



# DEFINICE METADATOVÝCH FORMÁTŮ

19. října

# 2014

Dokument verze 1.1.2

---

## Definice metadatových formátů pro digitalizaci monografických dokumentů (monografií, kartografických dokumentů, hudebnin)

### Autoři úprav:

Mgr. Pavla Švástová, Bc. Jaroslav Kvasnica, Mgr. Iveta Lodrová

Autor původního dokumentu, ze kterého současná verze vychází, je Mgr. Jan Hutař, Ph.D.

# Historie verzí

jméno	datum	verze dokumentu	provedené změny
Iveta Lodrová, Jaroslav Kvasnica	říjen 2014	1.1.2	<ul style="list-style-type: none"> <li>- PREMIS Event &lt;linkingObjectIdentifierType&gt; zrušení příkladu "Označení typu identifikátoru", jako příklad uvedena možnost "file"</li> <li>- &lt;fileSec&gt; vedlejšího záznamu AMD_METS: oprava atributu ADMID (před opravou AMDID)</li> <li>- Autorsko-právní metadata: u elementu &lt;mdWrap&gt; MDTYPE oprava na hodnotu "other" (z hodnoty "CopyrightMD") a MIMETYPE na hodnotu "text/xml" (z hodnoty "text/html")</li> <li>- Vyjádření logické strukturální mapy: &lt;div&gt; typu MONOGRAPH: doplněn atribut DMDID pro použití u vícesvazku</li> </ul>
Jaroslav Kvasnica	květen 2014	1.1.1	<ul style="list-style-type: none"> <li>- MODS přidána možná hodnota "gmd" u atributu authority elementu &lt;form&gt;</li> <li>- ALTO elementy "styles", "textstyle" a "paragraphstyle" změny na MA</li> <li>- MIX u elementu captureDevice přidány možné hodnoty a povinnost změněna na MA</li> <li>- MIX u elementu scannerSensor přidány možné hodnoty</li> <li>- MIX u elementů xOpticalResolution, yOpticalResolution přidány možné hodnoty</li> <li>- MIX u elementu opticalResolutionunit přidány možné hodnoty</li> <li>-</li> </ul>
Pavla Švástová, Jaroslav Kvasnica	říjen 2013	1.1	kompletní seznam změn je k nalezení na stránkách ndk.cz
Pavla Švástová, Jaroslav Kvasnica	leden - červenec 2013	1.1	<a href="http://goo.gl/6wclM">http://goo.gl/6wclM</a>
Pavla Švástová	4.4.2012	1.0	přidán element <structLink> (zrušena druhá logická mapa), zpřesnění využití identifikátorů, sjednocena vnitřní ID u METS

Pavla Švástová	16.3.2012	0.4	druhá logická mapa, drobné úpravy
Jan Hutař	12.2.2012	0.3	doplněna možnost pro obrázek, který není součástí článku do logické strukturální mapy, TYPE <div> je PAGE
Jan Hutař	6.12.2011	0.3	úprava atributu USE v <fileGroup> u ALTO XML
Jan Hutař, Přemysl Twrdý	22.-25.11.2011	0.3	další úpravy
Jan Hutař – NK	1.-21.11.2011	0.3	úpravy celého textu, logické struktury a specifikace jednotlivých metadatových schémat
Jan Hutař – NK	únor-říjen 2011	0.2	změny v závislosti na vývoji specifikace metadat pro periodika
Jan Hutař – NK	1.2.2011	0.1	první verze pro zadávací dokumentaci NDK projektu

Historie verzí je seřazena chronologicky. Mgr. Jan Hutař, Ph.D. je autorem původního dokumentu, ze kterého současné verze vychází.

# Obsah

<b>HISTORIE VERZÍ</b>	<b>2</b>
<b>1 POPIS STANDARDU</b>	<b>6</b>
1.1 Účel standardu	6
1.2 Zodpovědnost	6
1.3 Dohoda mezi dodavatelem dat a zadavatelem	6
1.4 Popis a verze použitých standardů a identifikátorů	7
1.5 Definice pojmů	8
<b>2 VÝSTUPY DIGITALIZACE</b>	<b>9</b>
<b>3 GRANULARITA METADATOVÉHO ZÁZNAMU</b>	<b>10</b>
<b>4 IDENTIFIKÁTORY</b>	<b>11</b>
<b>5 STRUKTURA PSP BALÍČKU</b>	<b>12</b>
5.1 soubor info.xml	13
5.2 složka [masterCopy]	14
5.3 složka [userCopy]	14
5.4 složka [ALTO]	14
5.5 složka [TXT]	14
5.6 složka [amdSec]	14
5.7 soubor Hlavni_METS.xml	15
5.8 soubor MD5	16
<b>6 NÁZVOVÁ KONVENCE SLOŽEK A SOUBORŮ</b>	<b>17</b>
<b>7 METADATA</b>	<b>19</b>
7.1 Kořenový element hlavního METS záznamu	19
7.2 METS hlavička <metsHdr>	19
7.3 METS část <dmdSec> - Bibliografická metadata – MODS a Dublin Core	20
7.3.1 Pole MODS a Dublin Core pro jednotlivé části monografických dokumentů	22
7.3.1.1 Pole MODS a DC pro vícesvazkovou monografii	24
7.3.1.2 Pole MODS a DC pro svazek monografie	26
7.3.1.3 Pole MODS a DC pro kartografické dokumenty	34
7.3.1.4 Pole MODS a DC pro hudebniny	41
7.3.1.5 Pole MODS a DC pro vnitřní část monografického dokumentu (vnitřní část - kapitola, obraz, mapa apod.)	47
7.3.1.6 Pole MODS a DC pro přílohu monografických dokumentů	51
7.4 METS část <amdSec> - Technická a administrativní metadata – formáty MIX a PREMIS	56
7.4.1 PREMIS Object	57
7.4.1 PREMIS Event	61
7.4.2 PREMIS Agent	63
7.4.3 Technická metadata MIX	64
7.5 METS část <fileSec>	73
7.5.1 <fileSec> hlavního záznamu METS	73
7.5.2 <fileSec> vedlejšího záznamu AMD_METS.xml	74

7.6	METS část <structMap> - Strukturální metadata	75
7.6.1	<structMap> hlavního záznamu METS	75
7.6.1.1	Vyjádření fyzické strukturální mapy	75
7.6.1.2	Vyjádření logické strukturální mapy	76
7.6.1.2.1	Vyjádření logické struktury pro kapitoly s vazbou na ALTO bloky	76
7.6.1.2.2	Vyjádření logické struktury pro kapitoly bez vazby na ALTO bloky (pouze s vazbami na strany)	80
7.6.1.2.3	Vyjádření logické struktury bez popisu kapitol	81
7.6.2	<structMap> vedlejšího záznamu METS (AMD_METS.xml)	81
7.7	METS část <structLink> - Výčet stran	82
7.7.1	Výčet stran v případě popisu vnitřních částí (s odkazy na bloky ALTA i bez nich)	82
7.7.2	Výčet stran bez popisu vnitřních částí	82
7.8	OCR (ALTO XML a TXT OCR)	83

## **8 AUTORSKO-PRÁVNÍ METADATA 98**

# 1 Popis standardu

## 1.1 Účel standardu

Definice metadatových formátů (dále DMF) slouží jako předpis pro výsledek procesu digitalizace v digitalizačních projektech v ČR. Původně byl vytvořen pro projekt Národní digitální knihovny (NDK), dále je předepsán v programu VISK 7 a byl použit i v krajských digitalizačních projektech. Slouží jako jednotný formát pro paměťové instituce, které chtějí svá data dlouhodobě archivovat v úložišti Národní knihovny (dále NK).

**Pro koho není určen?** DMF není určen pro malé lokální digitalizační aktivity, které mají jako jediný účel zpřístupnění dokumentů pro omezený okruh uživatelů. Ty si z něj mohou brát inspiraci, ale nemusí ho do důsledku dodržet, pokud nemají v plánu archivovat v úložišti NK.

V této specifikaci je blíže definován typ **monografie - text, kartografický dokument a hudebnina**. Ostatní monografické dokumenty<sup>1</sup> se popisují jako monografie (např. grafika, fotografie, rukopis, starý tisk). Definované typy se liší pouze použitím bibliografických metadat (např. typ identifikátoru, použití typických popisů pro kartografické dokumenty jako měřítko mapy atd.), strukturání, technická i administrativní metadata zůstávají totožná pro všechny.

## 1.2 Zodpovědnost

Za DMF, jeho správnost a rozvoj, je zodpovědná Národní knihovna ČR. Kontaktním oddělením je [Odbor digitálních fondů](#) resp. [Oddělení pro standardy](#).

S doplňujícími dotazy lze kontaktovat Pavlu Švástovou ([pavla.svastova@mzk.cz](mailto:pavla.svastova@mzk.cz)) nebo Jaroslava Kvasnicu ([jaroslav.kvasnica@nkp.cz](mailto:jaroslav.kvasnica@nkp.cz)).

## 1.3 Dohoda mezi dodavatelem dat a zadavatelem

DMF definuje podobu dat a metadat vyrobených dodavatelem a slouží jednak jako závazný dokument pro zadání digitalizace a jednak pro kontrolu dodaných dat. Tento dokument **neřeší** níže uvedené podrobnosti dodávky dat. Požadavky definuje zadávající instituce a ošetřuje s dodavatelem dat smlouvou.

- **granularita záznamů** - zadavatel si musí určit, zda trvá na dodání dat, která budou obsahovat vnitřní popis kapitol či článků. DMF obsahuje definici toho, jak vyrobit vnitřní články, ale záleží na zadavateli a typu projektu.
- **podrobnosti skenování** - např. použité rozlišení (300 nebo více)
- **podrobnosti o bibliografických metadatech** - zadavatel obvykle poskytuje dodavateli bibliografická metadata v MARC 21, potom záleží na domluvě, zda chce metadata pro digitální dokumenty dále obohacovat, co vše konvertovat apod.
- **způsob předání dat** (FTP, externí disky, zabalené jako ZIP / nezabalené apod.)

<sup>1</sup> „Monografie je neseriálová publikace, která systematicky, všestranně a podrobně pojednává o jednom, zpravidla úzce vymezeném tématu. Neseriálový dokument, tj. dokument skládající se z jedné části nebo z většího, předem stanoveného nebo předpokládaného, konečného počtu částí, pokud tvoří jednu bibliografickou popisnou jednotku.“ (MATUŠÍK, Zdeněk. Monografie. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-10-25]. Dostupné z:

[http://aleph.nkp.cz/F/?func=direct&doc\\_number=000001000&local\\_base=KTD](http://aleph.nkp.cz/F/?func=direct&doc_number=000001000&local_base=KTD))

## 1.4 Popis a verze použitých standardů a identifikátorů

V této verzi specifikace budou použity následující verze standardů:

Standard	verze	odkaz na webové stránky standardu
METS	<a href="#">verze 1.9.1</a>	<a href="http://www.loc.gov/standards/mets/">http://www.loc.gov/standards/mets/</a>
MODS	<a href="#">verze 3.5</a>	<a href="http://www.loc.gov/standards/mods/">http://www.loc.gov/standards/mods/</a>
Dublin Core	<a href="#">verze 1.1</a>	<a href="http://dublincore.org/documents/dces/">http://dublincore.org/documents/dces/</a>
MIX	<a href="#">verze 2.0</a>	<a href="http://www.loc.gov/standards/mix">http://www.loc.gov/standards/mix</a>
PREMIS	<a href="#">verze 2.2</a>	<a href="http://www.loc.gov/standards/premis/">http://www.loc.gov/standards/premis/</a>
ALTO	<a href="#">verze 2.0</a>	<a href="http://www.loc.gov/standards/alto/">http://www.loc.gov/standards/alto/</a>
CopyrightMD	<a href="#">verze 0.91</a>	<a href="http://www.cdlib.org/groups/rmg/">http://www.cdlib.org/groups/rmg/</a>

Pro převod metadat k titulu monografického dokumentu z MARC 21 resp. MARCXML do MODS lze využít a dle potřeb knihovny upravit oficiální mapování a konverzní šablonu ze stránek <http://www.loc.gov/standards/mods/mods-conversions.html>.

Pro vytváření metadat ve formátu Dublin Core lze využít a dle potřeb knihovny upravit oficiální mapování a konverzní šablonu z formátu MODS ze stránek <http://www.loc.gov/standards/mods/mods-dcsimple.html>.

V této verzi specifikace mohou být použity následující identifikátory:

UUID	<a href="http://www.ietf.org/rfc/rfc4122.txt">http://www.ietf.org/rfc/rfc4122.txt</a>	
čČNB	<a href="http://www.caslin.cz/spoluprace/sluzby/ccnb/">http://www.caslin.cz/spoluprace/sluzby/ccnb/</a>	identifikátor entity tak jak odpovídá katalogizačnímu záznamu v bázi ČNB
URN:NBN	<a href="http://resolver.nkp.cz">http://resolver.nkp.cz</a>	
ISBN	<a href="http://www.nkp.cz/sluzby/sluzby-pro/isbn-ismn-issn">http://www.nkp.cz/sluzby/sluzby-pro/isbn-ismn-issn</a>	pouze pro titul monografického dokumentu nebo pro soubor monografických dokumentů, které mají pouze jeden souborný záznam, ISBN není přiděleno vždy
ISMN	<a href="http://www.nkp.cz/sluzby/sluzby-pro/isbn-ismn-issn">http://www.nkp.cz/sluzby/sluzby-pro/isbn-ismn-issn</a>	podobně jako ISBN, ale pro hudebniny
ISSN	<a href="http://www.techlib.cz/cs/295-issn">http://www.techlib.cz/cs/295-issn</a>	osmimístný číselný kód, kterým se jednoznačně identifikují názvy periodik

**Povinné předpoklady:**

- veškerá metadata musí pro zápis používat kódování **UTF-8**
- velikost písmen v názvech souborů a složek - **všechna malá**

**Další doporučení:**

- úpravy obrazu, které vedou ke změně rozměrů obrazu, rozlišení apod., se musí dělat před tím, než se vytvoří OCR, tj. budou se dělat zpravidla na TIFF souborech;
- OCR (ALTO XML) bude vznikat z uživatelské kopie - OCR je lepší ze souborů s kompresí (méně šumu)
- je nutné zachovat velikost obrazu uživatelských a archivních kopií stejnou (počet pixelů, rozlišení) tak, aby ALTO XML odpovídalo
- jako výchozí SW pro vytváření souborů JPEG2000 se bude používat Kakadu

**Význam pole "Povinnost":**

Pole "povinnost" uvádí, zda je plnění jednotlivých elementů povinné, doporučené nebo volitelné. Může nabývat následujících hodnot:

- **M - mandatory** (povinně plnit – element je součástí každého záznamu)
- **MA - mandatory if available** (povinně plnit pokud je to možné, pokud lze apod.)
- **R - recommended** (plnění hodnot elementu je doporučeno, není ovšem povinné)
- **RA - recommended if available** (doporučeno pokud lze plnit)
- **O - optional** (plnění hodnot elementu závisí na konkrétních potřebách instituce)

**1.5 Definice pojmů**

- **UC** = user copy = uživatelské kopie
- **MC** = master copy = archivní kopie
- **PS** = původní sken = obrazový soubor vzniklý při digitalizaci, který se po zpracování (ořez, narovnání apod.) maže a dále se neukládá
- základní intelektuální entita ve workflow digitalizace a následně i v LTP systému = **svazek monografického dokumentu** (viz dále)
- **balíček** = složka s definovaným názvem a obsahem
- **PSP balíček** – Producer Submission Package - balíček dat a metadat, který přichází od producenta dat (např. z workflow digitalizace), PSP balíček bude obsahovat kompletní intelektuální entitu, tj. svazek monografického dokumentu
- **SIP balíček** – Submission Information Package – je balíček dat a metadat v podobě, ve které je akceptovatelný pro LTP systém nebo pro aplikaci zpřístupnění. Obvykle vzniká z PSP balíčku po konverzích a kontrolách dat a metadat.



## 2 Výstupy digitalizace

1. archivní kopie (1 MC pro každou stránku)
  - 1.1. doporučené rozlišení je 300 PPI a více, v plné barvě a ve formátu bezztrátový JPEG2000
  - 1.2. Pro dokumenty (knihy, periodika) .... v rozlišení minimálně 300 PPI. Barevná hloubka pro tyto dokumenty bude z důvodu urychlení produkce fixně nastavena na 24 bitů, barva (RGB).
2. uživatelské kopie (1 UC pro každou vzniklou MC, tedy stránku)
  - 2.1. uživatelská kopie se ukládá ve ztrátovém formátu JPEG2000
3. OCR - [ALTO XML](#) soubor pro každou stránku
  - 3.1. OCR musí existovat i pro kartografické dokumenty a hudebniny. Účelem je převést veškeré zdigitalizované textové informace, přestože mohou být v menšině proti obrazovým.
4. OCR [TXT soubor](#) - pro možnost stáhnout si jen text dokumentu (tam kde kvalita OCR je odpovídající), vyhledávání/indexace.
5. metadata pro MC
  - 5.1. bibliografická metadata – [MODS a DC](#)
  - 5.2. strukturální metadata – [METS](#)  
 Obrazové soubory UC jsou pouze součástí strukturální mapy [<structMap>](#) a [<fileSec>](#). Součástí PSP balíčku jsou obrazy UC ve složce [userCopy].
  - 5.3. technická metadata – [MIX](#), [PREMIS](#)  
 METS nebude obsahovat technická metadata pro UC, jen pro MC.
  - 5.4. administrativní metadata – [PREMIS](#), [METS](#), [copyrightMD](#)
6. kontrolní metadatové soubory ([s kontrolními součty](#) a [údaji o vzniku dat](#) apod.)

*pozn. odkazy vedou na kapitulu s příslušnou problematikou*

## 3 Granularita metadatového záznamu

### Monografie

- základní intelektuální entitou pro monografie je 1 svazek
- pokud má monografie pouze jeden svazek, vznikne jeden metadatový popis (= METS záznam)
- pokud má monografie svazků více, např. vícesvazkový slovník, tak každý svazek je jedna intelektuální entita. Vznikne tedy ke každému svazku METS záznam a PSP balíček
- v knihovních katalogích jsou někdy vícesvazkové monografie katalogizovány jako jeden soubor, tj. mají jeden záznam v katalogu, někdy jsou jednotlivé díly vedeny jako jednotlivé záznamy v katalogu; v obou případech musí vzniknout metadatový popis ke každému svazku jako základní intelektuální entitě a také PSP balíček pro každý svazek
- každý METS záznam musí obsahovat metadata o nadřazené intelektuální entitě (např. soubor monografie) tak, aby bylo možné obě entity virtuálně spojit a bylo jasné uživateli, že se jedná o jeden soubor/titul, který má dva svazky.

### Kartografické dokumenty

- základní intelektuální entitou je 1 svazek = 1 METS záznam
- jedním svazkem se rozumí i mapa na 1 listu
- pro atlasy a jiné vícestránkové kartografické dokumenty je popis totožný s monografií (liší se pouze bibliografická metadata)

### Hudebniny

- základní intelektuální entitou pro hudebniny je 1 svazek = 1 METS záznam
- dále popis totožný s monografií (liší se pouze bibliografická metadata)

## 4 Identifikátory

### Producent dat musí dogenerovat do balíčku dle této specifikace:

- **identifikátor UUID** musí být přidělen ke každé úrovni bibliografických metadat (nenachází se v bibl. záznamu v knihovním katalogu)
- **URN:NBN** pro svazek monografického dokumentu, které bude přidělováno během digitalizace
  - URN:NBN může být přidělováno také nižším logickým úrovním (entitám) – např. článek ve sborníku, příloha
  - syntax URN:NBN musí odpovídat specifikaci identifikátoru URN:NBN pro resolver NK (např. urn:nbn:cz:nk-12345a pro výstupy z projektu NDK)

### Identifikátory pro úroveň TITLE (pouze u vícesvazkových monografií):

1. **UUID** (vygeneruje dodavatel) - **povinné**
2. **Číslo České národní bibliografie – ČČNB** (povinné, pokud je možné je záznamu přidělit)
3. **ISBN** (pokud je možné je záznamu přidělit)
4. **ISMN** (pokud je možné je záznamu přidělit)
5. Jiný identifikátor – lze využít různé identifikátory, které nějakým způsobem propojí digitální záznam s fyzickou jednotkou – např. **čárový kód**, **systémové číslo** nebo **pole 001** z knihovního katalogu atd. (Podrobnější info v [kapitole 7.3.1.1](#) - MODS a DC)

### Identifikátory pro úroveň VOLUME:

1. **UUID** (vygeneruje dodavatel) - **povinné**
2. **URN:NBN** (bude přiděleno aplikací "Resolver URN:NBN") – **povinné**
3. **Číslo České národní bibliografie – ČČNB** (povinné, pokud je možné je záznamu přidělit)
4. **ISBN** (pokud je možné je záznamu přidělit)
5. **ISMN** (pokud je možné je záznamu přidělit)
6. Jiný identifikátor

### Identifikátory pro úroveň vnitřní části:

1. **UUID** (vygeneruje dodavatel) - **povinné**
2. **URN:NBN** (bude přiděleno aplikací "Resolver URN:NBN", odlišné od URN:NBN na ostatních úrovních) – nepovinné, doporučené
3. Jiné identifikátory

### Identifikátory pro úroveň přílohy:

1. **UUID** (vygeneruje dodavatel) – **povinné**
2. **URN:NBN** (bude přiděleno aplikací "Resolver URN:NBN", odlišné od URN:NBN na ostatních úrovních) – povinné, pokud lze přidělit
3. Jiný identifikátor

## 5 Struktura PSP balíčku

Souborová struktura PSP balíčku:

SLOŽKA >	OBSAHUJE >>	OBSAHUJE >>>
Monografie	info.xml	
	masterCopy (složka)	obrazy JPEG 2000 bezztrátový (nebo jiný formát podporovaný LTP úložištěm NK ČR)
	userCopy (složka)	obrazy JPEG 2000 ztrátový
	ALTO (složka)	soubory ALTO.xml pro každou stranu
	TXT (složka)	soubory OCR.TXT pro každou stranu
	amdSec (složka)	AMD_METS.xml pro každou stranu
	hlavní_METS.xml	
	soubor.md5	

Technická a administrativní metadata nesmí být obsažena v hlavním METS záznamu. Každá stránka musí mít technická a administrativní metadata ve svém souboru s METS záznam (amd\_mets.xml).

**PSP balíček = 1 adresář pro svazek monografického dokumentu. V případě, že má monografický dokument 2 nebo více svazků/dílů, tak 1 svazek = 1 PSP.**

## Hlavní složka PSP balíčku obsahuje následující složky a soubory:

### 5.1 soubor info.xml

Soubor info.xml musí obsahovat každý PSP balíček, budou zde velmi krátce zaznamenány údaje o jeho vzniku.

