



**DEFINICE
METADATOVÝCH
FORMÁTŮ**

13.listopadu

2024

Dokument verze 0.1

Definice metadatových formátů pro datové disky

Autoři aktuální verze: Květa Fremrová, Veronika Ježková, Pavlína Kočišová, Vojtěch Kopský, Lenka Kravecová, Filip Pavčík

Dokument vychází z původního návrhu Lenky Kravecové.

Obsah

Obsah	2
1 Popis standardu	3
1.1 Účel standardu	3
1.2 Zdpovědnost	3
1.3 Dohoda mezi dodavatelem dat a zadavatelem	3
1.4 Popis a verze použitých standardů a identifikátorů	4
1.4.1 Použité verze metadatových formátů	4
1.4.2 Použité standardy pro ISO obrazy	4
1.4.3 Použité identifikátory	5
1.5 Definice pojmů	5
2 Výstupy digitalizace	7
3 Granularita metadatového záznamu	8
4 Identifikátory	9
5 Struktura SIP balíčku	10
5.1 Soubor info.xml	10
5.2 Složka pro obrazy disku isoimage	11
5.3 Složka pro obrazové soubory mastercopy	11
5.4 Složka pro obrazové soubory usercopy	12
5.5 Složka alto	12
5.6 Složka txt	12
5.7 Složka amdsec	12
5.8 Složka fifoid	12
5.9 Složka pro katalogizační záznam catalog_entry	13
5.10 Soubor mets	13
5.11 Soubor MD5 pro SIP balíček	13
6 Názevová konvence složek a souborů	15
7 Metadata	17
7.1 Kořenový element hlavního METS záznamu	17
7.2 METS hlavička <metsHdr>	18
7.3 METS část <dmdSec> – Bibliografická metadata – MODS a DC	19
7.3.1 Pole MODS a DC pro úroveň kolekce dat	22
7.3.2 Pole MODS a DC pro úroveň disku	32
7.3.3 Pole MODS a DC pro obrazové informace (přílohy)	40
7.4 METS část <amdSec> - Technická a administrativní metadata – formáty MIX a PREMIS	44
7.4.1 PREMIS Object	46
7.4.2 PREMIS Event	51
7.4.3 PREMIS Agent	53
7.4.4 Technická metadata MIX	55
7.4.5 Technická metadata File Systémů	63
7.5 METS část <fileSec>	68
7.5.1 <fileSec> hlavního záznamu METS	68
7.5.2 <fileSec> vedlejšího záznamu amd_mets.xml, resp. amd_mets_isoimg.xml	69
7.6 METS část <structMap> – Strukturální metadata	70
7.6.1 <structMap> hlavního záznamu METS	70
7.6.2 <structMap> vedlejšího záznamu METS (amd_mets.xml a amd_mets_isoimg.xml)	71
7.7 METS část <structLink>	72
7.8 OCR (ALTO XML a TXT OCR)	73
7.9 Autorsko-právní metadata	85
Historie verzí	89

1 Popis standardu

1.1 Účel standardu

Definice metadatových formátů (dále DMF) slouží jako předpis pro zpracování datových disků určených pro dlouhodobé uložení. Definuje podobu SIP balíčku.

Tento předpis je zaměřen jak na datové disky a soubory datových disků s vlastním katalogizačním záznamem, tak na přílohové disky, které spadají pod katalogizační záznam mateřského dokumentu (nejčastěji monografie nebo periodika). Předpis je zaměřen na dlouhodobé uchování CD a DVD, na kterých jsou soubory uloženy podle standardu ISO 9660 nebo ve formátu UDF a jejich rozšířeních. Specifikace předpokládá zpracování disků primárně s textovým či obrazovým obsahem, nezahrnuje audio a video disky.

1.2 Zdpovědnost

Za DMF, jeho správnost a rozvoj, je zodpovědná Národní knihovna ČR. Standard vznikl ve spolupráci s Moravskou zemskou knihovnou v Brně.

Kontaktním oddělením je [Odbor novodobých digitálních sbírek](https://standardy.ndk.cz/), resp. [Oddělení standardů digitálních sbírek](https://standardy.ndk.cz/) (<https://standardy.ndk.cz/>). S doplňujícími dotazy lze kontaktovat oddělení pro standardy na adrese metadata.ndk@nkp.cz. Pro hlášení chyb, dotazy nebo návrhy na rozšíření standardů lze také využít https://github.com/NLCR/Standard_NDK.

1.3 Dohoda mezi dodavatelem dat a zadavatelem

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

- **podrobnosti skenování** - např. použité rozlišení (300 PPI 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.)

1.4 Popis a verze použitých standardů a identifikátorů

1.4.1 Použité verze metadatových formátů

V této verzi specifikace budou použity následující verze metadatových standardů:

Standard	Verze	Odkaz na webové stránky standardu
METS	verze 1.11	http://www.loc.gov/standards/mets/
MODS	verze 3.8	http://www.loc.gov/standards/mods/
Dublin Core	verze 1.1	http://dublincore.org/documents/dces/
MIX	verze 2.0	http://www.loc.gov/standards/mix
PREMIS	verze 2.2	http://www.loc.gov/standards/premis/
ALTO	verze 2.0	http://www.loc.gov/standards/alto/
CopyrightMD	verze 0.91	https://cdlib.org/groups/rights-management-group-copyrightmd/

Pro převod metadat 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>.

1.4.2 Použité standardy pro ISO obrazy

Vzhledem k tomu, že pro všechny varianty formátů ISO obrazů¹ neexistuje jednotný standard, dle této specifikace mohou být využity následující verze standardů (a jejich rozšíření) pro ISO obrazy:

Standard	Odkaz na webové stránky standardu
ISO 9660 ²	https://www.iso.org/iso-9660-images-for-computer-files.html https://www.ecma-international.org/wp-content/uploads/ECMA-119_4th_edition_june_2019.pdf
UDF (Universal Disk Format)	http://www.osta.org/specs/pdf/udf260.pdf

¹ ISO Disk image file format (zkráceně ISO image nebo též ISO obraz) - pod tímto termínem se původně označovaly jenom archivační ISO obrazy vytvořené dle standardu ISO 9660:1988. Od 90. let se však pojem ISO obrazy začal používat i pro počítačová data vytvořené dle specifikace UDF (Universal Disk Format), která byla normalizována standardem ISO/EIC 13346. UDF se využívá nejčastěji (ale ne výlučně) pro ukládání dat z DVD nebo Blu-ray disků.

Viz podrobněji: <https://www.loc.gov/preservation/digital/formats/fdd/fdd000348.shtml#specs>

² Pod standard ISO 9660 spadají i jeho následující rozšíření: Rock Ridge Interchange Protocol (Standard IEEE P1282 z roku 1994), specifikace Joliet z roku 1995 a standard El Torito určený pro bootable CD disky.

1.4.3 Použité identifikátory

V této verzi specifikace mohou být použity následující identifikátory:

UUID	http://www.ietf.org/rfc/rfc4122.txt	
čČNB	http://www.caslin.cz/caslin/spoluprace/sluzby-souborneho-katalogu-cr/cislo-cnb-v-sk-cr	identifikátor entity tak, jak odpovídá katalogizačnímu záznamu v bázi ČNB
URN:NBN	http://resolver.nkp.cz	
ISBN	https://text.nkp.cz/sluzby/sluzby-pro/isbn-ismn-issn	pouze pro titul monografického dokumentu nebo pro soubor monografických dokumentů, které mají pouze jeden souborný záznam, ISBN není přiděleno vždy
ISSN	https://www.techlib.cz/cs/2844-ceske-narodni-stredisko-issn	osmimístný číselný kód, kterým se jednoznačně identifikují názvy periodik
ISMN	https://www.nkp.cz/sluzby/sluzby-pro/isbn-ismn-issn	podobně jako ISBN, ale pro hudebniny
Další identifikátor		je možné použít další standardizovaný identifikátor jako RFID, čárový kód atp., v případě nejasností kontaktujte Oddělení standardů digitálních sbírek

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
- je nutné zachovat velikost obrazu uživatelských a archivních kopií stejnou (počet pixelů, rozlišení) tak, aby ALTO XML odpovídalo OCR
- jako výchozí SW pro vytváření souborů JPEG 2000 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 je zcela dle konkrétních potřeb)

1.5 Definice pojmů

- **ISO image** = bitová kopie dat z optického disku podle standardu ISO 9660 nebo formátu UDF, datový soubor vytvořený zkopírováním obsahu disku; též obraz disku
- **MC** = master copy = archivní kopie obrazového souboru (MC)

- **UC** = user copy = uživatelská kopie obrazového souboru (UC)
- **PS** = původní sken – obrazový soubor vzniklý při digitalizaci, který se po zpracování (ořez, narovnání apod.) maže a dále se neukládá
- **kolekce dat** = soubor jednoho nebo více datových disků v jednom nebo více obalech pod jedním bibliografickým záznamem; jedna kolekce dat je základní intelektuální entitou
- **bibliografický záznam** = užíváno ve významu metadatového popisu obsahu dokumentu
- **booklet** = brožura k datovému nosiči
- **cover/obálka** = obal datového nosiče
- **imgdisc** = sken vrchní strany disku datového nosiče
- **datový disk** = počítačový disk obsahující datové soubory³
- **balíček** = složka s definovaným názvem a obsahem
- **SIP balíček** – Submission Information Package – je balíček dat a metadat v podobě, ve které je akceptovatelný pro LTP systém nebo pro aplikaci zpřístupnění; balíček bude obsahovat kompletní intelektuální entitu tj. jednu kolekci dat.

³„Soubor obsahující množinu dat uspořádanou určitým způsobem (např. text, tabulka, obrázek), která vzniká jako výstup určitého programu (např. textového editoru).“ (POTÁČEK, Jiří. Datový soubor. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha: Národní knihovna ČR, 2003- [cit. 2024-09-10]. Dostupné z: https://aleph.nkp.cz/F/?func=direct&doc_number=000000016&local_base=KTD).

2 Výstupy digitalizace

1. obraz disku (ISO image)
 - jeden nebo více ISO obrazů, záleží na počtu disků v kolekci
2. archivní kopie obrazových souborů (JPEG 2000)
 - vrchní strana datových disků, booklet, obal
 - vrchní strana disku se bude skenovat vždy; booklet, pokud existuje; pokud obal nenesé žádnou původní informaci není nutné jej skenovat
 - doporučené rozlišení je 300 PPI a více, v plné barvě (24 bitů, RGB) a ve formátu bezztrátový JPEG 2000
3. uživatelské kopie obrazových souborů (JPEG 2000)
 - vrchní strana datových disků, booklet, obal
 - uživatelská kopie se ukládá ve ztrátovém formátu JPEG 2000
4. metadata obrazu disku (ISO image)
 - bibliografická metadata – [MODS a DC](#)
 - strukturální metadata – [METS](#)
 - technická metadata – [PREMIS](#)
 - administrativní metadata – [PREMIS](#), [METS](#), [copyrightMD](#)
5. metadata obrazových souborů
 - bibliografická metadata – [MODS a DC](#)
 - strukturální metadata – [METS](#)
 - technická metadata – [MIX](#), [PREMIS](#)
 - METS nebude obsahovat technická metadata pro UC, jen pro MC
 - administrativní metadata – [PREMIS](#), [METS](#), [copyrightMD](#)
6. [OCR - ALTO XML](#) soubor pro každý sken
7. [OCR - TXT](#) soubor pro každý sken
 - pro možnost stáhnout si jen text dokumentu (tam kde kvalita OCR je odpovídající), vyhledávání/indexace
8. kontrolní metadatové soubory (s [kontrolními součty](#) a [údaji o vzniku dat](#) apod.)
9. Identifikace souborových formátů (XML)
 - jeden nebo více výstupů identifikace formátů (předpokládá se použití nástroje DROID); pro každý datový disk jeden soubor

pozn. odkazy vedou na kapitulu s příslušnou problematikou

3 Granularita metadatového záznamu

Datový disk obsahuje úrovně:

- **kolekce dat (datacollection)** - tvořena jedním nebo více datovými disky patřícími k jednomu mateřskému dokumentu (svazku monografie či číslu periodika)⁴
 - monografie - kolekci dat tvoří jeden svazek monografie (na samostatném disku nebo s jedním či více přílohovými disky)
 - kolekce dat = bibliografický záznam svazku monografie nebo monografického disku
 - periodika - kolekci dat tvoří vždy jedno číslo, v rámci kterého byly datové disky vydány (číslo periodika vydané na jednom disku nebo číslo periodika s jedním či více přílohovými disky)
 - kolekce dat = bibliografický záznam titulu periodika doplněného o údaje ročníku a čísla
- **disk (disc)** - jeden datový disk, jehož data jsou obsažena v jednom souboru ISO image (obraz disku)
- **související obrazové informace (supplement)**
 - sken vrchní strany disku (imgdisc)
 - obal datového disku (cover)
 - brožura k datovému disku (booklet)

Další informace:

- základní intelektuální entitou je kolekce jednoho nebo více datových disků patřících k jednomu svazku monografie nebo k jednomu číslu periodika
- všechny metadatové záznamy jsou zakotveny v METS formátu, jedna kolekce = jeden METS dokument
- SIP balíček je tvořen pro jednu kolekci datových disků (zahrnuje tedy všechny informace týkající se celku, ISO image a souvisejících obrazových informací)

⁴ Digitalizace mateřského dokumentu není předmětem tohoto předpisu a bude se řídit příslušnou aktuální specifikací (DMF pro monografie nebo DMF pro periodika). V případě digitalizace textového dokumentu s přílohovým diskem tedy vzniknou 2 SIP balíčky: 1 pro mateřský dokument (tj. monografii nebo číslo periodika) a 1 pro přílohové disky.

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 úroveň základní intelektuální entity
 - syntax URN:NBN musí odpovídat specifikaci identifikátoru URN:NBN pro resolver NK, viz [Metodika pro přidělování a správu životního cyklu unikátních perzistentních identifikátorů digitálních dokumentů podle standardu URN:NBN](#)

Požadované identifikátory:

- **strojově čitelný identifikátor** (např. RFID, čárový kód EAN apod.) - identifikuje samostatný disk nebo celou kolekci dat, **povinné**
- **UUID** - lze použít pro všechny úrovně digitálního objektu, **povinné**
- **URN:NBN** - pro dokument, který má samostatný bibliografický záznam, bude se přidělovat během digitalizace; **povinné** je přidělován na vrcholové úrovni popisu
- **Číslo České národní bibliografie – čČNB** - povinné, pokud je možné záznamu přidělit
- **ISBN** - pokud je možné záznamu přidělit
- **ISSN** - pokud je možné záznamu přidělit
- **ISMN** - pokud je možné záznamu přidělit

Kromě platných identifikátorů lze do popisných metadat zaznamenávat i ty již neplatné či nahrazené. **Neplatné identifikátory nelze ze zápisu MODS přepisovat do zápisu v Dublin Core**, jelikož v tomto případě nelze odlišit platné od neplatných.

Příklad v rozdílném zápisu validního a nevalidního identifikátoru v MODS:

```
<mods:identifier type="ccnb">cnb000529747</mods:identifier>
<mods:identifier invalid="yes" type="ccnb">cnb000390254</mods:identifier>
```

5 Struktura SIP balíčku

Souborová struktura SIP balíčku:

SLOŽKA/SOUBOR	OBSAHUJE	POVINNOST	POJMENOVÁNÍ SLOŽKY
soubor info.xml	popis obsahu balíčku (manifest)	ANO	
složka isoimage	obrazy disku	ANO	isoimage
složka mastercopy	archivní kopie obrazových souborů	ANO	mastercopy
složka usercopy	uživatelské kopie obrazových souborů	ANO	usercopy
složka alto	soubory OCR ALTO pro každý sken	ANO	alto
složka txt	soubory OCR TXT pro každý sken	ANO	txt
složka amdsec	vedlejší METS soubory pro každý ISO image a každý sken	ANO	amdsec
složka fifoid (FileFormatIdentification)	výstupy z identifikace souborových formátů (ve formátu XML)	ANO	fifoid
složka catalog_entry	katalogizační záznam ve formátu marcxml v samostatné složce	NE	catalog_entry
soubor mets.xml	hlavní METS záznam	ANO	
soubor *.md5 pro checksum	kontrolní součet balíčku	ANO	

Technická a administrativní metadata nesmí být obsažena v hlavním METS záznamu. Každý sken a každý ISO image soubor musí mít technická a administrativní metadata ve svém vlastním souboru s METS záznamem (amd_mets.xml pro skeny a amd_mets_isoimg.xml pro ISO image).

SIP balíček = 1 adresář pro kolekci jednoho nebo více datových disků.

Hlavní složka SIP balíčku obsahuje následující složky a soubory:

5.1 Soubor info.xml

Soubor info.xml musí obsahovat každý SIP balíček, budou zde velmi krátce zaznamenány údaje o jeho vzniku. Soubor info vždy musí v itemlistu odkazovat sám na sebe.

