

*Treść dokumentu: Warunki przekazywania obiektów cyfrowych do Biblioteki Narodowej.
Opis standardów przekazywania obiektów zdigitalizowanych oraz informacji towarzyszących w zakresie metadanych.*

Nazwa dokumentu: warunki_wplywu_obiektow_cyfrowych_do_BN_metadane_2014-03-11

Data dokumentu: 2014-03-11

Uwagi do treści dokumentu: Poniższe wymagania zastępują poprzednie wersje i stanowią całość wraz z warunkami podanymi w innych dokumentach odnoszących się do przekazywania obiektów cyfrowych do Biblioteki Narodowej opublikowanych w serwisie internetowym BN.

Warunki przekazywania obiektów cyfrowych do Biblioteki Narodowej

Metadane

A. Zakres

1. Niniejszy dokument opisuje warunki przyjmowania obiektów cyfrowych wtórnych, czyli powstałych w wyniku digitalizacji w zakresie metadanych. Cyfrowe odwzorowania obiektów powstałe w wyniku digitalizacji w dalszej części tego dokumentu określane będą jako zawartość cyfrowa. Niniejszy dokument nie precyzuje szczegółowych wymagań co do formatu plików zawierających zawartość cyfrową. Wymagania odnośnie formatów plików są określone w innym dokumencie.

B. Warunki podstawowe

2. Cała zawartość cyfrowa każdego obiektu musi być umieszczona w jednym osobnym folderze nazywanym dalej folderem obiektu.
3. W każdym folderze obiektu musi być jeden i tylko jeden plik METS opisujący ten obiekt nazywany dalej plikiem METS obiektu.
4. Plik METS obiektu musi zawierać wyłącznie lokalne odniesienia do innych plików w postaci ścieżek typu URN definiowanych względem położenia pliku METS.
5. Wszystkie dostarczone pliki zawartości cyfrowej obiektu muszą być wskazane w pliku METS obiektu (pozostała zawartość będzie traktowana jako niedostarczona).
6. Wszystkie metadane obiektu muszą być wskazane w pliku METS obiektu (pozostałe będą traktowane jako niedostarczone).
7. Akceptowane są wyłącznie poprawnie sformułowane pliki formacie METS w XML, które poprawnie przechodzą walidację na podstawie schematu XSD: www.loc.gov/standards/mets/mets.xsd (w Internecie dostępne są bezpłatne walidatory również on-line). Wszystkie używane w pliku METS przestrzenie nazw powinny być jawnie zaimportowane w nagłówku pliku oraz mieć publicznie dostępne w Internecie pliki XSD schematów pozwalających na ich walidację. Zalecane jest używanie standardu METS w wersji 1.9.1. Opis standardu METS dostępny jest pod adresem: <http://www.loc.gov/standards/mets/>

C. Organizacja zawartości cyfrowej

8. Jeśli przekazywana jest jednocześnie zawartość cyfrowa w kilku postaciach (np. pliki źródłowe TIFF, pochodne JPG, PDF, ALTO) to pliki każdej z postaci można umieścić w osobnym podfolderze folderu obiektu.
9. Jeśli obiekt ma hierarchiczną strukturę elementów składowych, można ją odwzorować umieszczając elementy w osobnych podfolderach folderu obiektu.
10. Jeśli obiekty zorganizowane są w hierarchiczną strukturę (nadrzędną w stosunku do obiektów), można ją odwzorować umieszczając podfoldery obiektów w folderach grupujących.

