

The Dublin Core Metadata Element Set, Ver. 1.1 a potrzeby i oczekiwania bibliotekarzy cyfrowych - analiza przypadków

Joanna Potęga
Biblioteka Narodowa



Agnieszka Wróbel
Biblioteka Uniwersytecka
w Warszawie





Kto używa Dublin Core?

38 bibliotek cyfrowych w Federacji Bibliotek Cyfrowych,
w tym:

- dLibra – **37**
- własne oprogramowanie – **1**

Podstawa opisu (dla materiałów digitalizowanych):

- Kopia cyfrowa dokumentu oryginalnego
(cechy formalne dok. analogowego wraz z informacjami
charakterystycznymi dla dokumentu cyfrowego) – **36**
- Rozróżnienie wydawcy i dat wyd. (dok. oryg., dok. cyfrowego)
– **2** (+ *1 własny schemat*)
(*Nowy dokument elektroniczny - 1 – MARC21*)



Teoria głosi, że...

Dublin Core simple

Title (tytuł)

Creator (twórca)

Subject and Keywords (temat i słowa kluczowe)

Description (opis)

Publisher (wydawca)

Contributor (współtwórca)

Date (data)

Resource Type (typ zasobu)

Format (format)

Resource Identifier (identyfikator zasobu)

Source (źródło)

Language (język)

Relation (powiązania)

Coverage (zakres/zasięg)

Rights Management (zarządzanie prawami)

Zdefiniowane problemy z metadanymi:

Organizacja metadanych

- nazewnictwo elementów DC
- interpretacja elementów DC
- „organizacja danych” w atrybucie

Struktura metadanych

- doprecyzowania (podatrybuty)
- dodatkowe elementy poza DC (elementy lokalne)

Organizacja metadanych



Różnica w nazewnictwie elementów DC



Subject and Keywords (temat i słowa kluczowe)

Temat i słowa kluczowe – **31**

Słowa kluczowe – **2**

Hasło przedmiotowe – **2**

Hasło przedmiotowe KABA – **1**

Dziedzina – **1**



Różnica interpretacyjna

przy zachowaniu tego samego nazewnictwa elementu DC

Relation (powiązania)

Odesłanie do źródeł pozostających w relacji z opisywanym. Rekomenduje się powiązanie głównego źródła ze źródłem związanym przy pomocy ciągu znaków lub numeru dostosowanego do formalnego systemu identyfikacyjnego np.: sygnatura mediów zawierających kopię źródła (np. mikrofilm, nośnik DVD itp.).

ePoradnik redaktora zasobów cyfrowych



Różnica interpretacyjna

przy zachowaniu tego samego nazewnictwa elementu DC


Przykłady wartości elementu Relation:

- DVD OSD00027
- (e-Biblioteka Historyczna ; 1)
- Mf. 10406
- Joachima Lelewela Bibliograficznych ksiąg dwoje T.1
- Tyt. poprz. Annalen der Chemie und Pharmacie.
- Biblioteczka Bonifraterska
- Fontes - Towarzystwo Naukowe w Toruniu; 13
- **Nasza Przeszłość t.106: 2006, s. 283-293**
- adligat: sygn.13278-13283
- bezpłatny dodatek
- Muzeum Narodowe w Rapperswilu

„Organizacja danych” w atrybucie

Rights Management

- Z rozbiem na podatrybuty:
 - = [atrybut Prawa] właściciel praw majątkowych
 - + [podatrybut Licencja] nazwa licencji, rodzaj licencji
 - + [podatrybut Tekst (Treść)] link do treści licencji

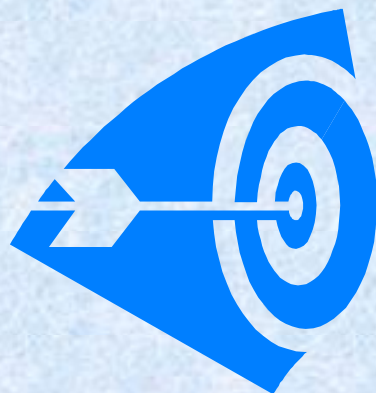
Prawa : Biblioteka Uniwersytecka we Wrocławiu
 Licencja : by 2.5 pl
 Treść licencji : Więcej informacji... 

