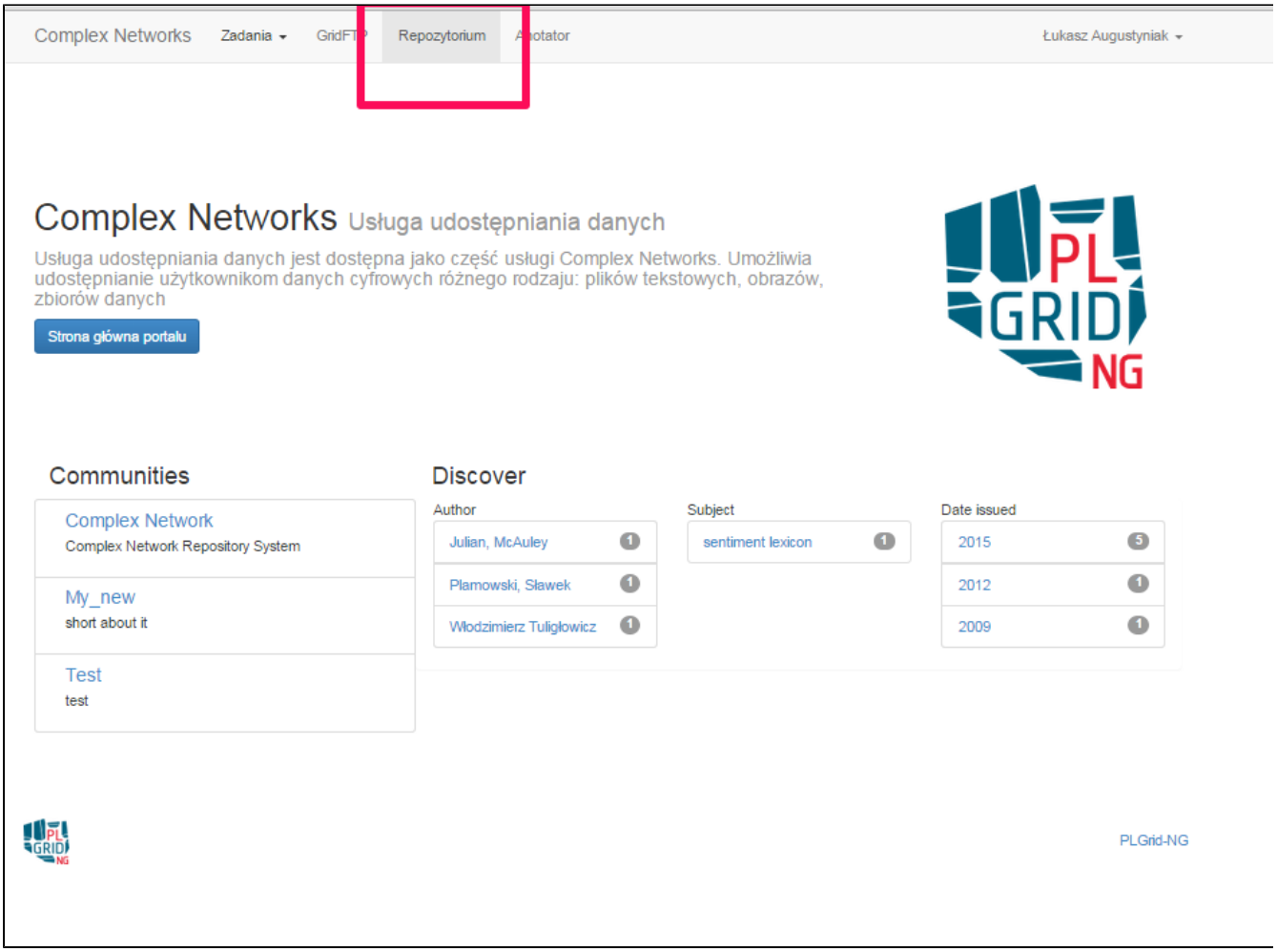
Aktywowanie usługi

Aby korzystać z usługi *Repozytorium* należy posiadać konto w infrastrukturze PLGrid, a następnie złożyć wniosek o dostęp do usługi w portalu https://portal.plgrid.pl/.

Pierwsze kroki

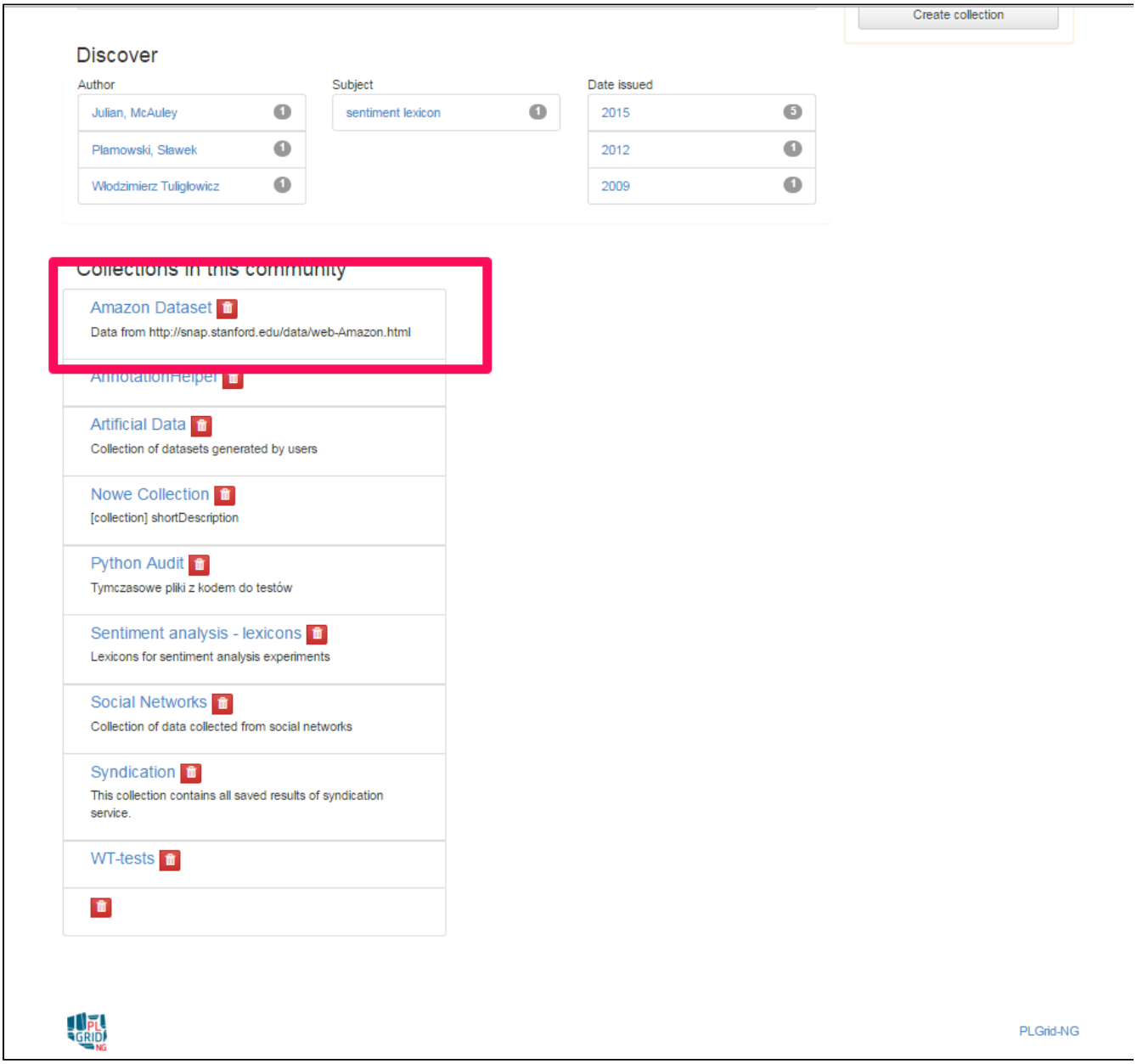
Repozytorium - główna strona

Chcąc przejść do repozytorium należy wybrać z menu głównego zakładkę **Repozytorium**