Aktuální verze XML schématu pro soubor info.xml je k nalezení na stránkách NDK

element	atributy	popis	Povinnost
<info>		kořenový element INFO záznamu	M
<created>		časový údaj o vzniku balíčku ve formátu ISO8601 na úroveň vteřin	M
<metadataversion>		verze metadatové specifikace, podle které byl balíček zpracován Povolené hodnoty pro monografie: <b>1.1; 1.1.1; 1.1.2</b>	M
<packageid>		název kořenového adresáře balíčku viz kap. 6	M
<mainmets>		název hlavního METS souboru včetně přípony	M
<validation>			M
	version	verze validačního nástroje (při použití validátoru NDK stačí jen verze, v případě použití jiného validačního nástroje zde vyplnit i jeho název)	M
		výstup validačního nástroje (např. OK; Valid)	R
<titleid>		soupis identifikátorů titulu - vypsát všechny možné (viz type), které titul má přidělen	M
	type	možné hodnoty: isbn, issn, ccnb, urnnbn	M
		identifikátor	M
<collection>		údaje o větším celku (projektu), pokud do některého balíček patří - např. digitalizace pro ANL	R
<institution>		název instituce, která je zadavatelem digitalizace - např. pokud je dodavatelem externí firma	R
<creator>		tvůrce balíčku – kód instituce (firmy), která balíček vytvořila	M
<size>		velikost balíčku v kB - bez souboru info.xml	M
<itemlist>		obsahuje seznam všech souborů v balíčku vč. souborů v podadresářích a souboru info.xml	M
	itemtotal	celkový počet souborů	M

<item>		obsahuje cestu k jednomu souboru např. "\\txt\txt_nk-00027x_0001.txt"  element je opakován podle nutnosti - tj. podle počtu souborů	M
<checksum>		odkaz na soubor s MD5 a jeho MD5 kontrolní součet	M
	type	bude vždy "MD5"	M
	checksum	kontrolní součet souboru - použít stejný algoritmus jako byl použit v MD5 souboru	M
		odkaz na samotný soubor MD5	M
<note>		poznámka	O

## 5.2 složka [masterCopy]

Složka s master kopiemi, obsahuje archivní kopie souborů např. ve formátu JPEG2000 v bezztrátové kompresi, 1 soubor = 1 stránka, tj. obsahuje všechny naskenované stránky monografického dokumentu.

## 5.3 složka [userCopy]

Složka s uživatelskými kopiemi, pro každou naskenovanou stránku monografického dokumentu obsahuje jeden soubor např. ve formátu JPEG2000 se ztrátovou kompresí.

## 5.4 složka [ALTO]

Obsahuje ke každé stránce 1 ALTO XML soubor, tj. tolik ALTO XML souborů, kolik je stránek svazku monografického dokumentu.

## 5.5 složka [TXT]

Obsahuje ke každé stránce 1 OCR soubor jako čistý text. Tj. tolik OCR.TXT souborů, kolik je stránek svazku monografického dokumentu.

## 5.6 složka [amdSec]

Složka s technickými metadaty – **musí obsahovat pro každou naskenovanou stránku monografického dokumentu 1 METS soubor (AMD\_METS.xml)**. Každý METS soubor AMD\_METS.xml obsahuje následující části METS formátu:

- **amdSec** – administrativní metadata – obsahuje části:
  - technická metadata **techMD**, která:
    - ve formátu **PREMISObject** popisují vlastnosti:
      - archivní kopie
      - ALTO XML
      - původního souboru, ze kterého vznikly archivní kopie např. soubory TIFF

- případná jiná reprezentace stránky
- ve formátu **MIX** popisují vlastnosti:
  - archivní kopie (veškeré generace archivních kopií)
  - případná jiná reprezentace stránky
- metadat o provenienci digitálních objektů **digiprovMD** – obsahuje části
  - **PREMISevent**
  - **PREMISagent**
- **fileSec** – sekce s odkazy na soubory – povinná část METS záznamu - v případě tohoto METS záznamu pro jednu stránku, který vzniká primárně k zachycení technických a administrativních metadat, bude odkazovat na soubory, které jsou s tou konkrétní stránkou spojeny, tj. archivní kopie, ALTO XML a OCR TXT.
- **structMap** – pouze fyzická strukturální mapa, povinná část METS záznamu. Bude ukazovat strukturu souborů k dané stránce, tj. opět archivní kopie, ALTO XML a OCR TXT.
- **copyrightMD** - autorsko-právní metadata je možné definovat i pro jednotlivé stránky. Možnosti a pravidla dědičnosti jsou uvedeny v příslušné kapitole (kapitola 8 - Autorsko-právní metadata).

## 5.7 soubor Hlavní\_METS.xml

Další částí PSP balíčku je hlavní METS dokument. Hlavní METS záznam tedy obsahuje:

- **dmdSec** – bibliografická metadata k svazku monografického dokumentu včetně popisu nadřazených entit (např. titul) nebo naopak částí (např. kapitola). Hlavním formátem musí být MODS, nutná je i přítomnost zkráceného záznam v Dublin Core.
- **amdSec** - autorsko-právní metadata k svazku monografického dokumentu popisují, zda nebo kdy může být dílo zveřejněno. Tato metadata **nejsou povinná**. Formátem bude copyrightMD.
- **fileSec** – hlavní část s linky na všechny digitální objekty (archivní kopie, ALTO XML a OCR TXT), které se váží k jednomu svazku monografického dokumentu. Obsahuje také linky na administrativní metadata AMD\_METS.xml do složky [amdSec].
- **structMap** – strukturální mapa pro celý dokument, tj. pro jeden svazek monografického dokumentu. Obsahuje:
  - logickou část – vyjadřuje logickou strukturu svazku s odkazy na ALTO XML
  - fyzickou část – obsahující informace o všech reprezentacích konkrétní stránky (archivní kopie, ALTO XML, OCR TXT a AMD\_METS.xml)
  - mapování na ALTO XML areas
- **structLink** – výčet stran jednotlivých úrovní monografického dokumentu na základě přidání vazeb mezi logickou a fyzickou strukturou

## 5.8 soubor MD5

Poslední částí PSP balíčku je soubor s kontrolními součty. PSP balíček musí obsahovat jeden soubor .md5. Tento soubor .md5 musí obsahovat kontrolní součet pro každý soubor obsažený v PSP balíčku (kromě info.xml a .md5 souboru samotného).

Obsah souboru je definován následující gramatikou ve formátu ABNF<sup>3</sup>

```

SOUBOR          = *RADEK
RADEK           = HODNOTA-MD5 MEZERA NAZEV-SOUBORU KONEC-RADKU
HODNOTA-MD5     = 32HEXDIG
MEZERA          = " " / TAB
NAZEV-SOUBORU  = *SEGMENT
KONEC-RADKU    = ( CR LF ) / LF
SEGMENT         = PATH-SEP FILENAME-CHARS
PATH-SEP        = "/" / "\"
TAB             = %x09
CR              = %x0D
LF              = %x0A
FILENAME-CHAR   = ALPHA / DIGIT / "." / "_" / "-"
FILENAME-CHARS = 1*FILENAME-CHAR

```

kde sémantika je následující:

- HODNOTA-MD5 - produkce tohoto pravidla reprezentuje výsledek výpočtu MD5 hashovací funkce, do které vstupuje obsah souboru označeného jménem souboru NAZEV-SOUBORU
- NAZEV-SOUBORU - produkce pravidla tvoří jméno souboru ve formě absolutní cesty v hierarchické struktuře balíku SIP. Cesta je absolutní vůči kořenové složce SIP balíčku

---

<sup>3</sup> <http://tools.ietf.org/html/rfc4234>



## 6 Názvová konvence složek a souborů

### Pojmenování PSP balíčku

- každý PSP balíček přicházející z digitalizace musí obsahovat pouze jedinou intelektuální entitu (svazek monografického dokumentu).
- název balíčku **musí vycházet** z identifikátoru této entity, tj. URN:NBN nebo UUID
- každý svazek monografického dokumentu musí mít svůj jednoznačný identifikátor, tím pádem pak každý PSP balíček a každý soubor v něm má vlastní jednoznačný identifikátor díky využití prefixů
- názvy nesmí obsahovat mezery a diakritiku, doporučené oddělovače jsou podtržítka a pomlčka

### Pojmenování složek

- viz návrh struktur PSP balíčku (kap. 5)
- názvy nesmí obsahovat mezery, diakritiku a dvojtečku, doporučené oddělovače jsou podtržítka a pomlčka.

Př.

Identifikátor URN:NBN přidělený dokumentu vypadá takto: **urn:nbn:cz:nk-00027x**. Z něho se použije pouze sufix **nk-00027x**, takže složka je nazvaná malými písmeny bez diakritiky a dvojteček takto: **nk-00027x**.

### Pojmenování souborů

- každý soubor musí obsahovat předponu a příponu jasně identifikující, o který soubor se jedná:
  - hlavní METS musí obsahovat předponu "mets" a příponu **.xml** (např. mets\_nk-00027x.xml)
  - soubor info.xml musí obsahovat předponu "info" a příponu **.xml** (např. info\_nk-00027x.xml)
  - vedlejší METS musí obsahovat předponu "amd\_mets" a příponu **.xml** (např. amd\_mets\_nk-00027x.xml)
  - archivní kopie obrazu musí obsahovat předponu "mc" a příponu **.jp2** (např. mc\_nk-00027x\_0001.jp2)
  - uživatelská kopie obrazu musí obsahovat předponu "uc" a příponu **.jp2** (např. uc\_nk-00027x\_0001.jp2)
  - soubor ALTO musí obsahovat předponu "alto" a příponu **.xml** (např. alto\_nk-00027x\_0001.xml)
  - soubor OCR TXT musí obsahovat předponu "txt" a příponu **.txt**. (např. txt\_nk-00027x\_0001.txt)
  - soubor MD5 musí obsahovat předponu "md5" a příponu **.md5** (např. md5\_nk-00027x.md5)
- názvy jakýchkoliv souborů náležejících k jedné základní entitě (svazek) musí být založeny na jednom typu identifikátoru
- názvy musí být pouze malými písmeny, nesmí obsahovat mezery a diakritiku, doporučené oddělovače jsou podtržítka a pomlčka

Složka jednoho balíčku PSP, který obsahuje jen jeden obrazový soubor k první stránce svazku monografie, jejíž URN:NBN je urn:nbn:cz:nk-00027x, pak musí vypadat následovně (příklad balíčku z digitalizace NDK):

nk-00027x		
	info_nk-00027x.xml	
	[masterCopy]	mc_nk-00027x_0001.jp2
	[userCopy]	uc_nk-00027x_0001.jp2
	[ALTO]	alto_nk-00027x_0001.xml
	[TXT]	txt_nk-00027x_0001.txt
	[amdSec]	amd_mets_nk-00027x_0001.xml
	mets_nk-00027x.xml	
	nk-00027x.md5	

## 7 Metadata

- veškerá metadata budou „zabalena“ pomocí kontejnerového formátu METS
- vložení metadatových formátů do kontejneru METS musí být vždy formou <mdWrap>

### 7.1 Kořenový element hlavního METS záznamu

Kořenový element hlavního METS záznamu k jednomu svazku monografického dokumentu musí obsahovat linky na specifikace jednotlivých použitých metadatových schémat (METS, MODS, Dublin Core, CopyrightMD).

element	atributy	popis	Povinnost
<mets>		kořenový element METS záznamu	M
	LABEL	LABEL – název titulu monografického dokumentu, včetně roku vydání, např. Honzíkova cesta, 1979	M
	TYPE	TYPE – hodnota vždy "Monograph"	M

### 7.2 METS hlavička <metsHdr>

Dokumentuje vznik a úpravy METS záznamu.

element	atributy	popis	Povinnost
<metsHdr>		hlavička METS záznamu	M
	LASTMODDATE	LASTMODDATE – datum poslední úpravy záznamu, musí být ve tvaru ISO 8601 (na úroveň vteřin)	M
	CREATEDATE	CREATEDATE – datum vytvoření záznamu, musí být ve tvaru ISO 8601 (na úroveň vteřin)	M
<agent>		údaje o tvůrci záznamu METS	M
	ROLE	ROLE – hodnota "CREATOR"	M
	TYPE	TYPE – hodnota "ORGANIZATION"	M
<name>		jméno jednotlivce nebo organizace; tvůrce záznamu, buď dodavatel (firma XY) nebo v případě tvorby záznamu v knihovně bude využita sigla knihoven, tj. pro NK ČR hodnota "ABA001"	M
<agent>		údaje o vlastníkovvi METS	M
	ROLE	ROLE – hodnota "ARCHIVIST"	M
	TYPE	TYPE – hodnota "ORGANIZATION"	M
<name>		jméno organizace; vlastník záznamu, v případě tvorby záznamu v knihovně bude vždy využita sigla knihoven, tj. pro NK ČR hodnota "ABA001"	M

## 7.3 METS část <dmdSec> - Bibliografická metadata – MODS a Dublin Core

- MODS a DC budou vloženy v METS části dmdSec
- z MARC 21 resp. MARCXML do MODS lze využít a dle potřeb knihovny upravit oficiální mapování a konverzní šablonu ze stránek <http://www.loc.gov/standards/mods/mods-conversions.html>
- pro vytvoření DC z MODS formátu musí být použito (a dle potřeb knihovny upraveno) oficiální mapování Kongresové knihovny – viz <http://www.loc.gov/standards/mods/mods-conversions.html>
- u digitalizovaných dokumentů je bibliografický popis vytvářen primárně z pohledu popisu fyzické předlohy, nejde o popis elektronického dokumentu

### Monografický dokument

#### a) Jednosvazkový dokument

- základní intelektuální entitou pro popis je svazek monografického dokumentu, tj. v jednom METS záznamu, který bude obsahovat metadata a strukturu jednoho svazku, budou MODS záznamy k tomuto svazku
- metadata budou popisovat entity<sup>4</sup>:
  - **svazek (volume)** - popis svazku u klasické monografické publikace (1 svazek = 1 záznam) odpovídá záznamu v katalogu
  - **vnitřní část (internalPart)** - typy "textový oddíl" (Chapter) a "obraz" (Picture)
    - bližší určení typů "kapitol" a "obrazu" (fotografie, tabulka, ilustrace, graf apod.) bude možné vyjádřit pomocí atributů a výrazů kontrolovaného slovníku v elementu <genre>
    - u popisu vnitřní části je i element <recordInfo> - důvodem je to, že popis bude vznikat manuálně a je žádoucí o vzniku záznamu držet kontrolní údaje; element je volitelný
  - **příloha (supplement)** - přílohou se rozumí volně vložená entita do jednotlivého svazku, např. mapa, klíč (řešení úloh), pracovní sešit, CD/DVD apod.

#### b) Vícesvazkový dokument

- **titul (title)** - popis nadřazené entity vícedílné monografické publikace
- **svazek (volume)** - popis jednoho svazku z vícedílné monografické publikace (1 svazek = 1 záznam) odpovídá záznamu v katalogu
- **vnitřní část (internalPart)** - typy "textový oddíl" (Chapter) a "obraz" (Picture)
  - bližší určení typů "kapitol" a "obrazu" (fotografie, tabulka, ilustrace, graf apod.) bude možné vyjádřit pomocí atributů a výrazů kontrolovaného slovníku v elementu <genre>
  - u popisu vnitřní části je i element <recordInfo> - důvodem je to, že popis bude vznikat manuálně a je žádoucí o vzniku záznamu držet kontrolní údaje; element je volitelný
- **příloha (supplement)** - přílohou se rozumí volně vložená entita do jednotlivého svazku, např. mapa, klíč (řešení úloh), pracovní sešit, CD/DVD apod.

### Rozlišujeme 3 druhy příloh monografického dokumentu:

**a) příloha, která se neskenuje**, ale chceme o ní vytvořit bibliografický záznam, dát najevo čtenáři, že existuje – to může být např. CD/DVD apod.

- digitální podoba přílohy (pokud existuje) není součástí balíčku PSP svazku

<sup>4</sup> toto pořadí nevyjadřuje logickou strukturu dokumentu, ta je popsána dále ve Specifikaci

- popis lze udělat v rámci popisu přílohy (Supplement) v MODS – viz specifikace níže
  - pokud existuje záznam v katalogu k této příloze (např. CD/DVD, mapa apod.), bude využit pro generování MODS záznamu přílohy
- taková příloha není součástí logické strukturální mapy formátu METS

**b) příloha** podobného typu, tvaru a velikosti jako je popisovaný svazek monografického dokumentu, **která se spolu s číslem skenuje**

- digitální podoba přílohy je, spolu se svazkem (Volume), součástí PSP balíčku svazku a je součástí hlavního METS záznamu
- popis lze udělat v rámci popisu přílohy (Supplement) v MODS – viz specifikace níže
- taková příloha může mít vnitřní části (InternalPart) stejně jako svazek (Volume) a jejich text je součástí ALTO XML, které je společné pro svazek (Volume) i přílohu (Supplement)
- taková příloha je součástí logické strukturální mapy formátu METS
- taková příloha je součástí fyzické strukturální mapy formátu METS (linky mezi jednotlivými soubory reprezentujícími stránky a popisnými metadaty)

**c) příloha** odlišného typu, tvaru a velikosti než je popisovaný svazek monografického dokumentu, **která se skenuje zvlášť** na svazku nezávisle

- může jít např. o mapu apod., k těmto přílohám vznikají metadata podobně jako pro jednotlivé svazky monografických publikací, ovšem na původním svazku, ke kterému příloha patřila, nezávisle - tj. pro „původní“ svazek, u kterého byla příloha, vznikne 1 popis (PSP balíček s jedním hlavním METS záznamem a ALTO XML souborem) a pro přílohu je vytvořen další 1 popis (a PSP balíček s METS záznamem)

**Obecná pravidla pro bibliografická metadata**

- stránka se nebude popisovat, její logické i fyzické číslování i typ stránky je obsaženo ve struktuře METS dokumentu (část structMap)
- typ stránky bude odpovídat přesně seznamu typů stran z nově dohodnutých pravidel popisu, které vycházejí z DTD monografie, ale jsou přidány nové typy stran – obálka (cover), reprezentace (frontJacket) a mapa (map)
  - <http://goo.gl/z5NA0Q> (advertisement, backCover, backEndSheet, blank, cover, flyLeaf, frontCover, frontEndSheet, frontJacket, index, listOfIllustrations, listOfMaps, listOfTables, map, normalPage, spine, table, tableOfContents, titlePage)
- pro každou entitu vznikne jeden MODS záznam s vlastním ID, které bude označovat i typ části (např. oddíl, ilustrace apod.) v případě opakování částí se bude opakovat odpovídající počet MODS záznamů v jednom PSP balíčku.
- každý MODS záznam má vlastní <dmdSec> část
- každý MODS záznam bude uložen ve vlastní METS části <dmdSec> pomocí mdWrap
- **všechny top elementy MODS formátu jsou opakovatelné, kromě <recordInfo>**
- **všechny elementy Dublin Core jsou opakovatelné**
- každá část <dmdSec> musí mít ID a vnořený element <mdWrap> s atributy MDTYPE, MIMETYPE

element	atributy	popis	povinnost
<dmdSec>		identifikátor <dmdSec> části METS záznamu	M
	ID	pro <dmdSec> s popisem vícesvazkové monografie hodnota "MODSMD_TITLE" a "DCMD_TITLE"	M
		pro <dmdSec> s popisem svazku (titulu) monografického dokumentu hodnota "MODSMD_VOLUME" a "DCMD_VOLUME"	
		pro <dmdSec> s popisem vnitřní části monografického dokumentu hodnota dle typů vnitřní části (kapitola, oddíl, obraz) - hodnoty "MODSMD_CHAP" a "DCMD_CHAP" pro články a hodnoty "MODSMD_PICT" a "DCMD_PICT" pro obraz	
		pro <dmdSec> s popisem přílohy monografického dokumentu hodnota "MODSMD_SUPPL" a "DCMD_SUPPL"	
<mdWrap>		element obsahující vložené záznamy MODS	M
	MDTYPE	MDTYPE – hodnota "MODS" pro záznamy v MODS, hodnota "DC" pro záznam v Dublin Core	M
	MDTYPEVERSION	číslo verze MODS, hodnota pro záznamy v MODS, např. "3.4"	O
	MIMETYPE	MIMETYPE – hodnota "text/xml"	Ö

### 7.3.1 Pole MODS a Dublin Core pro jednotlivé části monografických dokumentů

- ID u elementu <mods>: Identifikátory budou začínat prefixy: MODS\_TITLE, MODS\_VOLUME, MODS\_SUPPL, MODS\_CHAP, MODS\_PICT pro MODS, obdobně pro DC. Za ty se dále přidá podtržítka a číslo, identifikující pořadí identifikátoru, zarovnané a doplněné o nuly na 4 místa. Čtyřmístná pořadová čísla u ID jsou uvedena proto, aby byla v celém dokumentu jednotná. ID tedy vypadá následovně:
  - titul vícesvazkového dokumentu
    - MODS\_TITLE\_0001
    - DC\_TITLE\_0001
  - svazek monografie (vždy jeden)
    - MODS\_VOLUME\_0001
    - DC\_VOLUME\_0001
  - příloha (může být více)
    - MODS\_SUPPL\_XXXX, kde XXXX je pořadové číslo přílohy, např. MODS\_SUPPL\_0001 je první příloha atd.
    - DC\_SUPPL\_XXXX, kde XXXX je pořadové číslo přílohy, např. DC\_SUPPL\_0001 je první příloha atd.

- kapitola (může být více)
  - MODS\_CHAP\_XXXX, kde XXXX je pořadové číslo kapitoly, např. MODS\_CHAP\_0001 je první kapitola atd.
  - DC\_CHAP\_XXXX, kde XXXX je pořadové číslo kapitoly, např. DC\_CHAP\_0001 je první kapitola atd.
- obrázek (může být více)
  - MODS\_PICT\_XXXX, kde XXXX je pořadové číslo obrázku, např. MODS\_PICT\_0001 je první obrázek atd.
  - DC\_PICT\_XXXX, kde XXXX je pořadové číslo obrázku, např. DC\_PICT\_0001 je první obrázek atd.

### 7.3.1.1 Pole MODS a DC pro vícesvazkovou monografii

Vícesvazkové dokumenty musí být zpracovávány na dvou úrovních, budou obsahovat sekce MODS\_TITLE i MODS\_VOLUME. Tato úroveň není určena pro jednosvazkové dokumenty.

