

05. Import metadanych z formartu CSV

Aplikacja redaktora systemu dLibra umożliwia importowanie metadanych obiektu z pliku CSV. W tym celu należy utworzyć nowy plik z rozszerzeniem CSV, np. za pośrednictwem programu MS Excel.

Jeśli metadane obiektu mają być importowane dla większej liczby języków (np. polskiego, angielskiego, czy uniwersalnego), plik powinien mieć strukturę taką jak zaprezentowano na poniższym rysunku:

	A	B	C	D	E	F
1		pl	en	universal		
2	Title	Tytuł Polski	English Title			
3	Creator			Jan Kowalski		
4	Creator			Janusz Nowak		
5	Type	Artykuł	Article			
6						
7						
8	Title			ABC		
9	Subject	Literatura	Literature			
10	Subject	Humanizm	Humanism			
11	Creator	PCSS	PSNC			
12						
13	Title			Mój tytuł		
14	Subject	Import danych	Literature			
15	Subject	Import z pliku CSV	Humanism			
16	Creator	PCSS	PSNC			
17						
18						
19						

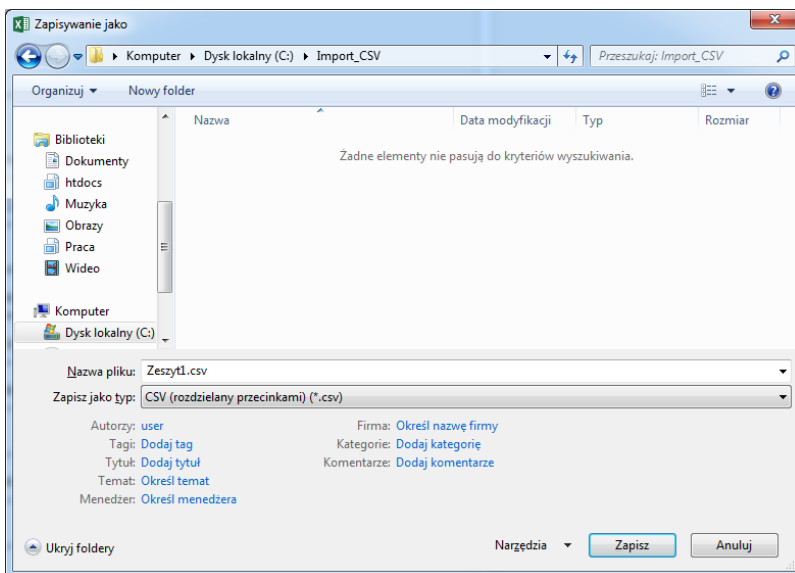
W pliku pierwsza kolumna zawiera nazwy RDF poszczególnych pól metadanych (nazwy RDF można zobaczyć w Aplikacji Redaktora i Administratora systemu dLibra). W kolejnych kolumnach znajdują się wartości przypisane do pól w odpowiednim języku. Na przykładzie druga kolumna zawiera wartości w języku polskim, trzecia w języku angielskim, a czwarta w języku uniwersalnym.

W jednym pliku może znajdować się kilka opisów obiektów, przy czym każdy opis obiektu powinien być oddzielony co najmniej jednym pustym wierszem. W przykładzie pokazano trzy niezależne opisy. Pierwszy opis zawiera wypełnione wartości dla pól: Tytuł (nazwa RDF "Title", wprowadzone dwie wartości - jedna w języku polskim "Tytuł Polski" i jedna w angielskim "English Title"), Twórca (nazwa RDF "Creator", wprowadzone dwie wartości dla języka uniwersalnego "Jan Kowalski" i "Janusz Nowak" - dwa wiersze oznaczone nazwą RDF "Creator"), Typ (nazwa RDF "Type", wprowadzone wartości - jedna w języku polskim "Artykuł" i jedna w angielskim "Article").