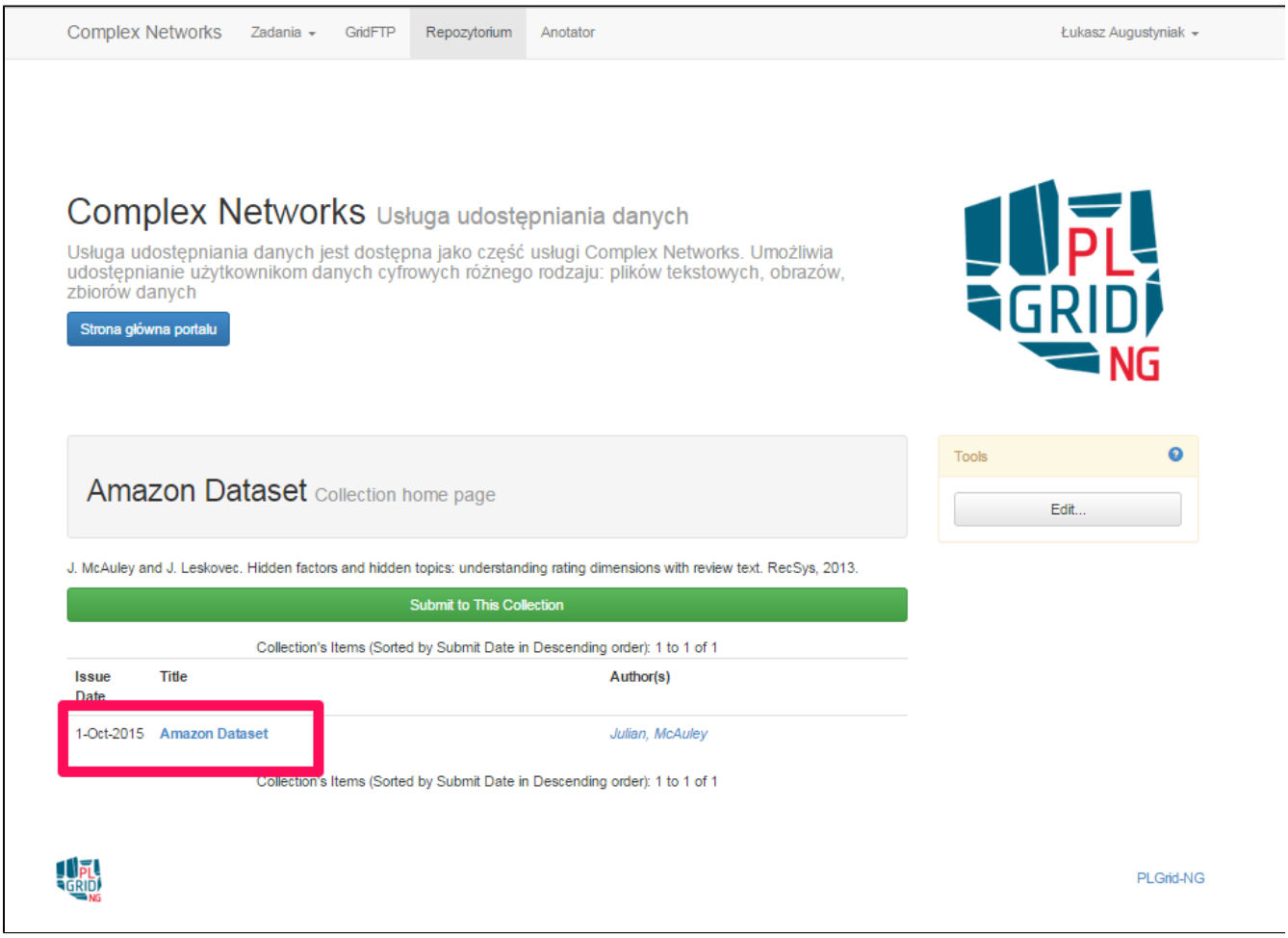


Dodanie danych do repozytorium - Amazon Data

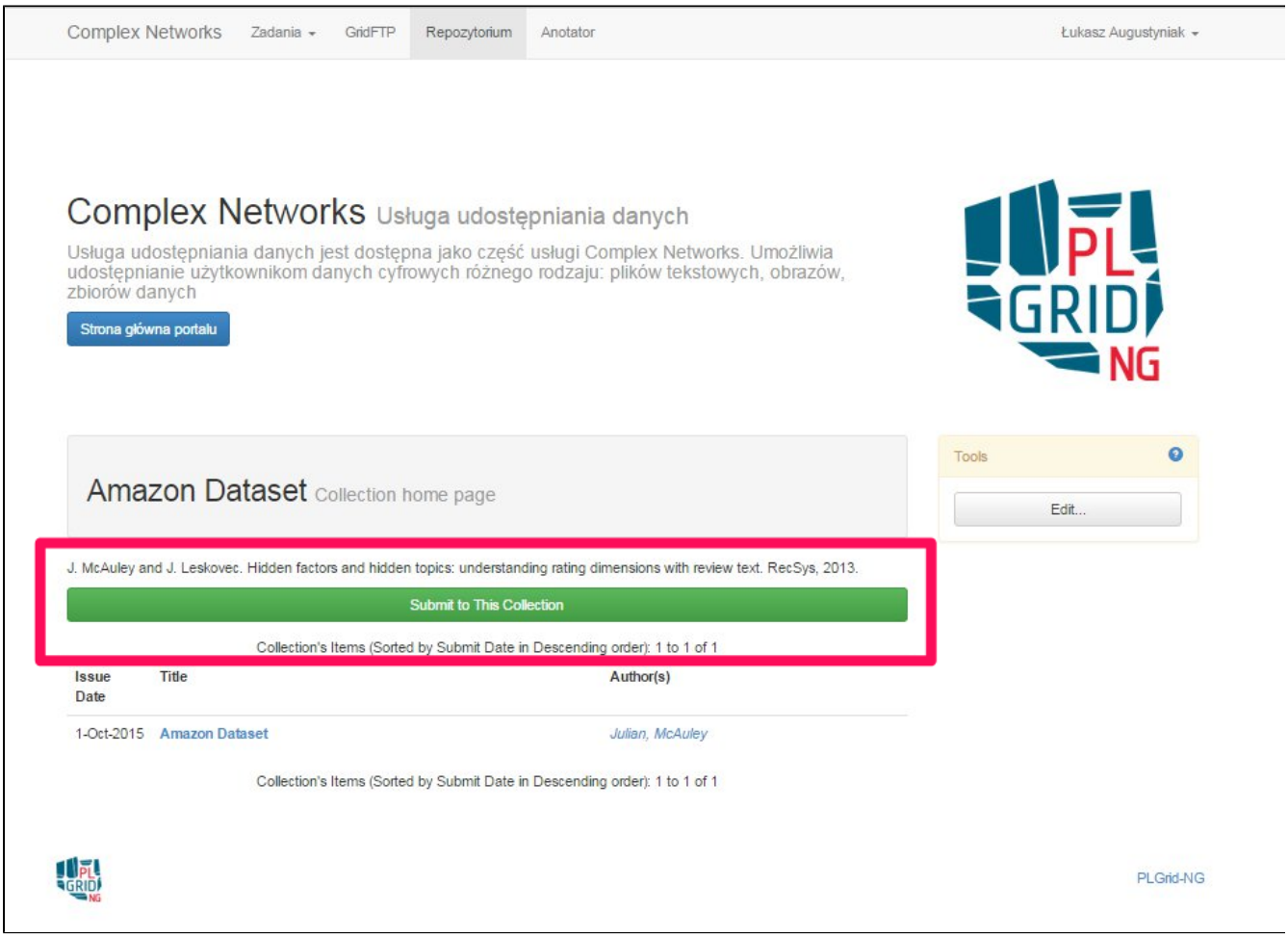
Chcąc dodać dane do repozytorium, należy wybrać kolekcję, w której dane zostaną umieszczone. Lista kolekcji jest widoczna na pierwszej stronie usługi w dolnej części strony.



Po wybraniu docelowej kolekcji, wyświetlana jest jej strona główna. Na tej stronie można przejrzeć paczki danych umieszczone w kolekcji. Dane są prezentowane w formie listy. Po kliknięciu na element listy, można przejść do wyświetlenia szczegółów paczki danych.

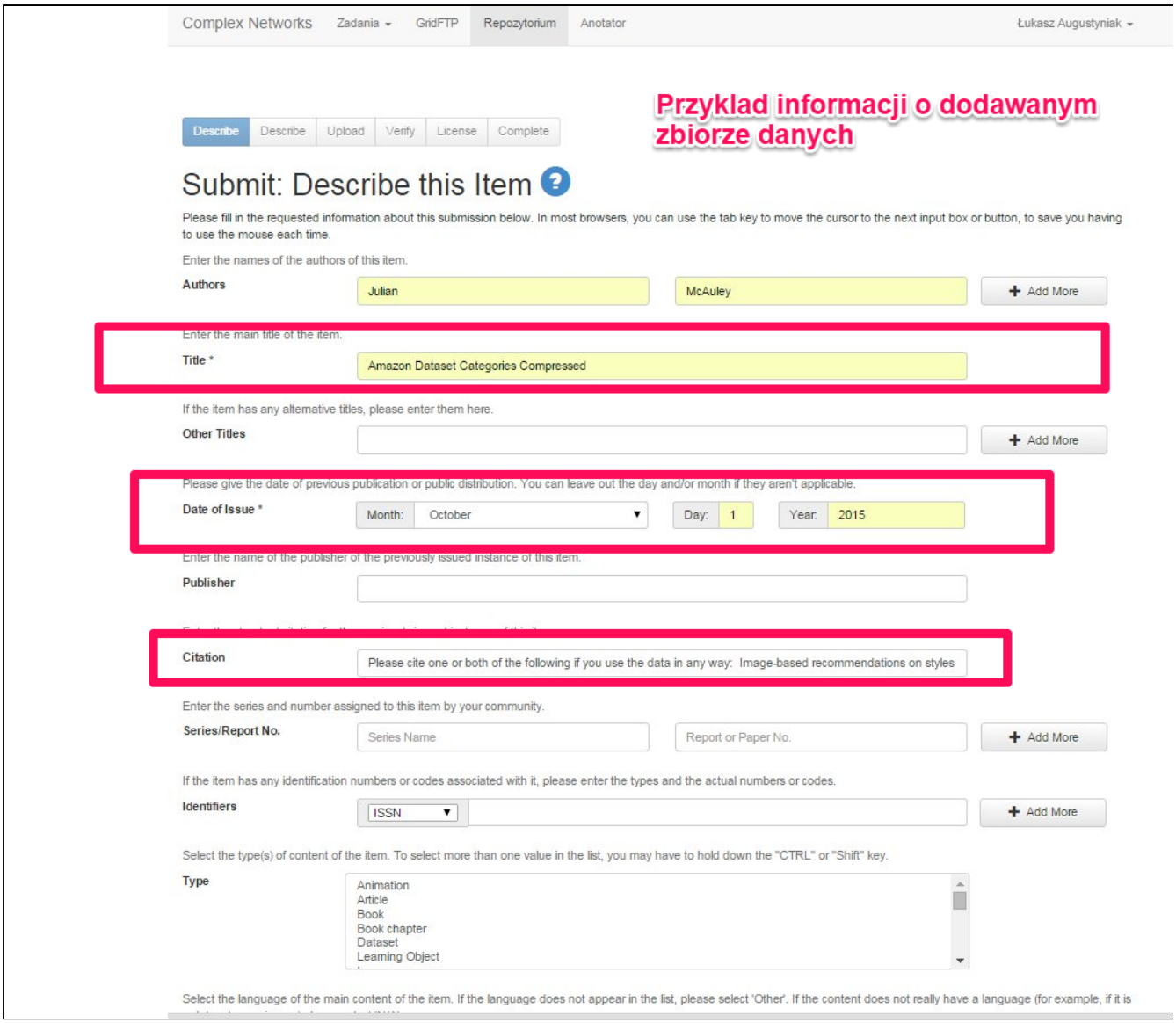


Pod tytułem kolekcji a nad listą paczek danych znajduje się link umożliwiający dodawanie nowych paczek danych do kolekcji *Submit to this collection*.

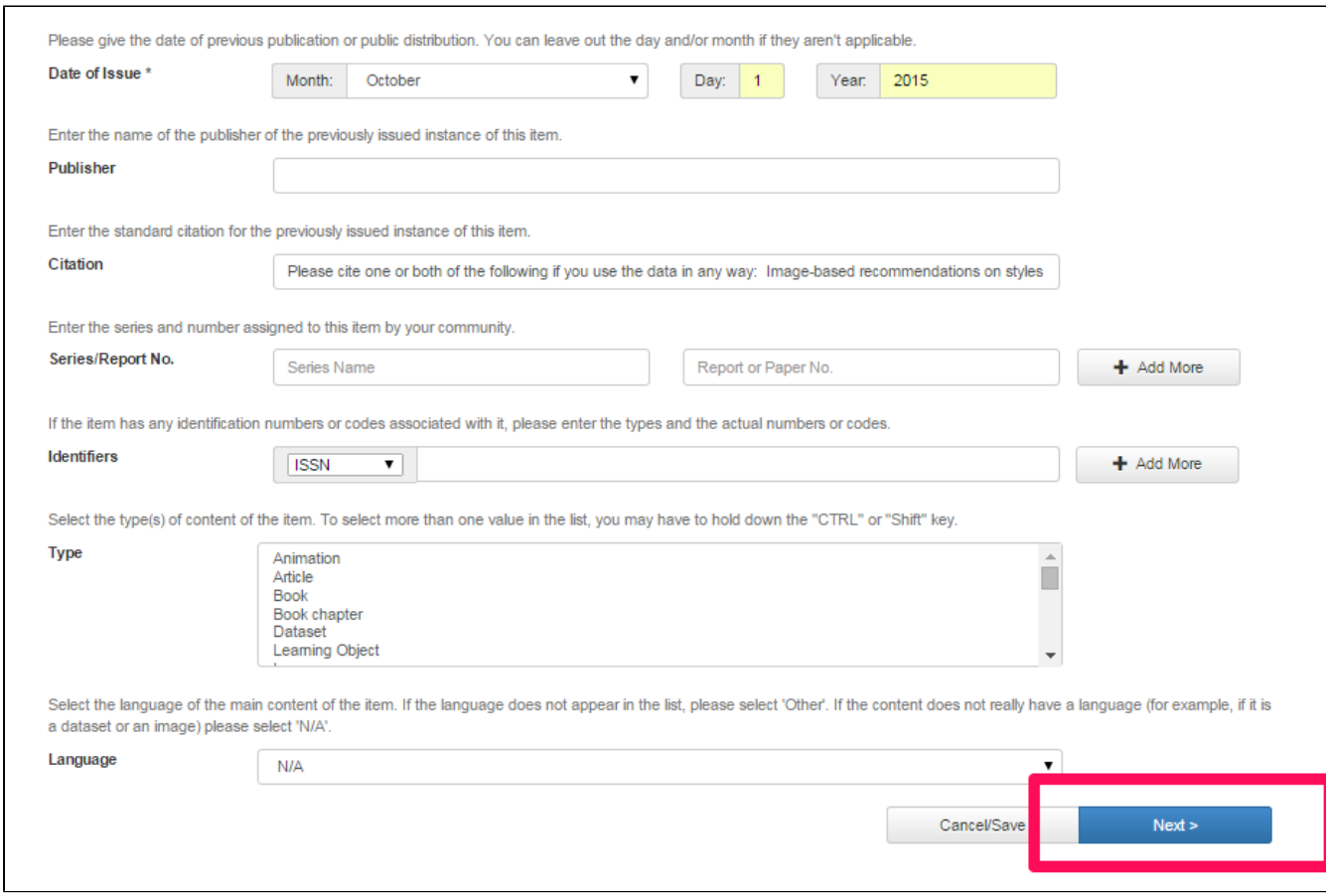


Aby dodać paczkę danych do kolekcji należy, oprócz wskazania lokalizacji pliku na dysku, podać informacje opisujące paczkę danych. Pola oznaczone gwiazdką (*****) są wymagane. Przy każdym z pól znajduje się opis wskazujący na znaczenie przekazywanych informacji.