Prawa : Janusz Stanisław Bień
Licencja : GNU Free Documentation License
Tekst : kliknij tutaj, żeby przejść 

- Bez rozbitcia na podatrybuty
 - = [atrybut Prawa] właściciel praw majątkowych, nazwa licencji, rodzaj licencji, link do treści licencji*
- *odrębne wartości

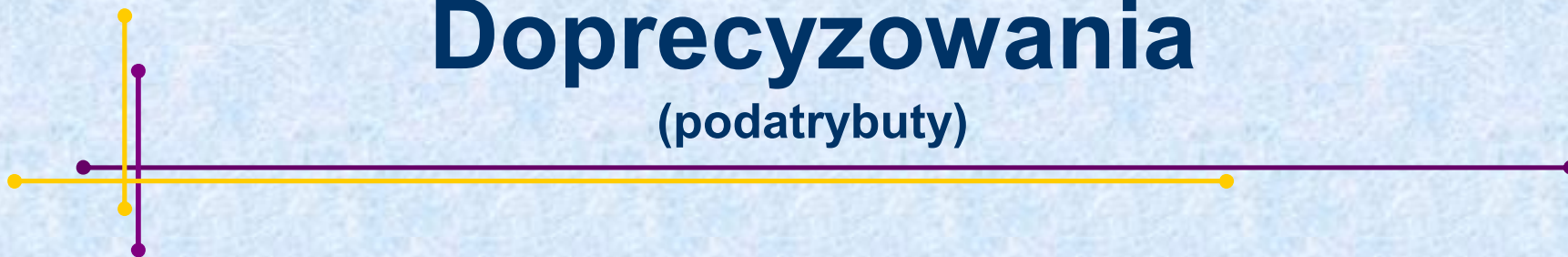
Prawa : by-nc 2.5 pl ; Creative Commons ; Więcej...  ;
Mansfeld, Bogusław

Struktura metadanych



Doprecyzowania

(podatrybuty)



- **Ile bibliotek stosuje podatrybuty?**
 - 37 dLibrowych – **24**
- **Ile atrybutów doprecyzujemy?**
 - z 15 w DC simple – **10**

Doprecyzowania

(podatrybuty)

- **W jakich atrybutach - w ilu bibliotekach - występują doprecyzowania?**
 - Title – 8
 - Subject and Keywords – 5
 - Description – 3
 - Publisher – 21
 - Contributor – 1
 - Date – 2
 - Resource Identifier – 6
 - Source – 1
 - Relation – 3
 - Rights Management – 7

Title – 8

- Wariant tytułu – 2
- Inny tytuł – 2
- Podtytuł – 2
- Tytuł oryginału – 1
- Gatunek sceniczny / Miejsce powstania / Data powstania – 1

Subject and Keywords – 5

- Hasło przedmiotowe – 3
- Hasło przedmiotowe / Słowa kluczowe – 1
- Słowa kluczowe użytkowników – 1

Description – 3

- Wymiary – 1
- Abstrakt – 1
- Komentarz – 1

Publisher – 21

- Miejsce wydania – 19
- Miejsce wydania / Drukarz – 1
- Drukarz – 1

Contributor – 1

- Instytucja sprawcza – 1

Date – 2

- Data druku – 1
- Miejsce wydania – 1

Resource Identifier – 6

- ISBN / ISSN – 2
- Identyfikator oryginału – 2
- Sygnatura oryginału / Identyfikator obiektu cyfrowego – 1
- Identyfikator oryginału / Identyfikator obiektu cyfrowego – 1

Source – 1

- Uwagi / Współprawy z / [Powiązania](#) / Prowienienca /
Linki – 1

Relation – 3

- Katalog GAIK – 1
- OPAC WWW kat. online – 1
- Seria – 1

Rights Management – 7

- Licencja / Tekst (Treść) – 2
- Prawa (w tym przypadku nazwa atrybutu = Prawa do dysponowania publikacją) – 1
- Digitalizacja / Lokalizacja oryginału – 1
- Lokalizacja / Lokalizacja oryginału – 1/1 (=2)
- Kontakt / Digitalizacja / Sygnatura / Zespół archiwalny / Sygn. – 1



Dodatkowe atrybuty poza DC

- TAGI / Słowa kluczowe użytkowników – **26/1 (=27)**
- Miejsce wydania – **4**
- Wariant tytułu / tytuł w j.ang. – **1/1 (=2)**
- Identyfikator obiektu cyfrowego – **2**
- Uwagi – **8**
 - Uwagi – **5**
 - Adnotacje – **1**
 - Inf.tech. – **1**
 - Uwagi archiwistów – **1**

Dodatkowe atrybuty poza DC cd.