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

Element	Atributy	Popis	Povinnost
<info>		kořenový element INFO záznamu	M
x<created>		časový údaj o vzniku balíčku ve formátu ISO 8601 na úrovni	M

		vteřin	
x<metadataversion>		verze metadatové specifikace, podle které byl balíček zpracován	M
x<packageid>		název kořenového adresáře balíčku (viz kap. 6)	M
x<mainmets>		název hlavního METS souboru včetně přípony	M
x<validation>		výstup validačního nástroje	R
	version	verze validačního nástroje	R
x<titleid>		identifikátory titulu - vypsát všechny (viz <i>type</i>), které má titul přidělen; každý identifikátor se zapíše do vlastního <titleid>, element lze podle potřeby opakovat	M
	type	možné hodnoty: isbn, issn, ismn, ccnb, urnnbn, uuid	M
x<collection>		údaje o větším celku (projektu), pokud do nějakého balíček patří	R
x<institution>		název instituce, která je zadavatelem digitalizace	R
x<creator>		tvůrce balíčku – kód instituce (firmy), která balíček vytvořila	M
x<size>		velikost balíčku v kB - bez souboru info.xml	M
x<itemlist>		obsahuje seznam všech souborů v balíčku vč. souborů v podadresářích a souboru info.xml	M
	itemtotal	celkový počet souborů	M
xx<item>		obsahuje cestu k jednomu souboru např. “\isoimage\isoimg_101ad2ad-31c2-4aa5-be84-d59e87dc735d_0001.iso” element je opakován podle nutnosti - tj. podle počtu souborů	M
x<checksum>		odkaz na soubor s MD5 a jeho MD5 kontrolní součet	M
	type	bude vždy "md5"	M
	checksum	kontrolní součet souboru - stejný algoritmus jako byl použit u md5 souboru	M
x<note>		poznámka	O

5.2 Složka pro obrazy disku isoimage

Složka obsahuje jeden či více souborů ISO image.

5.3 Složka pro obrazové soubory mastercopy

Složka s archivními kopiemi všech obrazových souborů např. v bezztrátovém JPEG 2000 (vrchní strana disku, booklet a obal).

5.4 Složka pro obrazové soubory usercopy

Složka s uživatelskými kopiemi všech obrazových souborů ve ztrátovém JPEG 2000.

5.5 Složka alto

Obsahuje ke každému skenu jeden ALTO XML soubor, tj. vznikne tolik ALTO XML souborů kolik je skenů patřících k jedné kolekci.

5.6 Složka txt

Obsahuje ke každému skenu jeden OCR soubor jako čistý text, tj. vznikne tolik OCR TXT souborů, kolik je skenů patřících k jedné kolekci.

5.7 Složka amdsec

Složka s technickými metadaty musí obsahovat pro každý ISO image a každý archivní obrazový soubor 1 vedlejší METS soubor. Ten obsahuje následující části METS formátu:

- **amdSec** - administrativní metadata - obsahuje části:
 - technická metadata **techMD**, která:
 - ve formátu **PREMIS Object** popisují vlastnosti:
 - ISO image (obrazu disku)
 - původního obrazového souboru (PS), ze kterého vznikly archivní obrazové kopie, např. ve formátu TIFF
 - archivní kopie obrazového souboru
 - ALTO XML
 - ve formátu **MIX** popisují vlastnosti:
 - archivní kopie obrazu (veškeré generace archivních kopii)
 - případné jiné reprezentace stránky
 - metadata o provenienci digitálních objektů **digiprovMD** - obsahuje části:
 - **PREMIS Event**
 - **PREMIS Agent**
- **fileSec** - sekce s výčty jednotlivých souborů a odkazy na ně; povinná část METS záznamu; v případě tohoto METS záznamu pro jeden ISO image nebo jednu stránku, který vzniká primárně k zachycení technických a administrativních metadat, bude odkazovat na soubory související s tím konkrétním ISO image nebo s tou konkrétní stránkou, tj. ISO soubor, archivní kopie obrazu, ALTO XML, OCR TXT
- **structMap** – pouze fyzická strukturální mapa; povinná část METS záznamu; bude ukazovat strukturu souborů k danému ISO image nebo stránce
- **copyrightMD** – autorsko-právní metadata lze definovat i pro konkrétní části dokumentu; možnosti a pravidla dědičnosti jsou uvedeny v příslušné kapitole ([kap. 7.9 - Autorsko-právní metadata](#)); nepovinné

5.8 Složka fifoid

Složka s výstupy z nástroje pro identifikaci souborových formátů. Pro každý datový disk bude vytvořen jeden soubor (ve formátu *.xml) s identifikací formátů souborů obsažených na disku.

5.9 Složka pro katalogizační záznam `catalog_entry`

Složka obsahuje dokument ve formátu *.xml, ve kterém je vložen původní katalogizační záznam transformovaný do formátu marcxml. Tato složka není povinná, je však doporučena.

5.10 Soubor mets

Hlavní METS záznam bude obsahovat části:

- **dmdSec** – bibliografická metadata ke každé úrovni datového disku ve formátu MODS a DC, včetně obrazových informací
- **amdSec** - autorsko-právní metadata, která popisují, zda nebo kdy může být dílo zveřejněno; tato metadata nejsou povinná; formátem bude copyrightMD
- **fileSec** – výčet všech digitálních objektů s odkazy na ně
- **structMap** – strukturální mapa pro celý dokument (fyzická i logická část)
- **structLink** - vazby mezi fyzickou a logickou částí strukturální mapy (tedy mezi jednotlivými ISO image a diskem a mezi obrázky a diskem či kolekcí).

5.11 Soubor MD5 pro SIP balíček

Pro kontrolní součty je využíván algoritmus MD5. Balíček musí obsahovat jeden soubor .md5. Tento soubor .md5 musí obsahovat kontrolní součet pro každý soubor obsažený v SIP balíčku (kromě souboru info.xml a souboru .md5 samotného).

Obsah souboru je definován následující gramatikou ve formátu ABNF⁵

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

⁵ <http://tools.ietf.org/html/rfc4234>

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 aboslutní cesty v hierarchické struktuře balíku SIP. Cesta je absolutní vůči kořenové složce SIP balíčku.

6 Názvová konvence složek a souborů

K jednoznačné identifikaci jedné kolekce datových disků slouží identifikátory UUID nebo URN:NBN. Tyto identifikátory budou také tvořit základ pojmenování jednotlivých souborů.

Pojmenování SIP balíčku:

- každý balíček přicházející z digitalizace musí obsahovat pouze jednu základní intelektuální entitu (kolekci jednoho nebo více datových disků)
- název balíčku **musí vycházet** z identifikátoru této entity, tj. z UUID nebo URN:NBN
 - z UUID se převezme celá část identifikátoru za dvojtečkou (tj. z uuid:21d5eff0-d9aa-11de-a7ba-000d606f5dc6 vznikne název balíčku 21d5eff0-d9aa-11de-a7ba-000d606f5dc6)
 - z URN:NBN se převezme část identifikátoru za urn:nbn:cz: (tj. z urn:nbn:cz:osa001-0001kl vznikne název balíčku osa001-0001kl)
- každá kolekce datových disků musí mít svůj jednoznačný identifikátor, tím pádem pak každý SIP balíček a každý soubor v něm má vlastní jednoznačný identifikátor díky využití prefixů
- názvy nesmí obsahovat mezery a diakritiku, povinné oddělovače jsou podtržítka, případně spojovník (dle syntaxe identifikátorů URN:NBN a UUID)

Pojmenování složek

- viz struktura SIP balíčku (kap. 5)
- názvy nesmí obsahovat mezery, diakritiku a dvojtečku, **přípustnými oddělovači jsou podtržítka, případně spojovník** (dle syntaxe identifikátorů URN:NBN a UUID)

Pojmenování souborů

Každý soubor musí obsahovat předponu a příponu jasně identifikující, o který soubor se jedná.

V případě použití UUID (analogicky i URN:NBN):

- **soubor ISO image** – pojmenované jako "druh-souboru_uuid_XXXX.iso", kde druh souboru je "isoimg" a XXXX je pořadové číslo souboru, př.:
 - isoimg_039c7140-ff21-11eb-ba6b-5dd2b079df66_0001.iso
 - isoimg_039c7140-ff21-11eb-ba6b-5dd2b079df66_0002.iso
- **obrazy vrchní strany disku, obalu kolekce a digitalizovaného bookletu** – skeny označené jako "druh-souboru_uuid_XXXX.jp2", kde druh souboru je "mc" nebo "uc" a XXXX je pořadové číslo skenu; stejné pojmenování ponese i uživatelská kopie
 - mc_039c7140-ff21-11eb-ba6b-5pdd2b079df66_0001.jp2
 - mc_039c7140-ff21-11eb-ba6b-5dd2b079df66_0002.jp2
 - uc_039c7140-ff21-11eb-ba6b-5dd2b079df66_0001.jp2
 - uc_039c7140-ff21-11eb-ba6b-5dd2b079df66_0002.jp2
- **hlavní mets** – bude mít označení mets_uuid.xml
 - mets_039c7140-ff21-11eb-ba6b-5dd2b079df66.xml
- **soubor info** – bude mít označení info_uuid.xml
 - info_039c7140-ff21-11eb-ba6b-5dd2b079df66.xml
- **vedlejší METS obsahující technická metadata obrazových souborů** – bude mít označení "amd_mets_uuid_XXXX.xml, kde "amd_mets" určuje druh souboru a XXXX pořadové číslo xml souboru
 - amd_mets_039c7140-ff21-11eb-ba6b-5dd2b079df66_0001.xml
- **vedlejší METS obsahující technická metadata souborů ISO image** – bude mít označení "amd_mets_isoimg_uuid_XXXX.xml, kde "amd_mets_isoimg" určuje druh souboru a XXXX pořadové číslo xml souboru

- amd_mets_isoimg_039c7140-ff21-11eb-ba6b-5dd2b079df66_0001.xml
- **soubor ALTO** – bude mít označení “alto_uuid_XXXX.xml”, kde “alto” určuje druh souboru a XXXX pořadové číslo xml souboru
 - alto_039c7140-ff21-11eb-ba6b-5dd2b079df66_0001.xml
- **soubor OCR TXT** – bude mít označení “txt_uuid_XXXX.txt”, kde “txt” určuje druh souboru a XXXX pořadové číslo txt souboru
 - txt_039c7140-ff21-11eb-ba6b-5dd2b079df66_0001.txt
- **soubor MD5** – bude mít označení “md5_uuid.md5”, kde “md5” určuje druh souboru
 - md5_039c7140-ff21-11eb-ba6b-5dd2b079df66.md5
- **soubor catalog_entry** – bude mít označení „cat_entry_uuid.xml“
 - cat_entry_039c7140-ff21-11eb-ba6b-5dd2b079df66.xml
- **soubor fifoid** - bude mít označení "fifoid_uuid_XXXX.xml", kde "fifoid" určuje druh souboru a XXXX pořadové číslo xml souboru
 - fifoid_039c7140-ff21-11eb-ba6b-5dd2b079df66_0001.xml

Názvy jakýchkoliv souborů náležejících k jedné intelektuální entitě musí být založeny na jednom typu identifikátoru. Názvy musí být pouze malými písmeny a nesmí obsahovat mezery ani diakritiku, doporučené oddělovače jsou podtržítka, případně spojovník (dle syntaxe identifikátorů URN:NBN a UUID).

7 Metadata

- veškerá metadata budou „zabalena“ pomocí kontejnerového formátu METS, kromě souboru info.xml, který má vlastní metadatový formát
- vložení metadatových formátů do kontejneru METS musí být vždy formou <mdWrap>

Legenda pro čtení specifikace

- sloupec **Element** obsahuje název elementu; počet znaků „x“ před názvem elementu značí stupeň zanoření elementu v zápisu
- sloupec **Atributy** obsahuje název atributu, pokud se k danému elementu nějaký váže
- sloupec **Popis** obsahuje vysvětlení a příklad užití příslušného elementu/atributu; kde je to možné, je uvedeno doporučené nebo předepsané plnění; kde je to třeba, je uvedeno mapování ke konkrétním polím záznamu MARC21; ve sloupci jsou použity dva různé druhy uvozovek následovně: pokud uzavírají konkrétní hodnotu např. z kontrolovaného slovníku, tak jak se má objevit v xml dokumentu, jsou použity „anglické uvozovky nahoře“; pokud uzavírají pouze ilustrativní příklad, jsou použity „běžné české uvozovky“
- sloupec **Povinnost** určuje povinnost použití elementu/atributu; povinnost platí jak pro elementy MODS, tak pro elementy Dublin Core; pokud je rodičovský element např. doporučený a dceřiný element povinný, znamená to, že je dceřiný element povinný pouze tehdy, pokud je použit element rodičovský
- sloupec **Element DC** uvádí element Dublin Core, ke kterému je třeba mapovat příslušný element MODS

Barevné kódování

- buňky s elementy, které mají povinnost M nebo MA, jsou podbarveny oranžovou barvou; to platí, pokud je povinný i jejich rodičovský element; bílé jsou buňky s volitelnými elementy, resp. s elementy, které jsou povinné pouze při použití volitelného rodičovského elementu
- top elementy a kořenový element <mods> jsou vyznačeny **tučně**

7.1 Kořenový element hlavního METS záznamu

Kořenový element hlavního METS záznamu 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	název titulu dokumentu včetně roku vydání (údaje oddělené čárkami), např.: „Angličtina v cestovním ruchu, 2005“	M
	TYPE	hodnota vždy "data_disc"	M

7.2 METS hlavička <metsHdr>

Dokumentuje vznik a úpravy METS záznamu.

Element	Atributy	Popis	Povinnost
<metsHdr>		hlavička METS záznamu	M
	LASTMODDATE	datum poslední úpravy záznamu, musí být ve tvaru ISO 8601 (alespoň na úroveň vteřin)	M
	CREATEDATE	datum vytvoření záznamu, musí být ve tvaru ISO 8601 (alespoň na úroveň vteřin)	M
x<agent>		údaje o tvůrci METS	M
	ROLE	hodnota "CREATOR"	M
	TYPE	hodnota "ORGANIZATION"	M
xx<name>		jednoznačný identifikátor instituce - tvůrci metadat, v případě tvorby metadat v knihovně bude užita sigla knihovny, např. pro NK ČR tedy "ABA001".	M
x<agent>		údaje o vlastníkově METS	M
	ROLE	hodnota "ARCHIVIST"	M
	TYPE	hodnota "ORGANIZATION"	M
xx<name>		jméno organizace; vlastník záznamu, v případě tvorby záznamu v knihovně bude vždy využita sigla knihoven, tj. např. pro NK ČR hodnota "ABA001"	M

7.3 METS část <dmdSec> – Bibliografická metadata – MODS a DC

- MODS a DC budou vloženy v METS části <dmdSec>
- pole MODS se plní na základě konverze příslušných polí katalogizačního záznamu ve formátu MARC 21; na tyto pole a jejich indikátory se pak odkazuje i u příslušných elementů MODS
- z MARC 21 resp. MARCXML do MODS lze využít a dle potřeb knihovny upravit oficiální mapování a konverzní šablonu ze stránek <https://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>

Datový disk - kolekce dat

- základní intelektuální entitou pro popis je kolekce dat, tj. v jednom METS záznamu, který bude obsahovat metadata a strukturu jednoho a více datových disků, budou MODS a DC záznamy k této kolekci
 - kolekci dat se rozumí vrcholná úroveň popisu datového disku, ať už se týká jen jednoho disku, nebo souboru mnoha disků; kolekce dat odpovídá bibliografickému záznamu v katalogu; metadata budou popisovat entity:
 - **kolekce dat (datacollection)** – datový nosič nebo soubor datových nosičů vydaný jako celek/komplet; může obsahovat jeden či více datových disků v jednom obalu nebo ve více obalech, které jsou popsány jedním bibliografickým záznamem v katalogu a jejichž obsah spolu souvisí nebo na sebe navazuje; kolekce dat je nadřazená entita pro všechny následující
 - **disk (disc)** - popisuje zvláště každý disk v kolekci
 - **příloha (supplement)** - přílohou se rozumí volně vložená entita k intelektuální entitě, např. volně vložený booklet, a dále skeny obalu a vrchní strany disku.
- Rozlišujeme tři druhy příloh:
1. **obal (cover)** - popis pro fyzický obal kolekce
 2. **booklet** - popis pro doprovodný text v podobě knížečky či obrazovou přílohu, volně vloženou či sešitou s obalem
 3. **skan disku (imgdisc)** - popis pro sken vrchní strany datového disku

Přílohy se pomocí elementu <relatedItem> odkazují k nejvyšší popisované úrovni, tj. kolekci dat. Tyto druhy příloh pak budou zohledněny i v rámci fyzické mapy dokumentu v atributu TYPE.