#### Obsah sloupce "Popis":

- vysvětlení a příklad
- doporučené plnění tam, kde je to možné uvést
- povinnost platí pro elementy MODS i pro elementy Dublin Core stejně
- pokud je rodičovský element např. doporučený, a dceřiný element povinný, znamená to, že dceřiný element je povinný pouze tehdy, pokud je použit element rodičovský

Element MODS	Atributy	Popis	povinnost	Element DC
<mods>	ID	ID musí vyjadřovat název úrovně, tj. např. "MODS_TITLE_0001"	M	
<titleInfo>		Název titulu, souborný název pro plnění použít katalogizační záznam	M	
<title>		názvová informace – název monografického dokumentu hodnoty převzít z katalogu,	M	<dc:title>
<subTitle>		podnázev svazku monografie	MA	
<partNumber>		číslo svazku souborného záznamu, pokud existuje	R	
<partName>		název svazku souborného záznamu, pokud existuje	R	
<originInfo>			M	
<publisher>		jméno entity, která dokument vydala, vytiskla nebo jinak vyprodukovala; odpovídá poli 260, \$b katalogizačního záznamu. Pokud existuje více vydavatelů, přebírají se ze záznamu všichni	MA	



<edition>		údaj o pořadí vydání, odpovídá poli 250 \$a katalogizačního záznamu.	MA	
<genre>		bližší údaje o typu dokumentu <b>hodnota "title"</b>	M	<dc:type>model:monograph</dc:type>
<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální, které svazek monografického dokumentu má – viz přehled typů atributů níže  Uvádějí se i neplatné resp. zrušené identifikátory – atribut <b>invalid="yes"</b>	M	<dc:identifier>
	type	type: budou se povinně vyplňovat následující hodnoty, pokud existují:	M	
		<ul style="list-style-type: none"> <li>• <b>uuid</b> – vygeneruje dodavatel</li> </ul>	M	
		<ul style="list-style-type: none"> <li>• <b>isbn</b> - převzít z katalogizačního záznamu z pole 020, \$a, \$z - celého souboru</li> </ul>	MA	
		<ul style="list-style-type: none"> <li>• <b>ismn</b> – převzít z katalogizačního záznamu z pole 024 (1. ind.="2"), \$a, \$z - celého souboru</li> </ul>	MA	
		<ul style="list-style-type: none"> <li>• <b>ccnb</b> – čČNB - převzít z katalogizačního záznamu z pole 015, \$a, \$z - celého souboru</li> </ul>	MA	

### 7.3.1.2 Pole MODS a DC pro svazek monografie

Tato úroveň je povinná pro jednosvazkové i vícesvazkové dokumenty.

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	ID musí vyjadřovat název úrovně, tj. např. "MODS_VOLUME_0001"	M	
<titleInfo>		název svazku monografie pro plnění použít katalogizační záznam  pokud má monografie více typů názvů, element se opakuje podle potřeby	M	
	type	hlavní název bez <i>type</i> – pole 245 a \$a  type: alternative – pole 246 translated – pole 242 uniform – pole 130 resp. 240	MA	
<title>		názvová informace – název svazku monografie hodnoty převzít z katalogu,  odpovídající pole a podpole podle typu, viz <i>type</i>	M	<dc:title>
<subTitle>		podnázev svazku monografie  odpovídající pole a podpole podle typu, viz <i>type</i>	MA	<dc:title>
<partNumber>		číslo části  v případě, že se jedná o vícesvazkovou monografii, je zde uvedeno číslo svazku	MA	<dc:description>
<partName>		název části  v případě, že se jedná o vícesvazkovou monografii, je zde uveden název svazku  odpovídající pole a podpole podle typu, viz <i>type</i>	MA	<dc:description>
<name>		údaje o odpovědnosti za svazek POZOR – údaje o odpovědnosti nutno přebírat z polí 1XX a 7XX MARCu21 pokud má monografie autora a ilustrátora, element <name> se opakuje s různými rolemi	MA	

	type	použít jednu z hodnot: - personal - corporate - conference - family	MA	
	usage	usage: hodnota "primary" pro označení primární autority	O	
<namePart>		údaje o křestním jméně a příjmení apod. pokud je to možné, tak vyjádřit pro křestní jméno i příjmení pokud nelze rozlišit křestní jméno a příjmení, nepoužije se "type" a jméno se zaznamená v podobě jaké je, do jednoho elementu <namePart> pokud známe datum narození a úmrtí autora, vyplnit ve tvaru RRRR-RRRR s atributem type="date".	M	<dc:creator> nutno do jednoho pole DC spojit jméno i příjmení
	type	použít jednu z hodnot: - date (RA) - family (MA) - given (MA) - termsOfAddress (RA)		
<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	M	
<roleTerm>		popis role - nutno použít kontrol. slovník, např. z MARC21	M	
	type	type: code – kód role z kontrolovaného slovníku rolí ( <a href="http://www.loc.gov/marc/relators/relaterm.htm">http://www.loc.gov/marc/relators/relaterm.htm</a> )	M	
	authority	authority – údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority="marcrelator";		
<typeOfResource>		<b>pro monografie hodnota "text";</b>  mělo by se vyčítat z MARC21 katalogizačního záznamu, z pozice 06 návěští	R	<dc:type>
<genre>		bližší údaje o typu dokumentu hodnota " <b>volume</b> "	M	<b>PRO ÚROVEŇ VOLUME PRO VÍCESVAZEK:</b> <dc:type>model:monographunit</dc:type> <b>PRO SVAZEK MONOGRAFIE:</b> <dc:type>model:mo

				nograph</dc:type>
<originInfo>		informace o původu předlohy	M	
<place>		údaje o místě spojeném s vydáním, výrobou nebo původem popisovaného dokumentu	MA	<dc:coverage>
<placeTerm>		konkrétní určení místa a země vydání, např. Praha, resp. "xr" pro ČR  odpovídá hodnotám z katalogizačního záznamu, pole 260, \$a resp. pole 008/15-17	MA	<dc:coverage>
	type	Pokud má dokument více míst vydání v poli 260, \$a, přebírají se ze záznamu všechna místa		
	authority	authority - hodnota "marccountry" jen u údaje z pole 008		
<publisher>		jméno entity, která dokument vydala, vytiskla nebo jinak vyprodukovala odpovídá poli 260 \$b katalogizačního záznamu v MARC21; Pokud má monografie více vydavatelů, přebírají se za záznamu všichni (jsou v jednom poli 260).	MA	<dc:publisher>
<dateIssued>		datum vydání předlohy  přebírat z katalogu; odpovídá hodnotě z katalogizačního záznamu, pole 260, \$c a pole 008/07-10	M	<dc:date>
	encoding	encoding – hodnota "marc" jen u údaje z pole 008	R	
	point	point – hodnoty "start" resp. "end" jen u údaje z pole 008, pro rozmezí dat	M	
	qualifier	qualifier – možnost dalšího upřesnění, hodnota "approximate" pro data, kde nevíme přesný údaj	R	
<dateCreated>		datum vydání předlohy pro rukopisy  přebírat z katalogu; odpovídá hodnotě z katalogizačního záznamu, pole 260, \$c pokud je LDR/06="d", "f", "t"	R	

<edition>		údaj o pořadí vydání, odpovídá poli 250 \$a katalogizačního záznamu	R	
<issuance>		údaje o vydávání odpovídá hodnotě uvedené v návěští MARC21 na pozici 07  Možné hodnoty: – "monographic" <b>nebo</b> – "multipart monograph" – "single unit"	M	
<language>		údaje o jazyce dokumentu; v případě vícenásobného výskytu nutno element <language> opakovat	M	
	objectPart	objectPart: možnost vyjádřit jazyk konkrétní části svazku; možné hodnoty: "summary" – odpovídá poli 041 \$b "table of contents" - odpovídá poli 041 \$f "accompanying material" - odpovídá poli 041 \$g "translation" - odpovídá poli 041 \$h	MA	
<languageTerm>		přesné určení jazyka – kódem nutno použít kontrolovaný slovník ISO 639-2, <a href="http://www.loc.gov/standards/iso639-2/php/code_list.php">http://www.loc.gov/standards/iso639-2/php/code_list.php</a>  odpovídá poli 008/35-37, resp. 041	M	<dc:language>
	type	type: použít hodnotu "code"	M	
	authority	authority: použít hodnotu "iso639-2b"	M	
<physicalDescription>		obsahuje údaje o fyzickém popisu zdroje/předlohy	M	
<form>		údaje o fyzické podobě dokumentu, např. print, electronic, microfilm apod.  odpovídá hodnotě v poli 008/23	M	<dc:format>
	authority	authority: hodnota "marcform" nebo "gmd"	M	

<extent>		údaje o rozsahu (stran, svazků nebo rozměrů)  odpovídá hodnotě v poli 300, \$a, \$b a \$c  počet stránek bude vyjádřen ve fyzické strukturální mapě	RA	<dc:format>
<note>		poznámka o fyzickém stavu dokumentu; pro každou poznámku je nutno vytvořit nový <note> element	RA	
<abstract>		shrnutí obsahu jako celku odpovídá poli 520 MARC21	R	<dc:description>
<note>		obecná poznámka ke svazku monografie jako celku  odpovídá hodnotám v poli 245, \$c (statement of responsibility) a v polích 5XX (poznámky) katalogizačního záznamu	RA	<dc:description>
<subject>		údaje o věcném třídění předpokládá se přebírání z katalogizačního záznamu	R	
	authority	authority: vyplnit hodnotu „czenas“, „eczenas“ nebo prázdné. Odpovídá hodnotě v \$2	R	
<topic>		libovolný výraz specifikující nebo charakterizující obsah svazku monografie; použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (věcné téma) nebo obsah pole 650 záznamu MARC21	R	<dc:subject>
<geographic>		geografické věcné třídění použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (geografický termín) nebo obsah pole 651 záznamu MARC21	R	<dc:subject>
<temporal>		chronologické věcné třídění použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (chronologický údaj) nebo obsah pole 648 záznamu MARC21	R	<dc:subject>
<name>		jméno použité jako věcné záhlaví použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (jméno osobní) nebo obsah pole 600 záznamu MARC21  Struktura a atributy stejné jako pro údaje o původcích – viz element <name>	R	<dc:subject>

<classification>		klasifikační údaje věcného třídění podle Mezinárodního desetinného třídění  odpovídá poli 080 MARC21	R	<dc:subject>
	authority	authority: vyplnit hodnotu "udc"		
<relatedItem>		informace o dalších dokumentech/částech/zdrojích, které jsou ve vztahu k popisovanému dokumentu; Poznámka: element <relatedItem> může obsahovat jakýkoliv jiný element MODS – jejich použití se řídí pravidly popsányými pro tyto elementy	RA	
	type	type: např. hodnota "series"		
<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální, které svazek monografie má – viz přehled typů atributů níže  Uvádějí se i neplatné resp. zrušené identifikátory – atribut <b>invalid="yes"</b>	M	<dc:identifier>
	type	type: budou se povinně vyplňovat následující hodnoty, pokud existují:	MA	
		<b>uuid</b> – vygeneruje dodavatel	M	
		<b>urnnbn</b> - pro URN:NBN, např. zápis ve tvaru urn:nbn:cz:nk-123456 pro projekt NDK	M	
		<b>ccnb</b> – čČNB - převzít z katalogizačního záznamu z pole 015, \$a, \$z	MA	
		<b>isbn</b> - převzít z katalogizačního záznamu z pole 020, \$a, \$z	MA	
		<b>ismn</b> - převzít z katalogizačního záznamu z pole 024 (1. ind.="2"), \$a, \$z	MA	
		- jiný interní identifikátor - type = <b>barcode, oclc, sysno, permalink</b> apod.	R	
<location>		údaje o uložení popisovaného dokumentu, např. signatura, místo uložení apod.	MA	
<url>		pro uvedení lokace elektronického dokumentu	O	<dc:source>
	note	note: pro poznámku o typu URL (na plný text, abstrakt apod.)	O	

<physicalLocation>		údaje o instituci, kde je fyzicky uložen <b>daný konkrétní</b> popisovaný dokument, např. NK ČR nutno použít kontrolovaný slovník – sigly knihoven (ABA001 atd.) odpovídá poli 910 \$a v MARC21 pozn. u dokumentů v digitální podobě není možné vyplnit  <b>Neopakovatelný element</b>	M	<dc:source>
	authority	authority: hodnota "siglaADR"	O	
<shelfLocator>		signatura nebo lokační údaje o <b>daném konkrétním</b> dokumentu, který slouží jako předloha	M	<dc:source>
<part>		popis části, pokud je svazek částí souboru, element může být využit jen na zaznamenání <caption>	O	
	type	type: hodnota bude vždy "volume"	O	
<detail>			O	
<caption>		text před označením čísla, např. "č.", „část“, "No." apod.	RA	
<recordInfo>		údaje o metadatovém záznamu – jeho vzniku, změnách apod.	M	
<descriptionStandard>		Popis standardu, ve kterém je přebíraný katalogizační záznam  Odpovídá hodnotě návěští záznamu MARC21, pozice 18 - hodnota „aacr“ pro LDR/18 ="a"	O	
<recordContentSource>		kód nebo jméno instituce, která záznam vytvořila nebo změnila	R	
	authority	authority – hodnota "marcorg"	R	
<recordCreationDate>		datum prvního vytvoření záznamu, na úroveň minut	M	
	encoding	encoding: záznam bude podle normy ISO 8601 na úroveň minut, hodnota atributu tedy "iso8601"	M	
<recordChangeDate>		datum změny záznamu	MA	
<recordIdentifier>		identifikátor záznamu v katalogu, přebírá se z pole 001	R	
	source	source – hodnota se přebírá z katalogu pole 003	R	



<recordOrigin>		údaje o vzniku záznamu hodnoty: "machine generated" nebo "human prepared"	R	
<languageOfCataloging>		jazyk katalogového záznamu	R	
<languageTerm>		přebírá se z katalogu - pole 40 \$b	R	
	authority	authority – hodnota "iso639-2b"	R	

## 7.3.1.3 Pole MODS a DC pro kartografické dokumenty

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	ID musí vyjadřovat název úrovně, tj. např. "MODS_VOLUME_0001"	M	
<titleInfo>		název kartografického dokumentu pro plnění použít katalogizační záznam  pokud má kartografický dokument více typů názvů, element se opakuje podle potřeby s příslušným atributem	M	
	type	hlavní název bez <i>type</i> – pole 245 a \$a  type: alternative – pole 246 translated – pole 242 uniform – pole 130 resp. 240	MA	
<title>		názvová informace – název kartografického dokumentu hodnoty převzít z katalogu, odpovídající pole a podpole podle typu viz <i>type</i>	M	<dc:title>
<subTitle>		podnázev kartografického dokumentu  odpovídající pole a podpole podle typu viz <i>type</i>	MA	<dc:title>
<partNumber>		číslo části	MA	<dc:description>
<partName>		název části  odpovídající pole a podpole podle typu viz <i>type</i>	MA	<dc:description>
<name>		údaje o odpovědnosti za svazek POZOR – údaje o odpovědnosti nutno přebírat z polí 1XX a 7XX MARCu21 pokud má kartografický dokument jiné původce než je autor, element <name> se opakuje s různými rolemi (kartograf, litograf apod.)	MA	
	type	použít jednu z hodnot: – personal – corporate – conference – family		
	usage	usage: hodnota "primary" pro označení primární autority	O	

<namePart>		údaje o křestním jméně a příjmení autora či názvu korporace. pokud je to možné, tak vyjádřit pro křestní jméno i příjmení pokud nelze rozlišit křestní jméno a příjmení, nepoužije se "type" a jméno se zaznamená v podobě jaké je, do jednoho elementu <namePart>	M	<dc:creator> nutno do jednoho pole DC spojit jméno i příjmení
	type	použít jednu z hodnot: – date (RA) – family (MA) – given (MA) – termsOfAddress (RA)		
<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	M	
<roleTerm>		popis role - nutno použít kontrol. slovník, např. z MARC21	M	
	type	type: "code" – kód role z kontrolovaného slovníku rolí ( <a href="http://www.loc.gov/marc/relators/relaterm.html">http://www.loc.gov/marc/relators/relaterm.html</a> )	M	
	authority	authority – údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority="marcrelator";		
<typeOfResource>		<b>pro kartografické dokumenty hodnota "cartographic";</b>  mělo by se vyčítat z MARC21 katalogizačního záznamu z pozice 06 návěští (pro cartographic hodnoty e a f)	R	
<genre>		bližší údaje o typu dokumentu hodnota <b>"cartographic"</b>	M	<dc:type>model: map</dc:type>
<originInfo>		informace o původu předlohy	M	
<place>		údaje o místě spojeném s vydáním, výrobou nebo původem popisovaného dokumentu	MA	<dc:coverage>
<placeTerm>		konkrétní určení místa a země vydání, např. Praha resp. xr pro ČR odpovídá hodnotám z katalogizačního záznamu, pole 260, \$a resp. pole 008/15-17	MA	<dc:coverage>
<placeTerm>	type	– "code" pro údaj z pole 008 – "text" pro údaj z pole 260 Pokud má dokument více míst vydání v poli 260, \$a, přebírají se ze záznamu všechna místa		

	authority	authority - hodnota "marccountry" jen u údaje z pole 008		
<publisher>		jméno entity, která dokument vydala, vytiskla nebo jinak vyprodukovala odpovídá poli 260 \$b katalogizačního záznamu v MARC21; Pokud má kartografický dokument více vydavatelů, přebírají se za záznamu všichni (jsou v jednom poli 260)	MA	<dc:publisher>
<dateIssued>		datum vydání předlohy, přebírat z katalogu; odpovídá hodnotě z katalogizačního záznamu, pole 260, \$c a pole 008/07-10	M	<dc:date>
	encoding	encoding – hodnota "marc" jen u údaje z pole 008		
	point	point – hodnoty "start" resp. "end" jen u údaje z pole 008, pro rozmezí dat		
	qualifier	qualifier – možnost dalšího upřesnění, hodnota "approximate" pro data, kde nevíme přesný údaj	R	
<issuance>		údaje o vydávání odpovídá hodnotě uvedené v návěští MARC21 na pozici 07	M	
<language>		údaje o jazyce dokumentu; v případě vícenásobného výskytu nutno element <language> opakovat	M	
	objectPart	objectPart: možnost vyjádřit jazyk konkrétní části svazku; možné hodnoty: "summary" – odpovídá poli 041 \$b "table of contents" - odpovídá poli 041 \$f "accompanying material" - odpovídá poli 041 \$g "translation" - odpovídá poli 041 \$h		
<languageTerm>		přesné určení jazyka – kódem nutno použít kontrolovaný slovník ISO 639-2, <a href="http://www.loc.gov/standards/iso639-2/php/code_list.php">http://www.loc.gov/standards/iso639-2/php/code_list.php</a>  odpovídá poli 008/35-37, resp. 041	M	<dc:language>
	type	type: použít hodnotu "code"	M	
	authority	authority: použít hodnotu "iso639-2b"	M	
<physicalDescription>		obsahuje údaje o fyzickém popisu zdroje/předlohy	M	

<form>		údaje o fyzické podobě dokumentu, např. kartografický dokument, mapa apod. odpovídá hodnotě v poli 008/29	M	<dc:format>
	authority	authority: hodnota "marcform" nebo "gmd"	M	
<extent>		údaje o rozsahu (stran, svazků nebo rozměrů) odpovídá hodnotě v poli 300, \$a, \$b a \$c počet stránek bude vyjádřen ve fyzické strukturální mapě	RA	<dc:format>
<note>		poznámka o fyzickém stavu dokumentu; pro každou poznámku je nutno vytvořit nový <note> element	RA	
<abstract>		shrnutí obsahu jako celku odpovídá poli 520 MARC21	R	<dc:description>
<note>		obecná poznámka ke svazku monografie jako celku odpovídá hodnotám v poli 245, \$c (statement of responsibility) a v polích 5XX (poznámky) katalogizačního záznamu	RA	<dc:description>
<subject>		údaje o věcném třídění předpokládá se přebírání z katalogizačního záznamu	R	
	authority	authority: vyplnit hodnotu „czenas“, „eczenas“ nebo prázdné. Odpovídá hodnotě v \$2	R	
<cartographics>		kartografické údaje přebírá se ze záznamu MARC 21 pole 034 je žádoucí je vyplnit v případě, pokud se jedná o samostatnou mapu, pokud jde např. o atlas, vyplňuje se v nižší úrovni	MA	
<coordinates>		souřadnice obsah pole 034 \$d, \$e, \$f, \$g	MA	
<scale>		měřítko obsah pole 255 podpole a MARC21 záznamu	MA	
<geographic>		geografické věcné třídění použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (geografický termín) nebo obsah pole 651 záznamu MARC21	MA	
<topic>		libovolný výraz specifikující nebo charakterizující obsah kartografického dokumentu; použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (věcné téma) nebo obsah pole 650 záznamu MARC21	R	<dc:subject>
<temporal>		chronologické věcné třídění použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (chronologický údaj) nebo obsah pole 648 záznamu MARC21	R	<dc:subject>

<name>		jméno použité jako věcné záhlaví použit kontrolovaný slovník - např. z báze autorit AUT NK ČR (jméno osobní) nebo obsah pole 600 záznamu MARC21  Struktura a atributy stejné jako pro údaje o původcích – viz element <name>	R	<dc:subject>
<namePart>		celé jméno se запиše do tohoto elementu	M	
<classification>		klasifikační údaje věcného třídění podle Mezinárodního desetinného třídění odpovídá poli 080 MARC21	R	<dc:subject>
	authority	authority: vyplnit hodnotu "udc"		
<relatedItem>		informace o dalších dokumentech/částech/zdrojích, které jsou ve vztahu k popisovanému dokumentu; Poznámka: element <relatedItem> může obsahovat jakýkoliv jiný element MODS – jejich použití se řídí pravidly popsanými pro tyto elementy	RA	
	type	type: např. hodnota "series"		
<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální, které kartografický dokument obsahuje – viz přehled typů atributů níže  Uvádějí se i neplatné, resp. zrušené identifikátory – atribut <b>invalid="yes"</b>	M	<dc:identifier>
	type	budou se povinně vyplňovat následující hodnoty, pokud existují:	M	
		<b>urnnbn</b> - pro URN:NBN, např. zápis ve tvaru urn:nbn:cz:ndk-123456 pro projekt NDK	M	
		<b>uuid</b> - vygeneruje dodavatel	M	
		<b>ccnb</b> – čČNB - převzít z katalogizačního záznamu, z pole 015, \$a, \$z	MA	
		<b>isbn</b> - převzít z katalogizačního záznamu z pole 020, \$a, \$z	MA	
		- jiný interní identifikátor - type = <b>barcode, oclc, sysno, permalink</b> apod.	R	
<location>		údaje o uložení popisovaného dokumentu, např. signatura, místo uložení apod.	MA	
<url>		pro uvedení lokace elektronického dokumentu	O	<dc:source>

	note	note: pro poznámku o typu URL (na plný text, abstrakt apod.)	O	
<physicalLocation>		údaje o instituci, kde je fyzicky uložen <b>daný konkrétní</b> popisovaný dokument, např. NK ČR nutno použít kontrolovaný slovník – sigly knihoven (ABA001 atd.) odpovídá poli 910 \$a v MARC21 pozn. u dokumentů v digitální podobě není možné vyplnit <b>Neopakovatelný element</b>	M	<dc:source>
	authority	authority: hodnota "siglaADR"	O	
<shelfLocator>		signatura nebo lokační údaje <b>o daném konkrétním</b> dokumentu, který slouží jako předloha	M	<dc:source>
<part>		popis části, pokud je svazek částí souboru, element může být využit jen na zaznamenání <caption>	O	
	type	type: hodnota bude vždy "volume"	O	
<detail>			O	
<caption>		text před označením čísla, např. "č.", "část", "No." apod.	RA	
<recordInfo>		údaje o metadatovém záznamu – jeho vzniku, změnách apod.	M	
<descriptionStandard>		Popis standardu, ve kterém je přebíraný katalogizační záznam Odpovídá hodnotě návěští záznamu MARC21, pozice 18 - hodnota "aacr" pro LDR/18 ="a"	O	
<recordContentSource>		kód nebo jméno instituce, která záznam vytvořila nebo změnila	R	
	authority	authority – hodnota "marcorg"	R	
<recordCreationDate>		datum prvního vytvoření záznamu, na úroveň minut	M	
	encoding	encoding: záznam bude podle normy ISO 8601 na úroveň minut, hodnota atributu tedy "iso8601"	M	
<recordChangeDate>		datum změny záznamu	MA	
<recordIdentifier>		identifikátor záznamu v katalogu, přebírá se z pole 001	R	
	source	source – hodnota se přebírá z katalogizačního záznamu, pole 003	R	

	encoding	encoding: záznam bude podle normy ISO 8601 na úrovni minut, hodnota atributu tedy "iso8601"	R	
<recordOrigin>		údaje o vzniku záznamu hodnoty: "machine generated" nebo "human prepared"	R	
<languageOfCataloging>		jazyk katalogového záznamu	R	
<languageTerm>		přebírá se z katalogu - pole 40 \$b	R	
	authority	authority – hodnota "iso639-2b"	R	