- Lokalizacja oryginału – 7
- Digitalizacja / Reprodukacja cyfrowa – 6/1 (=7)
- Sygnatura BCPW / Sygnatura archiwalna / Identyfikator oryginału – 1/1/1 (=3)
- Numer zespołu archiwalnego – 2
- Nazwa zespołu archiwalnego – 1
- Sponsor digitalizacji – 2
- www – 2
- OPAC / WebOPAC BN / Pełny opis katalogowy – 1/1/1 (=3)
- INNE: Opis fizyczny / Seria / Wydanie / Publikacja / Tryb dostępu – 1/1/1/1/1/ (=5)

Dlaczego jest tak jak jest?

próba odpowiedzi na podstawie ankiety oraz analizy



Czynniki wpływające na jakość:

- różnorodność zasobów i ich specyfika
- różnorodność praktyk katalogowych instytucji współtworzących BC (także w BC konsorcyjnych)
- mechanizmy tworzenia metadanych przez bibliotekarzy / redaktorów cyfrowych

Ankieta październik/listopad 2009

Skierowana do 41 bibliotek – 29 odpowiedzi

- Dublin Core – 27
- MARC21 – 1
- własny schemat – 1

Sposoby tworzenia metadanych:

- tylko import bez ingerencji w dane – **1** (MARC21)
- tylko import z ingerencją w dane – **3**
- import i też samodzielne tworzenie opisów – **10**
- samodzielne tworzenie opisów przez redaktora – **13 / 1**
(własny)

Import – dane wyjściowe:

- MARC21 – **12** (+1 w przygotowaniu)
- MARC – **1**
- MARC BN – **1**
- *MASTER* – **1** (w planach)

W przypadku importu - najczęściej uzupełniane przez redaktora atrybuty:

- Temat i słowa kluczowe – **10**
- Typ zasobu – **10**
- Prawa – **9**
- Opis – **8**
- Źródło – **5**
 - + **1** (sygn. oryg.)
 - + **1** (import. sygn uzup. link do opisu)
- Format – **4**
- Digitalizacja (reprodukcja cyfrowa) – **4**
- Lokalizacja oryginału – **3**

Zalecenia do tworzenia opisów:

- TAK – 27
- NIE – 2

- Tylko zalecenia ogólne – 15
- Tylko zalecenia szczegółowe – 8
- Ogólne ze wskazówkami opisu dla poszczególnych typów dokumentów – 4

Zalecenia w bibliotekach tworzonych przez wielu partnerów (16):

- Wspólne dla wszystkich partnerów – 15
- Każdy z partnerów tworzy własne – 1

Wykorzystanie źródeł przy tworzeniu zaleceń:

- Tylko *ePoradnik* redaktora zasobów cyfrowych – **2**
- Tylko ustalenia własne – **3**
- Tylko inne źródła – **2**
- *ePoradnik* i własne ustalenia – **10**
- Ustalenia własne i inne źródła – **2**
- *ePoradnik*, własne ustalenia i inne źródła – **8**

(Wykorzystanie *ePoradnika* w 20 BC)

Inne źródła:

- **Library of Congress (mapping MARC21 to DC) – 2**
- **The European Library (TEL) – 2**
- **Inne BC polskie – 4**
- **BC zagraniczne – 1**
- **DCMI – 1**
- **Format MARC21 – 4**
- **Tłumaczenia DC (EBIB) – 1**
- **Normy opisu bibliograficznego – 2**
- **Instrukcja DBC – 1**



Podsumowanie

- BC nie jest katalogiem bibliotecznym
- znaczna dowolność przy wpisywaniu danych
 - większa elastyczność i intuicyjność danych (dla użytkownika, ale w konsekwencji: szum informacyjny)
 - brak wsparcia takiego jakie daje KHW w katalogach
- brak słowników kontrolowanych dla poszczególnych atrybutów (Indeks autorów/twórców, Indeks słów kluczowych)



Podsumowanie cd.