Po wypełnieniu danych, wciskamy **Next**, aby przejść do kolejnego ekranu.

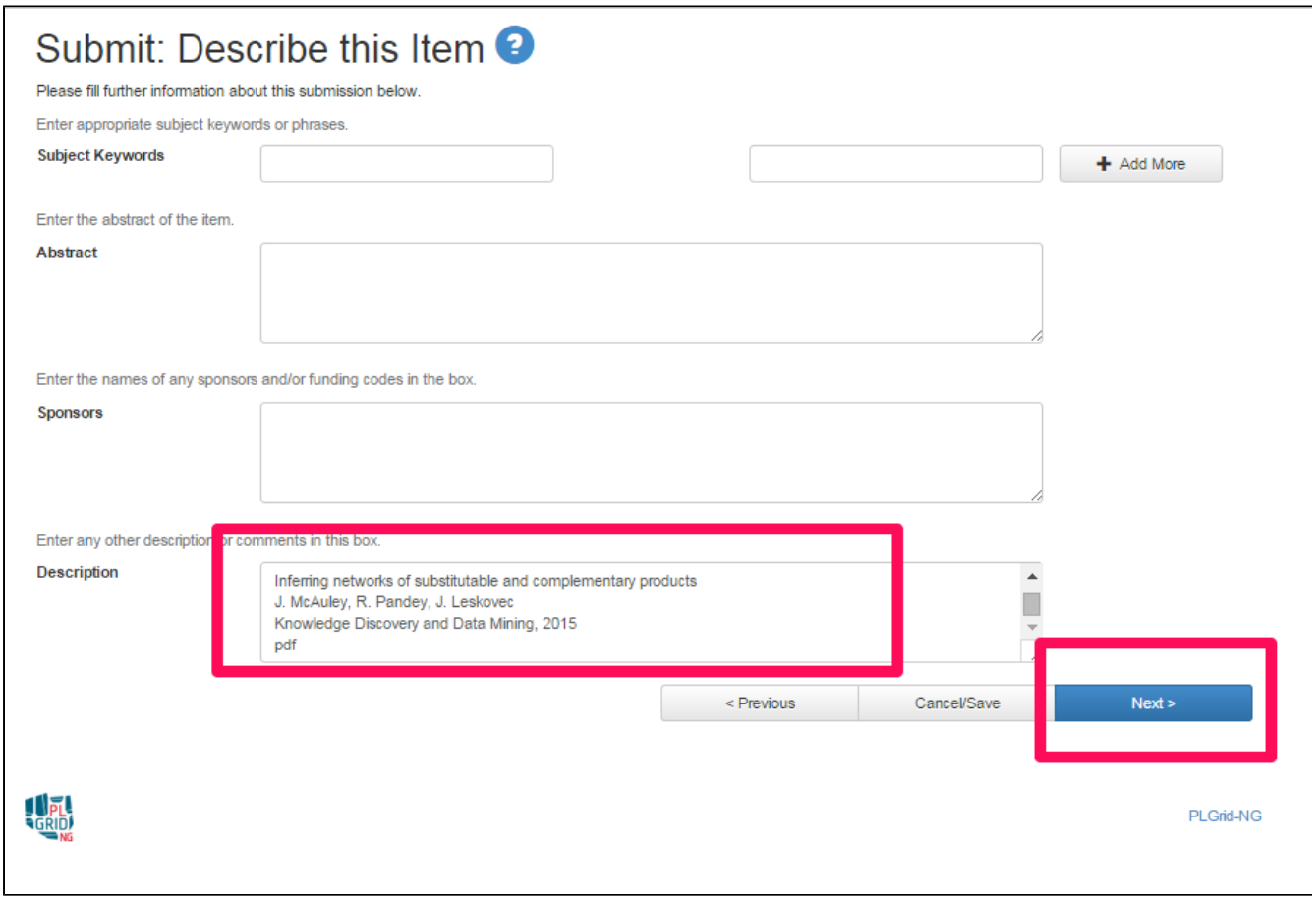


Na kolejnym ekranie należy podać pozostałe wymagane informacje. Po zakończeniu należy wcisnąć **Next**, aby przejść do następnego ekranu.



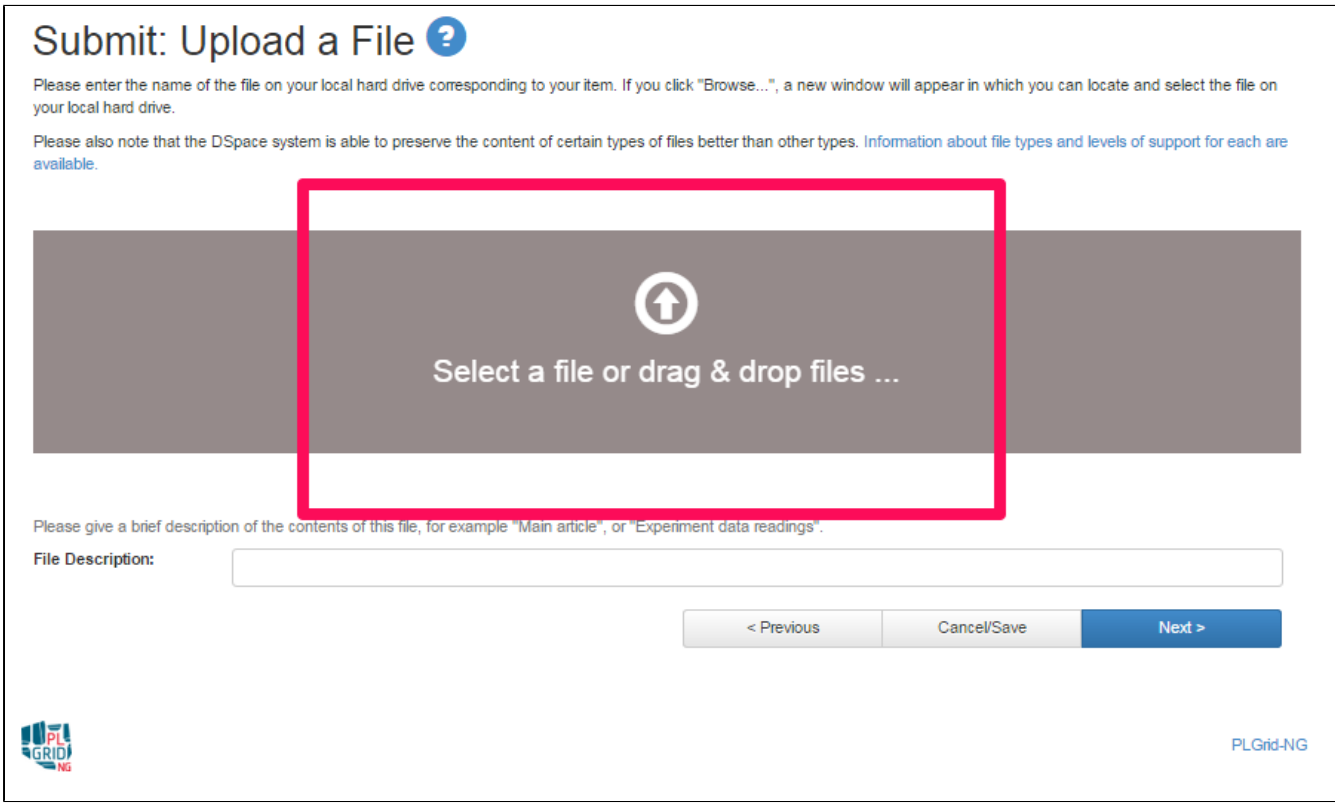
Na kolejnym ekranie należy wypełnić pole **Description**, w którym powinien się znaleźć krótki opis zawartości paczki danych dodawanych do kolekcji.

Po zakończeniu, należy wybrać przycisk **Next**, aby przejść do ekranu dodawania pliku z dysku.



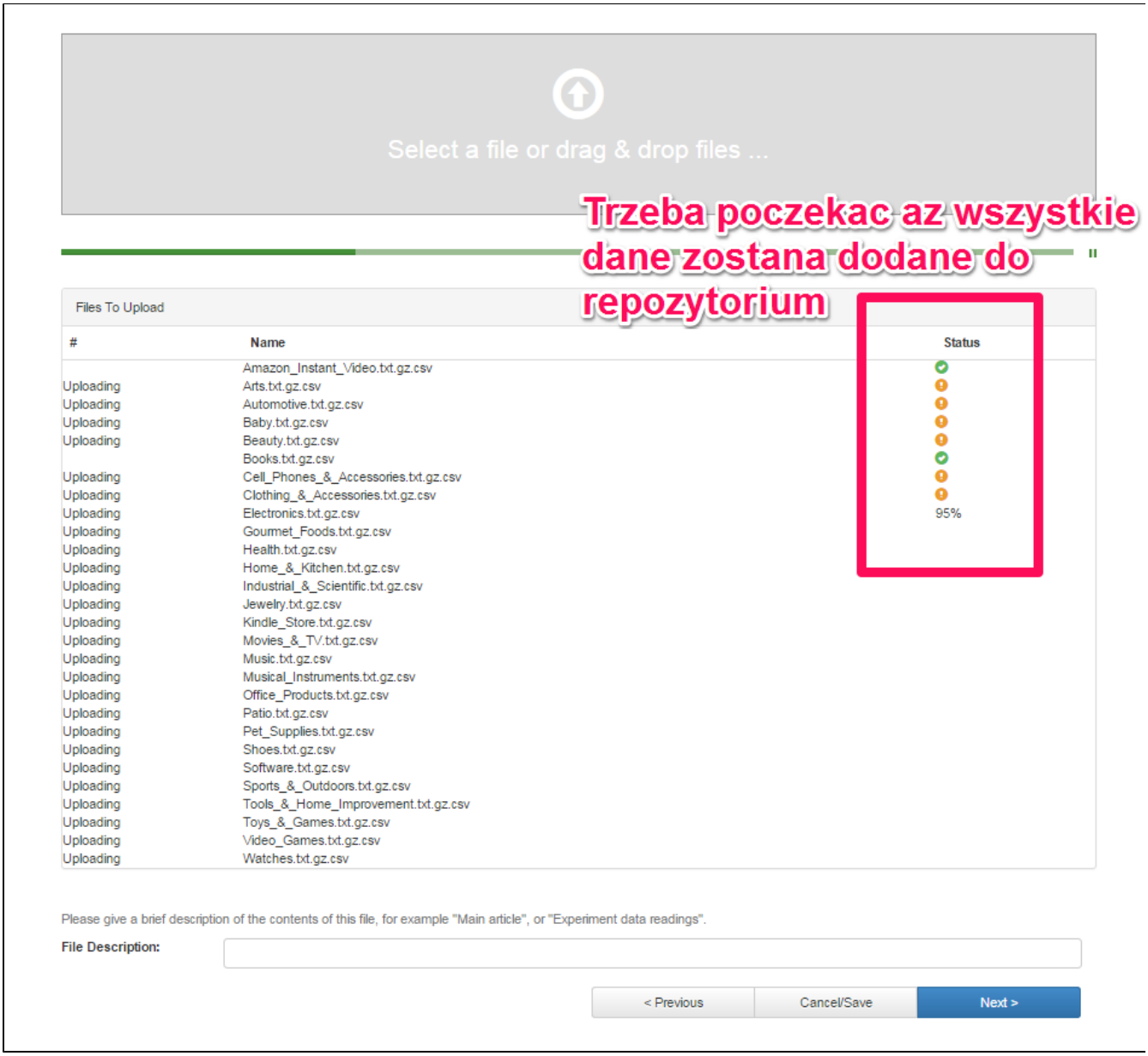
Pliki można dodawać na dwa sposoby:

1. Przez przeciągnięcie pliku z dysku na pasek *Select a file or drag & drop files...*, lub
2. Przez wskazanie lokalizacji pliku



Postęp przesyłania plików na serwer można śledzić na zielonym pasku postępu na środku ekranu.

Ważne Należy poczekać, aż wszystkie pliki zostaną dodane do repozytorium. W przeciwnym razie, może dojść do nieprzewidzianych błędów w czasie korzystania z repozytorium.