Obecná pravidla pro bibliografická metadata

- pro každou entitu vznikne jeden MODS záznam s vlastním ID, které bude označovat i typ části (datacollection, disc, supplement); v případě opakování částí stejného typu (disc, supplement) se bude opakovat tomu odpovídající počet MODS záznamů v jednom SIP balíčku
- jednotlivé MODS záznamy pro části (kolekce dat, disk, příloha) nejsou samopopisné, tj. neobsahují vždy údaje o vrchních entitách (příloha neobsahuje informace o kolekci dat apod.)
- 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 a MIMETYPE

- pokud je v převáděných záznamech použita mezinárodní Interpunkce standardního bibliografického popisu (ISBD), měla by být v MODS zachována, pouze pokud se vyskytuje uprostřed textové hodnoty jednoho elementu; interpunkce na začátku a na konci textové hodnoty elementu bude vypuštěna; např. pro ISBD „**Ithaca, NY : Cornell University Press, 1999.**“ by zápis v MODS vypadal takto:

```

<originInfo>
  <place>
    <placeTerm type="code" authority="marccountry">nyu</placeTerm>
    <placeTerm type="text">Ithaca, NY</placeTerm>
  </place>
  <agent>
    <namePart>Cornell University Press</namePart>
    <role>
      <roleTerm>publisher</roleTerm>
    </role>
  </agent>
  <dateIssued>1999</dateIssued>
</originInfo>

```

- kromě elementů, kde to může být součástí jejich plnění (např. <nonSort>), není žádoucí ponechávat na začátku a na konci textových hodnot mezery (bílé znaky)

Element	Atributy	Popis	Povinnost
<dmdSec>		element obsahující bibliografická metadata ve formátu MODS nebo DC	M
	ID	identifikátor <dmdSec> části METS záznamu pro <dmdSec> s popisem kolekce dat hodnota "MODSMD_DATACOLLECTION_0001" a "DCMD_DATACOLLECTION_0001"; pro <dmdSec> s popisem disku hodnota "MODSMD_DISC" a "DCMD_DISC"; pro <dmdSec> s popisem přílohy hodnota "MODSMD_SUPPL" a "DCMD_SUPPL"; tyto hodnoty budou následované podtržítkem a čtyřmístným číslem označujícím pořadí příslušné části (tj. identifikátory mají obdobný formát jako identifikátory pro kolekci dat) např.: "MODSMD_DISC_0001" a "DCMD_DISC_0001" pro první disk "MODSMD_SUPPL_0001" a "DCMD_SUPPL_0001" pro první přílohu	M
x<mdWrap>		element obsahující vložené záznamy MODS	M
	MDTYPE	hodnota "MODS" pro záznam v MODS; hodnota "DC" pro záznam v Dublin Core	M
	MDTYPEVERSION	číslo verze MODS, hodnota pro záznamy v MODS, např. "3.8 "	O
	MIMETYPE	hodnota "text/xml"	M
xx<xmldata>		kontejnerový element, do kterého se vkládají záznamy <mods>	M

Pokyny pro tvorbu identifikátorů ID

ID u elementu <mods>: Identifikátory budou začínat prefixy MODS_DATACOLLECTION, MODS_DISC a MODS_SUPPL (obdobně pro DC); za ty se dále přidá podtržítka a číslo identifikující pořadí identifikátoru, zarovnané a doplněné o nuly na 4 místa (kvůli jednotnosti s ostatními standardy)

ID tedy vypadá následovně:

- kolekce dat:
 - MODS_DATACOLLECTION_0001
 - DC_DATACOLLECTION_0001
- disk:
 - MODS_DISC_XXXX
 - DC_DISC_XXXX
- přílohy (skeny obalu a disků, případně bookletu):
 - MODS_SUPPL_XXXX
 - DC_SUPPL_XXXX

Kde XXXX je pořadové číslo disku, resp. přílohy.

7.3.1 Pole MODS a DC pro úroveň kolekce dat

[Legenda pro čtení specifikace](#)

Element MODS	Atributy	Popis	Povinnost	Element DC
<code><mods></code>	ID	ID musí vyjadřovat název úrovně, tj. pro úroveň kolekce dat "MODS_DATACOLLECTION_0001"	M	
	version	verze mods: "3.8"	R	
<code>x<titleInfo></code>		název kolekce dat, souborný název (pro plnění lze použít katalogizační záznam)	M	
	type	hlavní název bez <i>type</i> - pole 245 \$a může existovat pouze jeden hlavní název bez <i>type</i> hodnoty: "abbreviated" - pole 210 "alternative" - pole 246 "translated" - pole 242 "uniform" - pole 130, resp. 240	MA	
<code>xx<nonSort></code>		část názvu, která má být vynechána při řazení; za hodnotou je nutno ponechat mezeru pro správné zobrazování názvu v Krameriu např.: <nonSort>The </nonSort> <title>Beatles</title> pozor v případě členů, které k názvu přiléhají, v takovém případě se mezera nekládá: <nonSort>L' </nonSort> <title>Equipe</title>	O	
<code>xx<title></code>		názvová informace	M	<dc:title>
<code>xx<subTitle></code>		podnázev	MA	<dc:title>
<code>xx<partNumber></code>		číslo části vícesvazkové monografie; údaj se čerpá z polí 240 \$n a 245 \$n - týká se pouze monografií, pro číslování periodik se využije samostatný element <part> s příslušnými subelementy	MA	<dc:description>
<code>xx<partName></code>		název části vícesvazkové monografie; údaj se čerpá z pole 240 \$p nebo 245 \$p - týká se pouze monografií, pro název části periodika se využije samostatný element <part> s příslušnými subelementy	MA	<dc:description>
<code>x<name></code>		údaje o odpovědnosti přebírá se z 1XX a 7XX MARC21	MA	

		opakovatelný element pro více autorů/různé role		
	type	použít jednu z hodnot: <ul style="list-style-type: none"> • "personal" /pro osobu • "corporate" /pro společnost, instituci nebo organizaci • "conference" /pro konference, jednání a jiné pojmenované akce • "family" /rodina, rod 	MA	
xx<namePart>		údaje o křestním jméně a příjmení apod. nutno vyjádřit pro křestní jméno i příjmení; pokud nelze rozlišit křestní jméno a příjmení, nepoužije se <i>type</i> 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"	MA	<dc:creator> nutno do jednoho pole DC spojit jméno i příjmení
	type	použít jednu z hodnot: <ul style="list-style-type: none"> • "date" /životopisná data autora • "family" /příjmení • "given" / křestní jméno • "termsOfAddress" /akademický titul či hodnost 	RA MA MA RA	
xx<alternativeName>		kontejnerový element pro zachycení alternativního jména	O	
xxx<namePart>		alternativní jméno, např. pseudonym, přezdívka	M	
xx<nameIdentifier>		číslo národní autority	MA	<dc:creator>
xx<etal>		element indikující, že existuje více autorů, než pouze ti, kteří byli uvedeni v <name> elementu; v případě užití tohoto elementu je dále top element <name> neopakovatelný; <etal> je nutné umístit do samostatného top elementu <name>, ve kterém se nesmí objevit subelementy <namePart> a <nameIdentifier>; <etal> je neopakovatelný element, který se do zápisu vkládá ručně např.: <mods:name> <etal>a kol.</etal> </mods:name>	O	
xx<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
xxx<roleTerm>		popis role	M	
	type	hodnota "code" – kód role z kontrolovaného	M	

		slovníku (http://www.loc.gov/marc/relators/relaterm.html)		
	authority	údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority="marcrelator"	M	
x<typeOfResource>		bude se vyčítat z katalogizačního záznamu, z pozice 06 návěští	R	
x<genre>		hodnota "data collection"	M	<dc:type>model: datacollection</ dc:type
x<genre>		opakovaný element <genre> slouží k převodu hodnoty konkrétního žánru z pole 655 MARC21	R	<dc:type>
x<originInfo>		informace o původu kolekce	M	
xx<place>		údaje o místě spojeném s vydáním, výrobou nebo původem popisovaného dokumentu	MA	
xxx<placeTerm>		konkrétní určení místa vydání, např. Brno, resp. "xr" pro ČR odpovídá hodnotám z katalogizačního záznamu, pole 260 \$a, resp. pole 008/15-17	M	<dc:coverage>
	type	"code" pro údaj z pole 008 "text" pro údaj z pole 260	M	
	authority	hodnota "marccountry" jen u údaje z pole 008	MA	
xx<agent>		informace o entitě, která dokument vydala, vytiskla nebo jinak vyprodukovala	MA	
xxx<namePart>		jméno entity, která dokument vydala, vytiskla nebo jinak vyprodukovala, odpovídá poli 260 \$b v MARC21; informace o výrobcí je možné přebírat i údaje z pole 260 \$f	M	<dc:publisher>
xxx<role>		specifikace role organizace nebo osoby uvedené v elementu <agent>	M	
xxxx<roleTerm>		popis role hodnoty: "publisher" pro údaje z pole 260 \$b "manufacturer" pro údaje z pole 260 \$f	M	
xx<dateIssued>		datum vydání, odpovídá hodnotě z katalogizačního záznamu, pole 260 \$c a pole 008/07-10=publication	M	<dc:date>
	encoding	hodnota "marc" jen u údaje z pole 008	R	
	point	hodnoty "start", resp. "end" jen u údaje z pole 008,	MA	

		pro rozmezí dat		
	qualifier	možnost dalšího upřesnění, hodnota "approximate" pro data, kde nevíme přesný údaj	R	
xx<issuance>		údaje o vydávání odpovídá hodnotě uvedené v návěští MARC21 na pozici 07 hodnoty "monographic", "serial" nebo "integrating resource"	M	

SEKCE <originInfo> PRO ZÁZNAMY ZPRACOVANÉ PODLE RDA:

x<originInfo>		informace o původu kolekce: odpovídá poli 264	M	
	eventType	hodnoty dle druhého indikátoru pole 264: 264_0 "production" (MA) údaje o vytvoření zdroje v nezveřejněné podobě 264_1 "publication" (MA) údaje o nakladateli zdroje 264_2 "distribution" (R) údaje o distribuci zdroje 264_3 "manufacture" (R) údaje o tisku, výrobě zdroje ve zveřejněné podobě 264_4 "copyright" (R) údaje o ochraně podle autorského práva (copyright) Je nutné využít alespoň jednu z možných hodnot.		
xx<place>		údaje o místě spojeném s vytvořením, vydáním, distribucí nebo výrobou popisovaného dokumentu odpovídá hodnotě 264 \$a	MA	
xxx<placeTerm>		konkrétní určení místa vydání, např. Praha, resp. "xr" pro ČR odpovídá hodnotám z katalogizačního záznamu, pole 264 \$a, resp. pole 008/15-17	M	<dc:coverage>
	type	pokud má dokument více míst vytvoření/vydání/distribuce/výroby v poli 264 \$a, přebírají se ze záznamu všechna místa (v jednom poli 264) "code" pro údaj z pole 008 "text" pro údaj z pole 264 \$a	M	
	authority	hodnota "marccountry" jen u údaje z pole 008	MA	
xx<agent>		informace o entitě, která dokument vytvořila, vydala, distribuovala nebo vyrobila	MA	
xxx<namePart>		jméno entity, která dokument vytvořila, vydala, distribuovala nebo vyrobila odpovídá poli 264 \$b v MARC21	M	<dc:publisher>

		pokud má titul více vydavatelů/ distributorů/ výrobců, přebírají se ze záznamu všichni (jsou v jednom poli 264)		
xxx<role>		specifikace role organizace nebo osoby uvedené v elementu <agent>	M	
xxxx<roleTerm>		popis role, možné hodnoty podle druhého indikátoru pole 264: 264_0 "producer" 264_1 "publisher" 264_2 "distributor" 264_3 "manufacturer"	M	
xx<dateIssued>		datum vydání; odpovídá hodnotě z katalogizačního záznamu <dateIssued> lze použít POUZE PRO POLE 264_1 \$c nebo pole 008/07-10 !!! Pro ostatní výskyty v poli 264 \$c využít element <dateOther> s odpovídajícím atributem type. !!! Pro 264_4 použít element <copyrightDate>	M	<dc:date>
	encoding	hodnota "marc" jen u údaje z pole 008	R	
	point	hodnoty "start", resp. "end" jen u údaje z pole 008, pro rozmezí dat	MA	
	qualifier	možnost dalšího upřesnění, hodnota "approximate" pro data, kde nevíme přesný údaj	R	
xx<dateOther>		datum vytvoření, distribuce, výroby; tento element se využije v případě výskytu \$c v poli 264 s identifikátorem 0, 2 nebo 3	R	
	type	264_0: <dateOther type="production"> 264_2: <dateOther type="distribution"> 264_3: <dateOther type="manufacture">	M	
xx<copyrightDate>		vyžije se pouze v případě výskytu pole 264 s druhým indikátorem 4 a podpolem \$c 264_4: <copyrightDate>	R	<dc:date>
xx<issuance>		údaje o vydávání odpovídá hodnotě uvedené v návěštích MARC21 na pozici 07; hodnoty "monographic", "serial" nebo "integrating resource"	M	

x<physicalDescription>		obsahuje údaje o fyzickém popisu kolekce	MA	
xx<form>		Pro záznamy podle RDA: údaje o typu média a typu nosiče zdroje/předlohy odpovídá hodnotám z pole: 337 NEPOVINNÉ (hodnota "computer" - viz kontrolovaný slovník pole 337_ http://www.loc.gov/standards/valuelist/rdamedia.html) 338 POVINNÉ (hodnota "computerdisc" - viz kontrolovaný slovník pole 338_ http://www.loc.gov/standards/valuelist/rdacarrier.html)	MA	<dc:format>
	authority	Pro záznamy v RDA, hodnoty: pole 337: authority="rdamedia" (hodnota computer) pole 338: authority="rdacarrier" (hodnota computerdisc)	MA	
	type	pro záznamy v RDA pole 337: type="media" (hodnota computer) pole 338: type="carrier" (hodnota computerdisc)	MA	
xx<extent>		údaje o rozsahu - vyplní se počet disků v kolekci; u datových disků s vlastním katalogovým záznamem lze přebírat z pole 300 \$a	M	<dc:format>
x<note>		slouží k vepisování údajů z polí 5XX, 562, 245; pro každou poznámku musí být samostatný <note> element	RA	<dc:description>
	type	<ul style="list-style-type: none"> • pro obsah pole 538 \$a hodnota "system details" • pro obsah pole 546 \$a hodnota "language" • pro obsah pole 561 hodnota "ownership" • pro obsah pole 562 hodnota "version identification" • pro obsah pole 245 hodnota "statement of responsibility" 	MA	
x<subject>		údaje o věcném třídění, předpokládá se přebírání z katalogizačního záznamu	R	
	authority	vyplnit hodnotu "czenas", "eczenas" z pole 072 \$2, nebo jinou hodnotu z tohoto pole; případně lze ponechat prázdné	R	

xx<topic>		libovolný výraz specifikující nebo charakterizující obsah titulu; použít kontrolovaný slovník – např. z báze autorit AUT NK ČR (věcné téma) nebo obsah pole 650 a 653 s druhým indikátorem prázdným či hodnotou 0 záznamu MARC21 nebo obsah pole 072 \$x	M	<dc:subject>
xx<geographic>		geografické věcné třídění; použít kontrolovaný slovník – např. z báze autorit AUT NK ČR (geografický termín) nebo obsah pole 651 a 653 s druhým indikátorem s hodnotou 5 záznamu MARC21	R	<dc:subject>
xx<temporal>		chronologické věcné třídění; použít kontrolovaný slovník – např. z báze autorit AUT NK ČR (chronologický údaj) nebo obsah pole 648 a 653 s druhým indikátorem s hodnotou 4 záznamu MARC21	R	<dc:subject>
xx<name>		jméno použité jako věcné záhlaví; použít kontrolovaný slovník – např. z báze autorit AUT NK ČR (jméno osobní) nebo obsah pole 600 a 653 s druhým indikátorem s hodnotou 1, 2 či 3 záznamu MARC21	R	
xxx<namePart>		celé jméno se zapíše do tohoto elementu	R	<dc:subject>
xxx<nameIdentifier>		číslo národní autority	R	
xx<titleInfo>		pole použité jako vedlejší věcné záhlaví	R	
xxx<title>		unifikovaný název	M	<dc:subject>
x<classification>		klasifikační údaje věcného třídění podle Mezinárodního desetinného třídění	R	<dc:subject>
	authority	hodnota "udc"	M	
x<classification>		klasifikační údaje věcného třídění podle Konspektu odpovídá poli 072 \$a MARC21	R	<dc:subject>
	authority	hodnota "udc" (v případě 072 \$a) hodnota "Konspekt" (v případě 072 \$9)	M	
	edition	hodnota "Konspekt" (v případě 072 \$a)	M	
x<relatedItem>		V případě přílohových disků se element použije pro zaznamenání informací o digitalizovaném mateřském dokumentu (monografii, periodiku ap., jehož je disk přílohou).	RA	