- różnice w interpretacji poszczególnych elementów DC (przy zachowaniu jednorodnego nazewnictwa elementów)
- nierównoważne informacje w tym samym elemencie
- informacja tego samego typu w wielu różnych elementach
- przypadki umieszczania wielu wartości w jednym elemencie jako ciągłego tekstu (brak powtórzeń elementu – brak możliwości wyszukiwania po wartości)
- umieszczanie wielu informacji poza DC – lokalność informacji



Wnioski

- Różni użytkownicy
- Różne potrzeby
- Różne typy dokumentów (archiwalia)
- Różne praktyki opracowania dokumentów
- Większość uznaje, że opisujemy kopie dokumentów wersji papierowej
- DC dedykowany dla dokumentów *born digital*

The Dublin Core Metadata Element Set, Ver. 1.1 a potrzeby i oczekiwania bibliotekarzy cyfrowych...



DCMI Metadata Terms?



DCMI Metadata Terms, Dublin Core Metadata Initiative, 01.2008

Tłum. pol.: *Terminy metadanych DCMI*, Biblioteka Narodowa, 09.2009
(dostęp na www.bn.org.pl → *Normy Formaty Standardy*)

Wykaz wszystkich elementów metadanych DC, w podziale na:

- **własności** (*twórca, tytuł, data, temat, format ...*):
 - **56** własności w nowej przestrzeni nazw */terms/*
 - 15 własności w dawnej przestrzeni nazw */elements/1.1/*
- **klasy danych** (*agent, lokalizacja, typ nośnika ...*)
- **słownikowe schematy kodowania danych** – wykaz (*UKD, MeSH, TGN ...*)
- **syntaktyczne schematy kodowania danych** – wykaz (*URI, ISO 3166 ...*)
- **terminy związane z *DCMI Abstract Model***



DCMI Metadata Terms, Własności w przestrzeni nazw */terms/*

Nazwa terminu: Date

Definicja: Pojedyncza data lub okres związany ze zdarzeniem, mającym miejsce w cyklu istnienia zasobu.

Uszczegółowienia:

- **Nazwa terminu: available** Etykieta: Data udostępnienia
- **Nazwa terminu: created** Etykieta: Data utworzenia
- **Nazwa terminu: dateAccepted** Etykieta: Data przyjęcia (zatwierdzenia zasobu)
- **Nazwa terminu: dateCopyrighted** Etykieta: Data copyright
- **Nazwa terminu: dateSubmitted** Etykieta: Data złożenia
- **Nazwa terminu: issued** Etykieta: Data wydania
- **Nazwa terminu: modified** Etykieta: Data modyfikacji
- **Nazwa terminu: valid** Etykieta: Data ważności



DCMI Metadata Terms, Własności w przestrzeni nazw */terms/*

Nazwa terminu: Description

Etykieta: Opis

Definicja: Opis zasobu

Komentarz: Opis może zawierać m.in.: streszczenie, spis treści, graficzne przedstawienie treści lub opis zasobu w formie swobodnego tekst

Uszczegółowienia:

- **Nazwa terminu: Abstract**
Definicja: Streszczenie zasobu
- **Nazwa terminu: tableOfContents**
Definicja: Spis treści



DCMI Metadata Terms, Własności w przestrzeni nazw */terms/*

Nazwa terminu: identifier

Etykieta: Identyfikator

Komentarz: Najlepiej jest identyfikować zasób za pomocą sformalizowanych ciągów znaków, zgodnie z danym systemem identyfikacji

Uszczegółowienia:

- **Nazwa terminu: bibliographicCitation**
Etykieta: Cytata bibliograficzna



**Dziękujemy
za uwagę!**

Joanna Potęga
Biblioteka Narodowa

j.potega@bn.org.pl

Agnieszka Wróbel
Biblioteka Uniwersytecka
w Warszawie

wrobel.a@uw.edu.pl