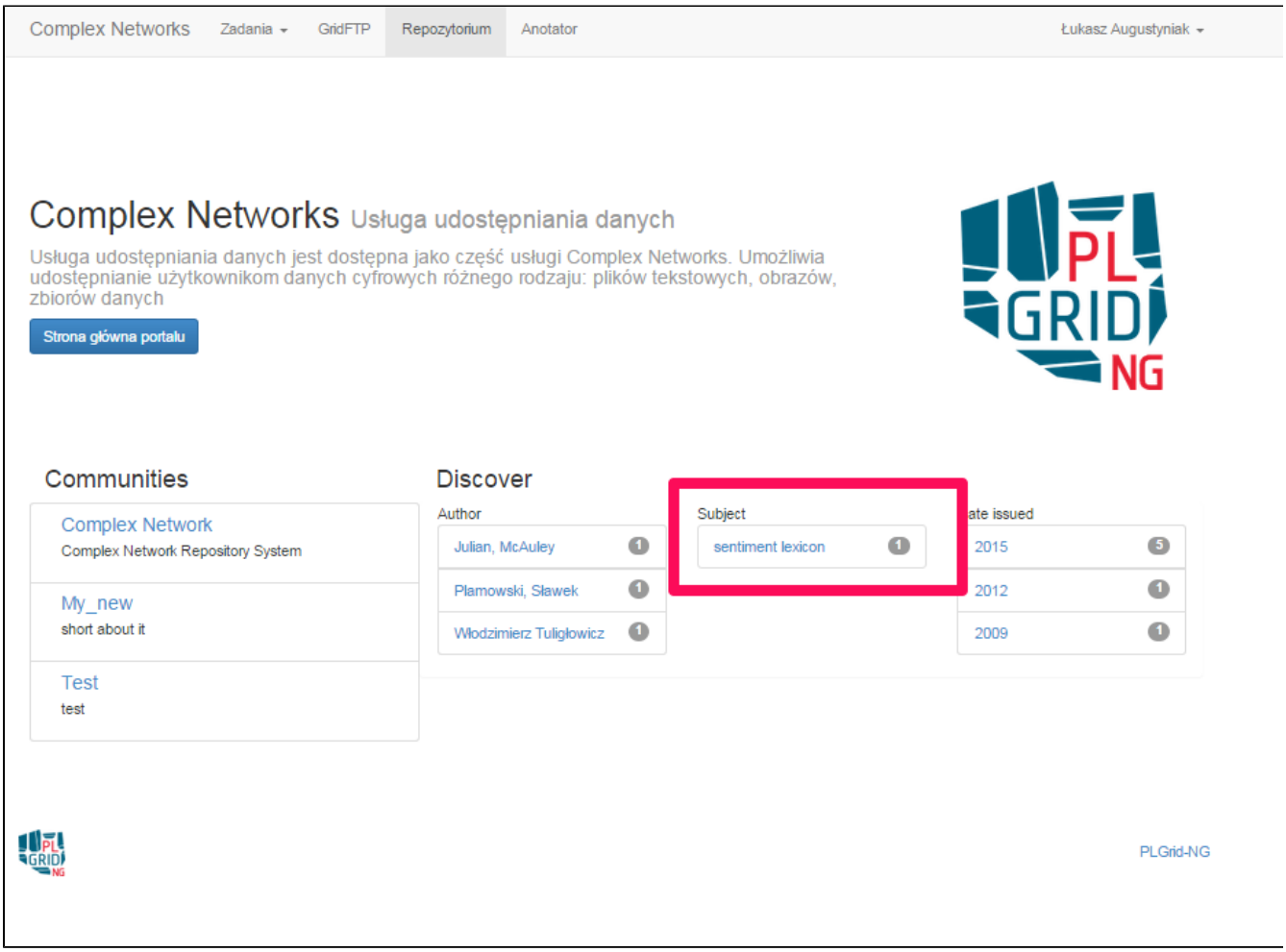
W celu szybkiego dodanie danych do repozytorium wystarczy kliknąć przycisk Next, sprawdzić poprawność wgranych plików oraz opisu tychże danych danych (zakładka **Verify)** oraz wyrazić zgodę na politykę prywatności i licencję repozytorium (**I grant the Licence**). Po wykonaniu tych kroków pliki zostaną dodane do repozytorium.

Pobranie zbioru danych z repozytorium - analiza wydźwięku leksykon

Po wejściu do Portalu przechodzimy do zakładki **Repozytorium**

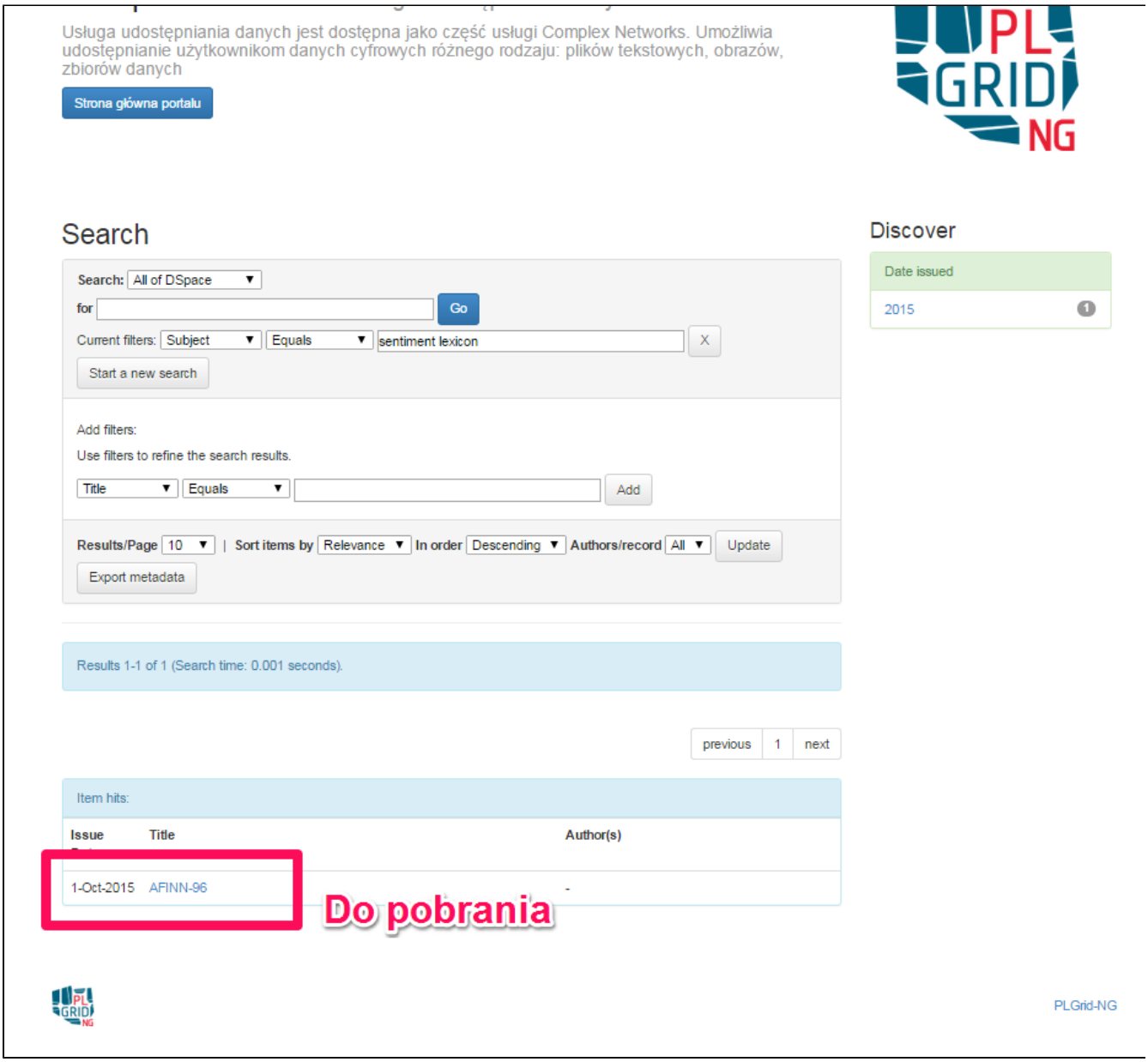
W dolnej części ekranu wyświetlane są informacje o paczkach danych przechowywanych w repozytorium pogrupowane wg. autorów, tematów i dat.

Wybieramy interesującą nas kategorię.

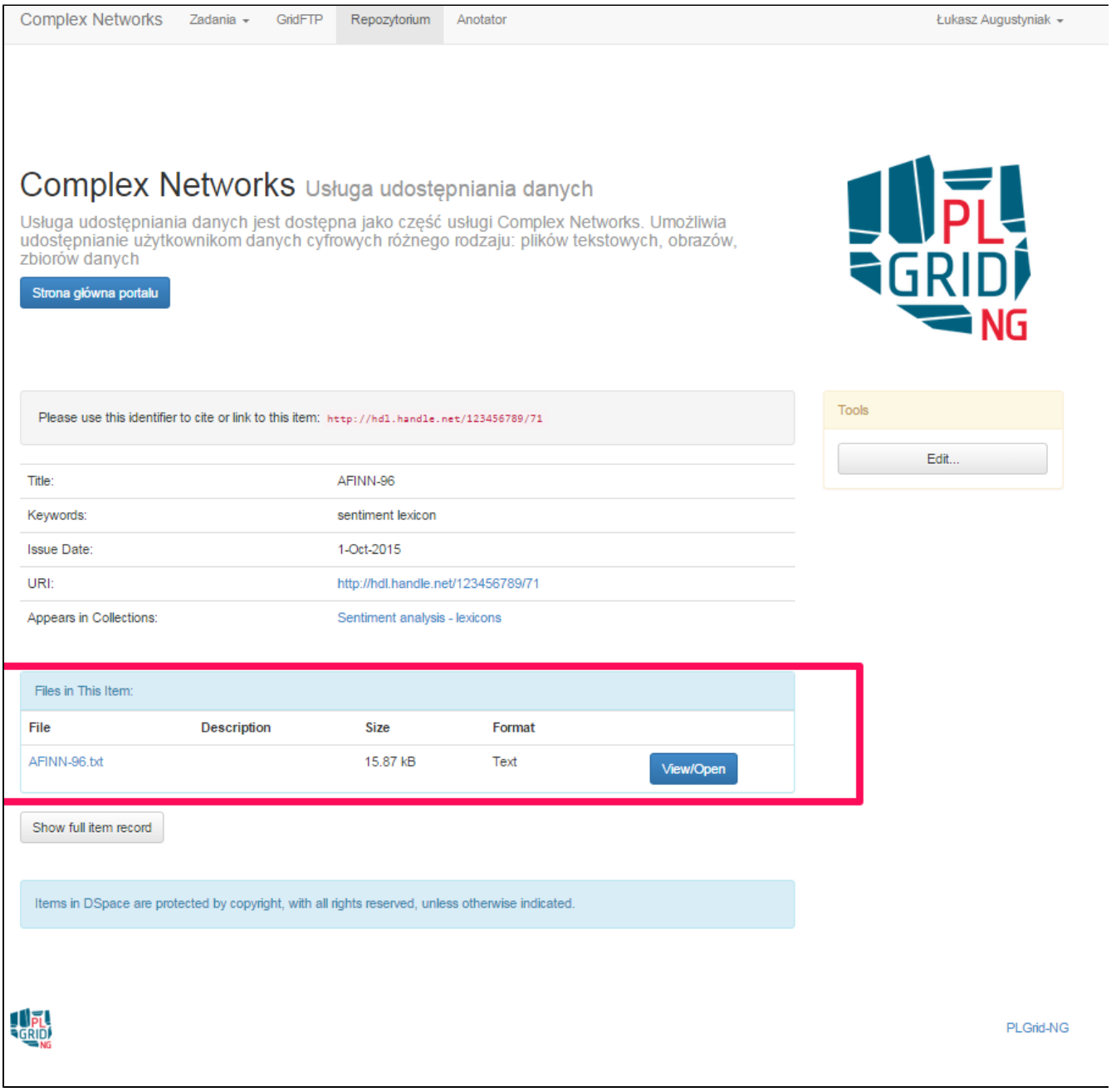


Na kolejnym ekranie możemy wyszukiwać lub filtrować wyniki wyszukiwania paczek danych. W naszym przypadku jest tylko jedna paczka dotycząca interesującego nas tematu.