Jeśli importowane będą jedynie metadane w języku polskim, plik może mieć poniższą strukturę (bez konieczności wprowadzania nazw kolumn dla poszczególnych języków, ponieważ jest tylko jeden wypełniany język). Tu również należy rozdzielić informacje o poszczególnych obiektach pustym wierszem:

	A	B	C
1	Title	Tytuł Polski	
2	Creator	Jan Kowalski	
3	Creator	Janusz Nowak	
4	Type	Artykuł	
5			
6	Title	Tytuł Polski 2	
7	Creator	Jan Kowalski	
8	Creator	Janusz Nowak	
9	Type	Artykuł	
10			
11	Title	Tytuł Polski 3	
12	Creator	Jan Kowalski	
13	Creator	Janusz Nowak	
14	Type	Artykuł	
15			
16	Title	Tytuł Polski 4	
17	Creator	Jan Kowalski	
18	Creator	Janusz Nowak	
19	Type	Artykuł	
20			

Po utworzeniu nowego pliku, należy zapisać go pod dowolną nazwą na dysku lokalnym komputera, wybierając odpowiedni typ pliku - Zapisz jako typ: CSV (rozdzielany przecinkami)(* .csv). Ekran poniżej przedstawia okno zapisywania pliku dla programu MS Excel.

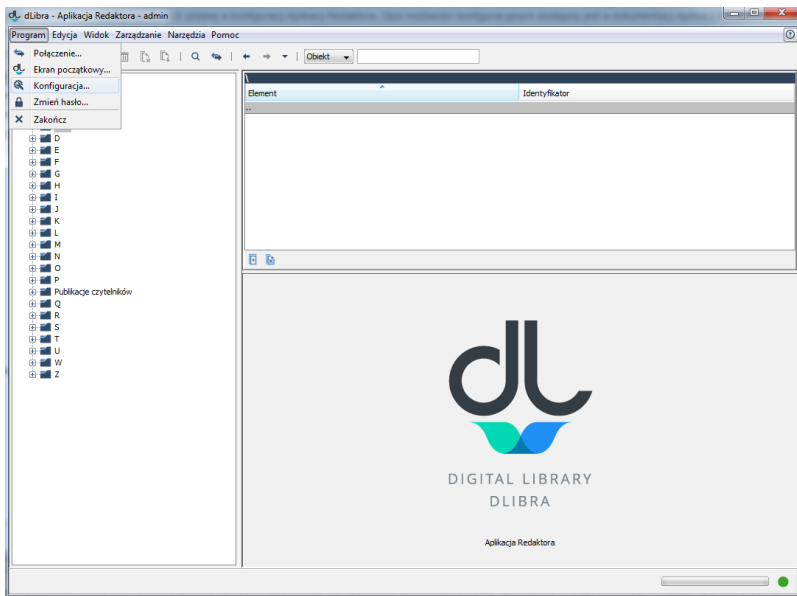


Plik przygotowany w MS Excel zapisany będzie z użyciem kodowania ANSI, z kolei Aplikacja Redaktora domyślnie pracuje w kodowaniu UTF-8. Dlatego też, aby poprawnie wczytać metadane z plików zapisanych z wykorzystaniem programu MS Excel, należy wprowadzić zmianę w konfiguracji Aplikacji Redaktora. Opis możliwości konfiguracyjnych dostępny jest w dokumentacji Aplikacji Redaktora i Administratora, w [sekcji dotyczącej konfiguracji aplikacji](#). W kontekście mechanizmu importu metadanych z CSV należy wprowadzić następującą konfigurację (linię) w pliku generalConfig.properties:

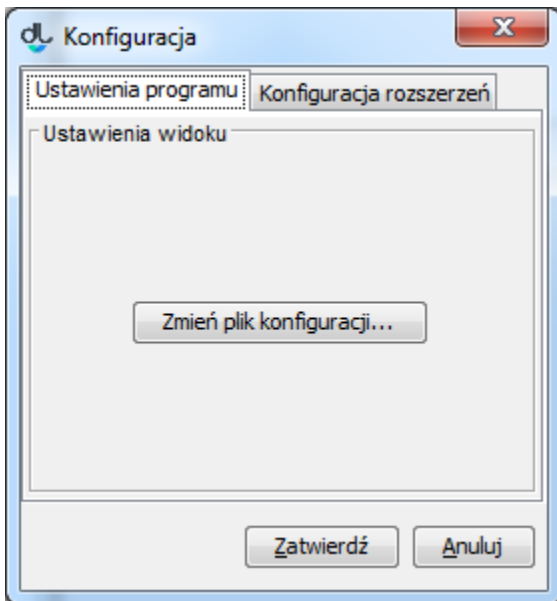
```
metadata.csv.encoding = Windows-1250
```

Jeśli tego nie zrobimy to pliki CSV będą importowane z wykorzystaniem kodowania UTF-8 (domyślne kodowanie używane w Aplikacji Redaktora i Administratora).