## 7.3.1.4 Pole MODS a DC pro hudebniny

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	ID musí vyjadřovat název úrovně, tj. např. "MODS_VOLUME_0001"	M	
<titleInfo>		název hudebniny pro plnění použít katalogizační záznam  pokud má hudebnina více typů názvů, element se opakuje podle potřeby s příslušným atributem	M	
	type	hlavní název bez <i>type</i> – pole 245 \$a type: alternative – pole 246 translated – pole 242 uniform – pole 130 resp. 240	MA	
<title>		názvová informace – název hudebniny, hodnoty převzít z katalogu, odpovídající pole a podpole podle typu viz <i>type</i>	M	<dc:title>
<subTitle>		podnázev hudebniny odpovídající pole a podpole podle typu viz <i>type</i>	MA	<dc:title>
<partNumber>		číslo části	MA	<dc:description>
<partName>		název části odpovídající pole a podpole podle typu viz <i>type</i>	MA	<dc:description>
<name>		údaje o odpovědnosti POZOR – údaje o odpovědnosti nutno přebírat z polí 1XX a 7XX MARCu21 pokud má hudebnina jiné původce než je autor, element <name> se opakuje s různými rolemi (skladatel, autor textu apod.)	MA	
	type	použít jednu z hodnot: – personal – corporate – conference – family		
	usage	usage: hodnota "primary" pro označení primární autority	O	
<namePart>		údaje o křestním jméně a příjmení autora či názvu korporace. pokud je to možné, tak vyjádřit pro křestní jméno i příjmení pokud nelze rozlišit křestní jméno a příjmení, nepoužije se "type" a jméno se zaznamená v podobě jaké je, do jednoho elementu <namePart>	M	<dc:creator> nutno do jednoho pole DC spojit jméno i příjmení

	type	použít jednu z hodnot: – date (RA) – family (MA) – given (MA) – termsOfAddress (RA)		
<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	M	
<roleTerm>		popis role - nutno použít kontrol. slovník, např. z MARC21	M	
	type	type: code – kód role z kontrolovaného slovníku rolí ( <a href="http://www.loc.gov/marc/relators/relaterm.html">http://www.loc.gov/marc/relators/relaterm.html</a> )	M	
	authority	authority – údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority="marcrelator";		
<typeOfResource>		<b>pro hudebniny hodnota "notated music";</b>  mělo by se vyčítat z MARC21 katalogizačního záznamu z pozice 06 návěští (pro notated music hodnoty c a d)	R	
<genre>		blíže údaje o typu dokumentu hodnota " <b>sheetmusic</b> "	M	<dc:type>model:sheetmusic</dc:type>
<originInfo>		informace o původu předlohy	M	
<place>		údaje o místě spojeném s vydáním, výrobou nebo původem popisovaného dokumentu	MA	<dc:coverage>
<placeTerm>		konkrétní určení místa a země vydání, např. Praha resp. xr pro ČR odpovídá hodnotám z katalogizačního záznamu, pole 260, \$a resp. pole 008/15-17	MA	<dc:coverage>
	type	– "code" pro údaj z pole 008 – "text" pro údaj z pole 260 Pokud má dokument více míst vydání v poli 260, \$a, přebírají se ze záznamu všechna místa		
	authority	authority - hodnota "marccountry" jen u údaje z pole 008		
<publisher>		jméno entity, která dokument vydala, vytiskla nebo jinak vyprodukovala odpovídá poli 260 \$b katalogizačního záznamu v MARC21; Pokud má hudebnina více vydavatelů, přebírají se ze záznamu všichni (jsou v jednom poli 260)	MA	<dc:publisher>

<dateIssued>		datum vydání předlohy, přebírat z katalogu; odpovídá hodnotě z katalogizačního záznamu, pole 260, \$c a pole 008/07-10	M	<dc:date>
	encoding	encoding – hodnota "marc" jen u údaje z pole 008		
	point	point – hodnoty "start" resp. "end" jen u údaje z pole 008, pro rozmezí dat		
	qualifier	qualifier – možnost dalšího upřesnění, hodnota "approximate" pro data, kde nevíme přesný údaj	R	
<issuance>		údaje o vydávání odpovídá hodnotě uvedené v návěští MARC21 na pozici 07	M	
<language>		údaje o jazyce dokumentu; v případě vícenásobného výskytu nutno element <language> opakovat	M	
	objectPart	objectPart: možnost vyjádřit jazyk konkrétní části svazku; možné hodnoty: "summary" – odpovídá poli 041 \$b "table of contents" - odpovídá poli 041 \$f "accompanying material" - odpovídá poli 041 \$g "translation" - odpovídá poli 041 \$h	MA	
<languageTerm>		přesné určení jazyka – kódem nutno použít kontrolovaný slovník ISO 639-2, <a href="http://www.loc.gov/standards/iso639-2/php/code_list.php">http://www.loc.gov/standards/iso639-2/php/code_list.php</a> odpovídá poli 008/35-37, resp. 041	M	<dc:language>
	type	type: použít hodnotu "code"	M	
	authority	authority: použít hodnotu "iso639-2b"	M	
<physicalDescription>		obsahuje údaje o fyzickém popisu zdroje/předlohy	M	
<form>		údaje o fyzické podobě dokumentu, odpovídá hodnotě v poli 008/23	M	<dc:format>
	authority	authority: hodnota "marcform" nebo "gmd"	M	
<extent>		údaje o rozsahu (stran, svazků nebo rozměrů) odpovídá hodnotě v poli 300, \$a, \$b a \$c počet stránek bude vyjádřen i ve fyzické strukturální mapě	RA	<dc:format>

<note>		poznámka o fyzickém stavu dokumentu; pro každou poznámku je nutno vytvořit nový <note> element	RA	
<abstract>		shrnutí obsahu jako celku odpovídá poli 520 MARC21	R	<dc:description>
<note>		obecná poznámka k hudebnině jako celku odpovídá hodnotám v poli 245, \$c (statement of responsibility) a v polích 5XX (poznámky) katalogizačního záznamu	RA	<dc:description>
<subject>		údaje o věcném třídění předpokládá se přebírání z katalogizačního záznamu	R	
<topic>		libovolný výraz specifikující nebo charakterizující obsah hudebniny; použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (věcné téma) nebo obsah pole 650 záznamu MARC21	R	<dc:subject>
<temporal>		chronologické věcné třídění použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (chronologický údaj) nebo obsah pole 648 záznamu MARC21	R	<dc:subject>
<name>		jméno použité jako věcné záhlaví použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (jméno osobní) nebo obsah pole 600 záznamu MARC21  Struktura a atributy stejné jako pro údaje o původcích – viz element <name>	R	<dc:subject>
<namePart>		celé jméno se запиše do tohoto elementu	M	
<classification>		klasifikační údaje věcného třídění podle Mezinárodního desetinného třídění odpovídá poli 080 MARC21	R	<dc:subject>
	authority	authority: vyplnit hodnotu "udc"		
<relatedItem>		informace o dalších dokumentech/částech/zdrojích, které jsou ve vztahu k popisovanému dokumentu; Poznámka: element <relatedItem> může obsahovat jakýkoliv jiný element MODS – jejich použití se řídí pravidly popsanými pro tyto elementy	RA	
	type	type: např. hodnota "series"		

<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální, které hudebnina obsahuje – viz přehled typů atributů níže Uvádějí se i neplatné resp. zrušené identifikátory – atribut <b>invalid="yes"</b>	M	<dc:identifier>
	type	type: budou se povinně vyplňovat následující hodnoty, pokud existují:	MA	
		<b>uuid</b> - vygeneruje dodavatel	M	
		<b>urnnbn</b> - pro URN:NBN, např. zápis ve tvaru urn:nbn:cz:ndk-123456 pro projekt NDK;	M	
		<b>isbn</b> - pokud existuje, převzít z katalogizačního záznamu z pole 020, \$a, \$z	MA	
		<b>ismn</b> - pokud existuje, převzít z katalogizačního záznamu z pole 024, \$a	MA	
		<b>ccnb</b> – čČNB - převzít z katalogizačního záznamu, z pole 015, \$a, \$z	MA	
		<b>jiný interní identifikátor - type = barcode, oclc, sysno, permalink apod.</b>	R	
<location>		údaje o uložení popisovaného dokumentu, např. signatura, místo uložení apod.	MA	
<url>		pro uvedení lokace elektronického dokumentu	O	<dc:source>
	note	note: pro poznámku o typu URL (na plný text, abstrakt apod.)	O	
<physicalLocation>		údaje o instituci, kde je fyzicky uložen <b>daný konkrétní</b> popisovaný dokument, např. NK ČR nutno použít kontrolovaný slovník – sigly knihoven (ABA001 atd.) odpovídá poli 910 \$a v MARC21  <b>Neopakovatelný element</b>	M	<dc:source>
	authority	authority: hodnota "siglaADR"	O	
<shelfLocator>		signatura nebo lokační údaje <b>o daném konkrétním</b> dokumentu, který slouží jako předloha	M	<dc:source>
<part>		popis části, pokud je svazek částí souboru, element může být využit jen na zaznamenání <caption>	O	
	type	type: hodnota bude vždy "volume"		

<detail>				
<caption>		text před označením čísla, např. "č.", "část", "No." apod.	RA	
<recordInfo>		údaje o metadatovém záznamu – jeho vzniku, změnách apod.	M	
<descriptionStandard>		Popis standardu, ve kterém je přebíraný katalogizační záznam Odpovídá hodnotě návěští záznamu MARC21, pozice 18 - hodnota "aacr" pro LDR/18 ="a"	O	
<recordContentSource>		kód nebo jméno instituce, která záznam vytvořila nebo změnila	R	
	authority	authority – hodnota "marcorg"		
<recordCreationDate>		datum prvního vytvoření záznamu, na úroveň minut	M	
	encoding	encoding: záznam bude podle normy ISO 8601 na úroveň minut, hodnota atributu tedy "iso8601"	M	
<recordChangeDate>		datum změny záznamu	MA	
<recordIdentifier>		identifikátor záznamu v katalogu, přebírá se z pole 001	R	
	source	source – hodnota se přebírá z katalogu pole 003	R	
	encoding	encoding: záznam bude podle normy ISO 8601 na úroveň minut, hodnota atributu tedy "iso8601"	R	
<recordOrigin>		údaje o vzniku záznamu hodnoty: "machine generated" nebo "human prepared"	R	
<languageOfCataloging>		jazyk katalogového záznamu	R	
<languageTerm>		prebírání se z katalogu - pole 40 \$b	R	
	authority	authority – hodnota "iso639-2b"	R	

### 7.3.1.5 Pole MODS a DC pro vnitřní část monografického dokumentu (vnitřní část - kapitola, obraz, mapa apod.)

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	ID musí vyjadřovat název úrovně, tj. např. "MODS_PICT_XXXX" pro obrázek v textu, "MODS_CHAP_XXXX" pro textový oddíl apod. "XXXX" je pořadové číslo kapitoly nebo obrázku např. "MODS_PICT_0001" bude v ID prvního obrázku atd.	M	
<titleInfo>		názvová informace vnitřní části	M	
<title>		vlastní název vnitřní části (oddílu, obrazu); u obrazu brát případně z popisku obrazu; pokud není titul, nutno vyplnit hodnotu "untitled"	M	<dc:title>
<subTitle>		podnázev vnitřní části (oddílu); např. podnázev kapitoly	MA	<dc:title>
<partNumber>		číslo vnitřní části	RA	<dc:title>
<partName>		název vnitřní části	RA	<dc:title>
<name>		údaje o odpovědnosti za vnitřní část (oddíl i obraz)	MA	<dc:creator> nutno do jednoho pole DC spojit jméno i příjmení
	type	type: použít jeden z typů: – personal – corporate – conference – family	MA	
<namePart>		údaje o křestním jméně a příjmení apod. nutno vyjádřit pro křestní jméno i příjmení pokud nelze rozlišit křestní jméno a příjmení, nepoužije se "type" a jméno se zaznamená v podobě jaké je do jednoho elementu <namePart>	MA	
	type	type: použít jednu z hodnot: - date – (RA) - family – (MA) - given – (MA) - termsOfAddress – (RA)		
<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
<roleTerm>		popis role nutno použít kontrol. slovník, např. z MARC21	MA	

	type	type: code – kód role z kontrolovaného slovníku rolí ( <a href="http://www.loc.gov/marc/relators/relaterm.html">http://www.loc.gov/marc/relators/relaterm.html</a> )	M	
	authority	authority – údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority="marcrelator"	M	
<genre>		bližší údaje o typu vnitřní části povinné <b>hodnota: "chapter" nebo "picture"</b>	M	<dc:type>model:internalpart</dc:type>
	type	type: doporučené	R	
		<b>hodnota pro chapter</b> – možnost vyplnit bližší určení typu oddílu (možnost použít DTD monografie, MonographComponentPart Types): - table of content - advertisement - abstract - introduction - review - dedication - bibliography - editorsNote - preface - chapter - article - index (použije se pro všechny typy seznamů mimo hlavní obsah; např. seznam obrazů, tabulek) - unspecified – pokud nepatří ani do jedné z výše uvedených kategorií		
		<b>hodnota pro picture</b> – možnost vyplnit další určení typu obrazu: - table - illustration - chart - photograph - graphic - map - advertisement - cover - unspecified – pokud nepatří ani do jedné z výše uvedených kategorií		
<language>		údaje o jazyce vnitřní části nelze plnit u obrazu; v případě vícenásobného výskytu nutno element <language> opakovat	MA	



<languageTerm>		přesné určení jazyka – kódem nutno použít kontrolovaný slovník ISO 639-2, <a href="http://www.loc.gov/standards/iso639-2/php/code_list.php">http://www.loc.gov/standards/iso639-2/php/code_list.php</a> nelze plnit u obrazu	M	<dc:language>
	type	type: použít hodnotu "code"		
	authority	authority: použít hodnotu "iso639-2b"		
<physicalDescription>		obsahuje údaje o fyzickém popisu vnitřní části; určeno spíše pro oddíly než pro obrazy	R	
<form>		údaje o fyzické podobě vnitřní části, např. print, electronic apod.	R	<dc:format>
	authority	authority: hodnota "marcform" nebo "gmd"		
<abstract>		shrnutí obsahu vnitřní části	R	<dc:description>
<note>		obecná poznámka k vnitřní části do poznámky by se měla dávat šifra autora vnitřní části, která se vyskytuje pod vnitřní částí	RA	<dc:description>
<subject>		údaje o věcném třídění	R	
<topic>		libovolný výraz specifikující nebo charakterizující obsah vnitřní části; lze (není ovšem nutno) použít kontrolovaný slovník	M	<dc:subject>
	authority	- např. z báze autorit AUT NK ČR (věcné téma) při použití autoritních záznamů použít AUT NK ČR a atribut authority: vyplnit hodnotu "czenas"; při použití volných klíčových slov atribut authority nepoužívat	O	
<cartographics>		kartografické údaje přebírá se ze záznamu MARC 21 pole 034 je žádoucí je vyplnit v případě, pokud se jedná o samostatnou mapu, která je v tomto případě vnitřní částí atlasu	MA	<dc:subject>
<coordinates>		souřadnice obsah pole 034 \$d, \$e, \$f, \$g	MA	<dc:subject>
<scale>		měřítko obsah pole 255 podpole a MARC21 záznamu	MA	<dc:subject>
<geographic>		geografické věcné třídění použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (geografický termín) nebo obsah pole 651 záznamu MARC21	MA	<dc:subject>
	authority	authority: vyplnit hodnotu "czenas"	R	

<temporal>		chronologické věcné třídění použit kontrolovaný slovník - např. z báze autorit AUT NK ČR (chronologický údaj)	R	<dc:subject>
	authority	authority: vyplnit hodnotu "czenas"	R	
<name>		jméno použité jako věcné záhlaví použit kontrolovaný slovník - např. z báze autorit AUT NK ČR (jméno osobní)	R	<dc:subject>
	authority	authority: vyplnit hodnotu "czenas"		
<namePart>		celé jméno se zapíše do tohoto elementu		
<classification>		klasifikační údaje věcného třídění podle Mezinárodního desetinného třídění plnit pouze pro oddíl odpovídá poli 080 MARC21	RA	<dc:subject>
	authority	authority: vyplnit hodnotu „udc“		
<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální, které vnitřní část má – viz přehled typů atributů níže  Uvádějí se i neplatné resp. zrušené identifikátory - atribut <b>invalid="yes"</b>	M	<dc:identifier> povinné
	type	type: budou se povinně vyplňovat následující hodnoty, pokud existují pro oddíl nebo obraz:	MA	
		uuid – vygeneruje dodavatel	M	
		urnbn - pro URN:NBN, u vnitřních částí monografií se s URN:NBN počítá primárně pro články ve sborníku, ne pro „obyčejné“ kapitoly	O	
		<b>jiný interní identifikátor - type = barcode, oclc, sysno, permalink apod.</b>	R	
<part>		vrchní element, který bude použit pouze na záznam rozsahu vnitřní části; nelze u obrazu	RA	
<extent>		upřesnění popisu části – rozsah na stránkách	MA	<dc:format>
<start>		první stránka, na které vnitřní část začíná	MA	<dc:coverage>
<end>		poslední stránka, na které vnitřní část končí	MA	<dc:coverage>
<recordInfo>		údaje o metadatovém záznamu vnitřní části – jeho vzniku, změnách apod.	M	
<recordContentSource>		kód nebo jméno instituce, která záznam vytvořila nebo změnila; nutno vytvořit kontrolovaný slovník	R	
<recordCreationDate>		datum prvního vytvoření záznamu vnitřní části	M	

	encoding	encoding: záznam bude podle normy ISO 8601 na úroveň minut, hodnota atributu tedy "iso8601"		
<recordChangeDate>		datum změny záznamu vnitřní části	R	
	encoding	encoding: záznam bude podle normy ISO 8601 na úroveň minut, hodnota atributu tedy "iso8601"		
<recordOrigin>		údaje o vzniku záznamu vnitřní části hodnoty: "machine generated" nebo "human prepared"	R	

### 7.3.1.6 Pole MODS a DC pro přílohu monografických dokumentů

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	ID musí vyjadřovat název úrovně, tj. "MODS_SUPPL_XXXX", kde XXXX je pořadové číslo přílohy, např. "MODS_SUPPL_0001" je první příloha atd.	M	
<titleInfo>		názvová informace přílohy použít názvové autority nebo katalogizační záznam		
<title>		názvová informace – název svazku monografie, jehož součástí příloha je; převzít z katalogu	M	<dc:title>
<partNumber>		číslo přílohy, pokud nějaké má povinné, pokud lze vyplnit	MA	<dc:description>
<partName>		název přílohy	MA	<dc:title>
<name>		údaje o odpovědnosti za přílohu	MA	
	type	type: použít jeden z typů: - personal - corporate - conference - family	M	
<namePart>		údaje o křestním jméně a příjmení apod. nutno vyjádřit pro křestní jméno i příjmení pokud nelze rozlišit křestní jméno a příjmení, nepoužije se "type" a jméno se zaznamená v podobě jaké je, do jednoho elementu <namePart>	M	<dc:creator> nutno do jednoho pole DC spojit jméno i příjmení

	type	použít jednu z hodnot: date – (RA) family – (MA) given – (MA) termsOfAddress – (RA)		
<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
<roleTerm>		popis role nutno použít kontrol. slovník, např. z MARC21	MA	
	type	type: "code" – kód role z kontrolovaného slovníku rolí ( <a href="http://www.loc.gov/marc/relators/relaterm.html">http://www.loc.gov/marc/relators/relaterm.html</a> )	M	
	authority	authority – údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority="marcrelator"	R	
<typeOfResource>		popis charakteristiky typu nebo obsahu přílohy jedna z hodnot: - <i>text</i> – např. pro přílohu typu <i>časopis, kniha, brožura</i> apod. - <i>cartographic</i> – pro mapy - <i>notated music</i> - <i>sound recording-musical</i> - pro hudební CD/DVD - <i>sound recording-nonmusical</i> - <i>sound recording</i> - <i>still image</i> – fotografie, plakáty apod. - <i>moving image</i> – pro filmová DVD - <i>three dimensional object</i> - <i>software, multimedia</i> – pro CD/DVD se SW - <i>mixed material</i>	R	<dc:type>
<genre>		bližší údaje o typu dokumentu hodnota: " <b>supplement</b> "	M	<dc:type>model:supplement</dc:type>
<originInfo>		informace o původu přílohy <i>plnit, pokud se liší od údajů v popisu svazku monografie (platí i pro jednotlivé sub-elementy)</i> Poznámka: Jeden nebo více výskytů elementů se předpokládá pro vydavatele, další výskyt v případě nutnosti popsat tiskaře. Pokud je nutno vyjádřit tiskaře (pole 260 \$f, \$e, \$g v MARC21), je nutno element <originInfo> opakovat s atributem transliteration="printer" a elementy <place>, <publisher>, <dateCreated>, které budou obsahovat údaje o tiskaři.	MA	
<place>		údaje o místě spojeném s vydáním, výrobou nebo původem přílohy	MA	<dc:coverage>