Wybieramy plik do pobrania przez kliknięcie w **Title**.

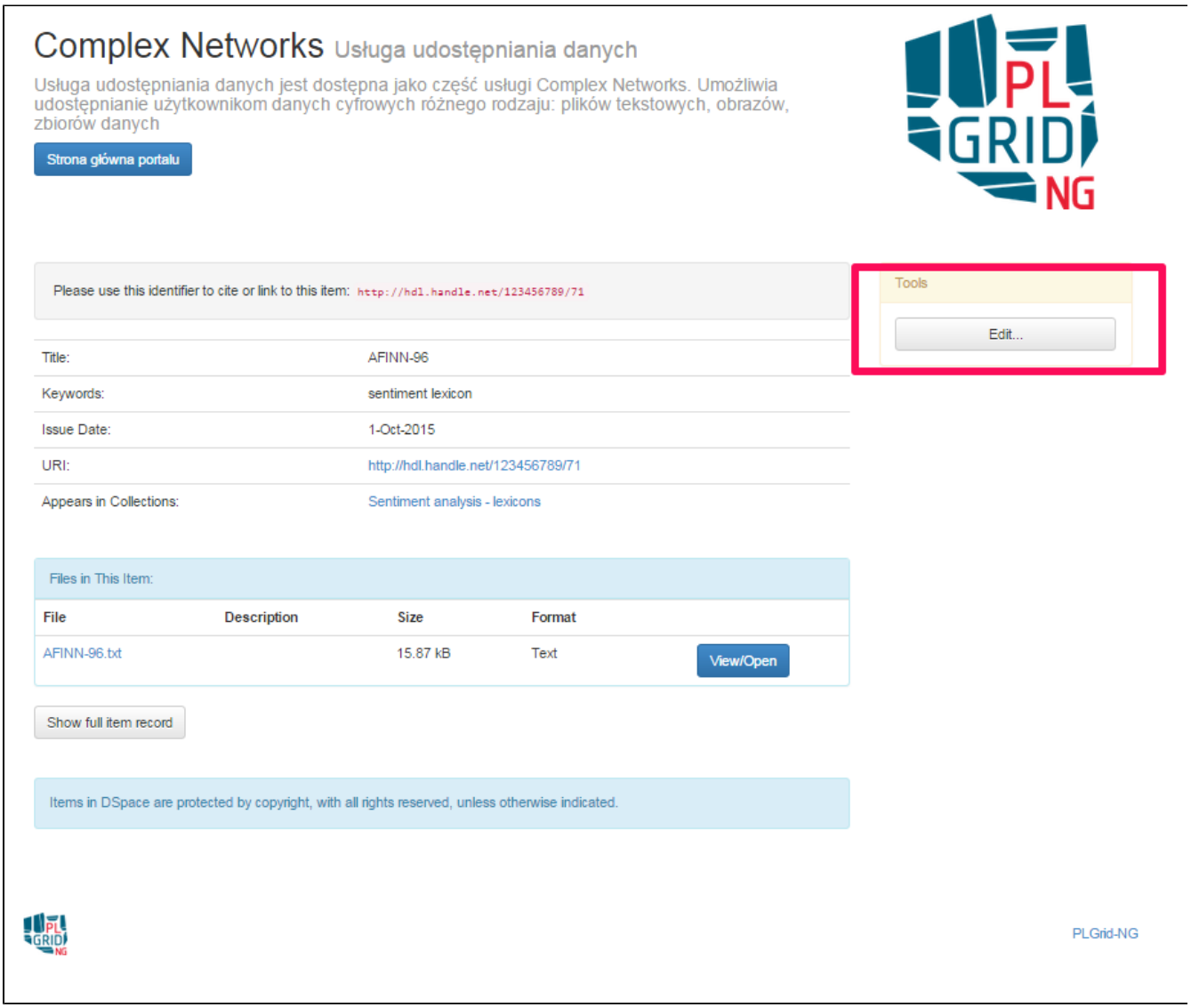


Plik zostanie pobrany na dysk po wybraniu **View/Open**.

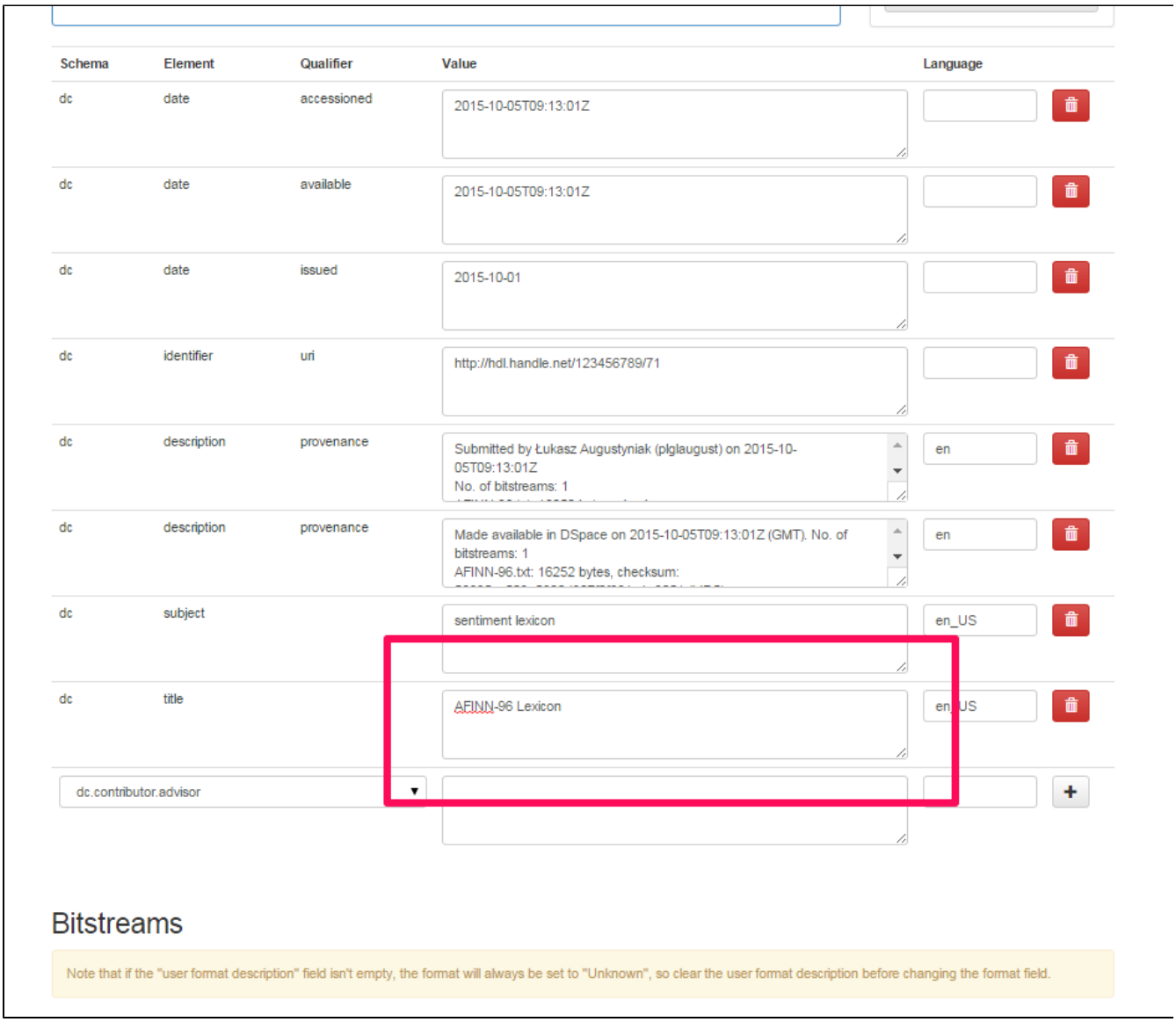


Edycja informacji o zbiorze danych - do zbioru danych z powyższego przykładu

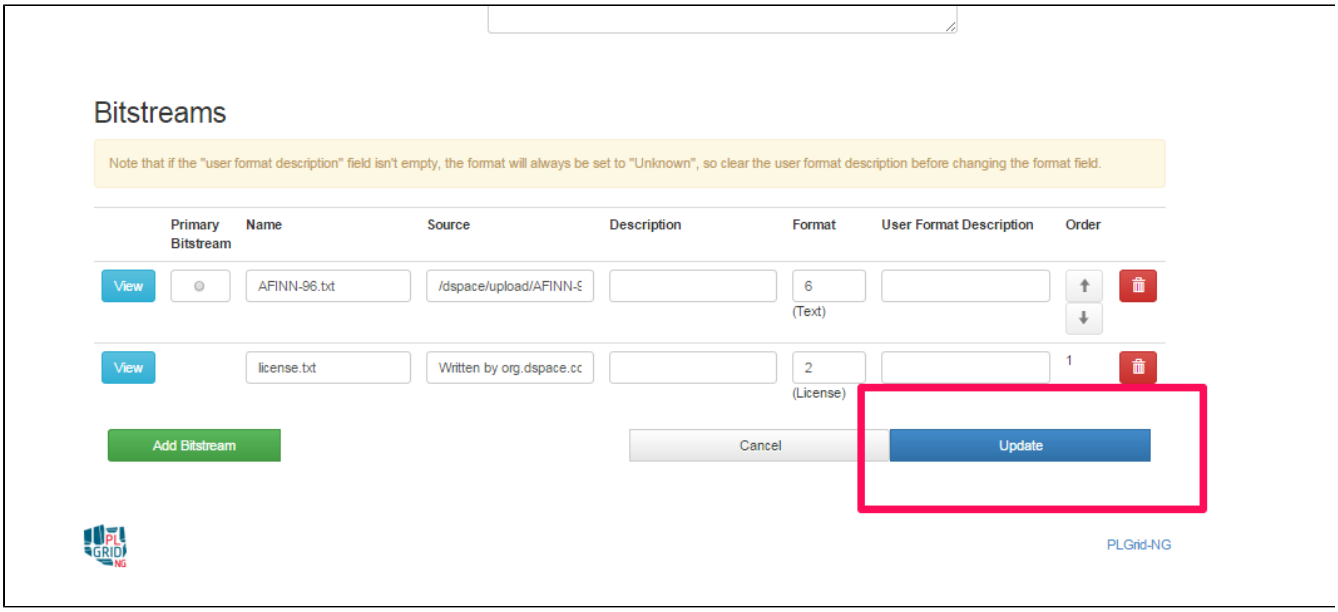
Z bocznego menu wybieramy przycisk **Edit [Edycja]**