Aby wczytać tak zmodyfikowany plik konfiguracyjny, z menu aplikacji redaktora należy wybrać opcję "Program", a następnie "Konfiguracja".



W nowym oknie należy kliknąć przycisk „Zmień plik konfiguracyjny” i wskazać lokalizację nowego pliku generalConfig.properties.



Po zatwierdzeniu zmian, zaimportowane dane zawierające polskie znaki diakrytyczne, będą widoczne w poprawny sposób.

Import metadanych z pliku CSV możliwy jest z poziomu edytora opisu bibliograficznego (ekran poniżej), który pojawia się, np. w kreatorze nowego obiektu cyfrowego (krok drugi) lub kreatorze planowanego obiektu (krok pierwszy).

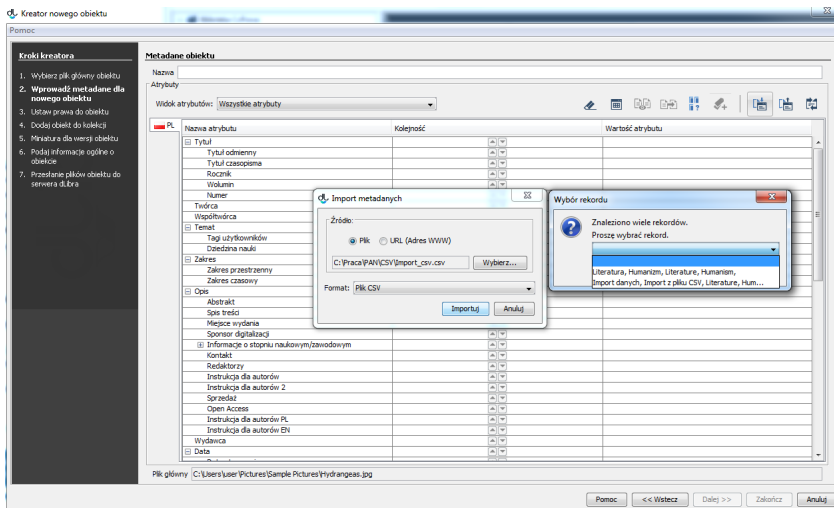
The screenshot shows the dLibra application interface. On the left is a tree view of the library's structure. The main window displays a list of bibliographic records with columns for '#', 'Element', and 'Identyfikator'. Below this, a metadata editor window is open, showing a table of attributes and their values for a selected record.

PL	Nazwa atrybutu	Kolejność	Wartość atrybutu
	Tytuł		Robotyczna arystokracja i biedni profesjonaliści. Imigranci z ...
	Tytuł odmienny		Central Eastern European Migration Review
	Tytuł czasopisma		2013
	Rocznik		Vol 2
	Wolumin		No 1
	Numer		Anna Sosnowska
	Twórca		
	Współtwórca		
	Temat		
	Tagi użytkownikowe		
	Dziedzina nauki		Nauki Humanistyczne i Społeczne
	Zakres		
	Zakres przestrzenny		
	Zakres czasowy		
	Opis		
	Abstrakt		
	Spis treści		
	Miejsce wydania		
	Sponsor digitalizacji		
	Informacje o stopniu naukowym/zawodowym		
	Uzyskany tytuł		
	Stronień struktów		

Aby zaimportować opis z poziomu edytora opisu bibliograficznego, należy wybrać znajdujący się na nim przycisk "Importuj...". W efekcie pojawi się okno importowania metadanych (ekran poniżej). W oknie tym należy wskazać lokalizację utworzonego pliku CSV oraz kliknąć przycisk „Importuj”.

The screenshot shows the 'Import metadanych' dialog box. It has a title bar with the dLibra logo and a close button. The main area contains a 'Źródło:' label, two radio buttons for 'Plik' (selected) and 'URL (Adres WWW)', a text field with the path 'C:\Praca\PAN\CSV\Import_csv2.csv', and a 'Wybierz...' button. Below this is a 'Format:' label and a dropdown menu set to 'Plik CSV'. At the bottom are 'Importuj' and 'Anuluj' buttons.

Jeżeli plik CSV zawiera więcej niż jeden opis obiektu to aplikacja poprosi o wskazanie opisu, którego dane mają zostać zaimportowane (ekran poniżej).



Po wskazaniu poprawnego rekordu i zatwierdzeniu wyboru metadane obiektu zostaną uzupełnione.