<placeTerm>		konkrétní určení místa, např. Praha odpovídá hodnotě katalogizačního záznamu, pole 260, \$a	MA	<dc:coverage>
	type	<ul style="list-style-type: none"> <li>- "code" pro údaj z pole 008</li> <li>- "text" pro údaj z pole 260</li> </ul> Pokud má dokument více míst vydání v poli 260 \$a, přebírají se ze záznamu všechna místa		
	authority	authority - hodnota "marccountry" jen u údaje z pole 008		
<publisher>		jméno entity, která přílohu vydala, vytiskla nebo jinak vyprodukovala odpovídá poli 260 \$b katalogizačního záznamu v MARC21	MA	<dc:publisher>
<dateIssued>		datum vydání přílohy, dle toho jaké údaje jsou k dispozici možno použít hodnotu z katalogizačního záznamu, pole 260 \$c jiná data než rok možno zapsat v následujících podobách: <ul style="list-style-type: none"> <li>- DD.MM.RRRR – pokud víme den, měsíc i rok vydání</li> <li>- RRRR – pokud víme pouze rok</li> <li>- MM.RRRR – pokud víme jen měsíc a rok vydání</li> <li>- DD.-DD.MM.RRRR – vydání pro více dní</li> <li>- MM.-MM.RRRR – vydání pro více měsíců</li> </ul>	MA	<dc:date>
	qualifier	qualifier – možnost dalšího upřesnění, hodnota "approximate" pro data, kde nevíme přesný údaj	O	
<dateCreated>		datum vytvoření přílohy bude použito pouze při popisu tiskaře, viz poznámka u elementu <originInfo> nebo např. u popisu CD/DVD apod. odpovídá hodnotě z katalogizačního záznamu, pole 260, \$g	R	
	qualifier	qualifier – možnost dalšího upřesnění, hodnota "approximate" pro data, kde nevíme přesný údaj	R	
<frequency>		údaje o pravidelnosti vydávání odpovídá údajům MARC21 v poli 310 nebo pozici 18 v poli 008	RA	
<language>		údaje o jazyce dokumentu	M	
<languageTerm>		přesné určení jazyka – kódem nutno použít kontrolovaný slovník ISO 639-2, <a href="http://www.loc.gov/standards/iso639-2/php/code_list.php">http://www.loc.gov/standards/iso639-2/php/code_list.php</a>	M	<dc:language>
	type	type: použít hodnotu "code"	M	

	authority	authority: použít hodnotu "iso639-2b"	M	
<b>&lt;physicalDescription&gt;</b>		obsahuje údaje o fyzickém popisu	M	
<form>		údaje o fyzické podobě dokumentu, např. print, electronic apod. povinné pro tištěné předlohy hodnota "print", pro elektronické přílohy "electronic" odpovídá hodnotám pozice 23 a 29 v poli 008 MARC21	M	<dc:format>
	authority	authority: hodnota "marcform" nebo "gmd"	R	
<extent>		údaje o rozsahu (stran, svazků nebo rozměrů) odpovídá hodnotám v poli 300 \$a, \$c MARC21, pokud jsou vyplněna obě pole, bude se element <extent> opakovat	RA	<dc:format>
<note>		poznámka o fyzickém stavu dokumentu; pro každou poznámku je nutno vytvořit nový <note> element	RA	
<b>&lt;abstract&gt;</b>		shrnutí obsahu dokumentu odpovídá poli 520 MARC21	RA	<dc:description>
<b>&lt;note&gt;</b>		obecná poznámka k dokumentu odpovídá poli 500 v MARC21	RA	<dc:description>
<b>&lt;subject&gt;</b>		údaje o věcném třídění	R	
	authority	authority: vyplnit hodnotu "czenas"	R	
<topic>		libovolný výraz specifikující nebo charakterizující obsah přílohy; použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (věcné téma)	MA	<dc:subject>
<geographic>		geografické věcné třídění použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (geografický termín)	R	<dc:subject>
<temporal>		chronologické věcné třídění použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (chronologický údaj)	R	<dc:subject>
<name>		jméno použité jako věcné záhlaví použít kontrolovaný slovník - např. z báze autorit AUT NK ČR (jméno osobní)	R	<dc:subject>
<namePart>		celé jméno se zapíše do tohoto elementu	R	
<b>&lt;classification&gt;</b>		klasifikační údaje věcného třídění podle Mezinárodního desetinného třídění odpovídá poli 080 MARC21	R	<dc:subject>
	authority	authority: vyplnit hodnotu "udc"		

<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální, které příloha má – viz přehled typů atributů níže  Uvádějí se i neplatné resp. zrušené identifikátory - atribut <b>invalid="yes"</b>	M	<dc:identifier>
	type	type: budou se povinně vyplňovat následující hodnoty, pokud existují:	M	
		<b>uuid</b> – vygeneruje dodavatel	M	
		<b>urnnbn</b> - pro URN:NBN, např. zápis ve tvaru urn:nbn:cz:nk-123456 pro projekt NDK;	MA	
		<b>ccnb</b> - čČNB - převzít z katalogizačního záznamu z pole 015, \$a, \$z	MA	
		<b>isbn</b> - převzít z katalogizačního záznamu z pole 020, \$a, \$z	MA	
		<b>ismn</b> - převzít z katalogizačního záznamu z pole 024 (1. ind.="2"), \$a, \$z	MA	
		<b>issn</b> - převzít z katalogizačního záznam NK ČR	MA	
	type	<b>jiný interní identifikátor - type = barcode, oclc, sysno, permalink apod.</b>	R	

## 7.4 METS část <amdSec> - Technická a administrativní metadata – formáty MIX a PREMIS

- pro všechna digitalizovaná data se bude využívat formát PREMIS (jeho části object, event a agent), pro obrazová data dále i formát MIX
- technická a administrativní metadata budou zabalena v části <amdSec> formátu METS ve vlastních formátech (MIX, PREMIS – části object; events; agent)
- technická a administrativní metadata budou vznikat i pro prvotní sken (většinou TIFF), který se po nutných úpravách maže a dále neuchovává
  - technická metadata jsou určena primárně pro zachycení technických informací o formátech souborů, o výsledcích validací a kontrol
  - administrativní metadata zachycují veškeré změny, procesy apod., které byly na datech i metadatach provedeny
- technická metadata přicházející z digitalizace jsou dále v maximální míře ukládána v LTP systému (po namapování do interního formátu LTP systému)
- všechny PREMIS záznamy budou ve vedlejším METS záznamu (AMD\_METS.xml), který je určen pro administrativní a technická metadata (spolu s MIX záznamy).
- celý METS záznam (AMD\_METS.xml) je linkován z hlavního METS záznamu dokumentu
- **plnění technických metadat se předpokládá z výstupů vzniklých využitím služeb třetích stran, jako jsou JHOVE2, PRONOM aj.)**
- **<amdSec> část bude existovat vždy jedna pro všechny reprezentace jedné stránky dokumentu (MC, ALTO XML, OCR.TXT) a bude obsahovat metadata v <techMD> a <digiprovMD> podčástech**
- část <amdSec> musí mít ID a vnořený element <techMD> nebo <digiprovMD>, oba s ID a vnořeným elementem <mdWrap> s atributem MDTYPE

Element	Atributy	Popis	Povinnost
<amdSec>		element obsahující technická metadata ve formátu PREMIS nebo MIX	M
	ID	ID – identifikátor konkrétní části <amdSec>, např. pro stránku 1 hodnota "PAGE0001", pro stránku 2 "PAGE0002" atd.	
<techMD> nebo <digiprovMD>		element rozlišující typy jednotlivých administrativních metadat	M
	ID	<b>ID pro část &lt;techMD&gt;:</b>	
		pro části obsahující PREMIS Object hodnota: <ul style="list-style-type: none"> <li>○ "OBJ_001" pro původní smazaný soubor, např. TIFF</li> <li>○ "OBJ_002" pro MC</li> <li>○ "OBJ_003" pro ALTO XML</li> </ul> počet PREMIS Object není omezen, číslování pokračuje OBJ 004 atd.	



		pro části obsahující MIX hodnota: <ul style="list-style-type: none"> <li>○ "MIX_001" = MIX metadata pro původní smazaný soubor, např. TIFF,</li> <li>○ "MIX_002" pro MC</li> </ul> číselná část ID odpovídá číselné části ID pro PREMIS Object, další záznamy MIX pro novou obrazovou reprezentaci stránky tedy přebírá číslování z techMD pro PREMIS Object (např. pro OBJ_004 v PREMIS Object by to byl MIX_004)	
		<b>ID pro část &lt;digiprovMD&gt;:</b>	
		pro části obsahující PREMIS Event hodnota: <ul style="list-style-type: none"> <li>○ "EVT_001" atd.</li> </ul> pro části obsahující PREMIS Agent hodnota <ul style="list-style-type: none"> <li>○ "AGENT_001" atd.</li> </ul>	
<mdWrap>		element obsahující vložené záznamy PREMIS, MIX	M
	MDTYPE	MDTYPE pro záznamy PREMIS , event i agent vždy hodnota „PREMIS“ pro záznamy MIX hodnota „NISOIMG“	

### 7.4.1 PREMIS Object

- popisovat se pomocí PREMIS object budou soubory, tj. dle specifikace PREMIS vždy úroveň tzv. file (ne reprezentace ani bitstream)
- záznam v PREMIS object se bude vytvářet pro každý soubor
  - vzniklý v procesu digitalizace (původní sken, který se dále maže);
  - archivní obrazové kopie (původní archivní obrazová kopie uložená do LTP),
  - ALTO XML,
  - případné další reprezentace stránky (například nová archivní obrazová kopie vytvořená migrací z původní MC)
- PREMIS object se nebude vytvářet pro OCR.TXT soubory ani pro UC
- pro každý záznam PREMIS object bude existovat vlastní podčást <techMD>
- záznam PREMIS Object pro jeden soubor bude obsahovat linky na události, které jsou popsány v PREMIS Events ve stejném METS metadatovém záznamu konkrétního dokumentu (svazku monografie) v části <digiprovMD>; přes <premis:relatedEventIdentification>, to samé platí pro objekty, které budou nalinkovány v případě vztahu (např. MC vznikla z PS) s popisovaným objektem přes <premis:relatedObjectIdentification>.
  - tj. např. PREMIS object popisující archivní soubor JPEG2000 je tímto způsobem nalinkován na původní sken např. ve formátu TIFF (resp. na jeho PREMIS object záznam) – pomocí tagu <relatedObjectIdentification>, který obsahuje ID původního objektu (např. TIFF)
  - zároveň pomocí tagu <relatedEventIdentification> je záznam PREMIS object archivního souboru JPEG2000 nalinkován na událost, během které vznikl

- **POZOR – Premis Object bude vznikat a uchovávat se i pro neexistující data (např. původní a posléze smazaný TIFF)**

#### Obsah pole „Popis“:

- vysvětlení a příklad
- doporučené plnění tam, kde je to možné
- výskyt elementu (jak je definováno formátem PREMIS – dle XSD)
  - 0-1 element je nepovinný, neopakovatelný
  - 0-n element je nepovinný, opakovatelný
  - 1-n element je povinný a opakovatelný
  - element je povinný a neopakovatelný

#### Obsah pole „Použití pro“

- použití jednotlivých elementů pro popis MC, PS (původní sken), XML (ALTO)

Element	Popis	Použití pro	Povinnost
<object>	kořenový element pro premis objekt; použít vždy s atributem podle typu objektu.  xsi:type="file" - pro soubor xsi:type="representation" - pro digitální reprezentaci xsi:type="bitstream" - pro bitstream  1-1	vše	M
<objectIdentifier>	identifikátor k jednoznačnému odlišení objektu v určitém kontextu; 1-n	MC, XML, PS	M
<objectIdentifierType>	popis kontextu, ve kterém je identifikátor unikátní, např. NDK, ANL nebo název repozitáře; nutno použít kontrolovaný slovník; 1-1	MC, XML, PS	M
<objectIdentifierValue>	vlastní hodnota identifikátoru, např. img0001-master, urn.nbn.cz-123465 apod.; 1-1	MC, XML, PS	M
<preservationLevel>	údaje o úrovni ochrany souboru, která se na něj vztahuje; některé soubory nejsou tak důležité jako jiné, mají menší úroveň ochrany; 0-n	MC, XML, PS	M
<preservationLevelValue>	hodnota úrovně ochrany, která je pro soubor relevantní, pro původní sken PS hodnota deleted, pro MC a XML hodnota preservation; 1-1	MC, XML, PS	M

<preservationLevelDateAssigned>	datum, kdy byla přiřazena hodnota úrovně ochrany, zápis v ISO 8601, na úroveň dne (DD-MM-RRRR) 0-1	MC, XML, PS	R
<b>&lt;objectCharacteristics&gt;</b>	technické údaje o souboru 1-n	MC, XML, PS	M
<compositionLevel>	údaj o tom, zda je nutné digitální objekt rozbalit nebo dekodovat; např. 0 (defaultně pro žádné zabalení nebo kodování); 1 pro jedno zabalení a kodování, podobně pak hodnota 2; 1-1	MC, XML, PS	M
<fixity>	údaje o kontrolním součtu 0-n	MC, XML, PS	M
<messageDigestAlgorithm>	použitý algoritmus kontrolního součtu, např. MD5 aj. 1-1	MC, XML, PS	M
<messageDigest>	hodnota kontrolního součtu 1-1	MC, XML, PS	M
<messageDigestOriginator>	agent (osoba, instituce, stroj, SW), který kontrolní součet vytvořil (např. JHOVE apod.) 0-1	MC, XML, PS	M
<size>	údaje o velikosti souboru v bytech 0-1	MC, XML, PS	M
<format>	údaje o formátu souboru pro soubory ALTO XML je možné vytvořit element dvakrát, jednou popisuje formát XML, podruhé obsahuje informace o použitém standardu ALTO 2.0 (viz příklad 2) 1-n	MC, XML, PS	M
<formatDesignation>	identifikace formátu souboru, výstup z JHOVE, PRONOM služeb apod. 0-1	MC, XML, PS	M
<formatName>	jméno formátu, např. image/tiff nebo Adobe PDF 1-1	MC, XML, PS	M
<formatVersion>	verze formátu, např. 6.0 0-1	MC, XML, PS	M
<formatRegistry>	identifikace formátu – dodatečná informace o záznamu formátů v registrech formátů (např. PRONOM aj.) 0-1	MC, XML, PS	M

<formatRegistryName>	jméno použitého registru formátů, např. UDFR, PRONOM aj. 1-1	MC, XML, PS	M
<formatRegistryKey>	unikátní identifikátor (označení) formátu v registru, např. fmt/155 z PRONOM 1-1	MC, XML, PS	M
<creatingApplication>	údaje o aplikaci, ve které byl popisovaný soubor vytvořen; nutno popsat skener, SW kde vzniklo ALTO XML/TXT, SW/kodek pro vytvoření JPEG2000 MC 0-n	MC, XML, PS	M
<creatingApplicationName>	název aplikace, např. ImageGear, Kakadu apod.; 0-1	MC, XML, PS	M
<creatingApplicationVersion>	verze aplikace, např. 15.03.000 0-1	MC, XML, PS	M
<dateCreatedByApplication>	datum a čas vytvoření, např. 2008-11-10T12:37:46; musí být ve tvaru ISO 8601 (na úroveň vteřin); 0-1	MC, XML, PS	M
<originalName>	původní jméno souboru , např. digibok_2007081301091_0011.jp2 0-1	MC, XML, PS	M
<relationship>	vyjádření vztahu popisovaného souboru k jiným souborům a událostem (eventům) 0-n	MC, XML	M
<relationshipType>	typ vztahu, doporučené hodnoty: derivation= vztah kde objekt je výsledkem změny jiného objektu; structural= vztah mezi částmi objektu; tj. např. ALTO vytvořené z TIFFU bude mít vztah derivation, podobně jako JPEG2000 z TIFFu vytvořený; 1-1	MC, XML;	M
<relationshipSubType>	upřesnění vztahu, doporučené hodnoty: created from; has source; is source of; has sibling; has part; is part of; has root; includes; is included in; apod.; tj. např. ALTO nebo JPEG2000 vytvořený z původního TIFFu budou mít vztah "created from" 1-1	MC, XML;	M
<relatedObjectIdentification>	identifikace souvisejícího souboru 1-n pro MC, XML pro vyjádření vztahu k původnímu objektu (skenu)	MC, XML	M
<relatedObjectIdentifierType>	specifikace kontextu, ve kterém je identifikátor souboru jedinečný, např. URN; temporary filepath; objectID 1-1	MC, XML	M

<relatedObjectIdentifierValue>	vlastní řetězec identifikátoru, např. URN:NBN:cz-1301091_011#0001 nebo název souboru, cesta k souboru apod. 1-1	MC, XML	M
<relatedEventIdentification>	identifikace s popisovaným souborem související události (eventu); seznam událostí viz PREMIS event 0-n	MC, XML	M
<relatedEventIdentifierType>	typ události, např. interní číslovací systém událostí jako no.nb.ev; NK repository event ID, UUID apod. 1-1	MC, XML	M
<relatedEventIdentifierValue>	hodnota identifikátoru události, např. NK_EVT_005 nebo hodnota UUID aj. 1-1	MC, XML	M
<relatedEventSequence>	pořadí události, např. 003; k určení pořadí lze určit datum události 0-1	MC, XML	R
<linkingEventIdentifier>	identifikátor události týkající původního skenu PS; typy událostí mohou být např. vytvoření, smazání 0-n pro PS nutný link na události vytvoření (digitalizace) a jeho vymazání	PS	M
<linkingEventIdentifierType>	typ identifikátoru události, např. UUID, NK_eventID, vlastní číslovací systém apod. 1-1	PS	M
<linkingEventIdentifierValue>	hodnota identifikátoru, např. event_01; img0001-master-event001 apod. 1-1	PS	M

### 7.4.1 PREMIS Event

- PREMIS Event záznamy shromažďují informace o procesech a událostech, které se týkají jednoho nebo více objektů, v našem případě souborů. Primární použití je k zaznamenání událostí, které popisovaný soubor mění nebo upravují.
- bude vznikat pro události, které se prováděly na obrazových datech
  - digitalizace – vytvoření prvního skenu (např. do TIFF)
  - vytvoření ALTO XML
  - vygenerování MC
  - vygenerování UC
  - vymazání PS
- popis událostí bude zachycovat informace o jejich výsledku/výstupu
- záznamy PREMIS event budou uloženy v METS záznamu určeném pro administrativní a technická metadata (AMD\_METS.xml) v jeho části <amdSec>, podčást <digiprovmD>
  - AMD\_METS.xml je linkován z hlavního METS záznamu dokumentu
- pro každou událost bude vytvořena jedna <digiprovmD> část

- každý záznam PREMIS Event je linkován na původce aktivity – tj. na PREMIS Agent záznam

### Obsah pole "Popis":

- vysvětlení a příklad
- doporučené plnění tam, kde je to možné
- výskyt elementu (jak je definováno formátem PREMIS – dle XSD)
  - 0-1 element je nepovinný, neopakovatelný
  - 0-n element je nepovinný, opakovatelný
  - 1-n element je povinný a opakovatelný
  - 1-1 element je povinný a neopakovatelný

### Pole záznamu PREMIS Event

Element	Popis	Pov.
<eventIdentifier>	údaje o identifikátoru události v kontextu digitalizace nebo repozitáře 1-1	M
<eventIdentifierType>	typ identifikátoru, např. no.nb.evt; NK_eventID, UUID apod. 1-1	M
<eventIdentifierValue>	hodnota identifikátoru, např. EVT_001; event_019 apod. 1-1	M
<eventType>	kategorizace události, nutno použít kontrolovaný slovník; typy událostí, které musí být zaznamenány: capture, migration, derivation, deletion 1-1	M
<eventDateTime>	datum a čas kdy byla událost provedena; nutno zapsat v ISO 8601 na úrovni vteřin 1-1	M
<eventDetail>	další údaje o události, doporučené hodnoty pro výše uvedené <eventType> následují za /: – capture/digitization – vznik prvního skenu – capture/XML_creation – capture/TXT_creation – migration/MC_creation – derivation/UC_creation – deletion/PS_deletion 0-1	M
<eventOutcomeInformation>	informace o výsledku události 0-n	R
<eventOutcome>	kategorizace výsledku události, např. slovy jako successful nebo failure, možno použít kódy – nutno používat kontrolovaný slovník nebo seznam kódů 0-1	M
<linkingAgentIdentifier>	identifikace jednoho nebo více agentů spojených s událostí 0-n	M

<linkingAgentIdentifierType>	označení typu identifikátoru, např. NK_AgentID, UUID apod. 1-1	M
<linkingAgentIdentifierValue>	hodnota identifikátoru, např. agent_softwareName_5.2; agent_novakJ apod. 1-1	M
<linkingAgentRole>	role agenta ve vztahu k události, např. software; SW component; operator; nutno používat kontrolovaný slovník 0-n	R
<linkingObjectIdentifier>	informace o objektu/souboru spojeného s událostí, link na něj 0-n	M
<linkingObjectIdentifierType>	označení typu identifikátoru, např. hodnota "file" 1-1	M
<linkingObjectIdentifierValue>	hodnota identifikátoru, např. URN:NBN:cz-_0011#0001 aj. 1-1	M

## 7.4.2 PREMIS Agent

- využití PREMIS Agent je spíše myšleno pro tzv. ochranné aktivity, které probíhají na archivních datech (AIP balíček) a je nutné pro každou událost na těchto datech mít přesnější informace o tom, kdo ji provedl (osoba administrátora nebo oprávněné osoby)
  - informace v PREMIS Event a PREMIS Object přicházející z procesu digitalizace v PSP balíčku jsou dostačující a dají nám dostatečné informace o události, kdy byla provedena, na jakém SW byla provedena (PREMIS object „creatingApplication“ + PREMIS event „eventDetail“ – tj. další upřesnění v PREMIS Agent není nutné
- záznam PREMIS agent obsahuje charakteristiku tzv. agenta, který je spojen s provedenou a zaznamenanou událostí (PREMIS Event)
  - agent může být osoba, organizace nebo software
- z PREMIS Event je linkováno na agenta, který určitou akci provedl, typ ID agenta a jeho hodnota jsou uvedené v Premis Events (<premis:linkingAgentIdentifier>), plný popis agenta je pak v PREMIS Agent
- záznamy PREMIS Agent budou uloženy v METS záznamu určeném pro administrativní a technická metadata (AMD\_METS.xml) v jeho části <amdSec>, podčást <digiprovmD>
  - AMD\_METS.xml je linkován z hlavního METS záznamu dokumentu
- pro každého agenta, tj. jeden PREMIS Agent záznam, bude vytvořena jedna <digiprovmD> část

### Obsah pole „Popis“:

- vysvětlení a příklad
- doporučené plnění tam, kde je to možné
- výskyt elementu (jak je definováno formátem PREMIS – dle XSD)
  - 0-1 element je nepovinný, neopakovatelný
  - 0-n element je nepovinný, opakovatelný
  - 1-n element je povinný a opakovatelný
  - 1-1 element je povinný a neopakovatelný

Element	Popis	Povinnost
<agentIdentifier>	popis identifikátoru, který jednoznačně označuje agenta v rámci jednoho kontextu (repozitář např.) 1-n	M
<agentIdentifierType>	označení typu identifikátoru, např. NK_AgentID, UUID apod. 1-1	M
<agentIdentifierValue>	hodnota identifikátoru, např. agent_softwareName_5.2; agent_novakJ apod. 1-1	M
<agentName>	textové upřesnění agenta, např. přesný název SW, plné jméno osoby apod. - FixImage1.3; Jan Novák; CCS docWorks 6.2.1; 0-n	R
<agentType>	obecné označení agenta – pro osoby např. osoba, pro SW např. software apod. hodnoty: organization; person; software 0-1	M
<agentNote>	použití pouze pokud je <agentType> Software a půjde o agenta souvisejícího s migrací původního souboru např. TIFF na JPEG2000 (creation/migration Event); bude obsahovat příkaz k výrobě JPEG2000 souboru v aplikaci Kakadu 0-n	MA