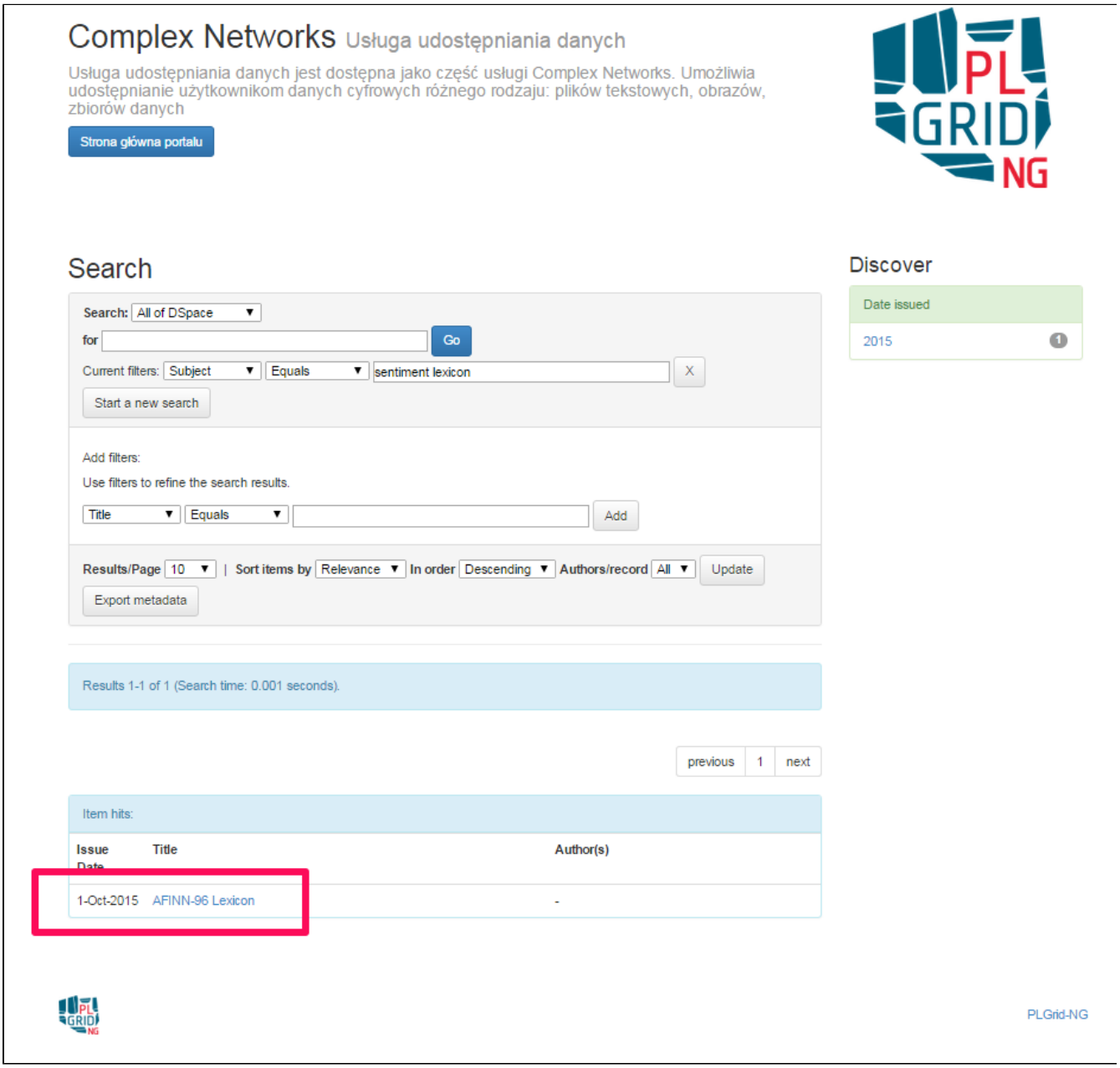
Chcemy zmienić tytuł dla zbioru danych z **AFINN-96** na **AFINN-96 Lexicon**



Zapisujemy informacje

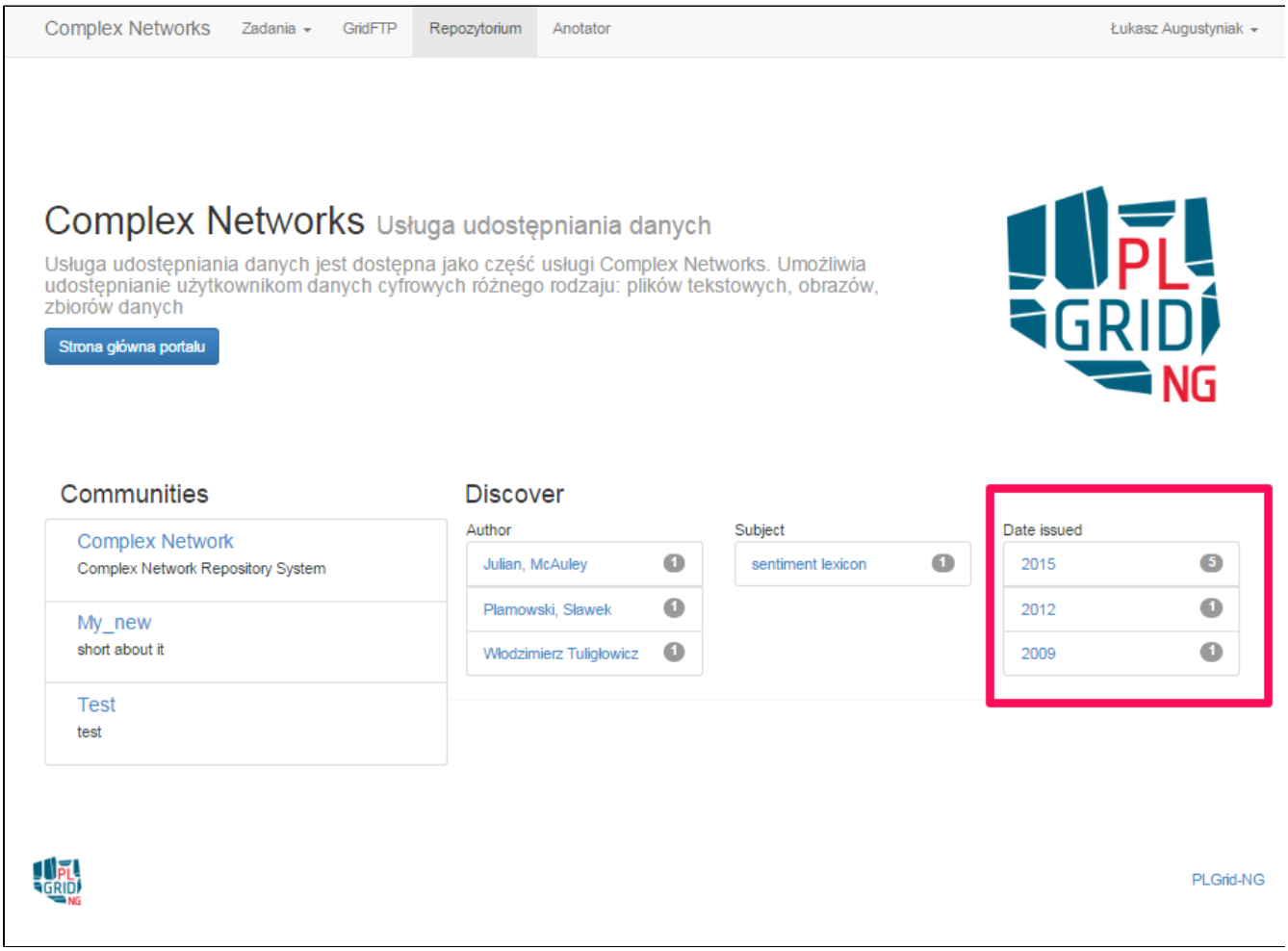


Po czym na stronie ze zbiorem widać zaktualizowany tytuł:

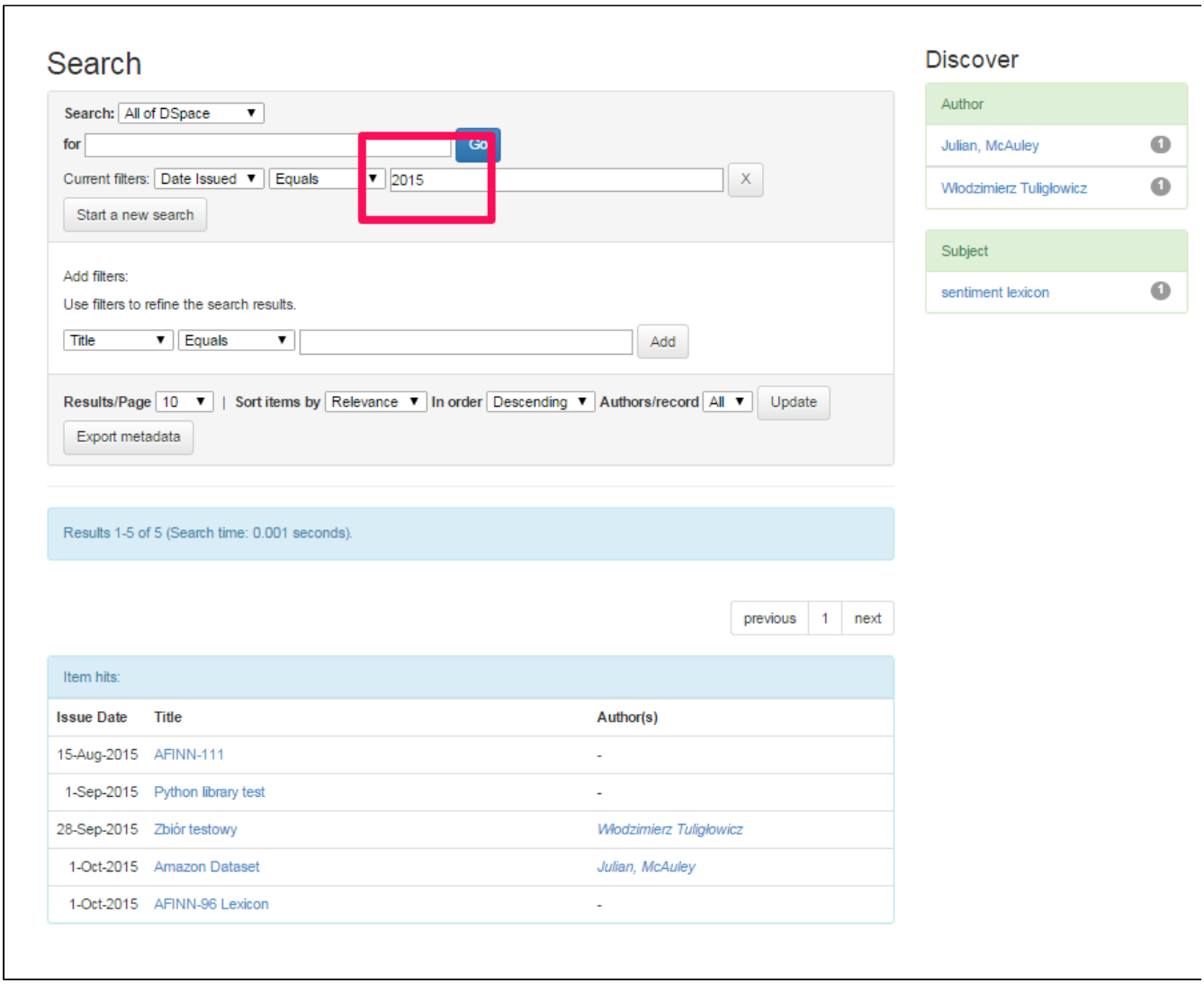


Filtrowanie/Wyszukiwanie danych

Podobnie, jak w przypadku użycia **Pobieranie danych**.