		Element <relatedItem> může obsahovat jakýkoliv jiný element MODS – jejich použití se řídí pravidly popsányými pro tyto elementy.		
	type	hodnota "host"	MA	
xx<titleInfo>		názvové informace mateřského dokumentu	R	
xxx<title>		název mateřského dokumentu	M	
xxx<subTitle>		podnázev mateřského dokumentu	MA	
xxx<partNumber>		číslo mateřského dokumentu (např. číslo výtisku periodika, ke kterému byl disk přiložen)	MA	
xx<identifier>		údaje o identifikátorech mateřského dokumentu	M	
	type	hodnoty: - "uuid" (M) - "urnnbn" (M)	M	
xx<location>		informace o uložení mateřského dokumentu	MA	
xxx<url>		odkaz na uložení digitalizovaného mateřského dokumentu v digitální knihovně	M	
	usage	hodnota "primary", pokud link vede k přímému zobrazení dokumentu		
x<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální – viz přehled typů atributů níže uvádějí se i neplatné resp. zrušené identifikátory - atribut invalid="yes"	M	<dc:identifier>
	type	budou se povinně vyplňovat následující hodnoty, pokud existují: - "uuid" (M) - "urnnbn" (M) - "ccnb" - čČNB - převzít z katalogizačního záznamu z pole 015 \$a, \$z - "isbn" - převzít z katalogizačního záznamu z pole 020 \$a, \$z - "ismn" - převzít z katalogizačního záznamu z pole 024 (1. identifikátor="2") \$a, \$z - "issn" - převzít z katalogizačního záznamu z pole 022 \$a, \$z - jiný interní identifikátor - hodnota atributu "barcode" - čárový kód (MA), podobně lze užít např. "sysno", "oclc", "permalink"	MA	

x<location>		údaje o uložení popisovaného dokumentu, např. signatura, místo uložení apod.	MA	
xx<physicalLocation>		údaje o instituci, kde je fyzicky uložen popisovaný dokument; nutno použít kontrolovaný slovník – sigly knihoven; odpovídá poli 910 \$a v MARC21 pozn. u dokumentů ze soukromých sbírek není možné vyplnit	MA	<dc:source>
	authority	hodnota "siglaADR"		
xx<shelfLocator>		signatura nebo lokační údaje	M	<dc:source>
x<part>		údaje o ročníku a čísle periodika - týká se pouze periodik, pro zápis vícesvazkové monografie se využijí elementy <partNumber> a <partName>	MA	
xx<detail>		údaje o číslování a typu části periodika - element se opakuje pro zachycení každého ročníku a čísla samostatně	M	
	type	hodnota "volume" pro ročník a hodnota "issue" pro číslo	M	
xxx<number>		- v případě <detail type="volume"> se uvede číslo ročníku - v případě <detail type="issue"> se uvede číslo výtisku (čísla)	M	
xxx<title>		název části	R	
xx<date>		datum vydání konkrétního čísla periodika (zápis data se řídí aktuálními Pravidly pro popis periodik)	R	
x<recordInfo>		údaje o metadatovém záznamu – jeho vzniku, změnách apod.	M	
xx<descriptionStandard>		popis standardu, ve kterém je přebíraný katalogizační záznam pro záznamy v AACR2: odpovídá hodnotě návěští záznamu MARC21, pozice 18 - hodnota "acr", tj. pro LDR/18="a" pro záznamy v RDA: odpovídá hodnotě záznamu MARC21 pole 040 a podpole \$e "rda" - bude použito pro primární rozlišení, zda jde o záznam vytvořený podle AACR2 nebo podle RDA	M	
xx<recordContentSource>		kód nebo jméno instituce, která záznam	R	

3.4.1 Pole MODS a DC pro úroveň kolekce dat

		vytvořila nebo změnila		
	authority	hodnota "siglaADR"	R	
xx<recordCreationDate>		datum prvního vytvoření záznamu, alespoň na úroveň minut podle normy ISO 8601	M	
	encoding	hodnota atributu "iso8601"		
xx<recordChangeDate>		datum změny záznamu, podle normy ISO 8601 alespoň na úroveň minut	R	
		hodnota atributu "iso8601"		
xx<recordIdentifier>		identifikátor záznamu v katalogu, přebírá se z pole 001	R	
	source	hodnota se přebírá z katalogu pole 003	R	
xx<recordOrigin>		údaje o vzniku záznamu; hodnoty: "machine generated" nebo "human prepared"	R	
xx<languageOfCataloging>		jazyk katalogového záznamu	R	
xxx<languageTerm>		přebírá se z katalogu - pole 40 \$b	M	
	authority	hodnota "iso639-2b"	R	

7.3.2 Pole MODS a DC pro úroveň disku

[Legenda pro čtení specifikace](#)

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	ID musí vyjadřovat název úrovně, tj. pro úroveň disku je to "MODS_DISC_0001", kde číslo je pořadové číslo disku v rámci kolekce dat	M	
	version	verze mods: "3.8"	R	
x<titleInfo>		názvové informace – pokud nemá disk svůj konkrétní název, kopíruje se informace z nadřazené struktury	M	
	type	hlavní název bez <i>type</i> - pole 245 \$a může existovat pouze jeden hlavní název bez <i>type</i> hodnoty: "abbreviated" - pole 210 "alternative" - pole 246 "translated" - pole 242 "uniform" - pole 130, resp. 240	MA	
xx<nonSort>		část názvu, která má být vynechána při řazení; za hodnotou je nutno ponechat mezeru pro správné zobrazování názvu v Krameriu např.: <nonSort>The </nonSort> <title>Beatles</title> pozor v případě členů, které k názvu přiléhají, v takovém případě se mezera nekládá: <nonSort>L' </nonSort> <title>Equipe</title>	O	
xx<title>		název	M	<dc:title>
xx<subTitle>		podnázev	MA	<dc:title>
xx<partNumber>		číslo disku či jeho pořadí	MA	<dc:description>
xx<partName>		název části	MA	<dc:description>
x<name>		údaje o odpovědnosti přebírá se z 1XX a 7XX MARC21 opakovatelný element pro více autorů/různé role	MA	
	type	použít jednu z hodnot: • "personal" /pro osobu • "corporate" /pro společnost, instituci nebo organizaci	MA	

		<ul style="list-style-type: none"> "conference" /pro konference, jednání a jiné pojmenované akce "family" /rodina, rod 		
xx<namePart>		<p>údaje o křestním jméně a příjmení apod. nutno vyjádřit pro křestní jméno i příjmení; pokud nelze rozlišit křestní jméno a příjmení, nepoužije se <i>type</i> a jméno se zaznamená v podobě, jaké je, do jednoho elementu <namePart></p> <p>pokud známe datum narození a úmrtí autora, vyplnit ve tvaru RRRR-RRRR s atributem <i>type="date"</i></p>	MA	<dc:creator> nutno do jednoho pole DC spojit jméno i příjmení
	type	<p>použít jednu z hodnot:</p> <ul style="list-style-type: none"> "date" /životopisná data autora "family" /příjmení "given" / křestní jméno "termsOfAddress" /akademický titul či hodnost <p>Pokud se nejedná o osobu, atribut <i>type</i> se nepoužije.</p>	RA MA MA RA	
xx<alternativeName>		kontejnerový element pro zachycení alternativního jména	O	
xxx<namePart>		alternativní jméno, např. pseudonym, přezdívka	M	
xx<nameIdentifier>		číslo národní autority	MA	<dc:creator>
xx<etal>		<p>element indikující, že existuje více autorů, než pouze ti, kteří byli uvedeni v <name> elementu; v případě užití tohoto elementu je dále top element <name> neopakovatelný;</p> <p><etal> je nutné umístit do samostatného top elementu <name>, ve kterém se nesmí objevit subelementy <namePart> a <nameIdentifier>;</p> <p><etal> je neopakovatelný element, který se do zápisu vkládá ručně</p> <p>např.:</p> <pre><mods:name> <etal>a kol.</etal> </mods:name></pre>	O	
xx<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
xxx<roleTerm>		popis role	M	
	type	<p>hodnota "code" – kód role z kontrolovaného slovníku</p> <p>(http://www.loc.gov/marc/relators/relaterm.html)</p>	M	

	authority	údaje o kontrolovaném slovníku využitém k popisu role, k popisu výše uvedeného MARC seznamu nutno uvést authority="marcrelator"	M	
x<typeOfResource>		hodnota "computer file"	R	
x<genre>		hodnota "disc"	M	<dc:type>model: disc</dc:type
x<originInfo>		informace o původu disku	M	
xx<place>		údaje o místě spojeném s vydáním, výrobou nebo původem popisovaného dokumentu	MA	
xxx<placeTerm>		konkrétní určení místa vydání, např. Brno, resp. "xr" pro ČR odpovídá hodnotám z katalogizačního záznamu, pole 260 \$a, resp. pole 008/15-17	M	<dc:coverage>
	type	"code" pro údaj z pole 008 "text" pro údaj z pole 260	M	
	authority	authority - hodnota "marccountry" jen u údaje z pole 008	MA	
xx<agent>		informace o entitě, která dokument vydala, vytiskla nebo jinak vyprodukovala	MA	
xxx<namePart>		jméno entity, která dokument vydala, vytiskla nebo jinak vyprodukovala, odpovídá poli 260 \$b v MARC21; informace o výrobci je možné přebírat i údaje z pole 260 \$f	M	<dc:publisher>
xxx<role>		specifikace role organizace nebo osoby uvedené v elementu <agent>	M	
xxxx<roleTerm>		popis role hodnoty: "publisher" pro údaje z pole 260 \$b "manufacturer" pro údaje z pole 260 \$f	M	
xx<dateIssued>		datum vydání; odpovídá hodnotě z katalogizačního záznamu, pole 260 \$c a pole 008/07-10=publication	M	<dc:date>
	encoding	hodnota "marc" jen u údaje z pole 008	R	
	point	hodnoty "start", resp. "end" jen u údaje z pole 008, pro rozmezí dat	MA	
	qualifier	možnost dalšího upřesnění, hodnota "approximate" pro data, kde nevíme přesný údaj	R	

SEKCE <originInfo> PRO ZÁZNAMY ZPRACOVANÉ PODLE RDA:

x<originInfo>		informace o původu disku; odpovídá poli 264	M	
	eventType	hodnoty dle druhého indikátoru pole 264: 264_0 "production" (MA) údaje o vytvoření zdroje v nezveřejněné podobě 264_1 "publication" (MA) údaje o nakladateli zdroje 264_2 "distribution" (R) údaje o distribuci zdroje 264_3 "manufacture" (R) údaje o tisku, výrobě zdroje ve zveřejněné podobě 264_4 "copyright" (R) údaje o ochraně podle autorského práva (copyright) Je nutné využít alespoň jednu z možných hodnot.		
xx<place>		údaje o místě spojeném s vytvořením, vydáním, distribucí nebo výrobou popisovaného dokumentu odpovídá hodnotě 264 \$a	MA	
xxx<placeTerm>		konkrétní určení místa vydání, např. Praha, resp. "xr" pro ČR odpovídá hodnotám z katalogizačního záznamu, pole 264 \$a, resp. pole 008/15-17	M	<dc:coverage>
	type	pokud má dokument více míst vytvoření/vydání/distribuce/výroby v poli 264 \$a, přebírají se ze záznamu všechna místa (v jednom poli 264) "code" pro údaj z pole 008 "text" pro údaj z pole 264 \$a	M	
	authority	hodnota "marccountry" jen u údaje z pole 008	MA	
xx<agent>		informace o entitě, která dokument vytvořila, vydala, distribuovala nebo vyrobila	MA	
xxx<namePart>		jméno entity, která dokument vytvořila, vydala, distribuovala nebo vyrobila odpovídá poli 264 \$b v MARC21 pokud má titul více vydavatelů,/ distributorů/ výrobců, přebírají se ze záznamu všichni (jsou v jednom poli 264)	M	<dc:publisher>
xxx<role>		specifikace role organizace nebo osoby uvedené v elementu <agent>	M	
xxxx<roleTerm>		popis role, možné hodnoty podle druhého indikátoru pole 264: 264_0 "producer" 264_1 "publisher" 264_2 "distributor" 264_3 "manufacturer"	M	

xx<dateIssued>		datum vydání; odpovídá hodnotě z katalogizačního záznamu <dateIssued> lze použít POUZE PRO POLE 264_1 \$c nebo pole 008/07-10 !!! Pro ostatní výskyty v poli 264 \$c využít element <dateOther> s odpovídajícím atributem type. !!! Pro 264_4 použít element <copyrightDate>	M	<dc:date>
	encoding	hodnota "marc" jen u údaje z pole 008	R	
	point	hodnoty "start", resp. "end" jen u údaje z pole 008, pro rozmezí dat	MA	
	qualifier	možnost dalšího upřesnění, hodnota "approximate" pro data, kde nevíme přesný údaj	R	
xx<dateOther>		datum vytvoření, distribuce, výroby; tento element se využije v případě výskytu \$c v poli 264 s identifikátorem 0, 2 nebo 3	R	
	type	264_0: <dateOther type="production"> 264_2: <dateOther type="distribution"> 264_3: <dateOther type="manufacture">	M	
xx<copyrightDate>		vyžije se pouze v případě výskytu pole 264 s druhým indikátorem 4 a podpolem \$c 264_4: <copyrightDate>	R	<dc:date>

x<physicalDescription>		obsahuje údaje o fyzickém popisu disku	MA	
xx<form>		hodnota "computer disc"	M	<dc:format>
xx<extent>		údaje o rozsahu - vyplní se celková velikost dat na konkrétním disku	M	<dc:format>
xx<note>		poznámka o fyzickém stavu disku; pro každou poznámku je nutno vytvořit nový <note> element	RA	
x<note>		slouží k vepisování údajů z polí 5XX, 562, 245; pro každou poznámku musí být samostatný <note> element	RA	<dc:description>
	type	<ul style="list-style-type: none"> pro obsah pole 561 hodnota "ownership" pro obsah pole 562 hodnota "version identification" pro obsah pole 245 hodnota "statement of responsibility" 	MA	

		<ul style="list-style-type: none"> pro obsah pole 508 hodnota "creation/production credits" pro obsah pole 546 \$a hodnota "language" 		
x<subject>		údaje o věcném třídění, předpokládá se přebírání z katalogizačního záznamu	R	
	authority	vyplnit hodnotu "czenas", "eczenas" z pole 072 \$2, nebo jinou hodnotu z tohoto pole; případně lze ponechat prázdné	R	
xx<topic>		libovolný výraz specifikující nebo charakterizující obsah titulu; použít kontrolovaný slovník – např. z báze autorit AUT NK ČR (věcné téma) nebo obsah pole 650 a 653 s druhým indikátorem prázdným či hodnotou 0 záznamu MARC21 nebo obsah pole 072 \$x	M	<dc:subject>
xx<geographic>		geografické věcné třídění; použít kontrolovaný slovník – např. z báze autorit AUT NK ČR (geografický termín) nebo obsah pole 651 a 653 s druhým indikátorem s hodnotou 5 záznamu MARC21	R	<dc:subject>
xx<temporal>		chronologické věcné třídění; použít kontrolovaný slovník – např. z báze autorit AUT NK ČR (chronologický údaj) nebo obsah pole 648 a 653 s druhým indikátorem s hodnotou 4 záznamu MARC21	R	<dc:subject>
xx<name>		jméno použité jako věcné záhlaví použít kontrolovaný slovník – např. z báze autorit AUT NK ČR (jméno osobní) nebo obsah pole 600 a 653 s druhým indikátorem s hodnotou 1, 2 či 3 záznamu MARC21	R	
xxx<namePart>		celé jméno se zapíše do tohoto elementu	R	<dc:subject>
xxx<nameIdentifier>		číslo národní autority	R	
xx<titleInfo>		pole použité jako vedlejší věcné záhlaví	R	
xxx<title>		unifikovaný název	M	<dc:subject>
x<classification>		klasifikační údaje věcného třídění podle Mezinárodního desetinného třídění	R	<dc:subject>
	authority	hodnota "udc"	M	
xx<classification>		klasifikační údaje věcného třídění podle Konspektu	R	<dc:subject>

		odpovídá poli 072 \$a MARC21		
	authority	hodnota "udc" (v případě 072 \$a) hodnota "Konspekt" (v případě 072 \$9)	M	
	edition	hodnota "Konspekt" (v případě 072 \$a)	M	
x<identifier>		údaje o identifikátorech, obsahuje unikátní identifikátory mezinárodní nebo lokální – viz přehled typů atributů níže uvádějí se i neplatné resp. zrušené identifikátory - atribut invalid="yes"	M	<dc:identifier>
	type	budou se povinně vyplňovat následující hodnoty, pokud existují: - "uuid" (M) - jiný interní identifikátor - hodnota atributu "barcode" - čárový kód (MA), podobně lze užít např. "sysno", "oclc", "permalink"	MA	
x<recordInfo>		údaje o metadatovém záznamu disku – jeho vzniku, změnách apod.	M	
xx<descriptionStandard>		popis standardu, ve kterém je přebíraný katalogizační záznam pro záznamy v AACR2: odpovídá hodnotě návěští záznamu MARC21, pozice 18 - hodnota "aacr", tj. pro LDR/18="a" pro záznamy v RDA: odpovídá hodnotě záznamu MARC21 pole 040 a podpole \$e "rda" - bude použito pro primární rozlišení, zda jde o záznam vytvořený podle AACR2 nebo podle RDA	M	
xx<recordContentSource>		kód nebo jméno instituce, která záznam vytvořila nebo změnila	R	
	authority	hodnota "siglaADR"	R	
xx<recordCreationDate>		datum prvního vytvoření záznamu, alespoň na úroveň minut podle normy ISO 8601	M	
	encoding	hodnota atributu "iso8601"		
xx<recordChangeDate>		datum změny záznamu, podle normy ISO 8601 alespoň na úroveň minut	R	
	encoding	hodnota atributu "iso8601"		
xx<recordOrigin>		údaje o vzniku záznamu; hodnoty: "machine generated" nebo "human prepared"	R	

xx<languageOfCataloging>		jazyk katalogového záznamu	R	
xxx<languageTerm>		přebírá se z katalogu - pole 40 \$b	M	
	authority	hodnota "iso639-2b"	R	

7.3.3 Pole MODS a DC pro obrazové informace (přílohy)

[Legenda pro čtení specifikace](#)