D. Format metadanych opisowych

11. Akceptowanymi przez BN (z zastrzeżeniem w podpunkcie 13) formatami metadanych opisowych są obecnie:
 - a. MARC21 (<http://www.bn.org.pl/dla-bibliotekarzy/nfs/marc/format-marc-21>)
 - b. Marc XML (<http://www.loc.gov/standards/marcxml/>)
 - c. MODS (<http://www.loc.gov/standards/mods/>)
 - d. DCMI Metadata Terms (<http://dublincore.org/documents/dcmi-terms/>)
 - e. Dlibra_AVIS (<http://www.wbc.poznan.pl/dlibra/attribute-schema.xsd>)
12. Jeśli biblioteka (lub inna instytucja przekazująca obiekty do BN) kataloguje obiekty źródłowe w formatach opisanych w punkcie 11, to wraz z przekazywaniem cyfrowych reprodukcji tych obiektów do BN należy przekazać wszystkie metadane obiektów źródłowych w formacie, w którym te metadane zostały oryginalnie wytworzone. Dostarczone pliki METS mają zawierać informacje pozwalające na identyfikację plików reprodukcji cyfrowych i odpowiadających im metadanych obiektów źródłowych.
13. Jeśli biblioteka (lub inna instytucja przekazująca obiekty do BN) kataloguje obiekty źródłowe w innych formatach niż opisanych w punkcie 11, to należy przekazać wszystkie metadane obiektów źródłowych po skonwertowaniu na jeden z formatów opisanych w punkcie 11 w sposób powodujący możliwie najmniejszą utratę informacji.
14. W szczególnych przypadkach, pod warunkiem wcześniejszego pisemnego uzgodnienia między instytucją przekazującą a BN kwestii technicznych i finansowych, w okolicznościach stosowania innego standardu metadanych opisowych istnieje możliwość przygotowania we współpracy z BN odpowiedniego konwertera.

E. Minimalny zestaw metadanych opisowych

15. Niezależnie od spełniania wymagań opisanych w sekcji D, przekazywany zestaw metadanych musi zawierać co najmniej następujące informacje, odpowiednio dla grup obiektów:
16. dokumenty drukowane - stare druki, wydawnictwa zwarte

- a. Autor/twórca
 - b. Tytuł
 - c. Miejsce wydania, wydawca, data wydania
 - d. Opis fizyczny
 - e. Sygnatura/numer inwentarzowy
17. dokumenty drukowane - wydawnictwa ciągłe
- a. Redaktor/instytucja sprawcza
 - b. Tytuł
 - c. Oznaczenie wydania
 - d. Numeracja (w zależności od przyjętej koncepcji digitalizacji wydawnictw ciągłych dane o numeracji każdej jednostki lub całego ciągu numeracyjnego)
 - e. Miejsce wydania, wydawca, data wydania
 - f. Opis fizyczny
 - g. Sygnatura/numer inwentarzowy
18. dokumenty rękopiśmienne
- a. Autor/twórca
 - b. Tytuł
 - c. Miejsce wydania, wydawca, data wydania
 - d. Opis fizyczny
 - e. Sygnatura/numer inwentarzowy
19. inkunabuły
- a. Dane do opisu
 - b. Autor/twórca
 - c. Tytuł
 - d. Miejsce wydania, wydawca, data wydania
 - e. Opis fizyczny
 - f. Standardowa literatura inkunabulistyczna – identyfikująca obiekt, nadająca tytuł
 - g. Sygnatura/numer inwentarzowy.

F. Struktury metadanych

20. W pliku METS obiektu powinna wystąpić przynajmniej jedna sekcja <METS:dmdSec> z metadanymi opisowymi dotyczącymi całego obiektu. Sekcja ta powinna być użyta (dowiązana) przy użyciu identyfikatora w najmniej zagłębionym elemencie <METS:div> elementu <METS:structMap>.
21. Zestawy metadanych obiektu muszą mieć postać XML i być osadzone w pliku METS obiektu lub wskazywane przez URN w zewnętrznych plikach XML umieszczonych w folderze obiektu.
22. Metadane deskryptywne (opisowe) odnoszące się do obiektu źródłowego mają być wskazane w sekcji <dmdSec> METS, opisane elementem <mdWrap> i zawarte w postaci XML w pliku METS lub opisane elementem <mdRef> wskazującym na zewnętrzny plik XML. W przypadku przekazywania metadanych w formacie binarnym MARC21, mają one być opisane elementem <mdRef> wskazującym na zewnętrzny plik, w którym mają być umieszczone formacie MARC ISO 2709.

23. Jeśli istnieje potrzeba wyróżnienia metadanych opisowych odnoszących się do obiektu źródłowego w stosunku do opisywanego obiektu cyfrowego (zawartości cyfrowej) metadane takie mogą zostać zawarte w sekcji <Mets:SourceMD>, jednak niezależnie od tego wszystkie metadane opisowe prawdziwe dla obiektu cyfrowego mają znaleźć się w podstawowej sekcji <METS:dmdSec> opisującej zawartość cyfrową.