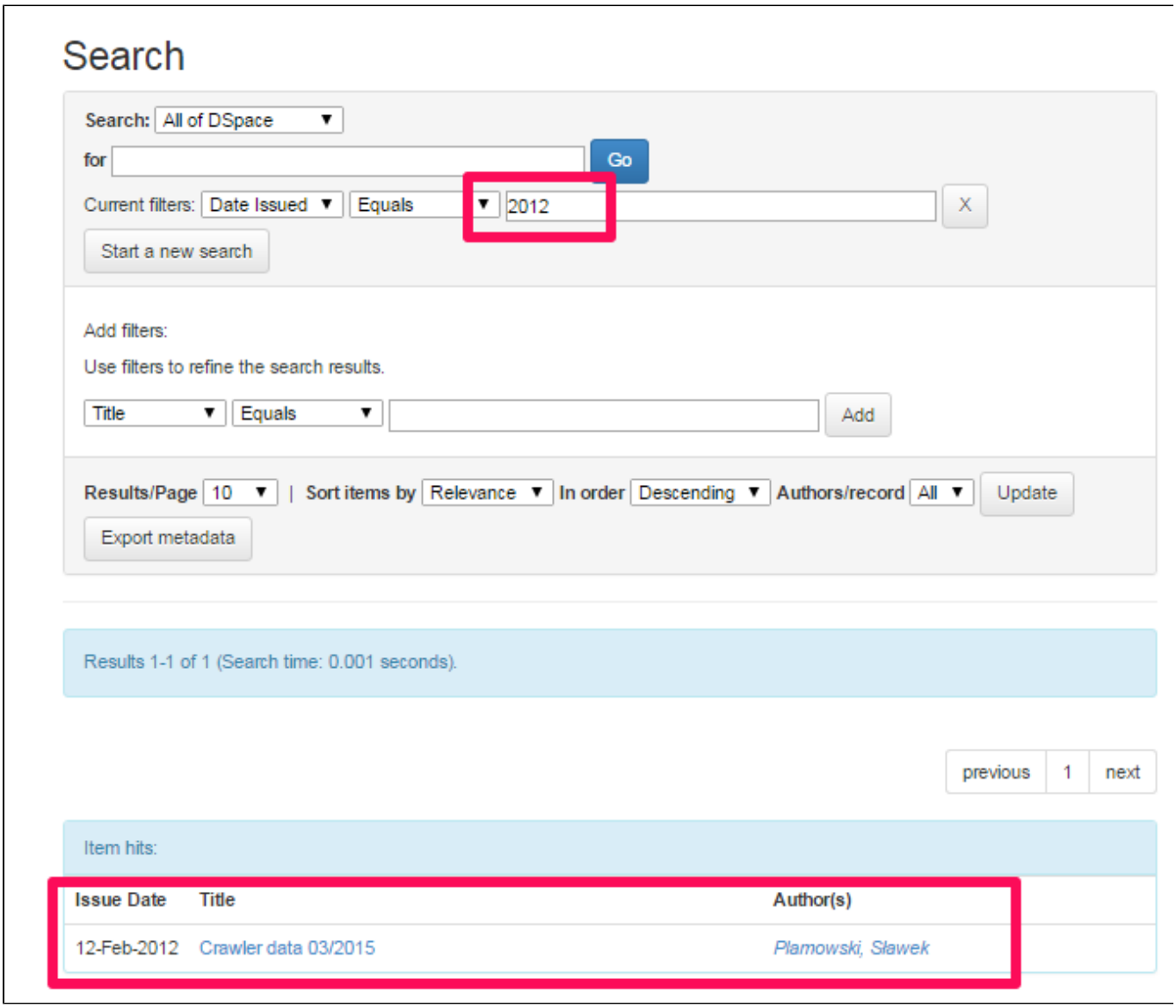


Wybieramy, np. publikacje w roku 2015



albo 2012.

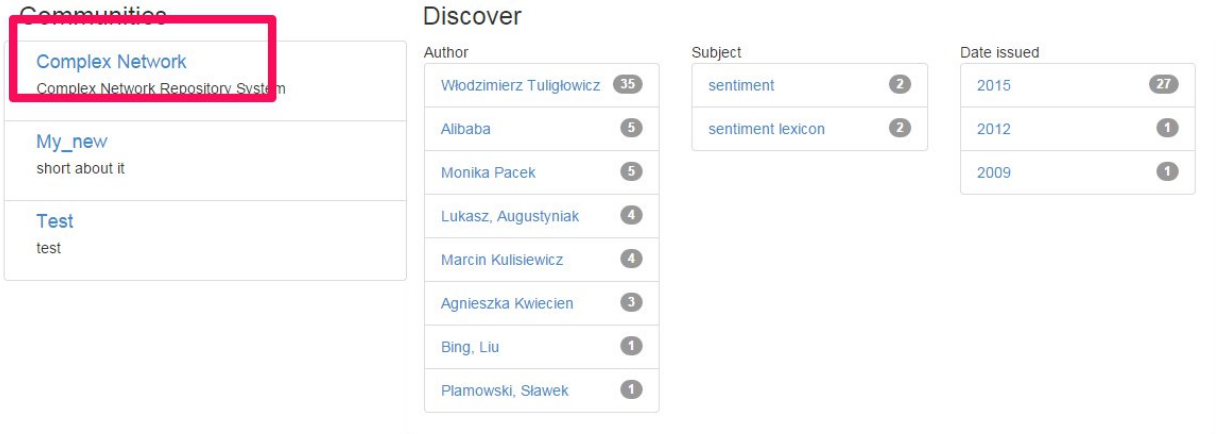
Filtrować możemy wg wielu parametrów, tworząc skomplikowane zapytania.



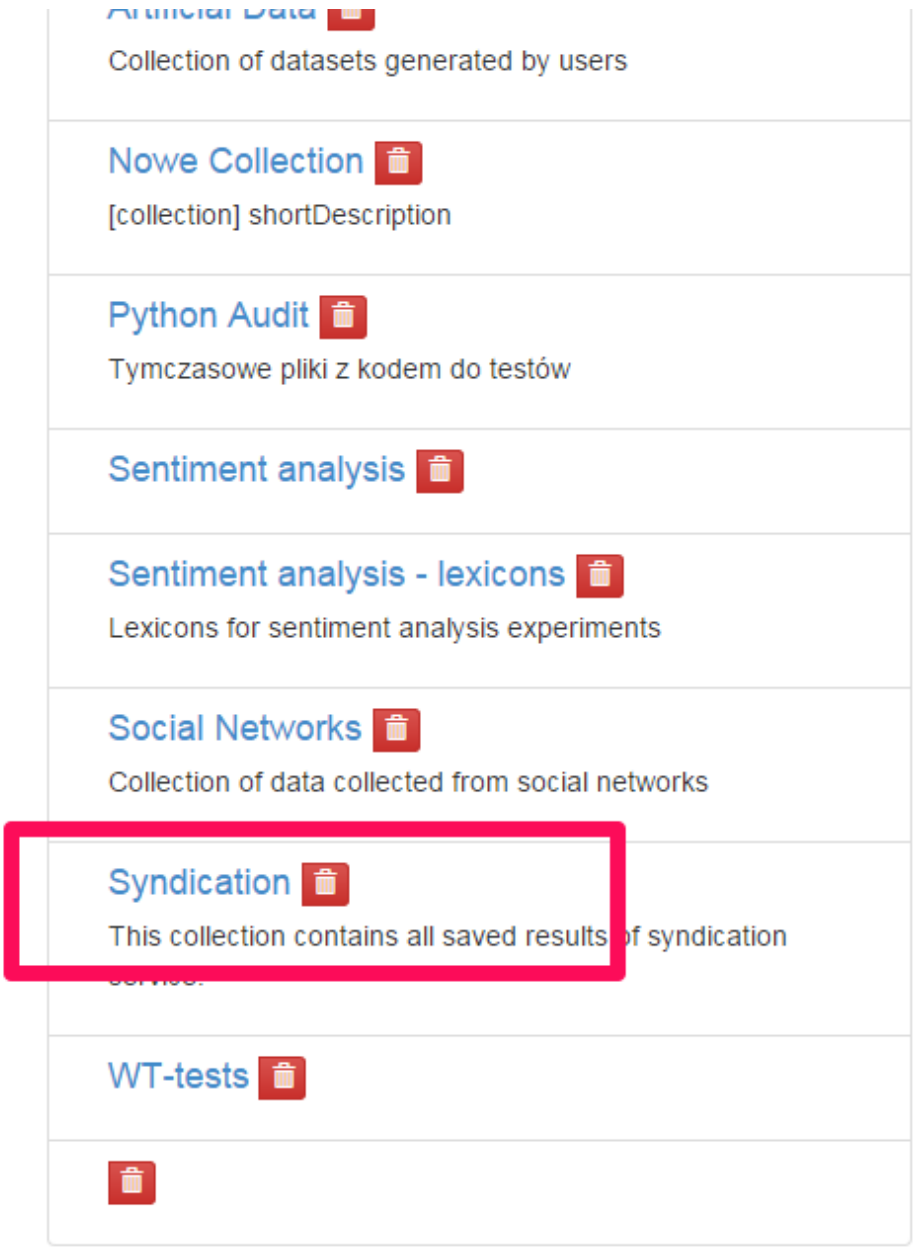
Usuwanie plików z repozytorium

Chcąc usunąć plik z repozytorium musimy najpierw wybrać Complex Networks, a następnie Kolekcję (ang. Collection), w której znajduje się plik do usunięcia. W poniższym przypadku chcemy usunąć plik z ostatnio zlecanego zadania syndykacji.

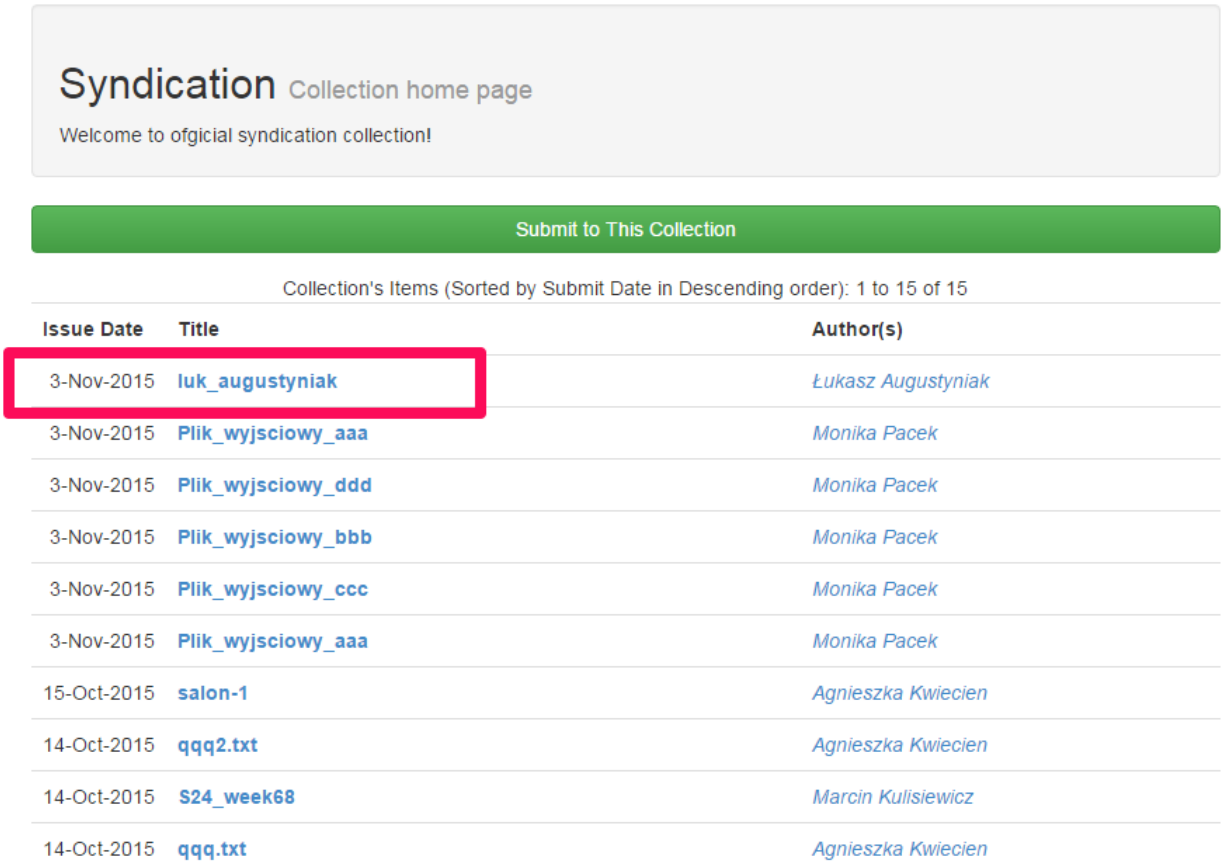
Krok 1. Należy wybrać Complex Networks.