Element MODS	Atributy	Popis	Povinnost	Element DC
<mods>	ID	ID obsahuje název úrovně, pro obrazovou přílohu je to "MODS_SUPPL_0001", kde čtyřmístné číslo je pořadové číslo přílohy	M	
	version	verze mods: "3.8"	R	
x<titleInfo>		názvové informace o příloze	M	
	type	hlavní název bez <i>type</i> hodnoty: "abbreviated" "alternative" "translated" "uniform"	MA	
xx<title>		hlavní název	MA	<dc:title>
xx<subTitle>		podnázev	O	<dc:title>
xx<partNumber>		číslo přílohy 01, 02 atd.	MA	<dc:description>
xx<partName>		slovní označení přílohy, označit druh přílohy: "imgdisc", "obal", "booklet"	M	<dc:description>
x<name>		autor přílohy	RA	
	type	Použít jednu z hodnot: <ul style="list-style-type: none"> • "personal" /pro osobu • "corporate" /pro společnost, instituci nebo organizaci • "conference" /pro konference, jednání a jiné pojmenované akce • "family" /rodina, rod 	MA	
xx<namePart>		údaje o křestním jméně a příjmení; pokud nelze rozlišit křestní jméno a příjmení, nepoužije se atribut <i>type</i> 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="data"	MA	<dc:creator> nutno do jednoho pole DC spojit jméno i příjmení

	type	použít jednu z hodnot: <ul style="list-style-type: none"> • "date" /životopisná data autora • "family" /příjmení • "given" / křestní jméno • "termsOfAddress" /akademický titul či hodnost pokud se nejedná o osobu, atribut <i>type</i> se nepoužije	RA MA MA RA	
xx<alternativeName>		kontejnerový element pro zachycení alternativního jména	O	
xxx<namePart>		alternativní jméno, např. pseudonym, přezdívka	M	
xx<nameIdentifier>		číslo národní autority	MA	<dc:creator>
xx<role>		specifikace role osoby nebo organizace uvedené v elementu <name>	MA	
xxx<roleTerm>		popis role	M	
	type	hodnota "code" – kód z kontrolovaného slovníku (http://www.loc.gov/marc/relators/relaterm.html)	M	
	authority	hodnota "marcrelator" – popis role dle slovníku MARC relator	R	
x<typeOfResource>		povolené hodnoty "text" a "still image" použití dle typu přílohy; hodnota "text" je aplikovatelná na typ booklet a cover, hodnota "still image" je aplikovatelná na imgdisc	M	
x<genre>		hodnota "supplement"	M	<dc:type>model :supplement</dc:type>
	type	hodnota podle druhu přílohy povolené hodnoty: "cover", "booklet", "imgdisc"	M	
x<language>		údaje o jazyce dokumentu v případě vícenásobného výskytu nutno element <language> opakovat	M	
xx<languageTerm>		přesné určení jazyka – kódem; nutno použít kontrolovaný slovník ISO 639-2, http://www.loc.gov/standards/iso639-2/php/code_list.php	M	<dc:language>

3.4.3 Pole MODS a DC pro obrazové informace (přílohy)

	type	použít hodnotu "code"	M	
	authority	použít hodnotu "iso639-2b"	M	
x<physicalDescription>		obsahuje údaje o fyzickém popisu přílohy	M	
xx<extent>		počet skenů přílohy	M	<dc:format>
xx<note>		poznámka o fyzickém stavu přílohy; pro každou poznámku je nutno vytvořit nový <note> element	RA	
x<note>		obecná poznámka k příloze	RA	<dc:description>
x<relatedItem>		lze použít k rozlišení, zda jde o přílohu společnou pro celou kolekci (např. obal či booklet společný pro několik disků) nebo přílohu jednotlivého disku (např. sken vrchní strany disku)	R	
	type	"host"	M	
xx<identifier>		identifikátor kolekce dat nebo disku	M	
	type	type="uuid" - uuid kolekce nebo konkrétního disku	M	
x<identifier>		údaje o identifikátorech uvádějí se i neplatné resp. zrušené identifikátory - atribut invalid="yes"	M	<dc:identifier>
	type	povinně se vyplní hodnota: - "uuid" (M)	M	
x<recordInfo>		údaje o metadatovém záznamu – jeho vzniku, změnách apod.	M	
xx<descriptionStandard>		popis standardu, ve kterém je přebíraný katalogizační záznam pro záznamy v AACR2: odpovídá hodnotě návěští záznamu MARC21, pozice 18 - hodnota "acr", tj. pro LDR/18="a" pro záznamy v RDA: odpovídá hodnotě záznamu MARC21 pole 040 a podpole \$e "rda" - bude použito pro primární rozlišení, zda jde o záznam vytvořený podle AACR2 nebo podle RDA	M	
xx<recordContentSource>		kód nebo jméno instituce, která záznam vytvořila nebo změnila; nutno vytvořit	R	

3.4.3 Pole MODS a DC pro obrazové informace (přílohy)

		kontrolovaný slovník		
	authority	hodnota "siglaADR"	R	
xx<recordCreationDate>		datum prvního vytvoření záznamu, alespoň na úroveň minut podle normy ISO 8601	M	
	encoding	hodnota atributu "iso8601"		
xx<recordChangeDate>		datum změny záznamu, podle normy ISO 8601 alespoň na úroveň minut	R	
	encoding	hodnota atributu "iso8601"		
xx<recordOrigin>		údaje o vzniku záznamu; hodnoty: "machine generated" nebo "human prepared"	R	
xx<languageOfCataloging>		jazyk katalogového záznamu	R	
xxx<languageTerm>		přebírá se z katalogu - pole 40 \$b	M	
	authority	hodnota "iso639-2b"	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 náležející k disku 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 u obrazových souborů 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 a procesy, které byly na datech i metadatach provedeny
- technická metadata přicházející z digitalizace jsou dále v maximální míře ukládána v LTP systému (po konverzi do interního formátu LTP systému)
- všechny PREMIS či MIX záznamy budou ve vedlejším METS záznamu (amd_mets.xml a amd_mets_isoimg.xml), který je určen pro administrativní a technická metadata
- tento vedlejší METS záznam (amd_mets.xml a amd_mets_isoimg.xml) je linkován z hlavního METS záznamu dokumentu (v části <fileSec>)
- plnění technických metadat se předpokládá z výstupů vzniklých využitím služeb třetích stran jakou jsou DROID, JHOVE, PRONOM, ISOLYZER aj.
- <amdSec> část bude existovat vždy jedna pro všechny reprezentace obrazové stránky dokumentu (MC, UC, u obrazu ALTO XML a OCR.txt) či jedné ISO image 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	identifikátor konkrétní části <amdSec> např. pro jeden ISO image „ISOIMG0001“, obrazovou stránku 1 hodnota "PAGE0001", pro obrazovou stránku 2 hodnota "PAGE0002" atd.	M
x<techMD> nebo x<digiprovMD>		element rozlišující typy jednotlivých administrativních metadat	M
	ID	ID pro část <techMD>: pro částí obsahující PREMIS Object hodnota: <ul style="list-style-type: none"> • "OBJ_001" pro původní smazaný soubor TIFF u obrazu a pro isoimg • "OBJ_002" pro MC • "OBJ_003" pro ALTO XML u obrazu počet PREMIS Object není omezen, číslování pokračuje OBJ_004 atd.	M

		<p>pro části obsahující MIX hodnota:</p> <ul style="list-style-type: none"> • "MIX_001" pro původní smazaný obrazový soubor, např. TIFF • "MIX_002" pro MC <p>číselná část ID v MIX odpovídá číselné části ID pro PREMIS Object, další záznamy MIX pro novou obrazovou reprezentaci stránky tedy přebírají číslování z techMD pro PREMIS Object (např. pro OBJ_004 v PREMIS Object by to byl MIX_004)</p> <p>ID pro část <digiprovMD>:</p> <p>pro části obsahující PREMIS Event hodnota:</p> <ul style="list-style-type: none"> • "EVT_001" atd. <p>pro části obsahující PREMIS Agent hodnota:</p> <ul style="list-style-type: none"> • "AGENT_001" atd. 	
xx<mdWrap>		element obsahující vložené záznamy PREMIS či MIX	M
	MDTYPE	pro záznamy PREMIS object, event i agent vždy hodnota "PREMIS" pro záznamy MIX hodnota "NISOIMG"	M
xxx<xmlData>		kontejnerový element, do kterého se vkládají záznamy PREMIS, resp. MIX	M

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:
 - ISO image vzniklý v procesu grabování, který se archivuje,
 - původní obraz (PS - primary scan) ve formátu TIFF, který se maže
 - archivní kopie obrazové stránky (MC - master copy) vzniklá migrací PS do JPEG 2000.
 - ALTO XML u obrazu
 - případné další reprezentace obrazové stránky (například nová archivní kopie vytvořená migrací z původní MC)
- PREMIS Object se nebude vytvářet pro UC (obraz) ani pro OCR.TXT u obrazových souborů
- 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 Event ve stejném METS záznamu konkrétního dokumentu v části <digiprovMD> přes <premis:relatedEventIdentification>; to samé platí pro objekty, které budou nalinkovány v případě vztahu (např. MC vzniklá z PS) s popisovaným objektem přes <premis:relatedObjectIdentification>.
 - tj. např. PREMIS Object popisující ISO image soubor je tímto způsobem nalinkován na původní data (resp. na jejich PREMIS Object záznam) – pomocí tagu <relatedObjectIdentification>, který obsahuje ID původního objektu
 - zároveň pomocí tagu <relatedEventIdentification> je záznam PREMIS Object souboru ISO image nalinkován na událost, během které vznikl
- POZOR – PREMIS Object bude vznikat a uchovávat se i pro neexistující data (např. původní a posléze smazaný TIFF u obrazových dokumentů)

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, opakovatelný
 - 1-1 element je povinný, neopakovatelný

Obsah pole "Použití pro":

- použití jednotlivých elementů pro popis MC, PS (původní sken), XML (ALTO) a ISO image.

Element	Popis	Použití pro	Povinnost
<object>	kořenový element pro PREMIS Object; použít vždy s atributem dle typu objektu xsi:type="file" – pro soubor 1-n	MC, XML, PS, ISOIMG	M
x<objectIdentifier>	identifikátor k jednoznačnému odlišení objektu v určitém kontextu; 1-n	MC, XML, PS, ISOIMG	M
xx<objectIdentifierType>	popis kontextu, ve kterém je identifikátor unikátní, např. „NDK“, „ANL“ nebo název repozitáře; nutno použít kontrolovaný slovník;	MC, XML, PS, ISOIMG	M

	1-1		
xx<objectIdentifierValue>	vlastní hodnota identifikátoru, např. „img0001-master“, „uuid:123465“ apod.; 1-1	MC, XML, PS, ISOIMG	M
x<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, ISOIMG	MA
xx<preservationLevelValue>	hodnota úrovně ochrany, která je pro soubor relevantní, pro původní sken PS hodnota "deleted", pro MC, XML a ISO image hodnota "preservation"; 1-1	MC, XML, PS, ISOIMG	M
xx<preservationLevelDateAssigned>	datum, kdy byla přiřazena hodnota úrovně ochrany, zápis v ISO 8601, na úroveň dne (RRRR-MM-DD) 0-1	MC, XML, PS, ISOIMG	R
x<objectCharacteristics>	technické údaje o souboru 1-n	MC, XML, PS, ISOIMG	M
xx<compositionLevel>	údaj o tom, zda je nutné digitální objekt rozbalit nebo dekódovat; např. "0" (defaultně pro žádné zabalení nebo kódování); "1" pro jedno zabalení a kódování, podobně pak hodnota "2"; 1-1	MC, XML, PS, ISOIMG	M
xx<fixity>	údaje o kontrolním součtu 0-n	MC, XML, PS, ISOIMG	M
xxx<messageDigestAlgorithm>	použitý algoritmus kontrolního součtu, např. „MD5“ aj. 1-1	MC, XML, PS, ISOIMG	M
xxx<messageDigest>	hodnota kontrolního součtu 1-1	MC, XML, PS, ISOIMG	M
xxx<messageDigestOriginator>	agent (osoba, instituce, stroj, SW), který kontrolní součet vytvořil (např. „JHOVE“ apod.) 0-1	MC, XML, PS, ISOIMG	M
xx<size>	údaje o velikosti souboru v bytech 0-1	MC, XML, PS, ISOIMG	M
xx<format>	údaje o formátu souboru pro soubory ALTO XML je možné vytvořit element	MC, XML, PS,	M

	dvakrát, jednou popisuje formát XML, podruhé obsahuje informace o použitém standardu ALTO 2.0 1-n	ISOIMG	
xxx<formatDesignation>	identifikace formátu souboru, výstup z JHOVE, PRONOM služeb apod. 0-1	MC, XML, PS, ISOIMG	M
xxxx<formatName>	jméno formátu dle MIME type (dle řízeného slovníku), "image/tiff" "image/jp2" "text/xml" "application/x-iso9660-image" 1-1	MC, XML, PS, ISOIMG	M
xxxx<formatVersion>	verze formátu, např. „6.0“; v případě formátu TIFF se zaznamená číslo revize formátu (obvykle "5.0" nebo "6.0") 0-1	MC, XML, PS, ISOIMG	M
xxx<formatRegistry>	identifikace formátu – dodatečná informace o záznamu formátů v registrech formátů (PRONOM) 0-1	MC, XML, PS, ISOIMG	M
xxxx<formatRegistryName>	jméno použitého registru formátů, vždy PRONOM. 1-1	MC, XML, PS, ISOIMG	M
xxxx<formatRegistryKey>	unikátní identifikátor (označení) formátu v registru, povinné použít PUID z registru PRONOM, další registry možné 1-1	MC, XML, PS, ISOIMG	M
xx<creatingApplication>	údaje o aplikaci, ve které byl popisovaný soubor vytvořen, element se používá pro software, tj. SW, kde vzniklo ALTO XML/TXT, SW/kodek pro vytvoření JPEG 2000, aplikace, která vytvořila ISO image apod. každá aplikace zvlášť v novém elementu 0-n	MC, XML, PS, ISOIMG	M
xxx<creatingApplicationName>	název aplikace, např. CDArcha, ImageGear, Kakadu apod.;; 0-1	MC, XML, PS, ISOIMG	M
xxx<creatingApplicationVersion>	verze aplikace, např. „15.03.000“ 0-1	MC, XML, PS, ISOIMG	M
xxx<dateCreatedByApplication>	datum a čas vytvoření, např. „2008-11-	MC, XML,	M

	10T12:37:46“; musí být ve tvaru ISO 8601 (na úroveň vteřin); např. 2008-11-10T12:37:46; 0-1	PS, ISOIMG	
x<originalName>	původní jméno souboru, např. „123456789_01.iso“ 0-1	MC, XML, PS, ISOIMG	M
x<relationship>	vyjádření vztahu popisovaného souboru k jiným souborům a událostem (eventům) 0-n	MC, XML, ISOIMG	M
xx<relationshipType>	typ vztahu, doporučené hodnoty: derivation = vztah, kde objekt je výsledkem změny jiného objektu; structural = vztah mezi částmi objektu; tj. např. uživatelská kopie skenu bookletu vytvořená z původního skenu bude mít vztah "derivation"; 1-1	MC, XML, ISOIMG	M
xx<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. uživatelská kopie skenu bookletu vytvořená z původního skenu budou mít vztah "created from"; 1-1	MC, XML, ISOIMG	M
xx<relatedObjectIdentification>	identifikace souvisejícího souboru 1-n pro MC, XML pro vyjádření vztahu k původnímu objektu	MC, XML, ISOIMG	M
xxx<relatedObjectIdentifierType>	specifikace kontextu, ve kterém je identifikátor souboru jedinečný, např. URN; temporary filepath; objectID 1-1	MC, XML, ISOIMG	M
xxx<relatedObjectIdentifierValue>	vlastní řetězec identifikátoru, např. „uuid:123456“ nebo název souboru, cesta k souboru apod. 1-1	MC, XML, ISOIMG	M
xx<relatedEventIdentification>	identifikace s popisovaným souborem související události (eventu); seznam událostí viz PREMIS event 0-n	MC, XML, ISOIMG	M
xxx<relatedEventIdentifierType>	typ události, např. interní číslovací systém událostí jako no.nb.evt; NK repository event ID, UUID apod. 1-1 (“NK_eventID”)	MC, XML, ISOIMG	M
xxx<relatedEventIdentifierValue>	hodnota identifikátoru události, např. „NK_EVT_00“5 nebo hodnota UUID aj.	MC, XML, ISOIMG	M

	1-1 (pro vytvoření (grabování) ISO img hodnota "CDArcha_001")		
xxx<relatedEventSequence>	pořadí události, např. „003“; k určení pořadí lze určit datum události 0-1	MC, XML, ISOIMG	R
x<linkingEventIdentifier>	identifikátor události týkající se původního skenu PS nebo grabování datového disku; typy událostí mohou být např. vytvoření, smazání aj. 0-n pro PS nutný link na události vytvoření (digitalizace) a jeho vymazání	PS ISO IMG	M
xx<linkingEventIdentifierType>	typ identifikátoru události, např. UUID, NK_eventID, vlastní číslovací systém apod. 1-1	PS ISO IMG	M
xx<linkingEventIdentifierValue>	hodnota identifikátoru, např. „masterCopy_001“, „grab_001“ apod. 1-1	PS ISO IMG	M

7.4.2 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 datech:
 - grabování – vytvoření iso image
 - digitalizace – vytvoření původního skenu
 - vygenerování MC
 - vytvoření ALTO XML pro obrazová data
 - vymazání PS v případě obrazových dat
- 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 a amd_mets_isoimg.xml) v jeho části <amdSec>, podčást <digiprovMD>
 - amd_mets.xml a amd_mets_isoimg.xml jsou linkovány z hlavního METS záznamu dokumentu (v části <fileSec>)
- 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

Element	Popis	Povinnost
<event>	kořenový element pro PREMIS Event	M
x<eventIdentifier>	údaje o identifikátoru události v kontextu digitalizace nebo repozitáře 1-1	M
xx<eventIdentifierType>	typ identifikátoru, např. no.nb.evt; NK_eventID, UUID apod. 1-1	M
xx<eventIdentifierValue>	hodnota identifikátoru, např. „EVT_001“; „event_019“ apod. 1-1	M
x<eventType>	kategorizace události, nutno použít kontrolovaný slovník; typy událostí, které musí být zaznamenány: "capture", "migration", "derivation", "deletion" 1-1	M
x<eventDateTime>	datum a čas kdy byla událost provedena; nutno zapsat v ISO 8601 na úroveň vteřin 1-1	M
<eventDetail>	další údaje o události, doporučené hodnoty pro výše uvedené <eventType> následují za /:	M