### 7.4.3 Technická metadata MIX

- **MIX záznam vzniká pouze pro obrazové soubory!**
  - tj. bude vznikat 1) pro archivní kopii, 2) další MIX záznam bude vznika pro původní soubor vzniklý prvotním skenováním (nejčastěji TIFF) a to i přesto, že tento TIFF se v průběhu výroby maže a není archivován
  - v případě vytvoření nové verze archivní kopie např. formátovou migrací, původní MIX záznam se zachovává a vytváří se navíc záznam nový pro aktuálně platnou verzi archivní kopie
  - tyto MIX záznamy budou součástí jednoho METS záznamu AMD\_METS.xml (v části <amdSec>, podčást <techMD>) pro administrativní a technická metadata, který vznikne ke každému obrazovému souboru a který je linkován z hlavního METS záznamu svazku monografie
- **MIX záznamy jednotlivých obrazových souborů se budou lišit – MIX záznam původního skenu nebude obsahovat např. element ImageProcessing, MIX záznam archivního souboru MC nebude naproti tomu obsahovat informace o procesu skenování, které se váží k původnímu skenu a budou v elementu ImageCaptureMetadata apod. – podrobnosti viz tabulka níže, sloupec „užití pro MC a PS“**
- pro každý záznam MIX bude vytvořena vlastní část <techMD>
- MIX může být také zapouzdřen v PREMIS Object <premis:objectCharacteristicsExtension>
- externí služby, jako např. JHOVE a PRONOM, budou využívány k plnění polí formátu MIX



- ve formátu MIX nebude uvedena informace o kontrolních součtech (fixity), která je obsažena v PREMIS Object a není nutno ji opakovat (viz MIX profily Nizozemí, Finska a Norska)
- <fileSize> je pouze doporučené, údaj o velikosti souboru je součástí popisu PREMIS Object

**Obsah pole "Popis":**

- vysvětlení a příklad
- doporučené plnění tam, kde je to možné
- výskyt elementu (jak je definováno formátem MIX – dle XSD)
  - 0-1 element je nepovinný, neopakovatelný
  - 0-n element je nepovinný, opakovatelný

**Obsah pole "Použití pro"**

- použití jednotlivých elementů pro MC, PS (původní sken) – určuje, který element je a který není součástí MIX záznamu MC nebo MIX záznamu popisujícího původní obrazový dokument ze skeneru

Element	Popis	Povinnost	Použití pro
<BasicDigitalObjectInformation>			
<ObjectIdentifier>	údaje o identifikátoru obrazového dokumentu, který je formátem MIX popsán; 0-n	R	MC, PS
<objectIdentifierType>	např. jméno souboru, nebo jiný identifikátor; 0-1	M	MC, PS
<objectIdentifierValue>	hodnota identifikátoru, např. 20110306_001.jp2 nebo urn:nbn:123456; 0-1	M	MC, PS
<fileSize>	velikost souboru 0-1	R	MC, PS
<FormatDesignation>	údaje o formátu obrazového souboru 0-1	M	MC, PS
<formatName>	název formátu, např. lze využít MIME types (Image/jp2 apod.) 0-1	M	MC, PS
<formatVersion>	verze formátu, např. 1.0 0-1	M	MC, PS
<byteOrder>	endianita, možnosti jsou little endian, middle (mix) endian a big endian 0-1	M	MC, PS
<Compression>	údaje o kompresi obrazového souboru 0-n	M	MC, PS

<compressionScheme>	informace o kompresním schématu, vyjádřeno číslem (např. 34712 je komprese JPEG2000) nebo slovy (např. JP2 Lossless) 0-1	M	MC, PS
<b>&lt;BasicImageInformation&gt;</b>	<b>základní technické údaje o obrazovém dokumentu</b> <b>0-1</b>	<b>M</b>	<b>MC, PS</b>
<b>&lt;BasicImageCharacteristics&gt;</b>	<b>0-1</b>	<b>M</b>	<b>MC, PS</b>
<imageWidth>	šířka obrazu v pixelech, např. 3987 0-1	M	MC, PS
<imageHeight>	výška obrazu v pixelech, např. 2345 0-1	M	MC, PS
<PhotometricInterpretation>	photometrická interpretace 0-1	M	MC, PS
<colorSpace>	barevný prostor, např. RGB 0-1	M	MC, PS
<ColorProfile>	údaje o barevném profilu 0-1 povinné pro dokumenty, kde je nutno uchovat přesnou reprezentaci barvy původního dokumentu a používá se ICC profil)	MA	MC, PS
<iccProfile>	ICC profil 0-1	M	MC, PS
<iccProfileName>	jméno profilu, např. sRGB, Adobe RGB aj. 0-1	M	MC, PS
<iccProfileVersion>	verze profilu, např. sRGB IEC61966-2.1 0-1	M	MC, PS
<iccProfileURL>	odkaz na profil, např. <a href="http://www.profil.cz/sRGB_v4_ICC_pref.icc">www.profil.cz/sRGB_v4_ICC_pref.icc</a> ; 0-1	R	MC, PS
<b>&lt;SpecialFormatCharacteristics&gt;</b>	<b>speciální technické údaje o obrazovém dokumentu, povinné použití pro formát JPEG2000</b> <b>0-1</b>	<b>MA</b>	<b>MC</b>
<b>&lt;JPEG2000&gt;</b>	<b>0-1</b>	<b>M</b>	<b>MC</b>
<CodecCompliance>	údaje o kodeku 0-1	M	MC
<codec>	název kodeku, např. Kakadu, LuraWave aj. 0-1	M	MC
<codecVersion>	verze kodeku, např. 3.1 0-1	M	MC

<codestreamProfile>	popis codestream profilu JPEG2000, např. P0 a P1 (viz ISO/IEC 15444-4); 0-1	M	MC
<complianceClass>	specifikace největší výšky, šířky a počtu komponentů, které dekodér dokáže dekodovat, lze použít hodnoty C0, C1 a C2; 0-1	M	MC
<EncodingOptions>	obsahuje informace o kodování JPEG2000 0-1	M	MC
<Tiles>	popis pixelové velikosti dlaždic formátu JPEG2000 0-1	M	MC
<tileWidth>	šířka dlaždice, např. 128 0-1	M	MC
<tileHeight>	výška dlaždice, např. 128 0-1	M	MC
<qualityLayers>	číselná hodnota počtu vrstev, do kterých byl JPEG2000 rozdělen, např. 12 0-1	M	MC
<resolutionLevels>	popis počtu nižších rozlišení, které lze z obrazu získat, např. 6 0-1	M	MC
<ImageCaptureMetadata>	<b>popis procesu skenování, je důležité vyplnit, protože tyto údaje nelze zjistit z finálního master/archivního souboru</b> 0-1	M	PS
<SourceInformation>	<b>informace o předloze</b> 0-1	R	PS
<sourceType>	Book, Newspaper aj.; nutno používat kontrolovaný slovník 0-1	M	PS
<SourceID>	identifikátor předlohy 0-n	R	PS
<sourceIDType>	typ identifikátoru, např. čČNB, URN:NBN 0-1	M	PS
<sourceIDValue>	vlastní hodnota identifikátoru 0-1 povinné	M	PS
<GeneralCaptureInformation>	<b>základní údaje o skenování</b> 0-1	M	PS
<dateTimeCreated>	údaj o datu a čase skenování, např. 2009-01-03T08:25:28; zapsat v ISO 8601 na úroveň vteřin 0-1	M	PS

<imageProducer>	entita provádějící skenování, např. The National Library of the Czech Republic, osoba apod. 0-1	M	PS
<captureDevice>	typ skenovacího zařízení  použít jedno z hodnot: <ul style="list-style-type: none"> <li>- reflection print scanner (nejčeflection printtyp zaříp zat</li> <li>- transmission scanner</li> <li>- digital still camera</li> <li>- still from video</li> </ul> 0-1	MA	PS
<b>&lt;ScannerCapture&gt;</b>	<b>údaje o skeneru</b> <b>0-1</b>	<b>M</b>	<b>PS</b>
<scannerManufacturer>	výrobce skeneru, např. 4DigitalBooks, Treventus, Zeutschel 0-1	M	PS
<ScannerModel>	údaje o konkrétním typu skeneru 0-1	M	PS
<scannerModelName>	jméno modelové řady skeneru, např. DL 0-1	M	PS
<scannerModelNumber>	číslo/označení modelu, např. 3000 0-1	M	PS
<scannerModelSerialNo>	výrobní číslo skeneru, např. E4R0003649 0-1	M	PS
<MaximumOpticalResolution>	údaje o maximálním optickém rozlišení skeneru 0-1	M	PS
<xOpticalResolution>	optické rozlišení na ose x, pouze číselné vyjádření např. 300 0-1	M	PS
<yOpticalResolution>	optické rozlišení na ose y, pouze číselné vyjádření např. 300 0-1	M	PS
<opticalResolutionUnit>	jednotka optického rozlišení, jedna z hodnot: no absolute unit; in.; cm 0-1	M	PS
<scannerSensor>	popis typu snímacího senzoru skenovacího zařízení, jedna z hodnot: undefined; MonochromeLinear; ColorTriLinear; ColorSequentialLinear; MonochromeArea; OneChipColourArea; TwoChipColorArea; ThreeChipColorArea; ColorSequentialArea; 0-1	M	PS
<ScanningSystemSoftware>	údaje o softwaru skenovacího zařízení 0-1	M	PS

<scanningSoftwareName>	název softwaru, např. Copinet 0-1	M	PS
<scanningSoftwareVersionNo>	číslo verze softwaru, např. 3.7 0-1	M	PS
<b>&lt;DigitalCameraCapture&gt;</b>	<b>údaje o snímacím zařízení (fotoaparát)</b> <b>0-1</b> <b>povinné, pokud je používán fotoaparát a není používán skener</b>	<b>MA</b>	<b>PS</b>
<digitalCameraManufacturer>	výrobce fotoaparátu, např. Canon 0-1	M	PS
<DigitalCameraModel>	popis modelu fotoaparátu 0-1	M	PS
<digitalCameraModelName>	název modelové řady, např. EOS 0-1	M	PS
<digitalCameraModelNumber>	označení modelu fotoaparátu, např. 1000D 0-1	M	PS
<digitalCameraModelSerialNo>	výrobní číslo přístroje, např. E12345 0-1	M	PS
<camerarSensor>	typ senzoru fotoaparátu, např. matrix aj. 0-1	M	PS
<CameraCaptureSettings>	údaje o nastavení fotoaparátu použitého ke snímání předloh 0-1	M	PS
<ImageData>	v rámci tohoto kontejnerového elementu budou použity následující sub-elementy: fNumber exposureTime isoSpeedRatings shutterSpeedValue apertureValue brightnessValue exposureBiasValue maxApertureValue subjectDistance meteringMode lightSource flash focalLength backLight exposureIndex sensingMethod cfaPattern autoFocus PrintAspectRatio  všechny hodnoty budou přebrány v případě použití fotoaparátu z údajů Exif	M	PS

<orientation>	popis orientace obrazu tak, jak je uložen vzhledem k jeho řádkům a sloupcům, např. normal*; normal, image flipper; normal, rotated 180°; unknown apod. 0-1	M	PS
<ImageAssessmentMetadata>	informace o digitálním obrazu pro jeho hodnocení a využití z hlediska dlouhodobé ochrany apod. 0-1	M	MC, PS
<SpatialMetrics>	rozměry obrázku, 2 rozměrná projekce objektů tak jak ji „vidí“ snímací zařízení 0-1	M	MC, PS
<samplingFrequencyPlane>	popis základní roviny, např. object plane (pro přímo ze předlohy digitalizované dokumenty), source object plane (pro digitalizaci mikrofilmů), camera/scanner focal plane (indikace sampl. frekvence fyzického senzoru); 0-1	R	MC, PS
<samplingFrequencyUnit>	jednotka měření sampl. frekvence  hodnoty: "no absolute unit of measurement"; "in."; "cm"; 0-1	M	MC, PS
<xSamplingFrequency>	údaje o počtu pixelů na jednotku samplovací frekvence pro šířku obrázku 0-1 povinné, pokud hodnota samplingFrequencyUnit je "in." nebo "cm"	MA	MC, PS
<numerator>	čitatel, číselné vyjádření, např. 300 0-1	M	MC, PS
<denominator>	jmenovatel, číselné vyjádření např. 1 0-1	M	MC, PS
<ySamplingFrequency>	údaje o počtu pixelů na jednotku samplovací frekvence pro výšku obrázku 0-1 povinné, pokud hodnota samplingFrequencyUnit je "in." nebo "cm"	MA	MC, PS
<numerator>	čitatel, číselné vyjádření, např. 300 0-1	M	MC, PS
<denominator>	jmenovatel, číselné vyjádření např. 1 0-1	M	MC, PS
<ImageColorEncoding>	doplňující údaje o barvě obrazu 0-1	M	MC, PS
<BitsPerSample>	počet bitů na kanál 0-1	M	MC, PS

<bitsPerSampleValue>	hodnota počtu bitů, např. 8, 1, 4 nebo 8,8,8 apod. 0-n POZOR – pro každou hodnotu je nutno element opakovat, tj. např. 3x element <bitsPerSampleValue> s hodnotou 8  <mix:BitsPerSample> <mix:bitsPerSampleValue>8</mix:bitsPerSampleValue> <mix:bitsPerSampleValue>8</mix:bitsPerSampleValue> <mix:bitsPerSampleValue>8</mix:bitsPerSampleValue> </mix:BitsPerSample>	M	MC, PS
<bitsPerSampleUnit>	specifikace jednotky, např. "integer" nebo "floating point" 0-1	R	MC, PS
<samplesPerPixel>	počet barevných komponentů na pixel, např. 1, 3, 4 0-1	M	MC, PS
<b>&lt;TargetData&gt;</b>	<b>informace o kalibračních tabulkách</b> <b>0-1</b> <b>povinné pro obrazy, kde se dělá kontrola oproti kalibrační tabulce</b>	<b>MA</b>	<b>MC</b>
<targetType>	typ kalibrační tabulky; 0= external (kalibrační tabulka se neobjeví na dig. obraze, je to oddělený dig. soubor); 1= internal (tabulka je naskenována spolu s přelohou a objeví se na dig. obraze); 0-n	M	MC
<targetID>	údaje o původu kalibrační tabulky 0-n	M	MC
<targetManufacturer>	výrobce/původce kalibrační tabulky, např. Eastman Kodak nebo NK ČR, oddělení kontroly kvality apod. 0-1	M	MC
<targetName>	název kalibrační tabulky, např. ColorChecker, MicrofilmScanTarget aj. 0-1	M	MC
<targetNo>	číslo nebo verze kalibrační tabulky 0-1	M	MC
<targetMedia>	údaj o tom, na jakém médiu je kalibrační tabulka, např. film, paper aj. 0-1	R	MC
<externalTarget>	údaje o externí kalibrační tabulce; např. link na <a href="http://skenservis.cz/target-00000001">http://skenservis.cz/target-00000001</a> nebo název a cesta ke konkrétnímu souboru 0-n povinné v případě, že byla použita externí kalibrační tabulka (targetType = 0)	MA	MC

<performaceData>	odkaz na soubor obsahující charakteristiku výkonu systému vzhledem k nastaveným hodnotám rozlišení atd.; možné hodnoty plnění – link URN nebo URL, nebo název souboru 0-n	R	MC
<ChangeHistory>	<b>dokumentace procesů provedených na obrazovém souboru v jeho životním cyklu</b> 0-1	<b>M</b>	<b>MC</b>
<ImageProcessing>	<b>údaje o zpracování obrazového souboru</b> 0-n	<b>M</b>	<b>MC</b>
<dateTimeProcessed>	2009-01-04T15:12:06; zapsat v ISO 8601 na úroveň vteřin 0-1	M	MC
<sourceData>	odkaz na původní zdrojová data, ze kterých byl vytvořen finální obrazový soubor; může to být např. URL nebo cesta do složky s původním skenem včetně názvu souboru; 0-1	M	MC
<processingAgency>	The National Library of the Czech Republic 0-n	R	MC



## 7.5 METS část <fileSec>

### 7.5.1 <fileSec> hlavního záznamu METS

#### file group

- pro obrazy i texty (ALTO XML/OCR.TXT) budou v hlavním METS záznamu použity elementy <fileGrp>, jeden element <fileGrp> bude existovat pro obrazy archivních kopií, další pro ALTO XML, další pro OCR.TXT soubory a další pro METS záznamy s technickými metadaty (AMD\_METS.xml)
1. **<fileGrp> pro obrazy archivních kopií**, bude mít tyto atributy: ID="MC\_IMGGRP" USE="Images"
    - každý soubor bude mít vlastní element <file> s následujícími atributy:
      - ID – identifikátor souboru jp2 jak je používán v METS záznamu
      - MIMETYPE – hodnota image/jp2
      - SIZE – velikost souboru jp2
      - CHECKSUMTYPE – hodnota MD5
      - CHECKSUM – hodnota kontrolního součtu
      - SEQ – pořadí souboru
      - CREATED – datum vytvoření, ISO8601 na úroveň vteřiny
    - subelementem pod <file> je element <Flocat>, který obsahuje link (ideálně v podobě nějakého identifikátoru) na obrazový soubor (xlink:href) a atribut LOCTYPE
  2. **<fileGrp> pro obrazy uživatelských kopií**, bude mít tyto atributy: ID="UC\_IMGGRP" USE="Images"
    - každý soubor bude mít vlastní element <file> s následujícími atributy:
      - ID – identifikátor souboru jp2 jak je používán v METS záznamu
      - MIMETYPE – hodnota image/jp2
      - SIZE – velikost souboru jp2
      - CHECKSUMTYPE – hodnota MD5
      - CHECKSUM – hodnota kontrolního součtu
      - SEQ – pořadí souboru
      - CREATED – datum vytvoření, ISO8601 na úroveň vteřiny
    - subelementem pod <file> je element <Flocat>, který obsahuje link (ideálně v podobě nějakého identifikátoru) na obrazový soubor (xlink:href) a atribut LOCTYPE
  3. **<fileGrp> pro ALTO XML** bude mít následující atributy: ID="ALTOGRP" USE="Layout"
    - každý ALTO XML soubor bude mít vlastní element <file> s následujícími atributy:
      - ID – identifikátor souboru ALTO XML jak je používán v METS záznamu
      - MIMETYPE – text/xml
      - SIZE – velikost souboru xml
      - CHECKSUMTYPE – hodnota MD5
      - CHECKSUM - hodnota kontrolního součtu
      - CREATED - datum vytvoření, ISO8601 na úroveň vteřiny

- subelementem pod <file> je element <Flocat>, který obsahuje link (ideálně v podobě nějakého identifikátoru) na xml soubor obsahující ALTO (xlink:href) a atribut LOCTYPE
4. **<fileGrp> pro soubory METS s technickými metadaty** AMD\_METS.xml bude mít následující atributy:
- i. ID="TECHMDGRP" USE="Technical Metadata"
- každý METS xml soubor bude mít vlastní element <file> s následujícími atributy:
    - ID - identifikátor souboru AMD\_METS.xml jak je používán v METS záznamu
    - MIMETYPE – text/xml
    - SIZE – velikost souboru xml
    - CHECKSUMTYPE – hodnota MD5
    - CHECKSUM - hodnota kontrolního součtu
    - SEQ – pořadí souboru
    - CREATED - datum vytvoření, ISO8601 na úroveň vteřiny
  - subelementem pod <file> je element <Flocat>, který obsahuje link (ideálně v podobě nějakého identifikátoru) na xml soubor AMD\_METS.xml (xlink:href) a atribut LOCTYPE
5. **<fileGrp> pro soubory OCR.TXT** bude mít následující atributy: ID="TXTGRP" USE="Text"
- každý OCR.TXT soubor bude mít vlastní element <file> s následujícími atributy:
    - ID - identifikátor souboru OCR.TXT jak je používán v METS záznamu
    - MIMETYPE – text/plain
    - SIZE - velikost souboru
    - CHECKSUMTYPE – hodnota MD5
    - CHECKSUM - hodnota kontrolního součtu
    - CREATED - datum vytvoření, ISO8601 na úroveň vteřiny
  - subelementem pod <file> je element <Flocat>, který obsahuje link (ideálně v podobě nějakého identifikátoru) na txt soubor (xlink:href) a atribut LOCTYPE

### 7.5.2 <fileSec> vedlejšího záznamu AMD\_METS.xml

- <fileSec> ve vedlejším METS záznamu AMD\_METS.xml bude obsahovat jeden element <fileGrp> s vnořenými elementy <file> pro každou reprezentaci stránky, tj. MC, ALTO XML a OCR.TXT
- atributy jednotlivých <file> elementů odpovídají atributům pro jednotlivé typy dokumentů uvedených výše pro <fileSec> hlavního METS záznamu
- navíc pro MC a ALTO XML (pro každou existující reprezentaci stránky s vlastním záznamem PREMIS Object nebo MIX) bude pro element <file> existovat atribut ADMID s ID těch <techMD> záznamů, které danou reprezentaci stránky popisují

## 7.6 METS část <structMap> - Strukturální metadata

### 7.6.1 <structMap> hlavního záznamu METS

- strukturální mapy v METS záznamu existují dvojího typu, fyzická a logická; fyzická zaznamenává hierarchické informace o dokumentu, včetně vazeb na fyzické soubory, ze kterých se skládají jednotlivé úrovně dokumentu
- 1 logická strukturální mapa v hlavním METS záznamu popisuje 1 svazek monografie a musí popisovat strukturu až na úroveň vnitřních částí (např. kapitol, nebo článků) apod.
  - součástí svazku monografie mohou být přílohy – pokud se skenují spolu se svazkem, popisuje strukturální mapa METS záznamu svazek včetně přílohy (bere se jako jeden svazek)
- strukturální mapa logická i fyzická včetně linků na ALTO XML bude v hlavním záznamu hlavni\_METS.xml
- pro každou stránku seskupuje METS logická strukturální mapa odkazy na textové bloky (nebo ilustrace), které jsou součástí té stránky. Informace o blocích textu nebo ilustracích na stránce jsou uloženy v 1 ALTO XML souboru, který stránce odpovídá. Každý blok a každá ilustrace má unikátní identifikátor, který je použit jako odkaz v METS strukturální mapě.