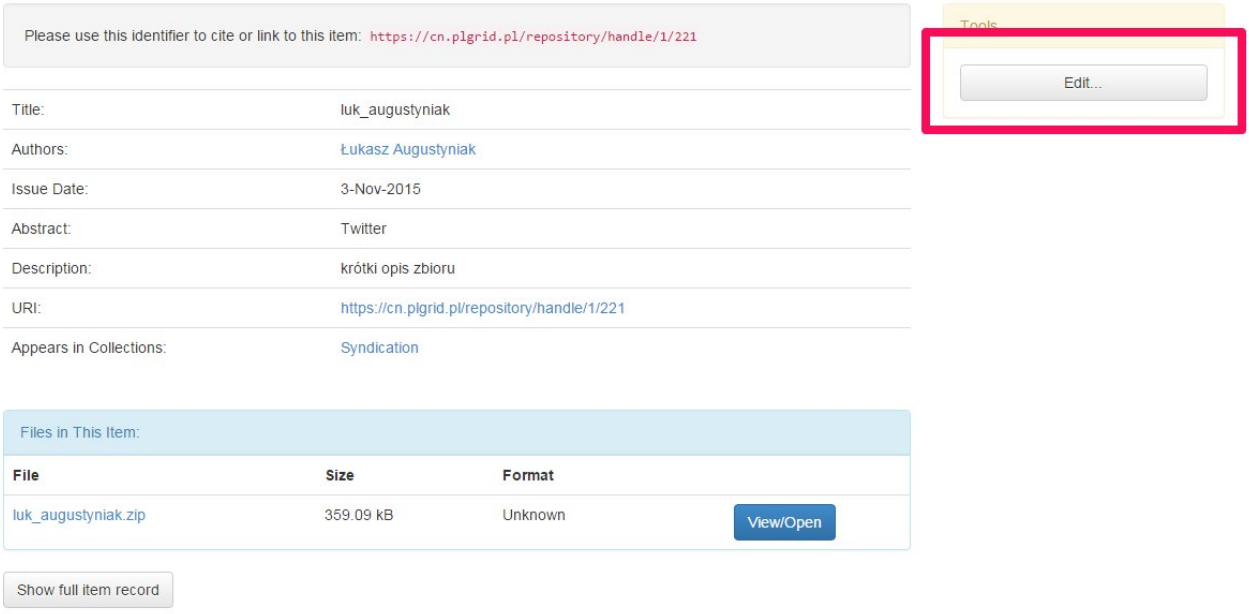
Krok 2. Po wykonaniu zadania syndykacji pliki trafiają automatycznie do kolekcji **Syndication,** zatem tutaj jest plik, który chcemy usunąć.



Krok 3. Należy wybrać ostatnio zlecone zadanie syndykacji pliki - interesującym nas zadaniem jest syndykacja konta Twitter Łukasz Augustyniak (@luk_augustyniak).

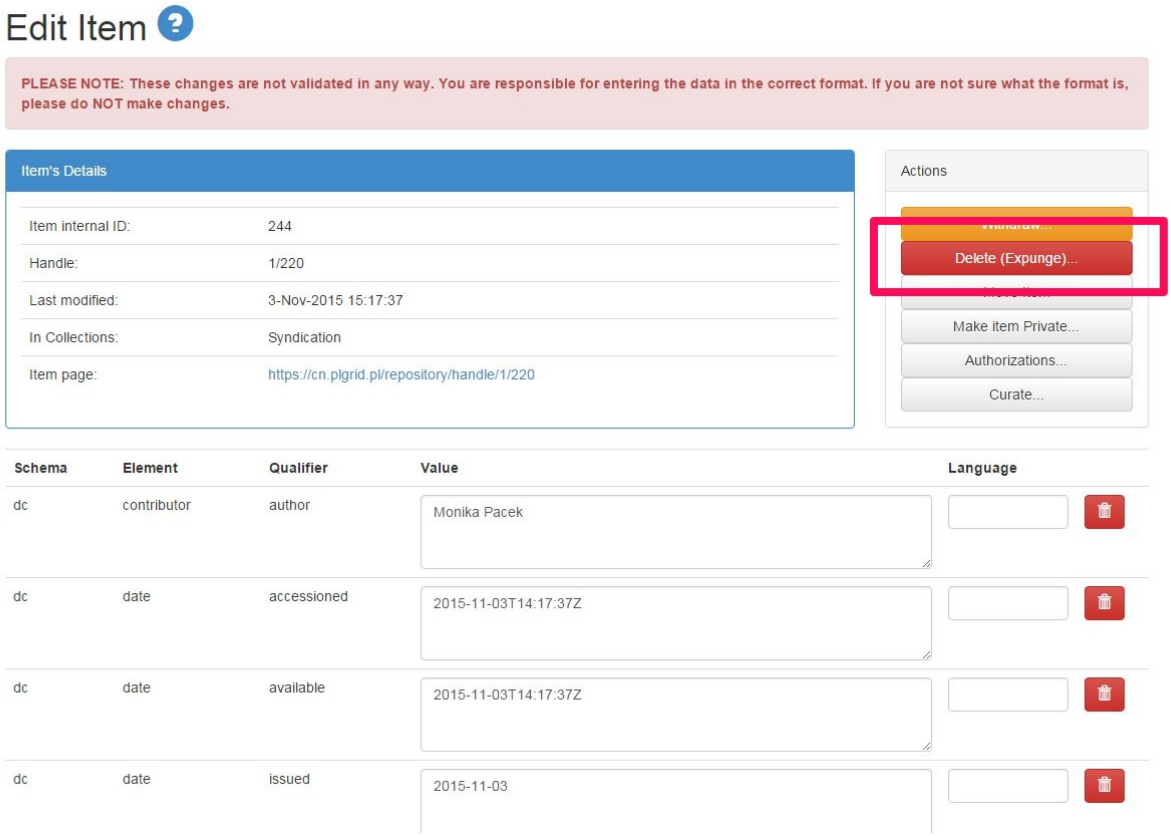


Krok 4. Należy kliknąć przycisk **Edit -** w celu edytowania zawartości zbioru plików.

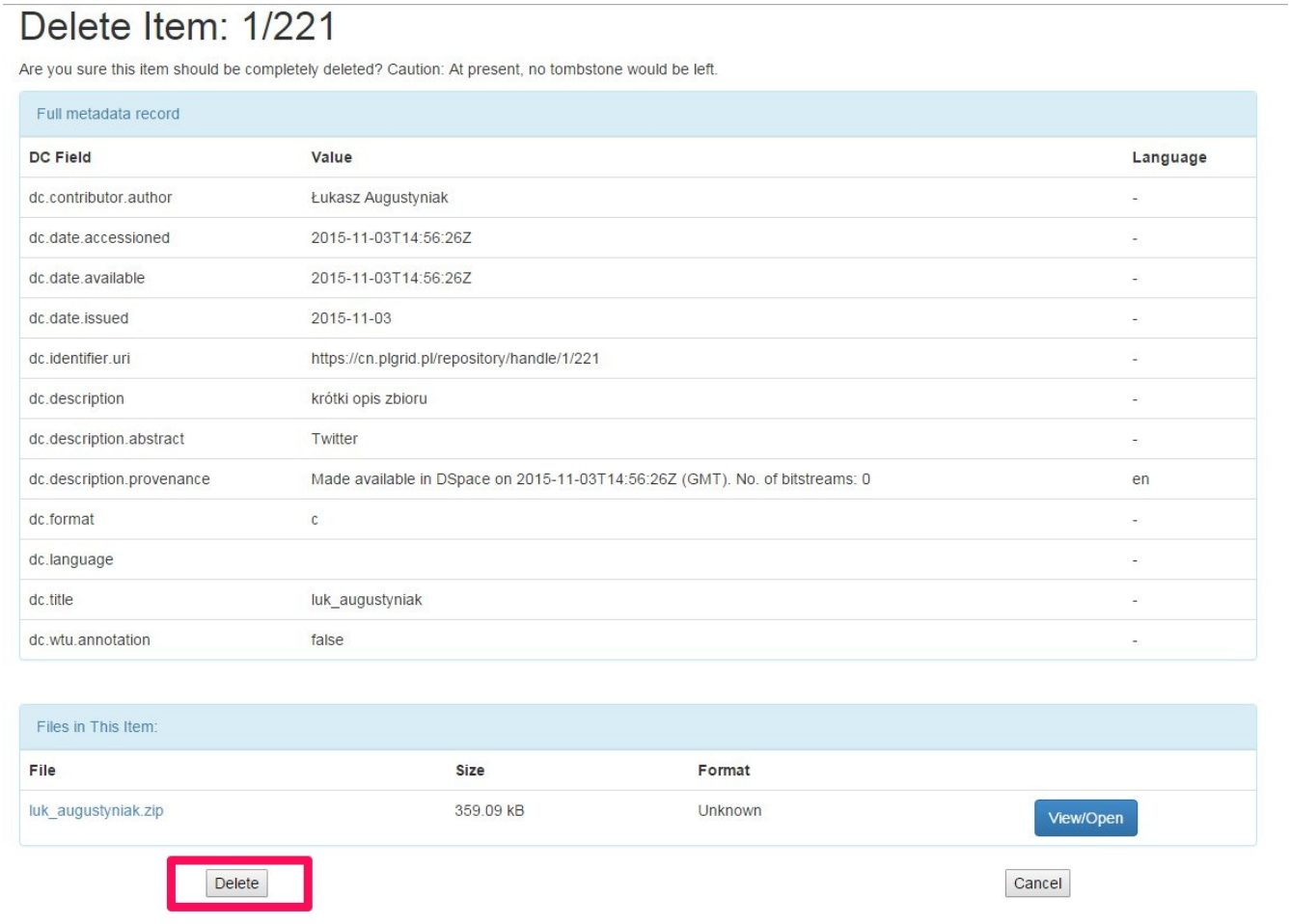


Krok 5. Usunąć pliki można na dwa sposoby:

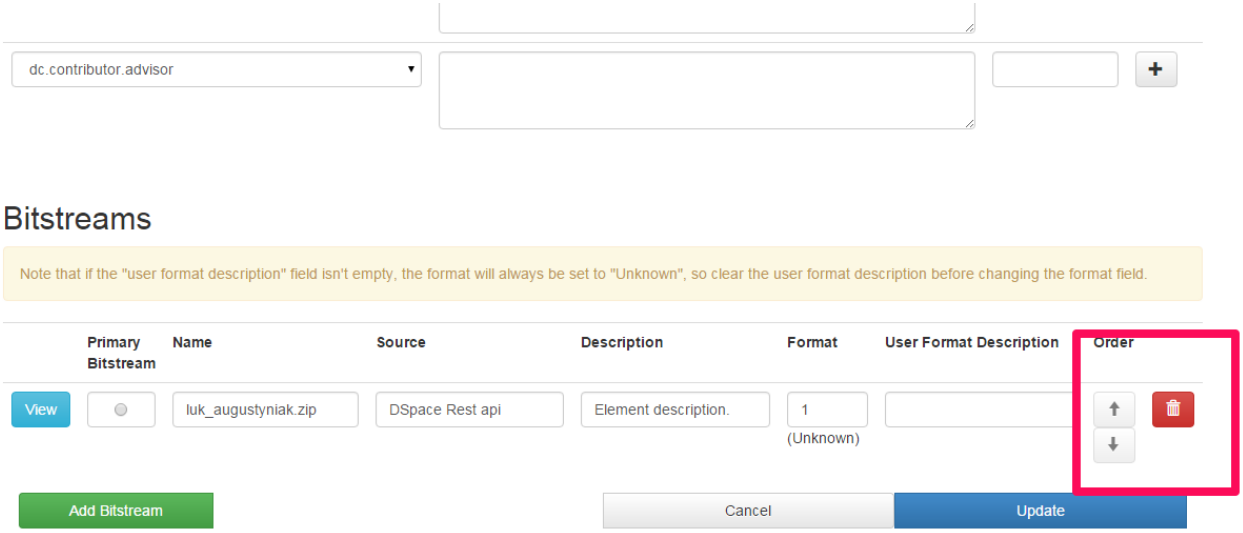
Krok 5.1.a Usuwanie całego zbioru plików, tj. w prezentowanym przypadku oznacza to usunięcie wszystkich plików skojarzonych z zadaniem syndykacji. W tym celu z menu bocznego należy wybrać **Delete (Expunge).**



Krok 5.1.b Należy **potwierdzić kliknięciem na przycisk Delete** chęć usunięcia całego zbioru plików. Należy pamiętać, że czynność ta jest **nieprzywracalna!**



Krok 5.2.a Usuwanie pojedynczego pliku. W tym celu w należy wrócić do widoku z kroku 5 i przewinąć stronę na sam jej dół. Proszę pamiętać, że po kliknięciu ikonki kosza (zrzut poniżej) plik zostaję automatycznie usunięty, nie jest potrzebne dodatkowe potwierdzanie tej czynności.



PLGrid-NG

Gdzie szukać dalszych informacji?

Więcej informacji można znaleźć na oficjalnej stronie usługi cn.plgrid.pl. Pytania proszę zadawać poprzez: helpdesk.plgrid.pl

Dodatkowo polecamy stronę Gdzie szukać pomocy?