	<ul style="list-style-type: none"> - capture/iso_creation – vznik iso image - capture/digitization – vznik PS (původního skenu u obrazových příloh) - capture/XML_creation - capture/TXT_creation - migration/MC_creation - deletion/PS_deletion <p>0-1</p>	
x<eventOutcomeInformation>	informace o výsledku události 0-n	R
xx<eventOutcome>	kategorizace výsledku události, např. slovy jako „successful“, „failure“; možno použít kódy dle kontrolovaného slovníku nebo seznamu kódů; 0-1	M
x<linkingAgentIdentifier>	identifikace jednoho nebo více agentů spojených s událostí 0-n	M
xx<linkingAgentIdentifierType>	označení typu identifikátoru, např. „NK_AgentID“, UUID apod. 1-1	M
xx<linkingAgentIdentifierValue>	hodnota identifikátoru, např. „agent_softwareName_5.2“; „agent_novakJ“ apod. 1-1	M
xx<linkingAgentRole>	role agenta ve vztahu k události, např. „software“; „SW component“; „operator“; bude sloužit pro zápis hardware u zvukových dokumentů (tj. role=player nebo AD nebo extraction workstation nebo preamp apod.) nutno používat kontrolovaný slovník 0-n	R
x<linkingObjectIdentifier>	informace o objektu/souboru spojeného s událostí, link na něj 0-n	M
xx<linkingObjectIdentifierType>	označení typu identifikátoru, např. hodnota "file" 1-1	M
xx<linkingObjectIdentifierValue>	hodnota identifikátoru, např. „URN:NBN:cz-_0011#0001“ aj. nebo „PS_1_0003_R“ 1-1	M

7.4.3 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 SIP balíčku jsou dostačující a dají nám dostatečné informace o události, kdy byla provedena a na jakém SW (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 Event (<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 a amd_mets_isoimg.xml) v jeho části <amdSec> a podčásti <digiprovMD>
 - amd_mets.xml a amd_mets_isoimg.xml jsou linkovány 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
<agent>	kořenový element pro PREMIS Agent 0-n	M
x<agentIdentifier>	popis identifikátoru, který jednoznačně označuje agenta v rámci jednoho kontextu (repozitář např.) 1-n	M
xx<agentIdentifierType>	označení typu identifikátoru, např. NK_AgentID, UUID apod. 1-1	M
xx<agentIdentifierValue>	hodnota identifikátoru, např. „agent_softwareName_5.2“; „agent_novakJ“ apod. 1-1	M
x<agentName>	textové upřesnění agenta, např. přesný název SW, plné jméno osoby apod. – „FixImage1.3“; „Jan Novák“; „CCS docWorks 6.2.1“; 0-n	R
x<agentType>	obecné označení agenta – pro osoby např. osoba, pro SW např. software apod. hodnoty: "organization"; "person"; "software", "hardware" 0-1	M

x<agentNote>	použití pouze pokud je <agentType> Software a půjde o agenta souvisejícího s migrací původního souboru např. TIFF na JPEG 2000 (creation/migration Event); bude obsahovat příkaz k výrobě JPEG 2000 souboru v aplikaci Kakadu nebo OpenJpeg 0-n	MA
---------------------------	--	----

7.4.4 Technická metadata MIX

- MIX záznam vzniká pouze pro obrazové soubory týkající se obalu, bookletu a skenu disků
- bude vznikat pro:
 - archivní obrazovou kopii
 - původní sken 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> a podčásti <techMD>) pro administrativní a technická metadata, který vznikne ke každému obrazovému souboru a který je linkován z hlavního METS záznamu
- 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.
- pro každý záznam MIX bude vytvořena vlastní část <techMD>
- MIX může být také zapouzdřen v PREMIS Object <premis:objectCharacteristicsExtension>
- externí služby jako JHOVE nebo 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
- <fileSize> je pouze doporučené, údaj je již součástí popisu u 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
 - 1-1 element je povinný, 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>			
x<ObjectIdentifier>	údaje o identifikátoru obrazového dokumentu, který je formátem MIX popsán; 0-n	R	MC, PS
xx<objectIdentifierType>	např. jméno souboru, nebo jiný identifikátor; 0-1	M	MC, PS
xx<objectIdentifierValue>	hodnota identifikátoru, např. „20110306_001.jp2“ nebo „urn:nbn:123456“; 0-1	M	MC, PS
x<fileSize>	velikost souboru 0-1	R	MC, PS

x<FormatDesignation>	údaje o formátu obrazového souboru 0-1	M	MC, PS
xx<formatName>	název formátu, např. lze využít MIME types („image/jp2“ apod.) 0-1	M	MC, PS
xx<formatVersion>	verze formátu, např. „1.0“ 0-1	M	MC, PS
x<byteOrder>	endianita, možnosti jsou little endian, middle (mix) endian a big endian 0-1	M	MC, PS
x<Compression>	údaje o kompresi obrazového souboru 0-n	M	MC, PS
xx<compressionScheme>	informace o kompresním schématu, vyjádřeno číslem (např. „34712“ je komprese JPEG 2000) nebo slovy (např. „JP2 Lossless“) 0-1	M	MC, PS
<BasicImageInformation>	základní technické údaje o obrazovém dokumentu 0-1	M	MC, PS
x<BasicImageCharacteristics>	0-1	M	MC, PS
xx<imageWidth>	šířka obrazu v pixelech, např. „3987“ 0-1	M	MC, PS
xx<imageHeight>	výška obrazu v pixelech, např. „2345“ 0-1	M	MC, PS
xx<PhotometricInterpretation>	fotometrická interpretace 0-1	M	MC, PS
xxx<colorSpace>	barevný prostor, např. „RGB“ 0-1	M	MC, PS
xxx<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
xxxx<IccProfile>	ICC profil 0-1	M	MC, PS
xxxxx<iccProfileName>	jméno profilu, např. „sRGB“, „Adobe RGB“ aj. 0-1	M	MC, PS
xxxxx<iccProfileVersion>	verze profilu, např. „sRGB IEC61966-2.1“ 0-1	M	MC, PS
xxxxx<iccProfileURI>	identifikátor profilu	R	MC, PS

	0-1		
x<SpecialFormatCharacteristics>	speciální technické údaje o obrazovém dokumentu, povinné použití pro formát JPEG 2000 0-1	MA	MC
xx<JPEG2000>	0-1	M	MC
xxx<CodecCompliance>	údaje o kodeku 0-1	M	MC
xxxx<codec>	název kodeku, např. „Kakadu“, „LuraWave“ aj. 0-1	M	MC
xxxx<codecVersion>	verze kodeku, např. „3.1“ 0-1	M	MC
xxxx<codestreamProfile>	popis codestream profilu JPEG 2000, např. „P0“ a „P1“ (viz ISO/IEC 15444-4); 0-1	M	MC
xxxx<complianceClass>	specifikace největší výšky, šířky a počtu komponentů, které dekodér dokáže dekodovat, lze použít hodnoty „C0“, „C1“ a „C2“; 0-1	M	MC
xxx<EncodingOptions>	obsahuje informace o kodování JPEG 2000 0-1	M	MC
xxxx<Tiles>	popis pixelové velikosti dlaždic formátu JPEG 2000 0-1	M	MC
xxxxx<tileWidth>	šířka dlaždice, např. „128“ 0-1	M	MC
xxxxx<tileHeight>	výška dlaždice, např. „128“ 0-1	M	MC
xxxx<qualityLayers>	číselná hodnota počtu vrstev, do kterých byl JPEG 2000 rozdělen, např. „12“ 0-1	M	MC
xxxx<resolutionLevels>	popis počtu nižších rozlišení, které lze z obrazu získat, např. „6“ 0-1	M	MC
<ImageCaptureMetadata>	popis procesu skenování, je důležité vyplnit, protože tyto údaje nelze zjistit z finálního master/archivního souboru 0-1	M	PS
x<SourceInformation>	informace o předloze 0-1	R	PS
xx<sourceType>	Booklet, Cover aj.; nutno používat kontrolovaný slovník 0-1	M	PS

xx<SourceID>	identifikátor předlohy 0-n	R	PS
xxx<sourceIDType>	typ identifikátoru, např. „URN:NBN“ 0-1	M	PS
xxx<sourceIDValue>	vlastní hodnota identifikátoru 0-1	M	PS
x<GeneralCaptureInformation>	základní údaje o skenování 0-1	M	PS
xx<dateTimeCreated>	údaj o datu a čase skenování, např. „2009-01-03T08:25:28“; zapsat v ISO 8601 na úroveň vteřin 0-1	M	PS
xx<imageProducer>	entita provádějící skenování, např. „The National Library of the Czech Republic“, osoba apod. 0-n	M	PS
xx<captureDevice>	typ skenovacího zařízení použít jedno z hodnot: <ul style="list-style-type: none"> - "reflection print scanner" - "transmission scanner" - "digital still camera" - "still from video" 0-1	MA	PS
x<ScannerCapture>	údaje o skeneru 0-1	MA	PS
xx<scannerManufacturer>	výrobce skeneru, např. „4DigitalBooks“, „Treventus“, „Zeutschel“ 0-1	M	PS
xx<ScannerModel>	údaje o konkrétním typu skeneru 0-1	M	PS
xxx<scannerModelName>	jméno modelové řady skeneru, např. „DL“ 0-1	M	PS
xxx<scannerModelNumber>	číslo/označení modelu, např. „3000“ 0-1	M	PS
xxx<scannerModelSerialNo>	výrobní číslo skeneru, např. „E4R0003649“ 0-1	M	PS
xx<MaximumOpticalResolution>	údaje o maximálním optickém rozlišení skeneru 0-1	M	PS
xxx<xOpticalResolution>	optické rozlišení na ose x, pouze číselné vyjádření např. „300“ 0-1	M	PS

xxx<yOpticalResolution>	optické rozlišení na ose y, pouze číselné vyjádření např. „300“ 0-1	M	PS
xxx<opticalResolutionUnit>	jednotka optického rozlišení, jedna z hodnot: "no absolute unit"; "in. "; "cm" 0-1	M	PS
xx<scannerSensor>	popis typu snímacího senzoru skenovacího zařízení, jedna z hodnot: "undefined"; "MonochromeLinear"; "ColorTriLinear"; "ColorSequentialLinear"; "MonochromeArea"; "OneChipColourArea"; "TwoChipColorArea"; "ThreeChipColorArea"; "ColorSequentialArea"; 0-1	M	PS
xx<ScanningSystemSoftware>	údaje o softwaru skenovacího zařízení 0-1	M	PS
xxx<scanningSoftwareName>	název softwaru, např. „Copinet“ 0-1	M	PS
xxx<scanningSoftwareVersionNo>	číslo verze softwaru, např. „3.7“ 0-1	M	PS
x<DigitalCameraCapture>	údaje o snímacím zařízení (fotoaparát) 0-1 povinné, pokud je používán fotoaparát a není používán skener	MA	PS
xx<digitalCameraManufacturer>	výrobce fotoaparátu, např. „Canon“ 0-1	M	PS
xx<DigitalCameraModel>	popis modelu fotoaparátu 0-1	M	PS
xxx<digitalCameraModelName>	název modelové řady, např. „EOS“ 0-1	M	PS
xxx<digitalCameraModelNumber>	označení modelu fotoaparátu, např. „1000D“ 0-1	M	PS
xxx<digitalCameraModelSerialNo>	výrobní číslo přístroje, např. „E12345“ 0-1	M	PS
xx<cameraSensor>	typ senzoru fotoaparátu, např. „matrix“ aj. 0-1	M	PS
xx<CameraCaptureSettings>	údaje o nastavení fotoaparátu použitého ke snímání předloh 0-1	M	PS
xxx<ImageData>	v rámci tohoto kontejnerového elementu budou použity následující sub-elementy:	M	PS

	<p><fNumber> <exposureTime> <isoSpeedRatings> <shutterSpeedValue> <apertureValue> <brightnessValue> <exposureBiasValue> <maxApertureValue> <subjectDistance> <meteringMode> <lightSource> <flash> <focalLength> <backLight> <exposureIndex> <sensingMethod> <cfaPattern> <autoFocus> <PrintAspectRatio></p> <p>všechny hodnoty budou přebrány v případě použití fotoaparátu z údajů Exif</p>		
x<orientation>	<p>popis orientace obrazu tak, jak je uložen vzhledem k jeho řádkům a sloupcům, např. "normal*"; "normal, image flipped"; "normal, rotated 180°"; "unknown" apod.</p> <p>0-1</p>	M	PS
<ImageAssessmentMetadata>	<p>informace o digitálním obrazu pro jeho hodnocení a využití z hlediska dlouhodobé ochrany apod.</p> <p>0-1</p>	M	MC, PS
x<SpatialMetrics>	<p>rozměry obrázku, 2 rozměrná projekce objektů tak jak ji „vidí“ snímací zařízení</p> <p>0-1</p>	M	MC, PS
xx<samplingFrequencyPlane>	<p>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);</p> <p>0-1</p>	R	MC, PS
xx<samplingFrequencyUnit>	<p>jednotka měření sampl. frekvence</p> <p>hodnoty: "no absolute unit of measurement"; "in."; "cm";</p> <p>0-1</p>	M	MC, PS
xxx<xSamplingFrequency>	<p>údaje o počtu pixelů na jednotku sámkovací frekvence pro šířku obrázku</p> <p>0-1</p> <p>povinné, pokud hodnota samplingFrequencyUnit je "in." nebo "cm"</p>	MA	MC, PS

xxxx<numerator>	čítatel, číselné vyjádření, např. „300“ 0-1	M	MC, PS
xxxx<denominator>	jmenovatel, číselné vyjádření např. „1“ 0-1	M	MC, PS
xxx<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
xxxx<numerator>	čítatel, číselné vyjádření, např. „300“ 0-1	M	MC, PS
xxxx<denominator>	jmenovatel, číselné vyjádření např. „1“ 0-1	M	MC, PS
x<ImageColorEncoding>	doplňující údaje o barvě obrazu 0-1	M	MC, PS
xx<BitsPerSample>	počet bitů na kanál 0-1	M	MC, PS
xxx<bitsPerSampleValue>	hodnota počtu bitů, např. 8, 1, 4 nebo 8,8,8 apod. 1-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
xxx<bitsPerSampleUnit>	specifikace jednotky, např. "integer" nebo "floating point" 0-1	R	MC, PS
xx<samplesPerPixel>	počet barevných komponentů na pixel, např. 1, 3, 4 0-1	M	MC, PS
x<TargetData>	informace o kalibračních tabulkách (calibration target); povinné pro obrazy, kde se dělá kontrola oproti kalibrační tabulce 0-1	MA	MC
xx<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
xx<TargetID>	údaje o původu kalibrační tabulky 0-n	M	MC

xxx<targetManufacturer>	výrobce/původce kalibrační tabulky, např. Eastman Kodak nebo NK ČR, oddělení kontroly kvality apod. 0-1	M	MC
xxx<targetName>	název kalibrační tabulky, např. „ColorChecker“, „MicrofilmScanTarget“ aj. 0-1	M	MC
xxx<targetNo>	číslo nebo verze kalibrační tabulky 0-1	M	MC
xxx<targetMedia>	údaj o tom, na jakém médiu je kalibrační tabulka, např. „film“, „paper“ aj. 0-1	R	MC
xx<externalTarget>	údaje o externí kalibrační tabulce; např. link na http://skenservis.cz/target-00000001 nebo název a cesta ke konkrétnímu souboru 0-n povinné v případě, že byla použita externí kalibrační tabulka (targetType = 0)	MA	MC
xx<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>	dokumentace procesů provedených na obrazovém souboru v jeho životním cyklu 0-1	M	MC
x<ImageProcessing>	údaje o zpracování obrazového souboru 0-n	M	MC
xx<dateTimeProcessed>	čas v ISO 8601 na úrovni vteřin, př. „2009-01-04T15:12:06“ 0-1	M	MC
xx<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
xx<processingAgency>	Instituce, př. „The National Library of the Czech Republic“ 0-n	R	MC

7.4.5 Technická metadata File Systémů

7.4.5.1 Technická metadata pro souborový systém ISO 9660

Tento záznam vzniká pouze pro disk image s file systémem ISO 9660. Metadata pro file systém UDF jsou popsána v [kapitole 7.4.5.2](#).

Záznam bude součástí METS záznamu amd_mets_isoimg.xml (v části <amdSec> a podčásti <techMD>) pro administrativní a technická metadata. Pro záznam bude vytvořena vlastní část <techMD>, která bude umístěna za záznamy PREMIS.

Plnění je předpokládáno z výstupu programu Isolyzer. U položky <originalFileSystem> je plnění předpokládáno z programu CD Archa.