24. Metadane opisowe mogą zostać użyte do opisu pojedynczego fragmentu zawartości cyfrowej (np. pojedynczego skanu). W takim przypadku należy użyć identyfikatora metadanych w elemencie <METS:file>.

Przykład

```
<METS:file ID="FILE001" MIMETYPE="image/jpeg" CREATED="2003-07-04T15:00:00"
SEQ="1" GROUPID="1" ADMID="ADM1" DMDID="dmd1">
```

25. Metadane opisowe mogą zostać użyte na bardziej zagłębionych poziomach <METS:div> będą wówczas dotyczyły tylko części struktury dokumentu (np. rozdziału, ilustracji, mapy, strony).

26. Dla obiektów, które powstały ze zdigitalizowanych numerów czasopism (w szczególności folderów zawierających zbiorcze pliki z wielu zeszytów), należy odpowiednio skonstruować sekcje METS fileSec oraz structMap tak, aby pozwalały zidentyfikować każdy numer i jeżeli występuje - dodatek do tego numeru.

27. Mocno zalecane jest używanie sumy kontrolnej MD5 dla elementów opisujących zawartość cyfrową. W przypadku użycia sumy kontrolnej pliki zawartości cyfrowej będą sprawdzane na wypadek błędów w oparciu o tę sumę.

Przykład

```
<METS:file ID="FID1" MIMETYPE="image/jpeg" SEQ="1" GROUPID="GID1"
CHECKSUMTYPE="MD5" CHECKSUM="A6C66A9F3832DECBACA2D0C7496F57FB">
<METS:FLocat LOCTYPE="URN" xlink:href="example2.jpg"/>
</METS:file>
```

28. Metadane techniczne dotyczące pojedynczego pliku obrazu powinny być osadzone w plikach binarnych z użyciem standardu EXIF (i) lub umieszczone w odpowiednich sekcjach pliku METS.

29. Doprecyzowania metadanych technicznych powinny być zamieszczane w pliku METS w sekcji <Mets:admSec> <Mets:TechMD> na przykład przy użyciu standardu NISOIMG (inaczej MIX).

Przykłady dostępne są pod adresem:

<http://www.loc.gov/standards/mets/METSPrimerRevised.pdf>

30. Podstawowe metadane dotyczące praw obiektu powinny być umieszczone w sekcji <Mets:admSec> w podsekcji <mets:rightsMD>

31. Metadane dotyczące proveniencji oraz cyklu życia obiektu cyfrowego (jeśli istnieją) powinny zostać zawarte w sekcji <Mets:admSec> <Mets:Digiprov:MD> zaleca się użycie w tym celu standardu PREMIS. Opis standardu PREMIS dostępny jest pod adresem: <http://www.loc.gov/standards/premis/>

32. Jeśli istnieje potrzeba wyróżnienia metadanych opisowych odnoszących się do obiektu źródłowego w stosunku do opisywanego obiektu cyfrowego (zawartości cyfrowej) metadane takie mogą zostać zawarte w sekcji <Mets:SourceMD> wykorzystując MODS, Marc XML, lub DCMI Metadata Terms jednak wszystkie metadane opisowe prawdziwe dla obiektu cyfrowego mają znaleźć się w podstawowej sekcji <METS:dmdSec> opisującej zawartość cyfrową.
33. Metadane techniczne na temat struktury stron wytworzone w procesie OCR (o ile istnieją) powinny być zapisane zgodnie ze standardem ALTO. Jeśli zostało wykonane OCR w celu uzyskania tekstowej zawartości obiektu, to jego wyniki zawierające tekstową postać zawartości cyfrowej w postaci XML w formacie ALTO mają zostać umieszczone w oddzielnych plikach (względem pliku METS obiektu) wskazanych w pliku METS obiektu w taki sposób, jak zawartość cyfrowa obiektu.

Opis standardu ALTO dostępny jest pod adresem: <http://www.loc.gov/standards/alto/>

G. Literatura

<http://www.loc.gov/standards/mets/mets-examples.html>

<http://www.dlib.org/dlib/march08/pearce/03pearce.html>

<http://www.dlib.org/dlib/september08/dappert/09dappert.html>

<http://www.loc.gov/standards/premis/ObjectsExercise-UCSD.pdf>

<http://bcpw.bg.pw.edu.pl/Content/1262>