#### 7.6.1.1 Vyjádření fyzické strukturální mapy

- bude mít následující atributy <structMap LABEL="Physical\_Structure" TYPE="PHYSICAL">
- fyzická strukturální mapa obsahuje rodičovský <div>, který obsahuje tyto atributy:
  - LABEL- může obsahovat titul svazku monografie
  - TYPE – např. monograph
  - ID – identifikátor div
  - DMDID – identifikátor části popisných metadat náležející ke svazku monografie
- jednotlivé stránky jsou zanořeny do rodičovského elementu <div> jako dceřiné <div> elementy
  - <div> pro soubory stránky bude mít tyto atributy:
    - TYPE – bude se plnit typem stránky (viz nová pravidla popisu pro monografie)
    - ID – identifikátor div
    - ORDERLABEL – pořadové číslo stránky, jak je na ní vytištěno
    - ORDER – pořadí stránky ve svazku monografie
  - <div> pro soubory stránky vždy obsahují link <fptr> na soubor obrazu archivní kopie, uživatelské kopie, na ALTO XML, na OCR.TXT a na AMD\_METS.xml
    - link na obrazový soubor archivní kopie má v elementu <fptr> následující atributy: FILEID, který obsahuje ID souboru archivní kopie
    - link na obrazový soubor uživatelské kopie má v elementu <fptr> následující atributy: FILEID, který obsahuje ID souboru uživatelské kopie
    - link na ALTO XML má v elementu <fptr> následující atributy: FILEID, který obsahuje ID ALTO XML souboru
    - link na OCR.TXT soubor má v elementu <fptr> následující atributy: FILEID, který obsahuje ID souboru OCR.TXT
    - link na AMD\_METS.xml soubor má v elementu <fptr> následující atributy: FILEID, který obsahuje ID souboru AMD\_METS.xml

## 7.6.1.2 Vyjádření logické strukturální mapy

### 7.6.1.2.1 Vyjádření logické struktury pro kapitoly s vazbou na ALTO bloky

- bude mít následující atributy <structMap LABEL="Logical\_Structure" TYPE="LOGICAL">
- logická struktura na úroveň oddílů nebo např. ilustrací se popisuje pomocí do sebe zanořených elementů <div>
- pokud stránka obsahuje jen obraz a žádný text, pak je popsána jedním elementem <div> s atributem TYPE="PAGE" a link do souboru ALTO XML vede přímo na element <ComposedBlock>
  - <div TYPE="PAGE"> lze využít jako kontejner na obrazy a další části stránky, které nejsou součástí článku
  - pro obraz je možno využít atributy a typy podřízených elementů <div> jaké je specifikováno v tabulce níže pro PICTURE, který je součástí článku
- stránky obsahující více logických oblastí jsou popsány jedním <div> elementem, který má vnořené <div> elementy pro každou logickou oblast, která odpovídá např. textovému oddílu (např. kapitola, článek) nebo obraz.
  - pokud se jedná o jednoduchý, celistvý text na jedné straně, tak je popsán jen jedním <div> elementem s atributem TYPE="chapter"
    - v tomto <div> jsou dále jako další <div> elementy zanořeny jednotlivé textové bloky (odstavce, nadpisy, obrazy apod.)
      - u každého bloku je odkaz do ALTO XML souboru na příslušný textový blok <TextBlock> – pomocí tohoto odkazu se v ALTO XML souboru nalezne jak text, tak i informace o jeho umístění na stránce (souřadnice), toto je realizováno pomocí struktury <area> v elementu <fptr>
      - u bloku tvořeného obrazem je odkaz do ALTO XML na příslušný komponovaný blok <ComposedBlock>; je realizováno pomocí struktury <area> v elementu <fptr>
      - v případě použití atributu ORDER umožňuje tento princip u oddílů vyjádřit i tzv. pořadí čtení jeho částí, jako jsou např. nadpis, autor, obrázek apod.
  - výjimečně, pokud textový oddíl není celistvý a je rozdělen na více částí, které se vyskytují na jedné nebo více stránkách, které nemusejí jít za sebou, je možné určit pořadí čtení těchto částí, opět pomocí atributu ORDER
    - pro každou část oddílu existuje vlastní <div> element, podřízený hlavnímu <div> elementu oddílu
    - element <div> každé části má atribut TYPE hodnotu "chapter-part" a atribut ID musí vyjadřovat o jakou z částí se jedná, tj. např. ID="chapter5-1" odpovídá první části oddílu číslo pět
- **POZOR – u monografie se dělení oddílů běžně nepředpokládá (kapitoly jsou běžně na více stránkách, většinou po sobě jdoucích)**
  - tj. dělení oddílů není povinné a lze využít pouze struktury odstavců, jak je naznačeno v první části příkladu, tj. pokud kapitola/odstavec pokračuje na další straně, logická mapa uvádí, že poslední odstavec (NORMAL\_TEXT) např. na stránce 5 odkazuje na ALTO náležející ke stránce 5 a v něm na poslední textový blok; následující odstavec v logické mapě bude odkazovat na ALTO náležející ke stránce 6 a v něm na první textový blok (viz příklad)
- do logické struktury PSP balíčku může být v případě její existence zakomponována i příloha (Supplement), která má vlastní <div> element s atributem TYPE="SUPPLEMENT"

- vnořené <div> elementy pro obraz a textové oddíly i jejich použití je shodné se způsobem popisu logické struktury u elementu <div> s atributem TYPE="VOLUME"
- výčet stránek k jednotlivým kapitolám je obsažen v elementu <structLink> - popsáno v kapitole 7.7

<div> type	Atributy	Popis	Povinnost
MONOGRAPH	LABEL TYPE ID	<div> obsahuje údaje nadřazené svazku LABEL – název celé monografie TYPE - hodnota MONOGRAPH ID – identifikátor <div>, hodnota např. MONOGRAPH_0001 DMDID – obsahuje identifikátor popisné části MODS k nadřazenému svazku; využít pouze u vícesvazkové monografie	M
VOLUME nebo SUPPLEMENT	LABEL TYPE ID DMDID	<div> obsahuje údaje o svazku monografie nebo o jeho příloze ----- LABEL – název (titul) svazku monografie, tedy např. „Honzíkova cesta“ TYPE- hodnota VOLUME nebo SUPPLEMENT ID – identifikátor <div>, např. hodnota "VOLUME_0001" nebo "SUPPL_0001" DMDID – obsahuje identifikátor DMD popisné části MODS svazku/přílohy	M
CHAPTER	LABEL TYPE ID DMDID ORDER	<div> obsahující údaje o jednom textovém oddílu a jeho částech ----- LABEL – název textového oddílu (např. kapitola, článek ve sborníku apod.) TYPE – hodnota CHAPTER s pořadovým číslem, např. CHAPTER_0001 ID – identifikátor <div> elementu DMDID – identifikátor popisných metadat ORDER – pořadí oddílu	M
		<div> TYPE="CHAPTER" může obsahovat další vnořené <div> různých typů popisující různé části textového oddílu, rozlišujeme tyto části (typy): <ul style="list-style-type: none"> <li>- TITLE</li> <li>- SUBTITLE</li> <li>- AUTHOR</li> <li>- TRANSLATOR</li> <li>- NORMAL_TEXT – běžný text bez dalšího upřesnění</li> <li>- PICTURE</li> <li>- NOTE</li> <li>- CHAPTER_PART - u oddílů, které jsou rozděleny na více míst na jedné stránce nebo více stránkách (v případě článků ve sborníku např.)</li> </ul> tento <div> pro jednu součást rozděleného článku pak může obsahovat stejné části jako <div> pro oddíl, tj. (TITLE, SUBTITLE, AUTHOR, TRANSLATOR, NORMAL_TEXT,	

		PICTURE)	
TITLE	TYPE ID ORDER	<div> obsahující link na textový blok s nadpisem oddílu (tedy např. kapitoly) ----- TYPE – hodnota "TITLE" ID – identifikátor <div> elementu, který popisuje jednu část oddílu (nadpis), např. hodnota "CHAPTER_PART_0001" ORDER – pořadí části oddílu	MA
<fptr> <area>	FILEID BEGIN BETYPE	FILEID – ID ALTO XML souboru, např. "ALTO_PAGE_0001" BEGIN – ID textového bloku v ALTO XML souboru BETYPE – hodnota "IDREF"	
SUBTITLE	TYPE ID ORDER	<div> obsahující link na textový blok s podnadpisem ----- TYPE – hodnota "SUBTITLE" ID – identifikátor <div> elementu, který popisuje jednu část oddílu (podnadpis), např. hodnota "CHAPTER_PART_0002" ORDER – pořadí části oddílu	MA
<fptr> <area>	FILEID BEGIN BETYPE	FILEID – ID ALTO XML souboru, např. "ALTO_PAGE_0001" BEGIN – ID textového bloku v ALTO XML souboru BETYPE – hodnota IDREF	
AUTHOR	TYPE ID ORDER	<div> obsahující link na textový blok se jménem autora ----- TYPE – hodnota „AUTHOR“ ID – identifikátor <div> elementu, který popisuje jednu část oddílu (autor), např. hodnota "CHAPTER_PART_0003" ORDER – pořadí části oddílu	MA
<fptr> <area>	FILEID BEGIN BETYPE	FILEID – ID ALTO XML souboru, např. "ALTO_PAGE_0001" BEGIN – ID textového bloku v ALTO XML souboru BETYPE – hodnota IDREF	

TRANSLATOR	TYPE ID ORDER	<div> obsahující link na textový blok se jménem překladatele ----- TYPE – hodnota „TRANSLATOR“ ID – identifikátor <div> elementu, který popisuje jednu část oddílu (překladatel), např. hodnota "CHAPTER_PART_0003" ORDER – pořadí části oddílu	MA
<fptr> <area>	FILEID BEGIN BETYPE	FILEID – ID ALTO XML souboru, např. „ALTO_PAGE_0001“ BEGIN – ID textového bloku v ALTO XML souboru BETYPE – hodnota IDREF	
NORMAL_TEXT	TYPE ID ORDER	<div> obsahující link na textový blok (nejčastěji odstavec) s běžným textem ----- TYPE – hodnota "NORMAL_TEXT" ID – identifikátor <div> elementu, který popisuje jednu část oddílu (běžný text), např. hodnota "CHAPTER_PART_0004" ORDER – pořadí části oddílu	M
<fptr> <area>	FILEID BEGIN BETYPE	FILEID – ID ALTO XML souboru, např. "ALTO_PAGE_0001" BEGIN – ID textového bloku v ALTO XML souboru BETYPE – hodnota IDREF	
PICTURE	LABEL TYPE ID DMDID ORDER	<div> pro obraz náležející k textovému oddílu; plní se, pokud se obraz vyskytuje ----- LABEL – název obrazu pokud existuje TYPE - PICTURE ID – identifikátor <div> elementu, který popisuje jednu část oddílu (běžný text), např. hodnota "CHAPTER_PART_0003" DMDID – link na bibliogr. popis obrazu ORDER – pořadí obrazu	MA
		<div> element s typem PICTURE může obsahovat další <div> elementy s typy CAPTION, PICT_AUTHOR, PICT_TITLE a IMAGE; - CAPTION obsahuje text případného popisku k obrazu - PICT_AUTHOR obsahuje text se jménem případného autora obrazu - PICT_TITLE obsahuje text názvu obrazu, pokud nějaký název existuje - IMAGE – obsahuje link do souboru ALTO XML na blok popisující vlastní obraz	
CAPTION	TYPE ID	<div> obsahující link na textový blok s popisem obrazu ----- TYPE – hodnota CAPTION ID – identifikátor <div> elementu, např. "CHAPTER_PART_4"	MA
<fptr> <area>	FILEID BEGIN BETYPE	FILEID – ID ALTO XML souboru BEGIN – ID textového bloku v ALTO XML souboru BETYPE – hodnota IDREF	
PICT_AUTHOR	TYPE ID	<div> obsahující link na textový blok s autorem obrazu ----- TYPE – hodnota PIT_AUTHOR	MA

		ID – identifikátor <div> elementu, např. "CHAPTER_PART_5"	
<fptr> <area>	FILEID BEGIN BETYPE	FILEID – ID ALTO XML souboru BEGIN – ID textového bloku v ALTO XML souboru BETYPE – hodnota IDREF	
PICT_TITLE	TYPE ID	<div> obsahující link na textový blok s názvem obrazu ----- TYPE – hodnota PICT_TITLE ID – identifikátor <div> elementu, např. "CHAPTER_PART_6"	MA
<fptr> <area>	FILEID BEGIN BETYPE	FILEID – ID ALTO XML souboru BEGIN – ID textového bloku v ALTO XML souboru BETYPE – hodnota IDREF	
IMAGE	TYPE ID	<div> obsahující link na komponovaný blok ALTO XML obsahující souřadnice vlastního obrazu ----- TYPE – hodnota IMAGE ID – identifikátor <div> elementu, např. "CHAPTER_PART_7"	MA
<fptr> <area>	FILEID BEGIN BETYPE	FILEID – ID ALTO XML souboru BEGIN – ID komponovaného bloku v ALTO XML souboru BETYPE – hodnota IDREF	
NOTE	ID	<div> obsahující link na textový blok s poznámkami k textu ----- ID – identifikátor <div> elementu, např. "CHAPTER_PART_9"	
CHAPTER_PART	TYPE ID ORDER	<div> obsahující další vnořené <div> odkazující na jednotlivé části konkrétní části rozděleného textového oddílu; možnost použít pro dělený oddíl (typu články např. ve sborníku) Pozn: pod <div> TYPE="CHAPTER_PART" lze vnořit všechny typy <div> jako pod <div> TYPE="CHAPTER" ----- TYPE – hodnota „CHAPTER_PART“ ID – identifikátor <div> konkrétní části, pro první část děleného oddílu např. "CHAPTER_2-1", tj. první část oddílu 2 ORDER – pořadí konkrétní části děleného oddílu	MA

#### 7.6.1.2.2 Vyjádření logické struktury pro kapitoly bez vazby na ALTO bloky (pouze s vazbami na strany)

- bude mít následující atributy <structMap LABEL="Logical\_Structure" TYPE="LOGICAL">
- logická struktura na úroveň kapitol nebo např. ilustrací se popisuje pomocí do sebe zanořených elementů <div>, ale dále už neobsahuje elementy <fptr> a <area>
- existuje tedy jen výčet popsaných kapitol a samostatných obrázků (těch, které nejsou součástí kapitoly) jako v případě předchozí kapitoly (7.6.1.2.1), ale bez vazeb na ALTO bloky
- výčet stránek k jednotlivým článkům je obsažen v elementu <structLink> - popsáno v kapitole 7.7



- 

### 7.6.1.2.3 ***Vyjádření logické struktury bez popisu kapitol***

- bude mít následující atributy <structMap LABEL="Logical\_Structure" TYPE="LOGICAL">
- logická struktura na úroveň čísla a případně titulu u vícesvazků se vyjadřuje pomocí zanořených elementů <div>
- <div> musí obsahovat identifikátor na příslušný blok metadat, př. odkaz na úroveň volume DMDID="MODSMD\_VOLUME\_0001", u vícesvazků pak musí obsahovat odkaz na úroveň titulu DMDID="MODSMD\_TITLE\_0001"
- výčet stránek k číslu a příloze je obsažen v elementu <structLink> - popsáno v kapitole 7.7

## 7.6.2 <structMap> vedlejšího záznamu METS (AMD\_METS.xml)

- bude obsahovat pouze fyzickou strukturální mapu (TYPE="PHYSICAL")
- ta bude obsahovat pouze jeden <div> element s atributem TYPE="MONOGRAPH\_PAGE"
- do <div> budou vnořeny odkazy na jednotlivé reprezentace stránky svazku (MC, ALTO XML a OCR.TXT) pomocí elementu <fptr> s atributem FILEID

## 7.7 METS část <structLink> - Výčet stran

- element <structLink> obsahuje výčet stran jednotlivých úrovní monografie na základě přidání vazeb mezi logickou a fyzickou strukturální mapou
- element <structLink> obsahuje subelement <smLink>, který obsahuje atributy "xlink:from" a "xlink:to"
  - "xlink:from" obsahuje ID divu z logické strukturální mapy
  - "xlink:to" obsahuje ID divu stránky z fyzické strukturální mapy

### 7.7.1 Výčet stran v případě popisu vnitřních částí (s odkazy na bloky ALTA i bez nich)

- vztahuje se ke kapitolám 7.6.1.2.1 a 7.6.1.2.2
- element <structLink> obsahuje:
  - výčet stran jednotlivých kapitol a případných samostatných obrázků (ID divu z logické strukturální mapy pro úroveň kapitol a obrázků)
  - výčet stran celého titulu a případné přílohy (ID divu z logické strukturální mapy pro úroveň titulu a přílohy)

### 7.7.2 Výčet stran bez popisu vnitřních částí

- vztahuje se ke kapitole 7.6.1.2.3
- element <structLink> obsahuje:
  - výčet stran monografie a případné přílohy (ID divu z logické strukturální mapy pro úroveň monografie a přílohy)

## 7.8 OCR (ALTO XML a TXT OCR)

- bude použita poslední verze formátu ALTO XML aktuální v době implementace, nebo verze předchozí (prosinec 2010 verze 2 – viz <http://www.loc.gov/standards/alto/>)
- níže uvedená specifikace **neobsahuje všechny elementy a atributy formátu ALTO XML, obsahuje pouze ty, které jsou pro tuto konkrétní specifikaci relevantní – každý uvedený element má vyjádřenou míru relevance výrazy: povinné, doporučené a nepovinné**
- elementy a atributy, které v této specifikaci nejsou uvedeny, nepovažujeme pro účely specifikace za důležité
- ALTO XML i OCR TXT vzniknou pro všechny obrazové soubory náležející k jedné intelektuální entitě (svazku monografie) včetně prázdných stran, fotografií hřbetu, předsádky apod.
- ALTO XML i OCR TXT budou vznikat na úrovni stránky
- ALTO XML soubor pro zcela prázdné stránky bude obsahovat element `/alto/Layout/Page/PrintSpace`, ten ovšem nebude obsahovat podelementy:  
`/alto/Layout/Page/PrintSpace/TextBlock`; `/alto/Layout/Page/PrintSpace/TextBlock/Illustration`;  
`/alto/Layout/Page/PrintSpace/TextBlock/GraphicalElement` ani  
`/alto/Layout/Page/PrintSpace/TextBlock/ComposedBlock`
- struktura ALTO XML bude generovaná na úrovni rozpoznání slova generovaná OCR
- kvalita rozpoznání znaků bude akceptována do určité hranice, výstupy nebudou ručně opravovány
- struktura ALTO umožní vyhledávání textu a jeho zvýraznění na úrovni slova, pokud bude použit odpovídající prohlížeč
- obrazy reprezentující stránku, které budou použity jako UC, musí odpovídat rozměry, orientací a natočením obrazu, který byl použit pro vytvoření OCR
- OCR TXT bude vznikat z hotových ALTO XML během procesu digitalizace
- ALTO XML se bude vytvářet pouze pro novodobé dokumenty, nebo dokumenty s určitou hranicí kvality OCR
- jméno OCR souboru musí odpovídat jménu obrazového souboru, ke kterému náleží; např. `pr_0007.jp2` a `al_0007.xml` nebo např. `123456_006_alto.xml` a `123456_006_archiv.jp2`
- kódování ALTO XML i TXT OCR musí být v UTF-8
- souřadnice pozic (HPOS, VPOS, WIDTH, HEIGHT) musí být vyjádřeny v pixelech
- v této specifikaci ALTO XML se počítá s OCR i pro text mimo tzv. textové "zrcadlo", tj. mimo hlavní text, jako jsou např. čísla stránek, běžící nadpisy a jiné části vyskytující se na okrajích stránky (top, left, top a bottom margin)
  - elementy `topMargin`, `leftMargin`, `rightMargin`, `bottomMargin` budou obsahovat elementy `<TextBlock>`, pro které platí stejná pravidla, jako pro element `<textBlock>` pro hlavní text stránky
  - pozor: údaje z OCR mimo hlavní text stránky by neměly být vyhledávatelné v aplikaci zpřístupnění, docházelo by ke zmatení uživatele a výsledků (např. při hledání titulu kapitoly by byly zobrazeny výsledky pro každou stránku, která obsahuje běžící nadpis apod.)
- pokud je na konci věty dělicí znaménko, ALTO XML i OCR TXT musí obsahovat oba fragmenty slova s dělítkem a současně také kompletní slovo – je vysvětleno dále v tabulce

- ilustrace, reklamy a jiné grafické části stránky nebudou vyjádřeny v tazích `/alto/Layout/Page/PrintSpace/Illustration` ani `Layout/Page/PrintSpace/GraphicalElement`, tyto nejsou v popisu/tabulce níže vůbec uvedeny
- ilustrace, reklamy a jiné grafické části stránky budou vyjádřeny v tagu `/alto/Layout/Page/PrintSpace/ComposedBlock/` s vyjádřením atributu `TYPE`, který bude označovat typ bloku (illustration, advertisement aj.)
  - např. ilustrace bude popsána v elementu `/alto/Layout/Page/PrintSpace/ComposedBlock/GraphicalElement`, kde `ComposedBlock TYPE` je `Illustration`
  - reklama s textem v rámečku bude popsána v elementu `Layout/Page/PrintSpace/ComposedBlock/TextBlock`, kde `ComposedBlock TYPE` je `Advertisement`
  - tabulky, grafy obdobně
- elementy `/alto/Layout/Page/PrintSpace/ComposedBlock/Illustration` a `Layout/Page/PrintSpace/ComposedBlock/ComposedBlock` také nebudou využity
- `/alto/Layout/Page/PrintSpace/ComposedBlock/TextBlock` a `/alto/Layout/Page/PrintSpace/ComposedBlock/GraphicalElement` nebudou obsahovat elementy `<Shape>`; tvar těchto bloků je vyjádřen v elementu `<Shape>` samotného elementu `<ComposedBlock>`; logicky pak souřadnice tvaru `<TextBlock>` nebo `<GraphicalElement>` obsaženého v `/alto/Layout/Page/PrintSpace/ComposedBlock` jsou většinou shodné, pokud není tvarů nebo bloků v rámci `/alto/Layout/Page/PrintSpace/ComposedBlock` více
- všechny vyplněné hodnoty jsou příklady plnění, plnění v konkrétní instituci je nutno specifikovat vlastními pravidly a kontrolovanými slovníky
- ALTO XML bude využíváno pro tzv. pořadí čtení, tj. např. článek vyskytující se na více stránkách nebo na více různých místech jedné stránky bude možné zobrazit celý a ve správném pořadí. K tomu je nutno znát jeho strukturu. Struktura bude vyjádřena v korespondujícím METS záznamu v logické strukturální mapě. Ta bude obsahovat odkazy na jednotlivé textové bloky článku, pomocí ID textových bloků použitých v ALTO XML.