Obsah pole "Popis":

- vysvětlení a příklad
- doporučené plnění tam, kde je to možné
- výskyt elementu
 - 0-1 element je nepovinný, neopakovatelný
 - 1-1 element je povinný, neopakovatelný
 - 0-n element je nepovinný, opakovatelný

Element	Popis	Povinnost	Použití pro
<BasicDigitalObjectInformation>			
x<toolInfo>	údaje o programu použitém pro analýzu disk image 1-1	M	ISO
xx<toolName>	název programu 1-1	M	ISO
xx<toolVersion>	verze programu 1-1	M	ISO
x<image>	údaje o disk image 1-1	M	ISO
xx<fileInfo>	obecné údaje o disk image 1-1	M	ISO
xxx<fileName>	název souboru 1-1	M	ISO
xxx<fileSizeInBytes>	velikost souboru 1-1	M	ISO
xx<fileSystems>	údaje o file systému disk image 1-1	M	ISO
xxx<fileSystem>	pouze ISO 9660 (UDF má vlastní metadatové	M	ISO

	schéma) 1-1		
xxx<originalFileSystem>	souborový systém původního optického disku, z něhož je image pořizován 1-1	M	ISO
xxxx<primaryVolumeDescriptor>		M	ISO
xxxxx<volumeIdentifier>	identifikátor původního disku 1-1	M	ISO
xxxxx<volumeSpaceSize>	počet logických bloků 1-1	M	ISO
xxxxx<logicalBlocSize>	velikost logického bloku 1-1	M	ISO
xxxxx<pathTableSize>	velikost path table 1-1	M	ISO
xxxxx<typeLPathTableLocation>	LBA (logical block addressing) umístění path table (little endian) 0-1	MA	ISO
xxxxx<optionalTypeLPathTableLocation>	alternativní LBA umístění path table (little endian) 0-1	MA	ISO
xxxxx<typeMPathTableLocation>	LBA umístění path table (big endian) 0-1	MA	ISO
xxxxx<optionalTypeMPathTableLocation>	alternativní LBA umístění path table (big endian) 0-1	MA	ISO
xxxxx<dataPreparerIdentifier>	program použitý na přípravu dat pro výrobu disku 0-1	MA	ISO
xxxxx<applicationIdentifier>	program použitý na vypálení disku 0-1	MA	ISO
xxxxx<volumeCreationDateAndTime>	datum a čas vytvoření originálu 1-1	M	ISO
xxxxx<volumeModificationDateAndTime>	datum a čas poslední úpravy originálu 0-1	MA	ISO

7.4.5.2 Technická metadata pro souborový systém UDF

Tento záznam vzniká pouze pro disk image s file systémem UDF. Metadata pro file systém ISO 9660 jsou popsána v [kapitole 7.4.5.1](#).

Záznam bude součástí METS záznamu amd_mets_isoimg.xml (v části <amdSec> a podčásti <techMD>) pro administrativní a technická metadata. Pro záznam bude vytvořena vlastní část <techMD>, která bude umístěna za záznamy PREMIS.

Plnění je předpokládáno z výstupu programu Isolyzer. U položky <originalFileSystem> je plnění předpokládáno z programu CD Archa.

Obsah pole "Popis":

- vysvětlení a příklad
- doporučené plnění tam, kde je to možné
- výskyt elementu
 - 0-1 element je nepovinný, neopakovatelný
 - 1-1 element je povinný, neopakovatelný
 - 0-n element je nepovinný, opakovatelný

Element	Popis	Povinnost	Použití pro
<BasicDigitalObjectInformation>			
x<toolInfo>	údaje o programu použitém pro analýzu disk image 1-1	M	ISO
xx<toolName>	název programu (vždy „Isolyzer“) 1-1	M	ISO
xx<toolVersion>	verze programu (př. „1.4.0“) 1-1	M	ISO
x<image>	údaje o disk image 1-1	M	ISO
xx<fileInfo>	obecné údaje o disk image 1-1	M	ISO
xxx<fileName>	název disk image 1-1	M	ISO
xxx<fileSizeInBytes>	velikost souboru 1-1	M	ISO
xx<fileSystems>	údaje o file systému disk image 1-1	M	ISO
xxx<fileSystem>	souborový systém – zde pouze UDF (ISO 9660 má vlastní metadatové schéma) 1-1	M	ISO

xxx<originalFileSystem>	souborový systém originálního disku (plnění z CDArcha) 1-1	M	ISO
xxxx<logicalVolumeDescriptor>		M	ISO
xxxxx<tagIdentifier>	identifikátor deskriptoru (vždy 6) 1-1	M	ISO
xxxxx<descriptorVersion>	verze deskriptoru 1-1	M	ISO
xxxxx<tagSerialNumber>	pořadové číslo značky 1-1	M	ISO
xxxxx<volumeSequenceNumber>	Pořadové číslo deskriptoru 1-1	M	ISO
xxxxx<logicalVolumeIdentifier>	název původního disku 1-1	M	ISO
xxxxx<logicalBlockSize>	velikost logického bloku v bytech 1-1	M	ISO
xxxxx<domainIdentifier>	identifikátor souboru pravidel 1-1	M	ISO
xxxxx<mapTableLength>	délka polí map oddílů/sektorů v bytech 1-1	M	ISO
xxxxx<numberOfPartitionMaps>	počet map oddílů/sektorů 1-1	M	ISO
xxxxx<implementationIdentifier>	program použitý na přípravu dat pro výrobu disku 1-1	M	ISO
xxxxx<integritySequenceExtentLength>	délka integrity sequence 1-1	M	ISO
xxxxx<integritySequenceExtentLocation>	umístění Logical volume integrity descriptor 1-1	M	ISO
xxxx<logicalVolumeIntegrityDescriptor>	1-1	M	ISO
xxxxx<tagIdentifier>	identifikátor deskriptoru (vždy 9) 1-1	M	ISO
xxxxx<descriptorVersion>	verze deskriptoru 1-1	M	ISO
xxxxx<tagSerialNumber>	pořadové číslo značky (specifikace sady deskriptorů) 1-1	M	ISO
xxxxx<timeStamp>	časová značka záznamu (čas uložení)	M	ISO

	deskriptoru) 1-1		
xxxxx<integrityType>	typ integrity deskriptoru, (otevřené nebo uzavřené médium) 1-1	M	ISO
xxxxx<numberOfPartitions>	počet partition v přidruženém logickém svazku 1-1	M	ISO
xxxxx<lengthOfImplementationUse>	délka pole implementačního užití v bytech 1-1	M	ISO
xxxxx<freeSpaceTable>	počet volných bloků 1-1	M	ISO
xxxxx<sizeTable>	velikost partition v blocích 1-1	M	ISO
xxxx<partitionDescriptor>			
xxxxx<tagIdentifier>	identifikátor deskriptoru (vždy "5") 1-1	M	ISO
xxxxx<descriptorVersion>	verze deskriptoru 1-1	M	ISO
xxxxx<tagSerialNumber>	pořadové číslo značky 1-1	M	ISO
xxxxx<volumeDescriptorSequenceNumber>	pořadové číslo tohoto deskriptoru 1-1	M	ISO
xxxxx<partitionNumber>	numerický identifikátor partition 1-1	M	ISO
xxxxx<accessType>	povolené metody přístupu k sektorům partition, kterou popisuje tento deskriptor 1-1	M	ISO
xxxxx<partitionStartingLocation>	číslo logického sektoru, kterým partition začíná 1-1	M	ISO
xxxxx<partitionLength>	počet logických sektorů z nichž partition sestává 1-1	M	ISO

7.5 METS část <fileSec>

Část <fileSec> bude obsahovat výčet všech souborů, které patří do kolekce dat.

7.5.1 <fileSec> hlavního záznamu METS

file group

- pro soubory ISO image, obrazy i texty budou v hlavním METS záznamu použity elementy <fileGrp>, jeden element <fileGrp> bude existovat vždy pro jednu skupinu souborů stejného formátu a účelu, tedy pro ISO image, archivní kopie obrazů, uživatelské kopie obrazů, další pro ALTO XML, OCR TXT soubory a další pro METS záznamy s technickými metadaty (amd_mets.xml a amd_mets_isoimg.xml)

<fileGrp> pro soubory ISO image - atributy ID="ISOIMGGRP" USE="isoimage"

- každý soubor bude mít vlastní element <file> s následujícími atributy:
 - ID - kopíruje název souboru bez přípony
 - MIMETYPE - application/x-iso9660-image
 - SIZE - velikost souboru
 - SEQ - pořadí souboru ve fileGrp
 - CHECKSUMTYPE - hodnota MD5
 - CHECKSUM - hodnota kontrolního součtu
 - CREATED - datum vytvoření, ISO8601 na úroveň vteřiny
- subelementem pod <file> je element <FLocat>, který obsahuje atribut LOCTYPE a odkaz na soubor, tj. v atributu xlink:href bude uvedena cesta ke konkrétnímu souboru (vždy absolutní vzhledem k hlavnímu adresáři balíčku).

<fileGrp> pro obrazy MC - atributy ID="MC_IMGGRP" USE="master"

- každý soubor bude mít vlastní element <file> s následujícími atributy:
 - ID - kopíruje název souboru bez přípony
 - MIMETYPE - image/jpeg2000
 - SIZE - velikost souboru
 - SEQ - pořadí souboru ve fileGrp
 - CHECKSUMTYPE - hodnota MD5
 - CHECKSUM - hodnota kontrolního součtu
 - CREATED - datum vytvoření, ISO8601 na úroveň vteřiny
- subelementem pod <file> je element <FLocat>, který obsahuje atribut LOCTYPE a odkaz na soubor, tj. v atributu xlink:href bude uvedena cesta ke konkrétnímu souboru (vždy absolutní vzhledem k hlavnímu adresáři balíčku).

<fileGrp> pro obrazy UC - atributy ID="UC_IMGGRP" USE="user"

- každý soubor bude mít vlastní element <file> s následujícími atributy:
 - ID - kopíruje název souboru bez přípony
 - MIMETYPE - image/jpeg2000
 - SIZE - velikost souboru
 - SEQ - pořadí souboru ve fileGrp
 - CHECKSUMTYPE - hodnota MD5
 - CHECKSUM - hodnota kontrolního součtu
 - CREATED - datum vytvoření, ISO8601 na úroveň vteřiny
- subelementem pod <file> je element <FLocat>, který obsahuje atribut LOCTYPE a odkaz na soubor, tj. v atributu xlink:href bude uvedena cesta ke konkrétnímu souboru (vždy absolutní vzhledem k hlavnímu adresáři balíčku).

<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 – kopíruje název souboru bez přípony
 - MIMETYPE – text/xml
 - SIZE – velikost souboru xml
 - SEQ - pořadí souboru ve fileGrp
 - CHECKSUMTYPE – hodnota MD5
 - CHECKSUM - hodnota kontrolního součtu
 - CREATED - datum vytvoření, ISO8601 na úroveň vteřiny
- subelementem pod <file> je element <FLocat>, který obsahuje atribut LOCTYPE a odkaz na soubor, tj. v atributu xlink:href bude uvedena cesta ke konkrétnímu souboru (vždy absolutní vzhledem k hlavnímu adresáři balíčku).

<fileGrp> pro soubory METS s technickými metadaty amd_mets.xml a amd_mets_isoimg.xml bude mít následující atributy: 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; kopíruje název souboru bez přípony
 - 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 atribut LOCTYPE a odkaz na soubor, tj. v atributu xlink:href bude uvedena cesta ke konkrétnímu souboru (vždy absolutní vzhledem k hlavnímu adresáři balíčku).

<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; kopíruje název souboru bez přípony
 - MIMETYPE – text/plain
 - SIZE - velikost souboru
 - SEQ - pořadí souboru ve fileGrp
 - CHECKSUMTYPE – hodnota MD5
 - CHECKSUM - hodnota kontrolního součtu
 - CREATED - datum vytvoření, ISO8601 na úroveň vteřiny
- subelementem pod <file> je element <FLocat>, který obsahuje atribut LOCTYPE a odkaz na soubor, tj. v atributu xlink:href bude uvedena cesta ke konkrétnímu souboru (vždy absolutní vzhledem k hlavnímu adresáři balíčku).

7.5.2 <fileSec> vedlejšího záznamu amd_mets.xml, resp. amd_mets_isoimg.xml

- <fileSec> ve vedleším METS záznamu amd_mets.xml nebo amd_mets_isoimg.xml bude obsahovat jeden element <fileGrp> s vnořenými elementy <file> pro každou reprezentaci stránky nebo souboru ISO image, tj. ISO image, MC pro obraz, příp. ALTO XML a OCR TXT
- atributy jednotlivých <file> elementů odpovídají atributům pro jednotlivé typy dokumentů uvedených výše pro <fileSec> hlavního METS záznamu
- navíc u MC pro obraz 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 mapa zaznamenává zejména vazby na fyzické soubory, ze kterých se skládají jednotlivé úrovně dokumentu
- logická mapa zaznamenává hierarchickou strukturu dokumentu bez vazeb na fyzické soubory; popisuje jeden datový disk (kolekci dat) na úroveň disku a obrazové přílohy
- vazby mezi konkrétní kolekcí dat a disky či konkrétní přílohou a obrázky jsou popsány v METS sekci <structLink>
- strukturální mapa logická i fyzická bude v hlavním záznamu mets.xml

7.6.1.1 Vyjádření fyzické strukturální mapy

- fyzická strukturální mapa bude mít následující atributy: LABEL="Physical_Structure", TYPE="PHYSICAL"
- obsahuje element <div> první úrovně, který popisuje nejvyšší úroveň kolekce dat; v něm jsou dále vnořeny další elementy <div> druhé úrovně, které odkazují na všechny soubory patřící do datového dokumentu (ISO image, obrázky, alto, ocr)
- **<div> první úrovně pro kolekci dat** obsahuje atributy:
 - ID - identifikátor <div> této úrovně (např. DIV_0000)
 - LABEL - název datového disku (nejvyšší úroveň)
 - TYPE - datacollection
 - DMDID - odkazuje na část bibliografických metadat náležejících k datovému dokumentu
 - ADMID - identifikátor amdsec, ve kterém jsou obsažena (nepovinná) autorsko-právní metadata
- **<div> druhé úrovně pro disk** obsahuje tyto atributy:
 - ID - identifikátor <div> této úrovně (např. DIV_DISC_0000)
 - TYPE - typ objektu, tj. disc
 - ORDER - reálné pořadí disku v kolekci
 - <div> druhé úrovně obsahuje vždy subelement <fptr>, ve kterém je odkaz na soubory související s diskem (tj. ISO image, amd_mets_isoimg.xml)
 - link na soubor má v elementu <fptr> atribut FILEID, který obsahuje ID souboru
- **<div> druhé úrovně se zopakuje i pro obrazové přílohy**, bude mít atributy:
 - ID - identifikátor <div> této úrovně (např. DIV_PAGE_0000)
 - TYPE - typ skenu obrazové přílohy (booklet, imgdisc nebo cover)
 - ORDER - pořadí skenu v příloze
 - ORDERLABEL - pořadové číslo stránky, jak je na ní vytištěno
 - element <fptr> s linkem na konkrétní soubory (MC, UC, OCR TXT, ALTO XML, amd_mets.xml)
 - link na soubor má v elementu <fptr> atribut FILEID, který obsahuje ID souboru
 - pro element <fptr> obsahující ALTO XML platí, že má navíc tyto atributy: BEGIN="P1", kde P1 je ID elementu <fptr> z ALTO XML souboru; a atribut BETYPE="IDREF"

7.6.1.2 Vyjádření logické strukturální mapy

- bude mít následující atributy: LABEL="Logical_Structure" TYPE="LOGICAL"
- <div> musí obsahovat identifikátor na příslušný blok metadat, př. odkaz na úroveň kolekce dat DMDID="MODS_DATACOLLECTION_0001"
- obrazové přílohy a soubory ISO image budou zanořeny přímo pod nejvyšší úroveň, tedy kolekci dat
- element <div> první úrovně – kolekce dat (datacollection)
- element <div> druhé úrovně – disk (disc)
- element <div> druhé úrovně – obrazové informace (supplement či rozdělit na booklet, cover a imgdisc)
- výčet souborů ISO image na disku a výčet obrázků v příloze bude uveden v části <structLink>

<div> type	Atributy	Popis	Povinnost
DATACOLLECTION	ID LABEL TYPE DMDID	<div> obsahuje údaje o nadřazeném datovém disku (kolekci dat), ID - identifikátor <div>, hodnota např. DATACOLLECTION_0001 LABEL - název celé kolekce TYPE - hodnota DATACOLLECTION DMDID - obsahuje identifikátor popisné části MODS k datovému dokumentu	M
DISC	ID LABEL TYPE DMDID	<div> obsahuje údaje o disku ID - identifikátor <div>, hodnota např. DISC_0001 LABEL - název disku (číslo) TYPE - hodnota DISC DMDID - obsahuje identifikátor popisné části MODS k disku	M
SUPPLEMENT	ID LABEL TYPE DMDID	<div> obsahuje údaje o obrazové příloze ID - identifikátor <div>, hodnota např. SUPPL_0001 LABEL - název přílohy - hodnoty BOOKLET, COVER, IMGDISC TYPE - hodnota SUPPLEMENT DMDID - obsahuje identifikátor popisné části MODS k příloze	M

7.6.2 <structMap> vedlejšího záznamu METS (amd_mets.xml a amd_mets_isoimg.xml)

- bude obsahovat pouze fyzickou strukturální mapu (TYPE="PHYSICAL")
- ta bude obsahovat pouze jeden <div> element s atributem TYPE="DISC" nebo TYPE="SUPPLEMENT"
- do <div> budou vnořeny odkazy na jednotlivé reprezentace dat nebo obrazové přílohy pomocí elementu <ftpr> s atributem FILEID

7.7 METS část <structLink>

METS část <structLink> bude ukazovat vztah mezi fyzickou a logickou strukturou, tedy vztah mezi jednotlivými disky a soubory ISO image a mezi jednotlivými přílohami a obrázky.

- element <structLink> obsahuje subelement <smLink> pro každý ISO image nebo obrázek
- bude mít atributy xlink:from (odkaz na disk či přílohu) a xlink:to (odkaz na ISO image či obrázek)

Př.:

```
<mets:structLink>
```

```
<mets:smLink xlink:from="DISC_0001" xlink:to="DIV_ISOIMG_0001"/>
```

```
<mets:smLink xlink:from="DISC_0002" xlink:to="DIV_ISOIMG_0002"/>
```

```
<mets:smLink xlink:from="DISC_0003" xlink:to="DIV_ISOIMG_0003"/>
```

```
<mets:smLink xlink:from="SUPPL_0001" xlink:to="DIV_PAGE_0001"/>
```

```
<mets:smLink xlink:from="SUPPL_0001" xlink:to="DIV_PAGE_0002"/>
```

```
<mets:smLink xlink:from="SUPPL_0001" xlink:to="DIV_PAGE_0003"/>
```

```
</mets:structLink>
```


7.8 OCR (ALTO XML a TXT OCR)

- plnění OCR je dobrovolné
- bude použita poslední verze formátu ALTO XML aktuální v době implementace nebo verze předchozí (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 (datovému disku) včetně prázdných stran, 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, orientaci 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 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 dělící 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
 - 1-n element je povinný, opakovatelný
 - 1-1 element je povinný, neopakovatelný

Element	Atribut	Popis	Povinnost
<code><Description></code>			
<code>x<MeasurementUnit></code>		měřící jednotka pro souřadnice v ALTO XML; možné hodnoty – dpi, pixel, inch1200 a mm10); inch1200 = 1/1200 inche; doporučené plnění je „mm10“ nebo „pixel“; 1-1	M
<code>x<sourceImageInformation></code>		informace o obrazovém souboru, ze kterého vzniklo ALTO XML; 1-1	MA
<code>xx<fileName></code>		jméno obrazového souboru, ze kterého bylo ALTO XML vytvářeno; ideálně i s filesystem cestou jeho uložení; např. n1ImageSeq-33386-b.tif//produkce/OCR/digibok_XY/XY_011.tif 1-1	M
<code>xx<fileIdentifier></code>		jedinečný identifikátor obrazového souboru; 0-n	R
<code>x<OCRProcessing></code>	ID	popis procesu vzniku OCR;	M

		1-n ----- ID OCR procesu, např. <OCRProcessing ID="OCRPROCES_1">; povinné	
xx<preProcessingStep>		procesy před vznikem OCR, které provádí SW pro OCR (např. natočení obrazu) 1-n	MA
xxx<processingDateTime>		určení času procesu, který předcházal samotnému OCR; např. 2008-03-29T19:42:23 dle ISO 8601 na úrovni vteřin; 1-1	O
xxx<processingAgency>		jméno nebo kód instituce, např. MZK, název externí firmy apod.; doporučujeme použít kontrolovaný slovník hodnot; 0-1	R
xxx<processingStepDescription>		popis procesu (např. zarovnání, ořez apod.); 0-n	O
xxx<processingStepSettings>		nastavení kroku popsaného v <processingStepDescription>, např. CCS OCR Processing Filter 0-1	O
xxx<processingSoftware>		popis SW, který upravoval obrázek před vznikem OCR; 1-1	MA
xxxx<softwareCreator>		výrobce softwaru - např. CCS Content Conversion Specialists GmbH, Germany; 1-1	M
xxxx<softwareName>		jméno softwaru - např. CCS docWORKS; 1-1	M
xxxx<softwareVersion>		verze SW, např. 6.2-1.16; 1-1	M
<ocrProcessingStep>		popis procesu vzniku OCR 1-1 – povinné pole	M
xxx<processingDateTime>		okamžik kdy bylo OCR vytvořeno; nutno zapsat v ISO 8601 na úroveň vteřin; 0-1	M
xxx<processingAgency>		jméno nebo kód instituce, např. NK CZ doporučujeme použít kontrolovaný slovník hodnot; 0-1	M
xxx<processingSoftware>		popis SW, který dělal vlastní OCR; 1-1	M
xxxx<softwareCreator>		výrobce softwaru - např. ABBYY, Russia;	M

		1-1	
xxxx<softwareName>		jméno softwaru - např. FineReader; 1-1	M
xxxx<softwareVersion>		např. 8.0; 1-1	M
<Styles>		styly definují vlastnosti jednotlivých grafických prvků stránky. styl definovaný v elementu vrchní úrovně je použit jako výchozí pro podřízené elementy; 1-1	MA
x<TextStyle>	ID FONTSTYLE FONTFAMILY FONTSIZE	definuje font textu; 1-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
x<ParagraphStyle>	ID ALIGN	definuje formátování textových bloků; 1-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
x<Page>	ID ACCURACY POSITION QUALITY PHYSICAL_IM G_NR HEIGHT	element popisující jednu stránku dokumentu; 1-n ----- ID – vygenerovaný identifikátor stránky, např. PAGE1, nebo P1 apod.; povinné	M

	WIDTH PC	<p>ACCURACY – procentuální odhad přesnosti OCR (0-100); doporučené</p> <p>POSITION – pozice stránky; hodnoty k plnění: Left, Right, Foldout, Single, Cover; nepovinné</p> <p>QUALITY – krátký údaj o kvalitě předlohy stránky; hodnoty k plnění: OK, Missing, Missing in original, Damaged, Retained, Target, As in original; nepovinné</p> <p>PHYSICAL_IMG_NR - fyzické (pořadové) číslo stránky v dokumentu; vyjádřeno číslem, např. 1,2,3 apod.; povinné</p> <p>WIDTH – šířka stránky vyjádřená v pixelech; povinné</p> <p>HEIGHT – výška stránky vyjádřená v pixelech; povinné</p> <p>PC - Confidence level OCR souboru – hodnota mezi 0 (nejistá kvalita) a 1 (dobrá kvalita); nepovinné; pokud nevyplníte ACCURACY – tak je vyplnění doporučené</p>	
xx<TopMargin>	ID HPOS VPOS WIDTH HEIGHT	<p>horní okraj – prostor mezi vrchní hranou listu a vrchní linkou textu; 0-1</p> <p>-----</p> <p>ID: unikátní ID pro element <TopMargin>, např. P1_TM0001 (page 1, topMargin0001); povinné</p> <p>HPOS: horizontální pozice; povinné</p> <p>VPOS: vertikální pozice; povinné</p> <p>WIDTH – šířka vrchního okraje; povinné</p> <p>HEIGHT – výška vrchního okraje; povinné</p>	M
xxx<TextBlock>		stejné plnění a pravidla jako pro element <TextBlock> vnořený do elementu <PrintSpace>	MA
xx<LeftMargin>	ID HPOS	<p>levý okraj – prostor mezi levým okrajem stránky a textem; 0-1</p>	M

	VPOS WIDTH HEIGHT	----- 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	
xxx<TextBlock>		stejné plnění a pravidla jako pro element <TextBlock> vnořený do elementu <PrintSpace>	MA
xx<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
xxx<TextBlock>		stejné plnění a pravidla jako pro element <TextBlock> vnořený do elementu <PrintSpace>	MA
xx<BottomMargin>	ID HPOS VPOS WIDTH HEIGHT	pravý okraj – prostor mezi spodním okrajem stránky a textem; 1-1 ----- ID: unikátní ID pro element <BottomMargin>, např. P1_BM0001 (page 1, bottomMargin0001); povinné HPOS: horizontální pozice;	M

		<p>povinné</p> <p>VPOS: vertikální pozice; povinné</p> <p>WIDTH – šířka spodního okraje; povinné</p> <p>HEIGHT – výška spodního okraje; povinné</p>	
xxx<TextBlock>		stejné plnění a pravidla jako pro element <TextBlock> vnořený do elementu <PrintSpace>	MA
xx<PrintSpace>	<p>ID</p> <p>HPOS</p> <p>VPOS</p> <p>WIDTH</p> <p>HEIGHT</p>	<p>popis tvaru pokrývajícího textové pole stránky; 1-1</p> <p>-----</p> <p>ID: unikátní ID pro element <printSpace>, 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
xxx<TextBlock>	<p>ID</p> <p>STYLEREFS</p> <p>HPOS</p> <p>VPOS</p> <p>WIDTH</p> <p>HEIGHT</p>	<p>popisy textových bloků na konkrétní stránce; 1-n</p> <p>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 <ParagraphStyle>; povinné</p> <p>HPOS: horizontální pozice bloku; povinné</p> <p>VPOS: vertikální pozice bloku; povinné</p>	MA

		<p>WIDTH – šířka textového bloku; povinné</p> <p>HEIGHT – výška textového bloku; povinné</p>	
xxxx<Shape>		<p> tvar textového bloku; 0-1 – pro jeden výskyt <TextBlock> jeden nebo žádný výskyt <Shape>; plnit v případě, že je tvar textového bloku nestandardní (víceúhelník)</p>	RA
xxxxx<Polygon>	POINTS	<p> popis (souřadnice) tvaru víceúhelníku; 0-1 ----- POINTS – vyjádření jednotlivých bodů víceúhelníku; povinné</p>	M
xxxx<TextLine>	ID STYLEREFS HPOS VPOS WIDTH HEIGHT	<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 <ParagraphStyle>; 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
xxxxx<String>	ID CONTENT HEIGHT WIDTH HPOS VPOS CC WC V případě	<p> řetězec znaků – vlastní obsah OCR; znaky tvoří jednotlivá slova a více tagů <String> větu <TextLine>; 1-n v rámci <TextLine> ----- ID obsahuje unikátní sekvenční číslo řetězce na stránce, např. "P3_ST0001" (strana 3, řetězec 1); povinné</p> <p>CONTENT – ukládá vlastní řetězec znaků (slovo);</p>	Mpart

	<p>dělení slov také: SUBS_TYPE SUBS-CONTENT</p>	<p>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 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 <CONTENT> ve větě (stringu) první; HypPart2 se vyskytuje u následujícího tagu <CONTENT> 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>	
xxxxxx<ALTERNATIVE>		<p>alternativní hodnota OCR řetězce pro jednotlivá slova; 0-n lze použít v případě nejistoty rozpoznání řetězce;</p>	0

xxxxx<HYP>	CONTENT WIDTH HPOS VPOS	<p>zápis znaku rozdělovníku slov 0-1 pro jeden výskyt <TextLine>; vždy pro poslední <String>; 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
xxxxx<SP>	ID WIDTH HPOS VPOS	<p>prázdný prostor mezi řádky; 0-n v rámci jednoho <TextLine>; vždy mezi řádky, tj. mezi tagy <String>;</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
xxx<ComposedBlock>	ID TYPE HPOS VPOS WIDTH HEIGHT STYLEREFS	<p>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</p> <p>-----</p> <p>ID: unikátní ID komponovaný blok, např. P6_CB0001 (stránka 6, komponovaný blok 0001); povinné</p>	MA

		<p>TYPE – označení typu komponovaného bloku; nutné používat kontrolovaný slovník (illustration, Advertisement, apod.); povinné</p> <p>HPOS: horizontální pozice bloku; povinné</p> <p>VPOS: vertikální pozice bloku; povinné</p> <p>WIDTH – šířka komponovaného bloku; povinné</p> <p>HEIGHT – výška komponovaného bloku; povinné</p>	
xxxx<Shape>		<p> 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)</p>	RA
xxxxx<Polygon>	POINTS	<p> popis tvaru víceúhelníku; 1-1 ----- POINTS – vyjádření jednotlivých bodů víceúhelníku povinné</p>	M
xxxx<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 <ComposedBlock> 0 nebo více elementů /alto/Layout/Page/PrintSpace/ComposedBlock/TextBlock >; plnit pokud je v komponovaném bloku text ----- 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;</p>	MA

		<p>povinné</p> <p>VPOS: vertikální pozice bloku; povinné</p> <p>WIDTH – šířka textového bloku; povinné</p> <p>HEIGHT – výška textového bloku; povinné</p>	
xxxxx<TextLine>		<p>/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</p>	
xxxx<GraphicalElement>	<p>ID HPOS VPOS WIDTH HEIGHT</p>	<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 <GraphicalElement>; 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

7.9 Autorsko-právní metadata

- autorsko-právní metadata jsou kompletně nepovinná
- 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.xml`, resp. `amd_mets_isoimg.xml`); 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 hlavního záznamu METS; naopak to neplatí - hlavní záznam nepřebírá autorsko-právní metadata z vedlejších záznamů
- přítomnost autorsko-právních metadat je třeba zohlednit ve fyzické strukturální mapě

Opakovatelnost elementů:

- 0-1 – nepovinný a neopakovatelný
- 0-n – nepovinný a opakovatelný

Element	Atributy	Popis	Povinnost
<amdSec>		element obsahující autorsko-právní metadata ve formátu CopyrightMD	M
	ID	použití jak u hlavního METS, tak vedlejšího záznamu <code>amd_mets</code> , resp. <code>amd_mets_isoimg.xml</code> ; identifikátor konkrétní části <amdSec>	MA
x<rightsMD>		element pro typ autorsko-právních metadat	M
	ID	ID pro část <rightsMD>: "RIGHTS_001", "RIGHTS_002" atd.	M
xx<mdWrap>			M
	MDTYPE	hodnota "OTHER"	M
	OTHERMDTYPE	hodnota "CopyrightMD"	M
	MIMETYPE	hodnota "text/xml"	M
xxx<xmlData>		kontejnerový element, do kterého se vkládají záznamy CopyrightMD	M

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 list⁶) (viz příklady v následující tabulce)
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; povinné jsou i oba atributy:</p> <p>copyright.status</p> <ul style="list-style-type: none"> • copyrighted (autorsky chráněný document) • pd (volně dostupný dokument) • pd_expired (volně dostupný dokument, kterému vypršelo trvání majetkových práv) • pd_holder (dokument dedikovaný k volnému užití majitelem; možno využívat např. v případě licencí Creative Commons v kombinaci s elementem <general.note>, ve kterém je upřesněn konkrétní druh licence) • unknown (status neznámý) <p>publication.status</p> <ul style="list-style-type: none"> • published (publikovaný dokument) • unpublished (nepublikovaný dokument) • unknown (informace o publikování není známa) <p>0-1</p>	M
x<creation>	<p>Informace o vytvoření dokumentu.</p> <p>0-1</p>	O
xx<year.creation>	<p>Rok vytvoření dokumentu ve formátu RRRR.</p> <p>0-1</p>	RA
xx<country.creation>	<p>Země, ve které byl dokument vytvořen: vyjádřeno iso kódem (viz Nepovinné atributy) př.: <country.creation iso.code="cz">Česká republika</country.creation></p> <p>0-1</p>	RA
x<creator>	<p>Informace o tvůrci dokumentu.</p> <p>0-1</p>	MA
xx<creator.corporate>	<p>Název instituce, pokud je autorem dokumentu.</p> <p>0-n</p>	MA
xx<creator.person>	<p>Informace o autorovi dokumentu.</p> <p>0-n</p>	MA

⁶ Dostupné z: <http://www.iso.org/iso/en/prods-services/iso3166ma/02iso-3166-code-lists/list-en1.html>

xxx<name>	Jméno a příjmení autora. 0-1	MA
xxx<year.birth>	Datum narození autora. Formát RRRR. 0-1	MA
xxx<year.death>	Datum úmrtí autora. Formát RRRR. 0-1	MA
xx<note>	Doplňující informace k tvůrci dokumentu. 0-n	O
x<publication>	Informace o vydání dokumentu 0-1	MA
xx<country.publication>	Země vydání – vyjádřeno iso kódem (viz Nepovinné atributy), př: <country.publication iso.code="cz">Česká republika</copyrightMD.country.publication> 0-1	R
xx<publisher>	Nakladatel. 0-1	R
xx<year.publication>	Rok vydání. Formát RRRR. 0-1	MA
xx<year.copyright>	Rok copyrightu. Formát RRRR. 0-1	RA
xx<year.renewal>	Rok případného obnovení copyrightu. Formát RRRR. 0-1	RA
xx<note>	Doplňující informace k vydání dokumentu. 0-n	O
x<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
xx<contact>	Kontakt na držitele práv. 0-n	O
xx<name>	Jméno nebo název držitele práv. 0-1	O
xx<note>	Doplňující informace o držiteli práv. 0-n	O
x<notice>	Do pole se vkládá oznámení o copyrightu, tak jak je uvedeno v dokumentu. 0-1	O
x<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.	O

	<p>Toto pole je doporučeno užívat v případě, že je dokument pod licencí Creative Commons a je potřeba upřesnit konkrétní druh použité licence.</p> <p>Možné hodnoty:</p> <ul style="list-style-type: none">- BY (uvedte původ)- BY-SA (uvedte původ – zachovejte licenci)- BY-ND (uvedte původ – nezpracovávejte)- BY-NC (uvedte původ – neužívejte komerčně)- BY-NC-SA (uvedte původ – neužívejte komerčně – zachovejte licenci)- BY-NC-ND (uvedte původ – neužívejte komerčně – nezpracovávejte) <p>0-n</p>	
--	--	--

Historie verzí

Jméno	Datum	Verze	Provedené změny
Květa Fremrová, Veronika Ježková, Pavlína Kočišová, Vojtěch Kopský, Lenka Kravecová, Filip Pavčík	13. listopadu 2024	0.1	- první oficiální draft