#### Obsah pole "Popis":

- vysvětlení a příklad
- doporučené plnění tam, kde je to možné
- výskyt elementu (jak je definováno formátem ALTO XML – dle XSD)
  - 0-1 element je nepovinný, neopakovatelný
  - 0-n element je nepovinný, opakovatelný

Element	Atribut	Popis	Povinnost
<Description>			
<MeasurementUnit>		měřicí jednotka pro souřadnice v ALTO XML; možné hodnoty – dpi, pixel, inch1200 a mm10); inch1200 = 1/1200 inche; doporučené plnění je „mm10“ nebo „pixel“; 0-1	M
<sourceImageInformation>		informace o obrazovém souboru, ze kterého vzniklo ALTO XML; 0-1	M
<fileName>		jméno obrazového souboru, ze kterého bylo ALTO XML vytvářeno; ideálně i s filesystem cestou jeho uložení; např. n1alimageSeq-33386-b.tif//produkce/OCR/digibok_XY/XY_011.tif 0-1	M
<fileIdentifier>		jedinečný identifikátor obrazového souboru; 0-n	R
<OCRProcessing>	ID	popis procesu vzniku OCR; 0-n ----- ID OCR procesu, např. <OCRProcessing ID="OCRPROCES_1">; povinné	M
<preProcessingStep>		procesy před vznikem OCR, které provádí SW pro OCR (např. natočení obrazu) 0-n	M
<processingDateTime>		určení času procesu, který předcházela samotnému OCR; např. 2008-03-29T19:42:23 dle ISO 8601 na úroveň vteřin; 0-1	O
<processingAgency>		jméno nebo kód instituce, např. NK CZ, název externí firmy apod.; doporučujeme použít kontrolovaný slovník hodnot; 0-1	R
<processingStepDescription>		popis procesu (např. zarovnání, ořez apod.); 0-n	O
<processingStepSettings>		nastavení kroku popsaného v <processingStepDescription>, např. CCS OCR Processing Filter 0-1	O

<processingSoftware>		popis SW, který upravoval obrázek před vznikem OCR; 0-1	M
<softwareCreator>		výrobce softwaru - např. CCS Content Conversion Specialists GmbH, Germany; 0-1	M
<softwareName>		jméno softwaru - např. CCS docWORKS; 0-1	M
<softwareVersion>		verze SW, např. 6.2-1.16; 0-1	M
<ocrProcessingStep>		popis procesu vzniku OCR 1-1 – povinné pole	M
<processingDateTime>		okamžik kdy bylo OCR vytvořeno; nutno zapsat v ISO 8601 na úroveň vteřin; 0-1	M
<processingAgency>		jméno nebo kód instituce, např. NK CZ doporučujeme použít kontrolovaný slovník hodnot; 0-1	M
<processingSoftware>		popis SW, který dělal vlastní OCR; 0-1	M
<softwareCreator>		výrobce softwaru - např. ABBYY, Russia; 0-1	M
<softwareName>		jméno softwaru - např. FineReader; 0-1	M
<softwareVersion>		např. 8.0; 0-1	M
<Styles>		styly definují vlastnosti jednotlivých grafických prvků stránky. styl definovaný v elementu vrchní úrovně je použit jako výchozí pro podřízené elementy; 0-1	MA

<TextStyle>	ID FONTSTYLE FONTFAMILY FONTSIZE	definuje font textu; 0-n ----- ID pro každý text style použitý v OCR souboru – povinné  FONTSTYLE – např. bold, italics apod.; doporučujeme používat kontrolovaný slovník; doporučené  FONTFAMILY – např. arial, calibri apod.; doporučujeme používat kontrolovaný slovník; povinné  FONTSIZE – velikost fontu, např. 10, 12 apod.; povinné	MA
<ParagraphStyle>	ID ALIGN	definuje formátování textových bloků; 0-n ----- ID pro každý odstavec + zarovnání; např. PAR_01, PAR_02 apod. povinné  ALIGN – zarovnání; povolené hodnoty: Left, Right, Center, Block aj.; povinné	MA
<Layout>		layout - rozložení struktur (slov, odstavců apod.) na jedné stránce dokumentu; 1-1 povinný výskyt element není opakovatelný	M

<Page>	ID ACCURACY POSITION QUALITY PHYSICAL_IMG_ NR HEIGHT WIDTH PC	element popisující jednu stránku dokumentu; 1-n ----- ID – vygenerovaný identifikátor stránky, např. PAGE1, nebo P1 apod.; povinné  ACCURACY – procentuální odhad přesnosti OCR (0-100); doporučené  POSITION – pozice stránky; hodnoty k plnění: Left, Right, Foldout, Single, Cover; nepovinné  QUALITY – krátký údaj o kvalitě předlohy stránky; hodnoty k plnění: OK, Missing, Missing in original, Damaged, Retained, Target, As in original; nepovinné  PHYSICAL_IMG_NR - fyzické (pořadové) číslo stránky v dokumentu; vyjádřeno číslem, např. 1,2,3 apod.; povinné  WIDTH – šířka stránky vyjádřená v pixelech; povinné  HEIGHT – výška stránky vyjádřená v pixelech; povinné  PC - Confidence level OCR souboru – hodnota mezi 0 (nejistá kvalita) a 1 (dobrá kvalita); nepovinné; pokud nevyplníte ACCURACY – tak je vyplnění doporučené	M
<TopMargin>	ID HPOS VPOS WIDTH HEIGHT	horní okraj – prostor mezi vrchní hranou listu a vrchní linkou textu; 0-1 ----- ID: unikátní ID pro element TopMargin, např. P1_TM0001 (page 1, topMargin0001); povinné  HPOS: horizontální pozice; povinné  VPOS: vertikální pozice; povinné  WIDTH – šířka vrchního okraje; povinné  HEIGHT – výška vrchního okraje; povinné	M



<TextBlock>		stejné plnění a pravidla jako pro element <TextBlock> vnořený do elementu <PrintSpace>	MA
<LeftMargin>	ID HPOS VPOS WIDTH HEIGHT	levý okraj – prostor mezi levým okrajem stránky a textem; 0-1 ----- ID: unikátní ID pro element LeftMargin, např. P1_LM0001 (page 1, leftMargin0001); povinné  HPOS: horizontální pozice; povinné  VPOS: vertikální pozice; povinné  WIDTH – šířka levého okraje; povinné  HEIGHT – výška levého okraje; povinné	M
<TextBlock>		stejné plnění a pravidla jako pro element <TextBlock> vnořený do elementu <PrintSpace>	MA
<RightMargin>	ID HPOS VPOS WIDTH HEIGHT	pravý okraj – prostor mezi pravým okrajem stránky a textem; 0-1 ----- ID: unikátní ID pro element RightMargin, např. P1_RM0001 (page 1, rightMargin0001); povinné  HPOS: horizontální pozice; povinné  VPOS: vertikální pozice; povinné  WIDTH – šířka pravého okraje; povinné  HEIGHT – výška pravého okraje; povinné	M

<TextBlock>		stejné plnění a pravidla jako pro element <TextBlock> vnořený do elementu <PrintSpace>	MA
<BottomMargin>	ID HPOS VPOS WIDTH HEIGHT	<p>pravý okraj – prostor mezi spodním okrajem stránky a textem; 0-1 ----- ID: unikátní ID pro element BottomMargin, např. P1_BM0001 (page 1, bottomMargin0001); povinné</p> <p>HPOS: horizontální pozice; povinné</p> <p>VPOS: vertikální pozice; povinné</p> <p>WIDTH – šířka spodního okraje; povinné</p> <p>HEIGHT – výška spodního okraje; povinné</p>	M
<TextBlock>		stejné plnění a pravidla jako pro element <TextBlock> vnořený do elementu <PrintSpace>	MA
<PrintSpace>	ID HPOS VPOS WIDTH HEIGHT	<p>popis tvaru pokrývajícího textové pole stránky; 0-1 ----- ID: unikátní ID pro element &lt;printSpace&gt;, např. P1_PS0001 (page 1, printSpace0001); - povinné</p> <p>HPOS: horizontální pozice; povinné</p> <p>VPOS: vertikální pozice; povinné</p> <p>WIDTH – šířka textového pole; povinné</p> <p>HEIGHT – výška textového pole; povinné</p>	M

<TextBlock>	ID STYLEREFS HPOS VPOS WIDTH HEIGHT	<p>popisy textových bloků na konkrétní stránce; 0-n pokud je stránka prázdná, TextBlock není potřeba uvádět; pokud je na stránce text tak ano</p> <p>-----</p> <p>ID obsahuje identifikátor textového bloku na stránce, např. "BLOCK1" nebo P1_TB0002 (stránka 1, textový blok 2); povinné</p> <p>STYLEREFS: reference na ID definice formátování textových bloků &lt;ParagraphStyle&gt;; povinné</p> <p>HPOS: horizontální pozice bloku; povinné</p> <p>VPOS: vertikální pozice bloku; povinné</p> <p>WIDTH – šířka textového bloku; povinné</p> <p>HEIGHT – výška textového bloku; povinné</p>	MA
<Shape>		<p>tvar textového bloku; 0-1 – pro jeden výskyt &lt;TextBlock&gt; jeden nebo žádný výskyt &lt;Shape&gt;; plnit v případě, že je tvar textového bloku nestandardní (víceúhelník)</p>	RA
<Polygon>	POINTS	<p>popis (souřadnice) tvaru víceúhelníku; 0-1</p> <p>-----</p> <p>POINTS – vyjádření jednotlivých bodů víceúhelníku; povinné</p>	M

<TextLine>	<p>ID          STYLEREFS          HPOS          VPOS          WIDTH          HEIGHT</p>	<p>popis jedné řádky textu v rámci textového bloku;          1-n          nutný alespoň jeden výskyt v rámci textového bloku          -----          ID obsahuje identifikátor řádky textu v textovém bloku,          např. "P1_TL0002 (stránka 1, řádka 2);          povinné</p> <p>STYLEREFS: reference na ID definice formátování          textových bloků &lt;ParagraphStyle&gt;;          nepovinné</p> <p>HPOS: horizontální pozice řádky;          povinné</p> <p>VPOS: vertikální pozice řádky;          povinné</p> <p>WIDTH – šířka řádky;          povinné</p> <p>HEIGHT – výška řádky;          povinné</p>	M
<String>	<p>ID          CONTENT          HEIGHT          WIDTH          HPOS          VPOS          CC          WC</p> <p>V případě dělení          slov také:          SUBS_TYPE          SUBS-          CONTENT</p>	<p>řetězec znaků – vlastní obsah OCR;          znaky tvoří jednotlivá slova a více tagů &lt;String&gt; větu          &lt;TextLine&gt;;          1-n v rámci &lt;TextLine&gt;          -----          ID obsahuje unikátní sekvenční číslo řetězce na stránce,          např. "P3_ST0001" (strana 3, řetězec 1);          povinné</p> <p>CONTENT – ukládá vlastní řetězec znaků (slovo);          povinné</p> <p>HPOS: horizontální pozice řetězce;          povinné</p> <p>VPOS: vertikální pozice řetězce;          povinné</p> <p>WIDTH – šířka řetězce;          povinné</p> <p>HEIGHT – výška řetězce;          povinné</p> <p>CC – úroveň důvěry v přesnost OCR rozpoznání každého          znaku v řetězci;          jde o seznam čísel, každé z nich mezi hodnotami 0 (jistá) a</p>	M

		<p>9 (nejistá) pro každý znak; např. CC="0001" pro CONTENT="TEXT"; nepovinné</p> <p>WC – úroveň důvěry v přesnost OCR výstupu celého řetězce - slova (word confidence); hodnota mezi 0 (nejistá) a 1 (jistá); např. WC="0,99"; nepovinné</p> <p>SUBS_CONTENT – obsah chybějící části řetězce v případě, že je slovo na konci řádku rozdělené i do druhého řádku; obsahuje celý řetězec - aby byl vyhledatelný i v případě, že slovo se na stránce vyskytuje, ale je rozděleno; povinné</p> <p>SUBS_TYPE – označení typu substituce; možné hodnoty: HypPart1; HypPart2; Abbreviation; povinné - při výskytu SUBS_CONTENT</p> <p>HypPart1 se vyskytuje při rozdělení slova u jeho první OCR části (u první části tagu &lt;CONTENT&gt; ve větě (stringu) první; HypPart2 se vyskytuje u následujícího tagu &lt;CONTENT&gt; v následující větě (stringu), který obsahuje druhou část rozděleného slova/řetězce; Abbreviation – typ substituce používaný při rozepisování zkratk v textu na jejich plný text; při dělení slov v textu HypPart1 a HypPart2 povinné, abbreviation nepovinné</p>	
<ALTERNATIVE>		<p>alternativní hodnota OCR řetězce pro jednotlivá slova; 0-n lze použít v případě nejistoty rozpoznání řetězce;</p>	O

<HYP>	CONTENT WIDTH HPOS VPOS	<p>zápis znaku rozdělovníku slov 0-1 pro jeden výskyt &lt;TextLine&gt;; vždy pro poslední &lt;String&gt;; může se vyskytnout pouze na konci řádku (1x)</p> <p>-----</p> <p>CONTENT – obsahuje řetězec znaků, které jsou v textu použity na rozdělení slova, nejčastěji „-“; povinné</p> <p>WIDTH – šířka dělicího znaku; doporučené</p> <p>HPOS: horizontální pozice dělicího znaku; doporučené</p> <p>VPOS: vertikální pozice dělicího znaku; doporučené</p>	MA
<SP>	ID WIDTH HPOS VPOS	<p>prázdný prostor mezi řádky; 0-n v rámci jednoho &lt;TextLine&gt;; vždy mezi řádky, tj. mezi tagy &lt;String&gt;;</p> <p>-----</p> <p>ID: unikátní ID pro prázdný prostor mezi řádky, např. P1_SP0001 (stránka 1, prázdný prostor 0001); povinné</p> <p>HPOS: horizontální pozice; povinné</p> <p>VPOS: vertikální pozice; povinné</p> <p>WIDTH – šířka prázdného prostoru; povinné</p>	M

<ComposedBlock>	ID TYPE HPOS VPOS WIDTH HEIGHT STYLEREFS	blok sestávající z jiných bloků; může obsahovat PrintSpace/ComposedBlock/TextBlock, PrintSpace/ComposedBlock/Illustration, PrintSpace/ComposedBlock/GraphicalElement, /PrintSpace/ComposedBlock/ComposedBlock, tj. stejné elementy (bloky), které obsahuje samotný element /alto/Layout/Page/PrintSpace; 0-n povinné pro vyjádření bloků textu (např. orámovaný text, reklamy), pro vyjádření ilustrací, tabulek a grafik ----- ID: unikátní ID komponovaný blok, např. P6_CB0001 (stránka 6, komponovaný blok 0001); povinné  TYPE – označení typu komponovaného bloku; nutné používat kontrolovaný slovník (illustration, Advertisement, apod.); povinné  HPOS: horizontální pozice bloku; povinné  VPOS: vertikální pozice bloku; povinné  WIDTH – šířka komponovaného bloku; povinné  HEIGHT – výška komponovaného bloku; povinné	MA
<Shape>		tvar komponovaného bloku; 0-1 – pro jeden výskyt /alto/Layout/Page/PrintSpace/ComposedBlock jeden nebo žádný výskyt /alto/Layout/Page/PrintSpace/ComposedBlock/Shape; doporučeno – v případě, že je tvar komponovaného bloku nestandardní (víceúhelník)	RA
<Polygon>	POINTS	popis tvaru víceúhelníku; 0-1 ----- POINTS – vyjádření jednotlivých bodů víceúhelníku povinné	M

<TextBlock>	ID STYLEREFS HPOS VPOS WIDTH HEIGHT	<p>v případě, že komponovaný blok (např. orámovaný tvar) obsahuje text;          platí stejná pravidla jako pro normální element          /alto/Layout/Page/PrintSpace/TextBlock;          0-n (pro jeden výskyt &lt;ComposedBlock&gt; 0 nebo více elementů          /alto/Layout/Page/PrintSpace/ComposedBlock/TextBlock&gt;;          plnit pokud je v komponovaném bloku text</p> <p>-----</p> <p>ID obsahuje identifikátor textového bloku v komponovaném bloku, např. P1_CB0002_SUB (stránka 1, textový blok 2, SUB značí komponovaný blok);          povinné</p> <p>STYLEREFS: reference na ID definice formátování textových bloků /alto/Styles/ParagraphStyle;          povinné</p> <p>HPOS: horizontální pozice bloku;          povinné</p> <p>VPOS: vertikální pozice bloku;          povinné</p> <p>WIDTH – šířka textového bloku;          povinné</p> <p>HEIGHT – výška textového bloku;          povinné</p>	MA
<TextLine>		/alto/Layout/Page/PrintSpace/ComposedBlock/TextBlock/TextLine a ostatní elementy v rámci /alto/Layout/Page/PrintSpace/ComposedBlock/TextBlock mají stejná pravidla a výskyty jako jako ve vrchním elementu /alto/Layout/Page/PrintSpace/TextBlock	



<GraphicalElement>	ID HPOS VPOS WIDTH HEIGHT	<p>popis grafického tvaru; v případě využití v rámci /alto/Layout/Page/PrintSpace/ComposedBlock označuje rozměry tvaru v rámci něhož je tabulka, ilustrace, reklama apod.;</p> <p>0-1 - pro jeden výskyt /alto/Layout/Page/PrintSpace/ComposedBlock 0 nebo max. 1 výskyt &lt;GraphicalElement&gt;; plní se, pokud je na stránce a tedy v komponovaném bloku ilustrace, tabulka apod.;</p> <p>-----</p> <p>ID – identifikátor grafického tvaru; povinné</p> <p>HEIGHT – výška grafického tvaru; povinné</p> <p>WIDTH – šířka grafického tvaru; povinné</p> <p>HPOS – horizontální pozice grafického tvaru; povinné</p> <p>VPOS – vertikální pozice grafického tvaru; povinné</p>	MA
--------------------	---------------------------------------	---	----

## 8 Autorsko-právní metadata

- Autorsko-právní metadata jsou kompletně nepovinná, z důvodu zachování kompatibility s předchozí verzí specifikace.
- V případě rozhodnutí o vytvoření autorsko-právních metadat platí povinnost elementů uvedených ve sloupci povinnost.
- V případě nevytvoření autorsko-právních metadat je status defaultně považován za neznámý (*unknown*).
- Autorsko-právní metadata mohou být vyplněna jak pro celý dokument (hlavní záznam METS) tak i pro konkrétní části dokumentu (vedlejší záznam AMD\_METS). Platí zde pravidlo hierarchické dědičnosti: v případě, že má vedlejší záznam vyplněné autorsko-právní metadata, tak jsou relevantní právě ta. V případě, že je vedlejší záznam neobsahuje, tak je přebírá z hlavní záznamu METS. Naopak to neplatí. Hlavní záznam nepřebírá autorsko-právní metadata z vedlejších záznamů.

### Opakovatelnost elementů:

0-1 nepovinný a neopakovatelný

0-N nepovinný a opakovatelný

### Část METS pro autorsko-právní metadata

element	atributy	popis	Povinnost
<amdSec>		element obsahující autorsko-právní metadata ve formátu ContainerMD	M
	ID	Pouze v případě použití u vedlejšího záznamu AMD_METS  ID – identifikátor konkrétní části <amdSec>, např. pro stránku 1 hodnota "PAGE0001", pro stránku 2 "PAGE0002" atd.	
<rightsMD>		element pro typ autorsko-právních metadat	M
	ID	<b>ID pro část &lt;rightsMD&gt;: např. "RIGHTS_VOLUME"</b>	
<mdWrap>	MDTYPE	MDTYPE – hodnota "OTHER"	
	MIMETYPE	MIMETYPE – hodnota "text/xml"	

**Nepovinné atributy:**

U některých elementů mohou být použity nepovinné atributy, které upřesní informaci v elementu.

1. *iso.code*: elementy: <country.publication>, <country.creation> - do atributu se zapisuje standardizovaný kód země podle ISO3166-1 (alpha 2-letter code list<sup>5</sup>) (viz příklad)
2. *year.type*: elementy: <year.copyright> <year.renewal> <year.publication> <year.creation> <year.birth> <year.death> - do atributu se zapisuje přesnost uvedeného roku. Možné hodnoty *exact* (rok je známý), *approximate* (odhadovaný rok na základě jiných zdrojů), *unknown* (rok není znám a není možné jej odhadnout/dohledat)

Element	Popis	Povinnost
<copyright>	<p>Povinný kořenový element. Označuje status dokumentu. <b>Povinné jsou i oba atributy:</b></p> <p>copyright.status</p> <ul style="list-style-type: none"> <li>• Copyrighted (autorsky chráněný document)</li> <li>• pd (volně dostupný dokument)</li> <li>• pd_expired (volně dostupný dokument, kterému vypršelo trvání majetkových práv)</li> <li>• unknown (status neznámý)</li> </ul> <p>publication.status</p> <ul style="list-style-type: none"> <li>• published (publikovaný dokument)</li> <li>• unpublished (nepublikovaný dokument)</li> <li>• unknown (informace o publikování není známa)</li> </ul> <p><b>0-1</b></p>	<b>M</b>
<creation>	<p>Informace o vytvoření dokumentu.</p> <p>0-1</p>	O
<year.creation>	<p>Rok vytvoření dokumentu ve formátu YYYY.</p> <p>0-1</p>	RA
<country.creation>	<p>Země, ve které byl dokument vytvořen: vyjádřeno iso kódem (viz Nepovinné atributy) př.: &lt;country.creation iso.code="cz"&gt;Česká republika&lt;/country.creation&gt;</p> <p>0-1</p>	RA
<creator>	<p>Informace o tvůrci dokumentu.</p> <p>0-1</p>	MA
<creator.corporate>	<p>Název instituce, pokud je autorem dokumentu.</p> <p>0-N</p>	MA

<sup>5</sup>Dostupné z: <http://www.iso.org/iso/en/prods-services/iso3166ma/02iso-3166-code-lists/list-en1.html>

<creator.person>	Informace o autorovi dokumentu. 0-N	MA
<name>	Jméno a příjmení autora. 0-1	MA
<year.birth>	Datum narození autora. Formát YYYY. 0-1	MA
<year.death>	Datum úmrtí autora. Formát YYYY. 0-1	MA
<note>	Doplňující informace k tvůrci dokumentu. 0-N	O
<publication>	Informace o vydání dokumentu 0-1	MA
<country.publication>	Země vydání – vyjádřeno iso kódem (viz Nepovinné atributy), př: <country.publication iso.code="cz">Česká republika</country.publication> 0-1	R
<publisher>	Nakladatel. 0-1	R
<year.publication>	Rok vydání. Formát YYYY. 0-1	MA
<year.copyright>	Rok copyrightu. Formát YYYY. 0-1	RA
<year.renewal>	Rok případného obnoveního copyrightu. Formát YYYY. 0-1	RA
<note>	Doplňující informace k vydání dokumentu. 0-N	O
<rights.holder>	Informace o držiteli práv. Používá se v případě, že práva drží někdo jiný než autor nebo nakladatel. Např. občanské sdružení autorů apod. 0-1	O
<contact>	Kontakt na držitele práv. 0-N	O
<name>	Jméno nebo název držitele práv. 0-1	O
<note>	Doplňující informace o držiteli práv. 0-N	O
<notice>	Do pole se vkládá oznámení o copyrightu, tak jak je uvedeno v dokumentu. 0-1	O

<general.note>	Pole slouží k vložení jakékoliv další informace, která je relevantní vzhledem k autorským právům dokumentu a nemohla být zařazena do žádného jiného pole. 0-N	O
----------------	--	---